

ASTES

Advances in Science, Technology & Engineering Systems Journal

Special Issue

Multidisciplinary Sciences and Engineering

2020-21

www.astesj.com
ISSN: 2415-6698

EDITORIAL BOARD
(Special Issue)

Editor-in-Chief

Prof. Passerini Kazmerski
Pritzker School of Molecular Engineering, University of Chicago, USA

Guest Editors

Prof. Wang Xiu Ying
Chongqing University, China

Prof. Yu Xiao Yan
Chongqing Normal University, China

Prof. María Jesús Espinosa Trujillo
Universidad Tecnológica Metropolitana, Mexico

Prof. Ahmad Yusairi Bani Hashim
Universiti Teknikal Malaysia Melaka, Malaysia

Prof. Mohamed Abdelaziz Hassan Eleiwa
University of Hail, KSA

Prof. Nicolae Tudoroiu
John Abbott College, Canada

Editorial

The Special Issue on Multidisciplinary Sciences and Engineering (2020–21) in the *Advances in Science, Technology and Engineering Systems Journal (ASTES Journal)* reflects a period marked by both disruption and accelerated innovation. As global challenges grew increasingly complex spanning public health, environmental sustainability, digital transformation, and resilient infrastructure the need for integrative approaches across scientific and engineering domains became more evident than ever. This issue brings together a diverse collection of contributions that demonstrate how interdisciplinary thinking not only enriches academic inquiry but also drives practical solutions to real-world problems.

A defining feature of this special issue is its breadth. The selected papers traverse traditional disciplinary boundaries, combining insights from fields such as computer science, mechanical and electrical engineering, materials science, environmental studies, and applied mathematics. Several contributions highlight the growing role of data-driven methodologies, including artificial intelligence and machine learning, in optimizing systems and enhancing predictive capabilities. Others emphasize sustainable design principles, renewable energy technologies, and resource-efficient processes areas that are critical for long-term societal resilience. The convergence of these perspectives illustrates the strength of multidisciplinary collaboration in addressing challenges that no single field can adequately resolve in isolation.

Equally noteworthy is the methodological diversity present in this issue. Authors employ a range of approaches, from theoretical modeling and simulation to experimental validation and case study analysis. This combination of rigor and applicability underscores the journal's commitment to bridging the gap between theory and practice. In particular, several studies offer scalable solutions and frameworks that can be adapted across different contexts, reinforcing the value of engineering innovation as a driver of economic and social progress.

The 2020–21 timeframe also provides an important backdrop for this collection. The global pandemic served as both a challenge and a catalyst, prompting rapid advancements in digital systems, remote sensing, healthcare technologies, and resilient supply chains. Many of the works included here reflect this shift, either directly or indirectly, by proposing technologies and strategies that enhance adaptability and robustness in uncertain environments. This responsiveness to contemporary needs highlights the dynamic nature of multidisciplinary research and its capacity to evolve alongside emerging global priorities.

The editorial team extends its sincere appreciation to the authors for their contributions, as well as to the reviewers whose thoughtful evaluations ensured the quality and integrity of the published work. The collaborative effort behind this special issue exemplifies the academic community's dedication to advancing knowledge and fostering innovation, even under challenging circumstances.

This special issue stands as a testament to the transformative potential of multidisciplinary sciences and engineering. By integrating diverse perspectives, methodologies, and applications, it not only advances scholarly discourse but also provides meaningful pathways toward addressing some of the most pressing issues of our time.

Guest Editor

Prof. Ahmad Yusairi Bani Hashim

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Special Issue **August 2020**

CONTENTS

Environmental Acoustics Modelling Techniques for Forest Monitoring
by Svetlana Segarceanu, George Suciu and Inge Gavăt

Learning the Influence between Partially Observable Processes using Scorebased Structure Learning
by Ritesh Ajoodha and Benjamin Rosman

The Role of RFID in Green IoT: A Survey on Technologies, Challenges and a Way Forward
by Zainatul Yushaniza Mohamed Yusoff, Mohamad Khairi Ishak and Kamal Ali Alezabi

Electroencephalogram Based Medical Biometrics using Machine Learning: Assessment of Different Color Stimuli
by Md Mahmudul Hasan, Nafiul Hasan, Dil Afroz, Ferdaus Anam Jibon, Md. Arman Hossen, Md. Shahrier Parvage and Jakaria Sulaiman Aongkon

Curved Pyramidal Metamaterial Absorber: From Theory to an Ultra-Broadband Application in the [0.3 – 30] GHz Frequency Band
by Zeinab Fneish, Hussam Ayad, Moncef Kadi, Jalal Jomaa and Ghaleb Faour

Mechanical Characterization of Recycled Aggregates Concrete Based on its Compressive Strength
by Khaoula Naouaoui and Toufik Cherradi

An Anonymity Preserving Framework for Associating Personally Identifying Information with a Digital Wallet
by Qazi Mudassar Ilyas and Muhammad Mehboob Yasin

Super Resolution Based Underwater Image Enhancement by Illumination Adjustment and Color Correction with Fusion Technique
by Md. Ashfaqul Islam, Maisha Hasnin, Nayeem Iftakhar and Md. Mushfiqur Rahman

Switching Capability of Air Insulated High Voltage Disconnectors by Active Add-On Features
by Mariusz Rohmann and Dirk Schräder

The Analysis of Standard Uncertainty of Six Degree of Freedom (DOF) Robot
by Auttapoom Loungthongkam and Chana Raksiri

Green Blocks Made of Recycled Construction Waste using Recycled Wastewater
by Elgaali Elgaali and Adel Al Wazeer

Power Converters and EMS for Fuel Cells CCHP Applications: A Structural and Extended Review
by Nganyang Paul Bayendang, Mohamed Tariq Kahn and Vipin Balyan

Methodology for Calculating Shock Loads on the Human Foot
by Valentyn Tsapenko, Mykola Tereschenko, Vadim Shevchenko and Ruslan Ivanenko

Evaluation the Effects of Climate Change on the Flow of the Arkansas River – United States
by Elgaali Elgaali and Zeyad Tarawneh

Fuzzy Analytical Hierarchy Process and Fuzzy Comprehensive Evaluation Method Applied to Assess and Improve Human and Organizational Factors Maturity in Mining Industry
by Yousra Karim and Abdelghani Cherkaoui

SLIP-SL: Walking Control Based on an Extended SLIP Model with Swing Leg Dynamics
by Junho Chang, Mustafa Melih Pelit and Masaki Yamakita

Design and Implementation of an Ultrasonic Scanner Setup that is Controlled using MATLAB and a Microcontroller
by Kamel Fahmi Bou-Hamdan

New Neural Networks for the Affinity Functions of Binary Images with Binary and Bipolar Components Determining
by Valerii Dmitrienko, Serhii Leonov and Aleksandr Zakovorotniy

Designing and Applying a Moral Turing Test
by Hyeongjoo Kim and Sunyong Byun

Challenges in IoT Technology Adoption into Information System Security Management of Smart Cities: A Review
by Zarina Din, Dian Indrayani Jambari, Maryati Mohd Yusof and Jamaiah Yahaya

Web-based Remote Lab System for Instrumentation and Electronic Learning
by Jose María Sierra-Fernández, Agustin Agüera-Pérez, Jose Carlos Palomares-Salas, Manuel Jesús Espinosa-Gavira, Olivia Florancias-Oliveros and Juan José González de la Rosa

Spatial Multi-Layer Perceptron Model for Predicting Dengue Fever Outbreaks in Surabaya
by Siana Halim, Andreas Handojo, Ivan Enrico Widodo, Felecia and Tanti Octavia

Technology Adoption in Education: A Systematic Literature Review
by Kayode Emmanuel Oyetade, Tranos Zuva and Anneke Harmse

Analyzing the Adoption of E-payment Services in Smart Cities using Demographic Analytics: The Case of Dubai
by Raed Said, Anas Najdawi and Zakariya Chabani
A Model-Driven Digital Twin Framework Development for Sulfur Dioxide Conversion Units Simulation
by Amine Mounaam, Ridouane Oulhiq, Ahmed Souissi, Mohamed Salouhi, Khalid Benjelloun and Ahmed Bichri

A Novel Approach for Estimating the Service Lifetime of Transformers within Distributed Solar Photovoltaic (DSPV) Systems
by Bonginkosi Allen Thango, Jacobus Andries Jordaan and Agha Francis Nnachi

Peculiar Stray Gassing Occurrences in Solar Photovoltaic Transformers during Service
by Bonginkosi Thango, Jacobus Jordaan and Agha Nnachi

Application of a Reusability Approach in Simulation of Heritage Buildings Performance, Taif-Saudi Arabia
by Ali Alzaed

A Review on TAM and TOE Framework Progression and How These Models Integrate
by Julies David Bryan and Tranos Zuva

Visual Saliency Detection using Seam and Color Cues
by Sk. Md. Masudul Ahsan and Aminul Islam

Synchronization in a Class of Fractional-order Chaotic Systems via Feedback Controllers: A Comparative Study
by Juan Luis Mata-Machuca

Predicting Student Academic Performance Using Data Mining Techniques
by Lonia Masangu, Ashwini Jadhav and Ritesh Ajoodha

Multi-Objective Design of Current Conveyor using Optimization Algorithms
by Abdelaziz Lberni, Malika Alami Marktani, Abdelaziz Ahaitouf and Ali Ahaitouf

Inferring Topics within Social Networking Big Data, Towards an Alternative for Socio-Political Measurement
by Khalid Ait Hadi, Rafik Lasri and Abdellatif El Abderrahmani

Realization and Energy Optimization of a Recharging Station for Electric Vehicles with Fixed Storage and Photovoltaic Panels
by David Roszczypala, Christophe Batard, Nicolas Ginot and Frédéric Poitiers

A Large Empirical Study on Automatically Classifying Software Maintainability Concerns from Issue Summaries
by Celia Chen and Michael Shoga

Lifestyle in Nursing Students at a University of North Lima
by Yanet Cruz Flores, Tania Retuerto-Azaña, Jaquelin Nuñez-Artica, Brian Meneses-Claudio, Hernan Matta Solis and Lourdes Matta-Zamudio

Autonomous Robot Path Construction Prototype Using Wireless Sensor Networks
by José Paulo de Almeida Amaro, João Manuel Leitão Pires Caldeira, Vasco Nuno da Gama de Jesus Soares and João Alfredo Fazendeiro Fernandes Dias

Applications of the Heuristic Optimization Approach for Determining a Maximum Flow Problem Based on the Graphs' Theory

by Simona Kirilova Filipova-Petrakieva

Multi-Robot System Architecture Design in SysML and BPMN
by Ahmed R. Sadik and Christian Goerick

Level of Empathy in Nursing Students Attending Clinical Practices of the Universidad de Ciencias y Humanidades
by Walter Cervera-Flores, Yenifer Choque-Garibay, Nahuel Gonzalez-Cordero, Brian Meneses-Claudio, Hernan Solis-Matta and Lourdes Matta-Zamudio

Comparison of Learning Style for Engineering and Non-Engineering Students
by Mimi Mohaffyza, Jailani Md Yunos, Yee Mei Heong, Junita, Fahmi Rizal and Badaruddin Ibrahim

QoE-aware Bandwidth Allocation for Multiple Video Streaming Players over HTTP and SDN
by Pham Hong Thinh, Tran Thi Thanh Huyen, Nguyen Ngoc Quang, Pham Ngoc Nam, Truong Cong Thang, Nguyen Viet Hung and Truong Thu Huong

PrOMor: A Proposed Prototype of V2V and V2I for Crash Prevention in the Moroccan Case
by Zakaria Sabir and Aouatif Amine

Effectiveness and Suitability of the Automotive EHPS Software Reliability and Testing
by Yanshuo Wang, Jim (Jinming) Yang and Ngandu M. Mbiye

Heuristic Techniques as Part of Heuristic Methods and Interaction of Personality Types in their Application
by Viktor Ivanov, Lubomir Dimitrov, Svitlana Ivanova and Olena Olefir

XMDD as Key Enabling Technology for Integration and Organizational Collaboration: Application to E-Learning Based on NRENs
by Salim Saay and Tiziana Margaria

Analysis of Real-time Blockchain Considering Service Level Agreement (SLA)
by Minkyung Kim, Kangseok Kim and Jai-Hoon Kim

Vehicle Rollover Detection in Tripped and Untripped Rollovers using Recurrent Neural Networks
by Kailerk Treetipsounthorn, Thanisorn Sriudomporn, Gridsada Phanomchoeng, Christian Dengler, Setha Panngum, Sunhapos Chantranuwathana and Ali Zemouche

FPGA Acceleration of Tree-based Learning Algorithms
by Haytham Azmi

Hybridization of Improved Binary Bat Algorithm for Optimizing Targeted Offers Problem in Direct Marketing Campaigns
by Moulay Youssef Smaili and Hanaa Hachimi

Estimation of Land Degradation Loss by Water Erosion: Case of the Site of Biological and Ecological Interest of Ain Asmama (Western High Atlas, Morocco)

by Adnane Labbaci, Said Moukrim, Said Lahssini, Said Laaribya, Hicham Mharzi Alaoui and Jamal Hallam

Parametric Study for the Design of a Neutron Radiography Camera-Based Detector System
by Evens Tebogo Moraba, Tranos Zuva, Chunling Du and Deon Marais

Microstrip Patch Antenna Designs with Quarter-Circular and Semi-Circular Slots in Patches for Wireless Communication Applications in Frequency Range of 1.2 GHz-4.6 GHz
by Cihan Dogusgen Erbas

A Fuzzy Controller Based SAPF for Power Quality Enhancement of Distribution System Integrated with Wind Energy Source
by Vikas Kumar Sharma and Lata Gidwani

Pre-University Students' Learning Styles and Attitude towards Mathematics Achievements
by Nurhilyana Anuar, Nurashikin Abdullah and Sharifah Norasikin Syed Hod

Mathematical Modelling of Output Responses and Performance Variations of an Education System due to Changes in Input Parameters
by Najat Messaoudi, Jaafar Khalid Naciri and Bahloul Bensassi

Design and Implementation of Reconfigurable Neuro-Inspired Computing Model on a FPGA
by Basutkar Umamaheshwar Venkata Prashanth and Mohammed Riyaz Ahmed

Recent Impediments in Deploying IPv6
by Ala Hamarsheh, Yazan Abdalaziz and Shadi Nashwan

Enhanced Power Utilization for Grid Resource Providers
by Tariq Alwada'n, Thair Khdour, Abdulsalam Alarabeyyat and Ali Rodan

Study of latencies in ThingSpeak
by Vítor Viegas, J. M. Dias Pereira, Pedro Girão and Octavian Postolache

Deep Learning based Models for Solar Energy Prediction
by Imane Jebli, Fatima-Zahra Belouadha, Mohammed Issam Kabbaj and Amine Tilioua

Novel Infrastructure Platform for a Flexible and Convertible Manufacturing
by Javier Stillig and Nejila Parspour

Vietnamese Text Classification with TextRank and Jaccard Similarity Coefficient
by Hao Tuan Huynh, Nghia Duong-Trung, Dinh Quoc Truong and Hiep Xuan Huynh

Data Aggregation, Gathering and Gossiping in Duty-Cycled Multihop Wireless Sensor Networks subject to Physical Interference
by Lixin Wang, Jianhua Yang, Sean Gill and Xiaohua Xu

Decentralized Management System for Solid-State Voltage Regulators in Nodes of Distribution Power Networks
by Igor Polozov, Elena Sosnina, Vladimir Kombarov and Ivan Lipuzhin

Smart Collar and Chest Strap Design for Rescue Dog through Multidisciplinary Approach
by Fang-Lin Chao, Wei Zhong Feng and Kaiquan Shi

Method for Improving the Quality of the Product Obtained by Abrasive Treatment with Impregnated Tools
by Viktor Butenko, Liana Gusakova, Dmitry Durov, Boris Safoklov and Oleg Dolgov

Level of Resilience and Family Functionality in Adolescents of two Educational Institutions of a Vulnerable Area in Lima Province
by Rosa Perez-Siguas, Hernan Matta-Solis, Eduardo Matta-Solis, Melissa Yauri-Machaca, Anika Remuzgo-Artezano and Lourdes Matta-Zamudio

Assessing the Operator's Readiness to Perform Tasks of Controlling by the Unmanned Aerial Platforms
by Dmytro Kucherov, Olha Sushchenko, Andrii Kozub and Volodymyr Nakonechnyi

Application of Polynomial Regression Analysis in Evaluating the Techno-Economic Performance of DSPV Transformers
by Bonginkosi Allen Thango, Jacobus Andries Jordaan and Agha Francis Nnachi

Design Approach of an Electric Single-Seat Vehicle with ABS and TCS for Autonomous Driving Based on Q-Learning Algorithm
by Jason Valera and Sebastian Herrera

Interpretation of Machine Learning Models for Medical Diagnosis
by Nghia Duong-Trung, Nga Quynh Thi Tang and Xuan Son Ha

Teaching/Learning Strategies in Context of Education 4.0
by Irina Golitsyna, Farid Eminov and Bulat Eminov

A Novel Approach for Evaluating Eddy Current Loss in Wind Turbine Generator Step-Up Transformers
by Bonginkosi Allen Thango, Jacobus Andries Jordaan and Agha Francis Nnachi

Open Access Research Trends in Higher Education: A Literature Review
by Mariutsi Alexandra Osorio-Sanabria, Astrid Jaime, Tamara Alcantara-Concepcion and Piedad Barreto

Simulated Annealing for Traveling Salesman Problem with Hotel Selection for a Distribution Company Based in Mexico
by Raúl Jiménez-Gutiérrez, Diana Sánchez-Partida, José-Luis Martínez-Flores and Eduardo-Arturo Garzón-Garnica

Microcontroller Based Data Acquisition and System Identification of a DC Servo Motor Using ARX, ARMAX, OE, and BJ Models
by Mokhlis Salah-eddine, Said Sadki and Bahloul Bensassi

Practical Simulation of Grounded/Floating Lossy Inductors Based on Commercially Available Integrated Circuit LT1228
by Tattaya Pukkalanun, Pitchayanin Moonmuang, Sumalee Unhavanich and Worapong Tangsrirat

Recycling and Reuse of Wastewater Generated in Car-Washing Facilities
by Elgaali Elgaali and Majid Akram

A Model for the Application of Automatic Speech Recognition for Generating Lesson Summaries
by Phillip Blunt and Bertram Haskins

Complex Order PIa+jbDc+jd Controller Design for a Fractional Order DC Motor System
by Pritesh Shah, Ravi Sekhar, Iswanto Iswanto and Margi Shah

Using Supervised Classification Methods for the Analysis of Multi-spectral Signatures of Rice Varieties in Panama
by Javier E. Sánchez-Galán, Fatima Rangel Barranco, Jorge Serrano Reyes, Evelyn I. Quirós-McIntire, José Ulises Jiménez and José R. Fábrega

Homology Modeling of CYP6Z3 Protein of Anopheles Mosquito
by Marion Olubunmi Adebiyi and Oludayo Olufolorunsho Olugbara

Neural Networks and Fuzzy Logic Based Maximum Power Point Tracking Control for Wind Energy Conversion System
by Hayat El Aissaoui, Abdelghani El Ougli and Belkassem Tidhaf

A Framework for the Alignment of ICT with Green IT
by Manuel Landum, Maria Margarida Madeira e Moura and Leonilde Reis

CNN-LSTM Based Model for ECG Arrhythmias and Myocardial Infarction Classification
by Lana Abdulrazaq Abdullah and Muzhir Shaban Al-Ani

An Improved Approach for QoS Based Web Services Selection Using Clustering
by : Mourad Fariss, Naoufal El Allali, Hakima Asaidi and Mohamed Bellouki

Particle Swarm Optimization, Genetic Algorithm and Grey Wolf Optimizer Algorithms Performance Comparative for a DC-DC Boost Converter PID Controller
by Jesus Aguila-Leon, Cristian Chiñas-Palacios, Carlos Vargas-Salgado, Elias Hurtado-Perez and Edith Xio Mara Garcia

Node-Node Data Exchange in IoT Devices Using Twofish and DHE
by Bismark Tei Asare, Kester Quist-Aphetsi and Laurent Nana

Cardiovascular Risk in Patients who go to the Medical Office of a Private Health Center in North Lima
by Jairo Zegarra-Apaza, Sara Oliveros-Huerta, Santiago Vilela-Cruz, Rosita Chero-Benites, Gissett Marcelo-Ruiz, Leslie Yelina Herrera-Nolasco, Brian Meneses-Claudio, Hernan Matta-Solis and Eduardo Matta-Solis

Comparative Study of Control Algorithms Through Different Converters to Improve the Performance of a Solar Panel
by Zouirech Salaheddine, El Ougli Abdelghani and Belkassem Tidhaf

Modelling Human-Computer Interactions based on Cognitive Styles within Collective Decision-Making
by Nina Bakanova, Arsenii Bakanov and Tatiana Atanasova

Comparison of Machine Learning Parametric and Non-Parametric Techniques for Determining Soil Moisture: Case Study at Las Palmas Andean Basin
by Carlos López-Bermeo, Mauricio González-Palacio, Lina Sepúlveda-Cano, Rubén Montoya-Ramírez and César Hidalgo-Montoya

Overcome Discrimination: A Logistic Regression with 10-year Longitudinal Investigation of Emo Kids' Facebook Posts
by Proud Arunrangsiwed and Yothin Sawangdee

Performance Analysis and Enhancement of Spline Adaptive Filtering based on Adaptive Step-size Variable Leaky Least Mean Square Algorithm
by Sethakarn Prongnuch and Suchada Sitjongsataporn

A Novel Blockchain-Based Authentication and Access Control Model for Smart Environment
by Nakhoon Choi and Heeyoul Kim

Multiple Machine Learning Algorithms Comparison for Modulation Type Classification Based on Instantaneous Values of the Time Domain Signal and Time Series Statistics Derived from Wavelet Transform
by Inna Valieva, Iurii Voitenko, Mats Björkman, Johan Åkerberg and Mikael Ekström

Issues in File Caching and Virtual Memory Paging with Fast SCM Storage
by Yunjoo Park and Hyokyung Bahn

Design of a Flapping Wings Butterfly Robot based on Aerodynamics Force
by Kanjanapan Sukvichai and Kan Yajai
An Enhanced Artificial Intelligence-Based Approach Applied to Vehicular Traffic Signs Detection and Road Safety Enhancement
by Anass Barodi, Abderrahim Bajit, Taoufiq El Harrouti, Ahmed Tamtaoui and Mohammed Benbrahim

Clustering of Mindset towards Self-Regulated Learning of Undergraduate Students at the University of Phayao
by Pratya Nuankaew

Dismantle Shilling Attacks in Recommendations Systems
by Ossama Embarak

Automatic License Plate Detection and Recognition for Jordanian Vehicles

by Khalil Mustafa Ahmad Yousef, Bassam Jamil Mohd, Yusra Abd-Al-Haleem Al-Khalaileh, Ahlam Hani Al-Hmeadat and Bushra Ibrahim El-Ziq

Deep Learning Approach for Automatic Topic Classification in an Online Submission System
by Tran Thanh Dien, Nguyen Thanh-Hai and Nguyen Thai-Nghe

Analysis of Long-term Equilibrium Relationship Between KRW, RMB, JPY Exchange Rates and International Financial Market Variables: Comparative Analysis of KRW, RMB, JPY
by Moon-Kyum Kim and Woong Ryeol Kim

A Framework to Align Business Processes: Identification of the Main Features
by Joaquina Marchão, Leonilde Reis and Paula Ventura Martins

Development of an EEG Controlled Wheelchair Using Color Stimuli: A Machine Learning Based Approach
by Md Mahmudul Hasan, Nafiul Hasan and Mohammed Saud A Alsubaie

An Explanatory Review on Cybersecurity Capability Maturity Models
by Adamu Abdullahi Garba, Maheyzah Muhamad Siraj and Siti Hajar Othman

Frequency Scaling for High Performance of Low-End Pipelined Processors
by Athanasios Tziouvaras, Georgios Dimitriou, Michael Dossis and Georgios Stamoulis

Brain Tumor Classification Using Deep Neural Network
by Gökalp Çınarer, Bülent Gürsel Emiroğlu, Recep Sinan Arslan and Ahmet Haşim Yurttakal

BLDC Motor Vibration Identification by Finite Element Method and Measurements
by Jerzy Podhajecki and Stanisław Rawicki

Dependency Head Annotation for Myanmar Dependency Treebank
by Hnin Thu Zar Aye and Win Pa Pa

Advanced Multiple Linear Regression Based Dark Channel Prior Applied on Dehazing Image and Generating Synthetic Haze
by Binghan Li, Yindong Hua and Mi Lu

Dilated Fully Convolutional Neural Network for Depth Estimation from a Single Image
by Binghan Li, Yindong Hua, Yifeng Liu and Mi Lu

Discretisation of Second Order Generalized Integrator to Design the Control Algorithm of Unified Power Quality Conditioner
by Mashhood Hasan, Bhim Singh, Waleed Hassan Alhazmi and Sachin Devassy

Using Formal Methods to Model a Smart School System via TLA+ and its TLC Model Checker for Validation
by Nawar Obeidat and Carla Purdy

Design and Development of an Advanced Affordable Wearable Safety Device for Women: Freedom Against Fearsome

by Israt Humaira, Kazi Arman Ahmed, Sayantee Roy, Zareen Tasnim Safa, F. M. Tanvir Hasan Raian and Md. Ashrafuzzaman

A Study of Stirling Engine Efficiency Combined with Solar Energy
by Oumaima Taki, Kaoutar Senhaji Rhazi and Youssef Mejdoub

Diagnosis of Tobacco Addiction using Medical Signal: An EEG-based Time-Frequency Domain Analysis Using Machine Learning
by Md Mahmudul Hasan, Nafiul Hasan, Mohammed Saud A Alsubaie and Md Mostafizur Rahman Komol

Updated Analysis of Business Continuity Issues Underlying the Certification of Invoicing Software, Considering a Pandemic Scenario
by Nelson Russo and Leonilde Reis

Application-Programming Interface (API) for Song Recognition Systems
by Murtadha Arif Bin Sahbudin, Chakib Chaouch, Salvatore Serrano and Marco Scarpa

Multi-Layered Machine Learning Model For Mining Learners Academic Performance
by Ossama Embarak

Load Balancing Techniques in Cloud Computing: Extensive Review
by Ahmad AA Alkhatib, Abeer Alsabbagh, Randa Maraqa and Shadi Alzubi

Accounting Software in Modern Business
by Lesia Marushchak, Olha Pavlykivska, Galyna Liakhovych, Oksana Vakun and Nataliia Shveda

Detailed Assessment of Dissaving Risk Against Life Expectancy for Elderly People using Anonymous Data and/or Random Data: A Review
by Yuya Yokoyama and Yasunari Yoshitomi

Gene Selection for Cancer Classification: A New Hybrid Filter-C5.0 Approach for Breast Cancer Risk Prediction
by Mohammed Hamim, Ismail El Moudden, Hicham Moutachaouik and Mustapha Hain

Cyclic Evaluation of Capacity of Recovered Traction Battery after Short-Circuit Damage
by Matus Danko and Marek Simcak

Method of Technological Forecasting of Market Behaviour of R&D Products
by Vasyl Kozyk, Oleksandra Mrykhina, Lidiya Lisovska, Anna Panchenko and Mykhailo Honchar

Detection and Counting of Fruit Trees from RGB UAV Images by Convolutional Neural Networks Approach
by Kenza Aitelkadi, Hicham Outmghoust, Salahddine laarab, Kaltoum Moumayiz and Imane Sebari

Assessment of Electricity Industries in SADC Region Energy Diversification and Sustainability

by Kakoma Chilala Bowa, Mabvuto Mwanza, Mbuyu Sumbwanyambe, Kolay Ulgen and Jan-Harm Pretorius

Active Disturbance Rejection Control Design for a Haptic Machine Interface Platform
by Syeda Nadiah Fatima Nahri, Shengzhi Du and Barend Jacobus van Wyk

Using Big Data Analytics to Predict Learner Attrition based on First Year Marks at a South African University
by Gcobisile Matafeni and Ritesh Ajoodha

Empathy in Nursing Students that do the Non-Medical Internship in Three Universities in Lima, 2019
by Liseth Acuña-Medina, Yumira Arias-Quispe, Yackeline Espeza-Veláquez, Brian Meneses-Claudio, Hernan Matta-Solis and Eduardo Matta-Solis

Investigating the Optical Behavior of Single/Multi-Dimensional Photonic Crystal Structures for Photovoltaic Applications
by Gehad Ali Alsayed, Zahraa Ismail and Sameh O. Abdellatif

Standalone Operation of Modified Seven-Level Packed U-Cell Inverter for Solar Photovoltaic System
by Kishan Bhushan Sahay, Pankaj Kumar Singh and Rakesh Maurya

Total Family Risk in Families who go to Popular Dining Rooms in a Vulnerable Area of Collique, Comas
by Hernan Matta-Solis, Rosa Perez-Siguas, Eduardo Matta-Solis and Melissa Yauri-Machaca

Health-Related Quality of life in Students of an Education Institution of Ventanilla
by Lucia Silva-Bueno, Brian Meneses-Claudio, Hernan Matta-Solis and Lourdes Matta-Zamudio
Hand Gesture Classification using Inaudible Sound with Ensemble Method
by Jinwon Cheon and Sunwoong Choi

Docker-C2A : Cost-Aware Autoscaler of Docker Containers for Microservices-based Applications
by Mohamed Hedi Fourati, Soumaya Marzouk, Mohamed Jmaiel and Tom Guerout

NPC five level inverter using SVPWM for Grid-Connected Hybrid Wind-Photovoltaic Generation System
by Elamri Oumaymah, Oukassi Abdellah, Bouhali Omar and El Bahir Lhoussain

A Rectification Circuit with Co-Planar Waveguide Antenna for 2.45 GHz Energy Harvesting System
by Nuraiza Ismail and Ermeey Abd Kadir

Mobile Money Wallet Attack Resistance using ID-based Signcryption Cryptosystem with Equality Test
by Seth Alornyo, Mustapha Adamu Mohammed, Francis Botchey and Collinson Colin M. Agbesi

Assessment of the Municipal Solid Waste Transfer Stations Suitability in Harare, Zimbabwe
by Trust Nhubu, Edison Muzenda, Belaid Mohamed and Charles Mbohwa

A Review of Plastic Waste Management Practices: What Can South Africa Learn?
by Zvanaka S. Mazhandu, Edison Muzenda, Mohamed Belaid, Tirivaviri A. Mamvura and Trust Nhubu

Framework for Decentralizing Municipal Solid Waste Management in Harare, Zimbabwe
by Trust Nhubu, Edison Muzenda, Belaid Mohamed and Charles Mbohwa

Social Skills and Resilience in Adolescent of Secondary Level of a Public Educational Institution in Puente Piedra Lima – 2020
by Niurka Jacome-Olacua, Joselyne Rodríguez-Paucar, Prhitty Marin-Garcia, Brian Meneses-Claudio, Hernan Solis-Matta and Eduardo Matta-Solis

Closed Loop Capacitive Accelerometer Model using Simple Regression Test for Linearity
by Mamudu Hamidu and Jerry John Kponyo

Simulating Get-Understand-Share-Connect Model using Process Mining
by Shahrinaz Ismail and Faes Tumin

A Cavity Structure based Flexible Piezoelectric for Low-Frequency Vibration Energy Harvesting
by Khairul Azman Ahmad, Siti Noraini Sulaiman, Noramalina Abdullah and Muhammad Khusairi Osman

Formal Proof of Properties of a Syntax-Oriented Editor of Robotic Missions Plans
by Laurent Nana, François Monin and Sophie Gire

Gripper Finger Design for Automatic Bottle Opener
by Suchada Sitjongsataporn, Kornika Moolpho and Sethakarn Prongnuch

Modeling and Design of a Compact Metal Mountable Dual-band UHF RFID Tag Antenna with Open Bent Stub Feed for Transport and Logistics Fields
by Hajar Bouazza, Aarti Bansal, Mohsine Bouya, Azeddine Wahbi, Antonio Lazaro and Abdelkader Hadjoudja

Using a safety PLC to Implement the Safety Function
by Karol Rástočný, Juraj Ždánsky and Jozef Hrbček

Prioritization of Sustainable Supply Chain Management Practices in an Automotive Elastomer Manufacturer in Thailand
by Saruntorn Mongkolchaichana and Busaba Phruksaphanrat

Supporting the Management of Predictive Analytics Projects in a Decision-Making Center using Process Mining
by Marlene Ofelia Sanchez-Escobar, Julieta Noguez, Jose Martin Molina-Espinosa and Rafael Lozano-Espinosa
Design of Platform to Support Workflow Continuity in Multi-Device Applications

by Oscar Chacón-Vázquez, Luis G. Montané-Jiménez, Carlos Alberto Ochoa-Rivera and Betania Hernández-Ocaña

Evaluation of the Physico-Chemical Properties of Soil and Apple Leaves (Malus Domestica) in Beni Mellal-Khenifra Region, Morocco
by Berrid Nabyl, Lougraimzi Hanane, El-Khabbazi Houda, Abidli Zakaria, Hamidi Otman, Keltoum Rahali and El Mahjoub Aouane

Investigation of the LoRa Transceiver in Conditions of Multipath Propagation of Radio Signals
by Dmytro Kucherov, Andrei Berezkin, Volodymyr Nakonechnyi, Olha Sushchenko, Ihor Ogirko, Olha Ogirko and Ruslan Skrynkovskyy

Multi-Criteria Decision Analysis Coupled with GIS and Remote Sensing Techniques for Delineating Suitable Artificial Aquifer Recharge Sites in Tafilalet Plain (Morocco)
by Aicha Ousrhire, Hassane Oulidi Jarar and Abdessamad Ghafiri

Challenges and New Paradigms in Conservation of Heritage-based Villages in Rural India -A case of Pragpur and Garli Villages in Himachal Pradesh
by Preeti Nair, Devendra Pratap Singh and Navneet Munoth

Electronically Tunable Triple-Input Single-Output Voltage-Mode Biquadratic Filter Implemented with Single Integrated Circuit Package
by Natchanai Roongmuanpha, Taweepol Suesut and Worapong Tangsrirat

The Performance of Project Teams Selected Based on Student Personality Types: A Longitudinal Study
by Svitlana Ivanova, Lubomir Dimitrov, Viktor Ivanov and Galyna Naleva

Artificial Neural Network Approach using Mobile Agent for Localization in Wireless Sensor Networks
by Basavaraj Madagouda and R. Sumathi

Allocation of Total Congestion Cost and load participation to Generators for a PoolCo Market in Deregulated Power System
by Yashvant Bhavsar and Saurabh Pandya

Parameters Degradation Analysis of a Silicon Solar Cell in Dark/Light Condition using Measured I-V Data
by Dominique Bonkoungou, Toussaint Guingane, Eric Korsaga, Sosthène Tassembedo, Zacharie Koalaga, Arouna Darga and François Zougmore

An Operational Responsibility and Task Monitoring Method: A Data Breach Case Study
by Saliha Assoul, Anass Rabii and Ounsa Roudiès

Performance Evaluation of a Gamified Physical Rehabilitation Balance Platform through System Usability and Intrinsic Motivation Metrics
by Rosula Reyes, Justine Cris Borromeo and Derrick Sze

An algorithm for Peruvian counterfeit Banknote Detection based on Digital Image Processing and SVM
by Bryan Huaytalla, Diego Humari and Guillermo Kemper

A Novel Approach to Design a Process Design Kit Digital for CMOS 180nm Technology
by Thinh Dang Cong, Toi Le Thanh, Phuc Ton That Bao and Trang Hoang

Creativity in Prototypes Design and Sustainability – The case of Social Organizations
by Clara Silveira, Leonilde Reis, Vitor Santos and Henrique S. Mamede

Computer Vision for Industrial Robot in Planar Bin Picking Application
by Le Duc Hanh and Huynh Buu Tu

Crystallinity and Hardness Enhancement of Polypropylene using Atmospheric Pressure Plasma Discharge Treatment
by Oscar Xosocotla, Horacio Martinez and Bernardo Campillo

LEACH Based Protocols: A Survey
by Nour Najeeb Abdalkareem Qubbaj, Anas Abu Taleb and Walid Salameh

Robust Adaptive Feedforward Sliding Mode Current Controller for Fast-Scale Dynamics of Switching Multicellular Power Converter
by Rihab Hamdi, Amel Hadri Hamida, Ouafae Bennis and Fatima Babaa

Evaluation of Personal Solar UV Exposure in a Group of Italian Dockworkers and Fishermen, and Assessment of Changes in Sun Protection Behaviours After a Sun-Safety Training
by Alberto Modenese, Fabio Bisegna, Massimo Borra, Giulia Bravo, Chiara Burattini, Anna Grasso, Luca Gugliermetti, Francesca Larese Filon, Andrea Militello, Francesco Pio Ruggieri and Fabriziomaria Gobba

Combining ICT Technologies To Serve Societal Challenges
by Helen Leligou, Despina Anastasopoulos, Anita Montagna, Vassilis Solachidis and Nicholas Vretos

Analysis of qCON and qNOX Anesthesia Indices and EEG Spectral Energy during Natural Sleep Stages
by Joana Cañellas, Anaïs Espinoso, Juan Felipe Ortega, Umberto Melia, Carmen González and Erik Weber Jensen

SEA: An UML Profile for Software Evolution Analysis in Design Phase
by Akram Ajouli

Quality of Nursing Care in Hospitalized Patients of the Carlos Lanfranco La Hoz Hospital, 2019
by Amancio Izquierdo-Príncipe, Jaqueline Garcia-Núñez, Brian Meneses-Claudio, Hernan Solis-Matta and Lourdes Matta-Zamudio

Quality of Life in Patients with Type 2 Diabetes of the Central Hospital of the Peruvian Air Force, 2019

by Jared Zavala-Izaguirre, Fanny Mego-Llanos, Sarita Cornejo-Quispitongo, Brian Meneses-Claudio, Hernan Solis-Matta and Lourdes Matta-Zamudio

Event Modeller Data Analytic for Harmonic Failures
by Futra Zamsyah Md Fadzil, Alireza Mousavi and Morad Danishvar

Social skills and Resilience in Adolescents of Secondary Education of the Kumamoto I 3092 Educational Institution, of the Puente Piedra District – Lima 2019
by Evelyn Roncal-Cespedes, Gloria Castillo-Laban, Brian Meneses-Claudio, Hernan Matta-Solis, Lourdes Matta-Zamudio and Eduardo Matta-Solis

Social Skills and Resilience in Adolescents of an Educational Institution in North Lima, 2019
by Lili Sifuentes-Gomez, Doris Vega-Davila, Betty Flores-Paz, Brian Meneses-Claudio, Hernan Matta-Solis and Eduardo Matta-Solis

Improved Design and Recommendations for Street Lighting in Gitega City
by Ntawuhorakomeye Noel, Ndayiragije Leonidas and Belov Mikhail Petrovich

Towards a Hybrid Probabilistic Timing Analysis
by Haoxuan Li, Ken Vanherpen, Peter Hellinckx, Siegfried Mercelis and Paul De Meulenaere

Investigating Students' Computational Thinking through STEM Robot-based Learning Activities
by Sasithorn Chookaew, Suppachai Howimanporn and Santi Hutamarn
2
Texture Based Image Retrieval Using Semivariogram and Various Distance Measures
by Rajani Narayan and Anjanappa Sreenivasa Murthy

Effect of Starch Oxidation Degree on the Properties of Hydrogels from Dialdehyde Starch and Polyvinyl Alcohol
by Jahel Desire Carrera, Daniela Alejandra Viteri Narváez, Marco Leon and José Francisco Alvarez-Barreto

A-MnasNet and Image Classification on NXP Bluebox 2.0
by Prasham Shah and Mohamed El-Sharkawy

Image Tag Recommendation based on Ranked Categorical Nearest Neighbors and Weighted Tag Features
by Anupama D. Dondekar and Balwant A. Sonkamble

BrcLightning – Risk Analysis and Scaling for Protection against Atmospheric Discharge – Extender
by Biagione Rangel De Araújo

Exposure to Optical Radiation and Electromagnetic Fields at the Workplace: Criteria for Occupational Health Surveillance According to Current European Legislation
by Alberto Modenese and Fabriziomaria Gobba

Fusion of Optical and Microfabricated Eddy-Current Sensors for the Non-Destructive Detection of Grinding Burn

by Isman Khazi, Andras Kovacs, Ulrich Mescheder, Ali Zahedi and Bahman Azarhoushang
Comparative Study Between Three Methods for Optimizing the Power Produced from Photovoltaic Generator
by El hadji Mbaye Ndiaye, Mactar Faye and Alphousseyni Ndiaye

An Investigation of the Effect of Optimal Plane Spacing Between Electrode Planes for the EIT Industrial Applications
by Yew Lek Chong and Renee Ka Yin Chin

Application of Deep Belief Network in Forest Type Identification using Hyperspectral Data
by Xianxian Luo, Songya Xu and Hong Yan

A Computational Modelling and Algorithmic Design Approach of Digital Watermarking in Deep Neural Networks
by Revanna Sidamma Kavitha, Uppara Eranna and Mahendra Nanjappa Giriprasad

ESP2: Embedded Smart Parking Prototype
by Tarek Frikha, Hedi Choura, Najmeddine Abdennour, Oussama Ghorbel and Mohamed Abid

A Framework for Adoption and Diffusion of Mobile Applications in Africa
by Chinedu Wilfred Okonkwo, Magda Huisman and Estelle Taylor

FPGA-Based Homogeneous and Heterogeneous Digital Quantum Coprocessors
by Valeriy Hlukhov

Type 2 Diabetes Risk and Physical Activity in outpatients treated in Health Centers in a District of North Lima, 2020
by Deisy Chipana-Collahua, Mariluz Chipana-Collahua, Rosa Villegas-Ortiz, Brian Meneses-Claudio and Hernan Matta-Solis

Sentiment Analysis in English Texts
by Arwa Alshamsi, Reem Bayari and Said Salloum

A Machine Vision Approach for Underwater Remote Operated Vehicle to Detect Drowning Humans
by Yaswanthkumar S K, Keerthana M and Vishnu Prasath M S

An Evaluation of some Machine Learning Algorithms for the detection of Android Applications Malware
by Olorunshola Oluwaseyi Ezekiel, Oluyomi Ayanfeoluwa Oluwasola and Irhebhude Martins

Touristic's Destination Brand Image: Proposition of a Measurement Scale for Rabat City (Morocco)
by Abdellatif Elouali, Smail Hafidi Alaoui, Noura Ettahir, Abderrazzak Khohmimidi, Nadia Motii, Keltoum Rahali and Mustapha Kouzer

The Designing of Institute's Educational Mascots for Brand Identity
by Nop Kongdee, Suparada Prapawong and Manissaward Jintapitak

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 15-26 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Environmental Acoustics Modelling Techniques for Forest Monitoring

Svetlana Segarceanu[*,1], George Suciu[1], Inge Gavăt[2]

[1]*Beia Consult International, Bucharest, 041386, Romania*

[2]*University Politehnica of Bucharest, Applied Electronics Department, Bucharest, 060042, Romania*

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 29 March, 2021
Online: 05 May, 2021

Keywords:
Environment Sound Recognition
Deep Neural Network
Dynamic Time Warping
Gaussian Mixtures Modeling

ABSTRACT

Environmental sounds detection plays an increasing role in computer science and robotics as it simulates the human faculty of hearing. It is applied in environment research, monitoring and protection, by allowing investigation of natural reserves, and showing potential risks of damage that can be deduced from the environmental acoustic. The research presented in this paper is related to the development of an intelligent forest environment monitoring solution, which applies signal analysis algorithm to detect endangering sounds. Environmental sounds are processed using some modelling algorithms based on which the acoustic forest events can be classified into one of the categories: chainsaw, vehicle, genuine forest background noise. The article will explore and compare several methodologies for environmental sound classification, among which the dominant Deep Neural Networks, the Long Short-Term Memory, and the classical Gaussian Mixtures Modelling and Dynamic Time Warping.

1. Introduction

Environmental sound recognition (AESR) is a relatively new discipline of computer science destined to extend the field of speech-based applications, or the study of music sounds, by exploring the vast range of environmental non-speech sounds.

This paper is an extension of work originally presented in the 43rd International Conference on Telecommunications and Signal Processing, held in Milan in July 2020 [1]. It investigates several sound modelling and classification paradigms, in the context of a forest monitoring system.

The new research area of AESR is intended to simulate the important function of human hearing, and possibly, to overpower human perception. This is motivated by the fact that hearing is the human second most important sense after vision and should not be disregarded when trying to build a computer that simulates human behaviour and senses. Consequently, AESR recently became an essential field of computer science [2]-[4].

By environmental sounds we mean everyday sounds, natural or artificial, other than speech. Natural sounds may be leaves rustling, animal noise, birds chirping, water ripple, wind blowing through trees, waves crashing onto the shore, thunder. Artificial sounds include sounds produced by diverse engines like cars, cranes, ATVs, snowmobile, lawnmower, pneumatic hammer, chainsaw, etc., but also creaky doors or creaky floors, gunshots, crowd in a club, breaking glass, vehicle tyres or brake noise. The environments where this discipline is applied are also very diverse, falling into two broad categories:

- Natural environment like forest environment, ecological units, seashore environment; however, the degree of naturalness is variable, and purely natural environments are very rare as one may intercept voices, vehicle sounds, chainsaw in forests or even jungle.
- Built environment, also called human made environment, like the urban environment, the household, maybe agricultural environment, harbours; obviously in any of these environments, natural acoustic events, like thunder, rain falling, wind blowing may also happen.

In reality there are no purely natural or artificial sounds, neither purely speech nor non-speech acoustic events. Most of the application investigate mixed environments where natural and artificial, speech or non-speech acoustic events are equally possible.

The applicability of AESR is related to the advent of Internet of Things (IoT). IoT devices have sensors, possibly acoustic,

*Corresponding Author Svetlana Segărceanu, Email: svetlana.segarceanu@beia.ro

and software that enable the collection and exchange of data through the Internet. Automatic recognition of the surrounding environment allows devices to switch between tasks with minimum user interference [5]. For a robot, audio recordings may provide important information about location and direction of a moving vehicle, or environmental information, such as speed of wind.

AESR has an important role in security, environment protection or environment research. Among possible applications are the identification of deforestation threats and illegal logging activities, through automatic detection of specific sounds like several engines, chainsaws, or vehicles. Detecting other illegal activities like hunting in forest or ecological reserves, by spotting gun shots, or human voices would be a useful application [6]-[9]. In recent times, solutions based on environmental sound recognition are applied in early wildfire detection [10].

Another type of applications concerns scientific monitoring of the environment. Such applications are intended for instance, to detect species, by discrimination between different animals or birds' sounds [11], [12].

Computational Auditory Scene Analysis (CASA), is a very complex field of AESR, aimed to the recognition mixtures of sound sources by simulating human listening perception using computational means. It mainly addresses two important tasks, Environmental Audio Scene Recognition (EASR) and Sound Event Recognition (SER) and has a huge importance in environment audio observation and surveillance. EASR refers to recognition of indoor or outdoor acoustic scenes (e.g., cafes/restaurants, home, vehicle or metro stations, supermarkets, versus crowded or silent streets, forest landscape, countryside, beaches, gym halls, swimming pools). SER is intended to the investigation of specific acoustic events in the audio environments, like dog barking, gunshots, sudden brake sounds, or human non-speech events, like coughing, whistling, screaming, child crying, snoring, sneezing [13].

An emerging field is the investigation and detection of acoustic emissions, used in monitoring landslide phenomena [14]. Acoustic emissions (AE) are elastic waves generated by movement at particle-to-particle contacts and between soil particles and structural elements. They are not perceived by the human ear, are super audible, and therefore their frequencies are very high, expected to range between 15kHz and 40 kHz. The devices used to acquire these waves should ensure a sampling frequency of over 80 kHz. AE monitoring is an active area, not very well developed due to low energy levels of these waves, which make it challenging to detect and quantify [15], [16].

Likewise, recent studies have shown that moving avalanches emit a detectable sub-audible sound signature in the low frequency infrasonic spectrum [17].

The study of underwater acoustic infrasonic emissions, provided by hydrophones, is another field of AESR research [18].

Our paper explores forest acoustics aiming to find the suitable sound modelling and classification approaches. The focus of our research is the detection of logging risk by identification of specific classes of sounds: chainsaw, vehicles, or possibly speech. We extend an earlier research on acoustic signal processing, by exploring the dominant paradigm in data modelling and, the deep neural networks (DNN). We investigate two types of DNNs, the Deep Feed Forward Neural Networks (FFNN) and a popular version of Recurrent Neural Networks (RNNs), the Long Short-Term Memory (LSTM). The two neural networks will be run on two types of feature spaces: the Mel-cepstral and the Fourier log-power spectrum feature spaces. We will compare their results with the former performance obtained using the Gaussian Mixtures Modelling (GMM) and the Dynamic Time Warping (DTW) in the context of a closed-set identification system. One main goal is to stress the importance of feeding as input to DNN less processed features, like log-power spectrum, as compared to the more elaborate sets of features, e.g., Mel-frequency cepstral features. Another purpose of the paper is to clarify some issues concerning signal pre-processing framework, like length of the analysis window and the underlying frequency domain to be used in spectral analysis.

The paper is structured as follows: the next section describes the state-of-the-art in environmental sound recognition; the third section details our approach, the signal feature extraction and modelling methods we applied in sound recognition; the fourth section presents the setup of the experiments and evaluates the proposed methods; the last part presents the conclusions of the paper.

2. State-of-the-art

Early attempts [2], [19] to assess speech typical methods in the context of non-speech acoustics, analyse classical methods like Artificial Neural Networks (ANN), or Learning Vector Quantization (LVQ), on Fourier, or Linear Predictive Coding (LPC) feature spaces.

In [20], the authors make an overall investigation of recognition methodologies for different categories of sounds. The environmental sounds are classified as stationary and non-stationary. The framework used for stationary acoustic signals coincides to a great extent with the one used in voice-based applications (speech or speaker recognition) in what concerns the specific features and feature space modelling methods. For feature extraction, the spectral features, like those derived from Mel analysis – Mel Frequency Cepstral Coefficients (MFCCs), LPC, Code Excited Linear Prediction (CELP), or techniques based on signal autocorrelation, prevail. The modelling approaches are also shared with voice-based applications: GMM, k-Nearest Neighbours (k-NN), Learning Vector Quantisation (LVQ), DTW, Hidden Markov Models (HMM), Support Vector Machine (SVM), Neural Networks, and deep learning. Concerning non-stationary signals, some successful techniques are based on sparse representations like the Matching Pursuit (MP) and MP-Gabor features. Alternative approaches use fusion of MFCC and other parameters to boost the performance.

In [5], the authors review the current methodologies used in AESR and evaluate their performance, efficiency, and computational cost. The leading approaches of the moment are GMM, SVM and DNN or Recurrent Neural Networks (RNN) The paper describes open-set identification experiments on two types of acoustic events, baby cries and smoke alarm, and a very large range of complementary environment acoustic events. as impostor data. In this respect GMM, using the Universal Background Model

(GMM-UBM) and two neural network architectures were compared. The Deep Feed Forward Neural Network yielded the best identification rate, while the best computational cost is made by GMM. SVM has an intermediate identification rate, yet at a high computational cost, assessed accounting for four basic operations: addition, comparison, multiplication, and lookup table retrieval (LUT). The computational cost is a critical feature in the context of IoT, where sound analysis applications are required to run on embedded platforms with hard constraints on the available computing power.

In [13], the authors make a thorough and extensive investigation of the most recent achievements and tendencies in AESR, more precisely in EASR and SER fields of CASA. from the perspective of acoustic feature extraction, the modelling methodology, performance, available acoustic databases. Besides the conventional approaches, new classes of features were applied lately by several implementations. Such characteristics are the auditory image-based features, basically regarding the time-frequency spectrograms as bidimensional images. where the frequency is not necessarily in the linear domain, but possibly adapted to a perceptual scale. Besides the log-power spectrum, Mel or Bark-frequency log-scale spectrograms, Spectrogram Image Features (SIF) [21], such characteristics as Mel scale with Constant-Q-Transform [22], wavelet coefficients [23] are referred.

Another class of features are generated by learning approaches with the goal to provide lower and enhanced representations. Such features are obtained applying techniques like quantization, i-vector, non-negative matrix factorization (NMF) [24], sparse coding, Convolutional Neural Network–Label Tree Embeddings (CNN-LTE) [25], etc.

Concerning the experimented methodologies, the deep learning methods are predominant, with Feed-Forward Neural Networks and Convolutional Neural Networks (CNN) in the leading position. Many strategies are currently operating CNNs in conjunction with a variety of features, among which log-scaled mel-spectrograms [26]-[28], CNN-LTE [25] or in hybrid approaches [29]. These implementations outperformed the other attempts to approach EASR and SER tasks.

Avalanche or landslide monitoring applications use methodologies based on thresholds for acoustic emissions energy, depending on the hazard risk level.

Concerning the general framework applied in AESR, we draw on the ideas presented in [20]. The usual pre-processing of the acoustic signal, applied in AESR includes a framing step, possibly followed by sub-framing or sequential processing. In the "framing" stage the signal is processed continuously, frame by frame. A classification decision is made for each frame and successive frames may belong to different classes. This is illustrated in figure 1, where a chainsaw is detected in a forest environment. Framing can enhance the acoustic signal classification by structuring the stream into more homogeneous blocks to better catch the acoustic event. Yet, there is no way of setting an optimal frame length, as for stationary events a length of 3s is a reasonable choice, while for acoustic events like thunder or gunshots, a 3s window length might be too large, and contain other acoustic events, so they could be associated to inappropriate classes. Due to the latest advances in instrumentation, different frame lengths are used to streamline energy consumption during a monitoring process, based on detecting energy levels of environmental sounds.

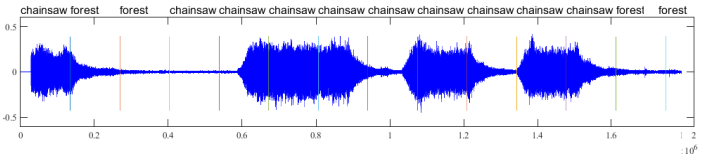


Figure 1: Framing process of a real-life sound and classification of each frame

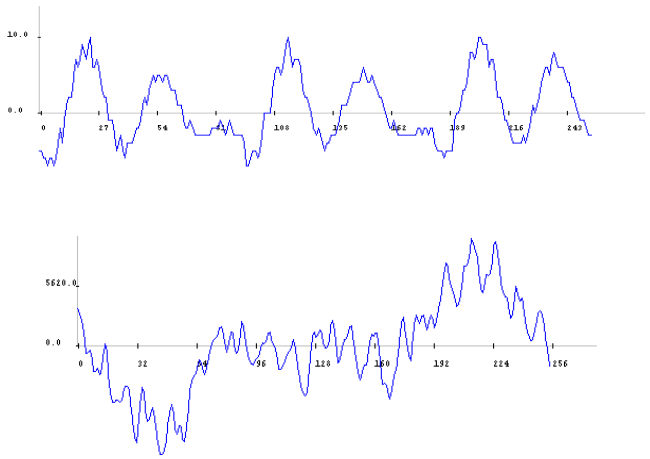


Figure 3: 22ms of chainsaw sound

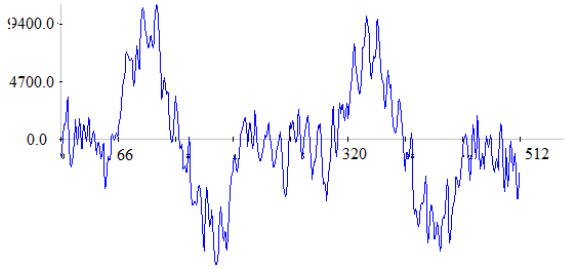


Figure 4: 44ms of chainsaw sound

Next, a sub-framing process is applied, by dividing the frame into usually overlapping, analysis subframes. The length of a subframe is explicitly set in [20] to 20-30ms. This length is suited for speech analysis, as it ensures a good resolution in time and frequency. Figure 2 presents 22ms of male speech which includes three fundamental periods of the respective voice. Whereas figure 4 represents 44ms of chainsaw sound which contains two periods of the chainsaw sound. However, we cannot infer anything about the signal periodicity from the segment of 22ms of chainsaw sound, represented in figure 3. Therefore, considering sub-frames of 44ms is a reasonable choice for chainsaw detection. However, a realistic setup must consider a value convenient to all sounds in the acoustic environment.

The further processing applied on the analysis frames is the same with that applied in speech signal analysis and its final goal is to extract characteristic features. The largely applied features are somehow derived from the Fourier features, and called spectral features. Non-spectral features are, for instance, energy, Zero-Crossing Rate (ZCR), Spectral Flatness (SF), all calculated in time domain. Concerning spectral analysis, the relevant frequency

interval for a signal sampling frequency of 44.1 kHz is not the whole frequency domain furnished by the Fourier transform, [0, 22.05] kHz, but a shorter range, as shown in figures 5,6. By setting the appropriate analysis frequency interval, we may improve the performance as accuracy and speed of execution, as another benefit of shortening the frequency interval is the decrease of the number of spectrum samples to be processed.

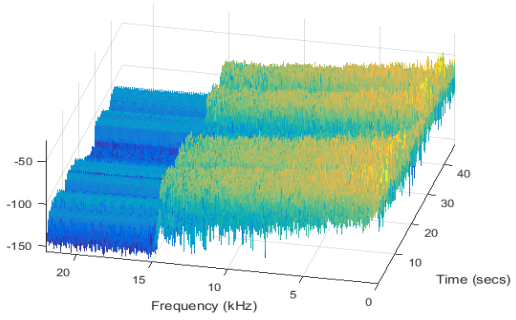


Figure 5: Power spectrum of a chainsaw sound

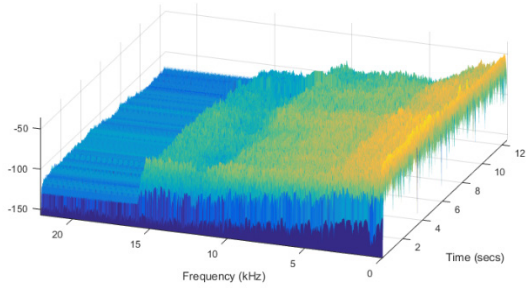


Figure 6: Power spectrum of a snowmobile sound

When choosing the spectral analysis, the usual pre-processing on the analysis frame involves (Hamming) windowing and pre-emphasis.

3. The Method

We have applied the framework mentioned above. At framing we divided the audio signal recordings into intervals of 3 seconds. We have used analysis frames of lengths of 22, 44 and 88ms with the usual pre-processing scheme as in speech-based applications. The modelling methods we have evaluated are GMM, DTW and FFNN. As baseline for the feature space, we used the Fourier spectrum coefficients and MFCCs. The set of MFCC parameters was possibly increased with Zero Crossing Rate (ZCR) or/and Spectral Flatness (SF). We have applied GMM on the MFCC feature space, and called this approach MFCC-GMM, DTW on the spectral features space, and FFNN on both the MFCC and spectral features spaces.

3.1. MFCC-GMM

GMM provides probabilistic weighted clustering that generates a coverage of the data space rather than a partition [30], [31]. Each cluster is modelled by a Gaussian distribution, usually called component, defined by the mean μ and the standard deviation $\boldsymbol{\Sigma}$ of cluster data, along with a weight w of the component inside the mixture. A data set belonging to the same class C, can be modelled by one or more Gaussian components, and the parameters of each component are calculated using the Estimation-Maximization algorithm, resulting in a model $\lambda_C = ((w_k, \mu_k, \boldsymbol{\Sigma}_k), k = 1\ldots, K)$, where K is the number of components. A key step of the algorithm is the initialization where initial values of the parameters (means, variances and weights) are defined. Poor initialization entails bad quality of classification or even impossibility to define the Gaussian parameters. We used at initialization a hierarchical algorithm, Pairwise Nearest Neighbour (PNN) [32] to ensure balanced data clustering, although other hierarchical algorithms such as Complete Linkage Clustering, or Average Linkage Clustering also provide good performance. We have evaluated different distance measures between hierarchy branches: Minkowski (Euclidean distance when square powers are considered), Chebyshev, Euclidian standardised distance. GMM modelling was applied on a feature space consisting of MFCCs and/or Spectral Flatness and ZCR, to generate models for C (C=3) classes of sounds. To classify a sequence of d dimensional features $X = \{x_1, x_2, \ldots, x_T\}$ into one of the classes its likelihood to belong to each class c is evaluated as

$$log(X, \lambda_c) = \sum_{t=1}^{T} p(x_t/\lambda_c) \quad (1)$$

where

$$p(x_t/\lambda) = \sum_{k=1}^{K} w_k \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t-\mu_k)^T \Sigma_k^{-1}(x_t-\mu_k)} \quad (2)$$

and the class with the maximum likelihood is selected:

$$\lambda_{best} = \underset{\lambda \in \{\lambda_1, \lambda_2, .., \lambda_C\}}{\operatorname{argmax}} (\log(X, \lambda)) \quad (3)$$

Mel frequency analysis [33] [34] is a perceptual approach to signal analysis based on human sensing of the frequency domain. We applied the MATLAB implementations of the Mel-scale frequency:

$$f_{mel}(f) = 1127 \ln(1 + f/700) \quad (4)$$

and bank of triangular filters for linear frequencies $f \in [f_{low}, f_{high}]$:

$$H_m[k] = \begin{cases} 0 & k < f[m-1] or\ k \geq f[m+1] \\ \frac{k - f[m-1]}{(f[m] - f[m-1])} & f[m-1] \leq k < f[m] \\ \frac{f[m+1] - k}{(f[m+1] - f[m])} & f[m] \leq k < f[m+1] \end{cases} \quad (5)$$

The power spectrum calculated on an analysis frame, is passed through the bank filter in (5), and the Mel Frequency Cepstral Coefficients (MFFCs) are derived by applying the Discrete Cosine Transform to the logarithm of the filtered spectrum [33].

Spectral Flatness (tonal coefficient) is meant to highlight noise from tonal sound and is calculated as ratio of geometrical and arithmetical means of spectral coefficients on analysis frames.

A zero crossing arises when two neighbouring samples have opposite signs, and its value on an analysis frame is:

$$Z_n = \sum_{m=-\infty}^{\infty} \left| sgn(x(m)) - sgn(x(m-1)) \right| * w(n-m) \quad (6)$$

3.2. Spectra Alignment using Dynamic Time Warping

Dynamic Time Warping (DTW) measures the similarity of two, usually time-varying, sequences, by optimally aligning them using a recurrent algorithm [35], complying with specific constraints, concerning boundary conditions, monotony, and continuity of the similarity function, and building an optimal path. One of the issues raised by the DTW algorithm is the long execution time, the main reason for which is the full calculus of DTW matrices, usually defined by distances between the elements of the sequencies. This can be contained by restraining the calculus to a low number of elements, the most likely to participate in the definition of the optimal path by applying an adjusting window as in figure 4, where the popular Sakoe-Chiba band [36] is applied. Figure 7 presents the graphical rendering of a DTW matrix for two sequencies *s(t)* and *r(t)*, the Sakoe-Chiba band, in grey, highlighting the optimal path, which does not lie entirely inside the band. When applying the Sakoe-Chiba band, the optimal path should lie inside the band and is figured in red in the image, accompanying only inside the band the real optimal path figured in black, thick where it lies within the band.

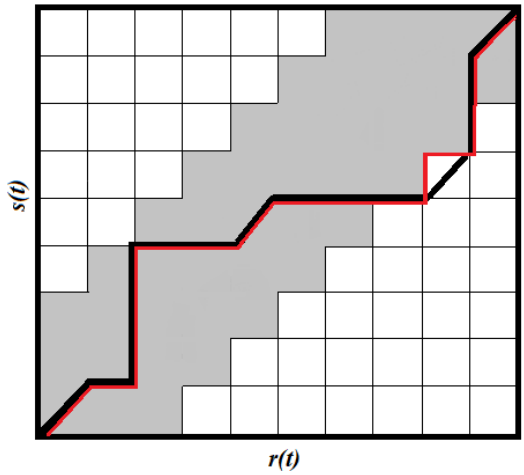


Figure 7: Alignment of sequences *r, s*, using the Sakoe- Chiba band, and the two optimal paths, the real one (black) and the one lying inside the band, which generally coincide.

The DTW algorithm was applied on power spectra of the signal. The power spectrum on an analysis window is calculated as sum of squared Fourier coefficients. One argument for using DTW to align spectral series is the fact that the acoustic signals received from devices of the same type share the same characteristics, such as sampling frequency, so, the generated spectra have the same lengths. Another premise is the fact that the interesting domain for this type of application is under 15 kHz, or even 10 kHz, which is demonstrated by figures 5, 6. This fact is used to align equal length spectra by the DTW algorithm. In the experiments the sampling frequency of the available audio files is equal to 44.1 kHz, the Fourier spectrum covers the domain [0, 22.05] kHz, but if the useful frequency domain is restricted to [0, 7.4] kHz and the analysis frame is of 22ms, the number of features is 171 instead of 512.

Classification of a sequence of feature vectors using the DTW algorithm consists in calculating the distortion between these vectors and the template (training) sequences, and to select the class whose templates show the smallest distortion with respect to the given feature sequence. The calculation involved in this process is based on the distances between individual vectors in the two sets to be compared. The distance may be evaluated in different ways. We have applied two distance measures in the calculus of the distortion measure, the Euclidian norm, and the 1-norm (sum of absolute values).

3.3. Deep feedforward networks (FFNN)

The artificial neural networks (ANNs) were intended to simulate human associative memory. They learn by processing known input examples, and corresponding expected results, creating weighted associations between them, stored within the network data structure. Deep feedforward networks or multilayer perceptrons (MLPs), are the quintessential deep learning models [37]. The basic unit of a FFNN is the artificial neuron, which expresses the biological concept of neuron [38]. They receive input data, combine the input through internal processing elements like weights and bias terms, and apply an optional threshold using an activation (transfer) function, as shown in figure 8. Transfer functions are applied to provide a smooth, differentiable transition as input values change. They are used to model the output to lie between 'yes' and 'no', mapping the output values between 0 to 1 or -1 to 1, etc. Transfer function are basically divided into linear and non-linear activation functions. Non-linear transfer functions are "S" – shaped functions like *arctg*, hyperbolic tangent, logistic functions as in (7):

$$f(x) = \sigma(x) = \frac{1}{1+e^{-x}} \tag{7}$$

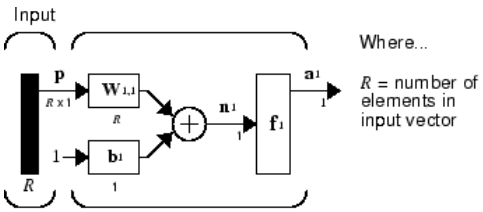


Figure 8: Structure of a neuron

The goal of a feedforward network for modelling and classification is to define a mapping $y = f(x, \theta)$ and learn the value of parameters θ to ensure the best approximation of the expected value y by the output of f, given the input x and parameters θ. FFNNs have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. A layer of neurons brings together the weight vectors and biases corresponding to its neurons, so it can be expressed by a matrix of weights and bias vectors, as in figure 9, 10. The transfer function is supposed to be the same for each neuron in the layer. The general diagram of a network is shown in figure 11, where the parameters to be tuned are the weight matrices and bias terms applied at the level of each layer, so that the output of the overall system would be close to expected values. These networks are called feedforward because the information flows in one direction through intermediate computations and there is no feedback connection. The number of neurons does not necessarily decrease with the layer level as presented in figure 10, but usually the goal is to reduce the dimensionality of the input layer, a process similar to feature extraction. The computation corresponding to figure 12 can be expressed by :

$$a^k = f^k(W_k a^{k-1} + b_k) =$$
$$f^k\left(W_k\left(f^{k-1}(W_{k-1}a^{k-2} + b_{k-1})\right) + b_k\right) = \quad (8)$$
$$f^k\left(W_k\left(f^{k-1}(W_{k-1}(f^{k-2}(W_{k-2}a^{k-3} + b_{k-2})) + b_{k-1})\right) + b_k\right) = \cdots$$

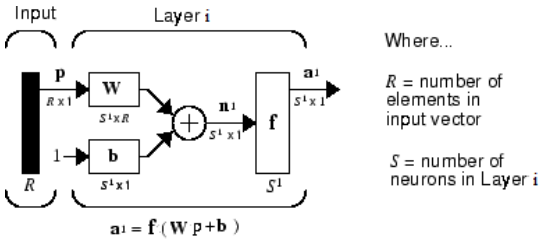


Figure 9: Structure of a layer of neurons

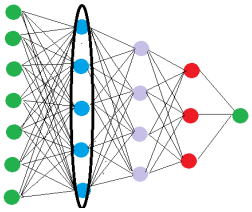


Figure 10: A Feed Forward Neural Network with three highlighted layers

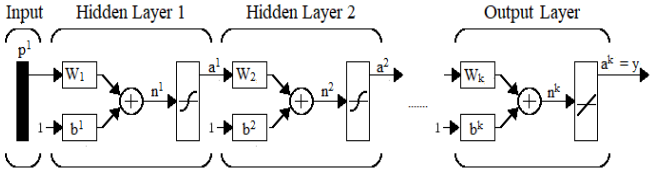


Figure 11: Flow of data in a feedforward network

In equations (8) the known information is

- The input parameters *p*, which might be measurements from sensors (wind speed, temperature, humidity), parameters coming from images (matrices of colours, or grey hues), or parameters coming from acoustic signals (Fourier spectrum on an analysis window, or more complicated parameters like cepstral, linear prediction coefficients),
- The expected output: for instance, to solve a three classes problem the output corresponding to each class input might be defined as either unidimensional (a scalar value for each class): (-1, 0, 1) or (0, 1. 2) or multidimensional (a vector for each class):((1, 0, 0), (0, 1, 0), (0, 0, 1)),
- The neural network architecture: number of hidden layers, number of neurons on each layer, etc.

Unknown parameters are:

- weights at layer *k*: W_k,
- bias terms at layer *k*: b_k.

Learning the unknown parameters is performed during the training process. Training of a FFNN can be made in batch mode or in incremental mode [38]. In batch mode, weights and biases are updated after all the inputs and targets are presented. Incremental networks receive the inputs one by one and adapt the weights according to each input. Usually, batch training is used. Equations (8) have as unknowns, the weight matrices and the bias terms, and a much more numerous training known data (all the input data and the corresponding expected values). This implicates the realistic conclusion that there will not be any solution of the equation system, so the training process looks for the values of the parameters, weights and biases, that make the error between the output value and expected output, minimal:

$$e(W, b) = \sum_{i=1}^{N}(y_i - a_i^k(p, W, b))^2 \quad (9)$$

where *N* is the number of (input, output) pair samples.

To minimize the least mean square (LMS) expression in (8) several schemes based on LMS algorithm using variants of the steepest descent procedure, are used. MATLAB has implemented and supports a range of network training algorithms among which: Levenberg-Marquardt Algorithm (LMA), Bayesian Regularization (BR), BFGS Quasi-Newton, Resilient Backpropagation, Scaled Conjugate Gradient, One Step Secant, etc. To start minimization of (9) using any of these algorithms, the user should provide an initial guess for the parameter vector θ=(W, b). The performance of the system depends on this initial guess. Most of the above algorithms try to optimize this process.

At the end of the training process, we get a FFNN model:

$$net = (W_k, b_k), \qquad k = 1,2...K. \quad (10)$$

where *K* is the number of layers in the network. To classify a vector of data $x = \{x_1, x_2, \ldots, x_d\}$, we "feed" it at the input of the network, perform all the operations applying the weights and biases to the input data, as in figure 11, and evaluate the output:

$$score = net\ (x) \quad (11)$$

If we code the output classes $y = \{y_1, y_2, \ldots, y_C\}$, *C* the number of classes, we compare the obtained output score to these values and if score is closest to y_c the input vector x will belong to class c.

We have applied the feedforward algorithm by feeding at input two types of features: power spectrum features and MFCCs.

3.4. Long Short-Time Memory (LSTM)

LSTM [39] is an artificial Recurrent Neural Network, and as any RNN is designed to handle sequences of events that occur in succession, with the understanding of each event based on information from previous events. They are able to handle tasks such as stock prediction or enhanced speech detection. One significant challenge for RNNs performance is that of the vanishing gradient which impacts RNNs long-term memory capabilities, restricted to only remembering a few sequences at a time. LSTMs proposes an architecture to overcome this drawback and allow to retain information for longer periods compared to traditional RNNs. Unlike standard feedforward neural networks, LSTM has feedback connections. It is capable of learning long-term dependencies, useful for certain types of prediction requiring the network to retain information over longer time periods, can process entire sequences of data (such as speech or video). It has been introduced in 1997 by the German researchers, Hochreiter and Schmidhuber.

The architecture of a LSTM Neural Network includes the cell (the memory part of the LSTM unit) and three "regulators", called gates, of the flow of information inside the LSTM unit:

- input gate to control the extent to which new values flow into the cell
- output gate to control the extent to which a value remains in the cell
- forget gate to control to what extent the value in the cell is used to compute the output activation of the LSTM unit

The LSTM is able to remove or add information to the cell state, through these gates. Some variations of the LSTM, like the Peephole LSTM or the Convolutional LSTM, ignore one or more of these gates.

The cell is responsible for keeping track of the dependencies between the elements in the input sequence. The activation function of LSTM gates is often the logistic sigmoid function. The connections to and from the LSTM gates, some recurrent, are weighted. The weights are learned during training, they determine how the gates operate. The diagram of a cell is presented in Figure 1(https://en.wikipedia.org/wiki/Long_short-term_memory#/media/File:The_LSTM_cell.png) and the LSTM flow is shown in figure 13 (Understanding LSTM Networks -- colah's blog) .

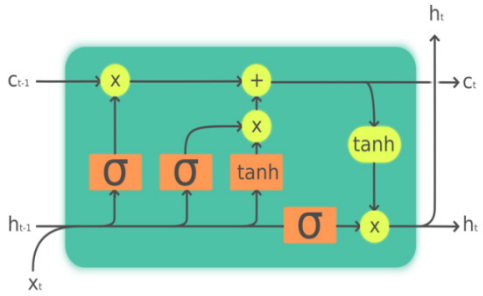


Figure 12: Structure of a LSTM cell

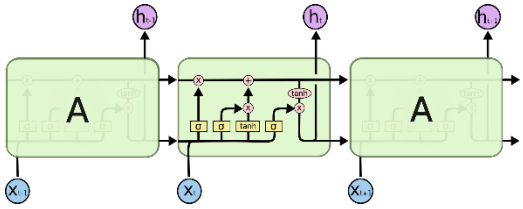


Figure 13: LSTM chain

The calculations that solve the LSTM paradigm are [39]:

$$\begin{aligned} f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \\ i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \\ o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \\ \widetilde{c_t} &= \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \\ c_t &= f_t \cdot c_{t-1} + i_t \cdot \widetilde{c_t} \\ h_t &= o_t \cdot \sigma_h(c_t) \end{aligned} \quad (12)$$

where the known data are:

- $x_t \in R^d$- input vector to the LSTM unit,
- d and h - number of input features and number of hidden units, respectively,

Unknowns :

- $f_t \in R^h$ - forget gate's activation vector
- $i_t \in R^h$ - input/update gate's activation vector
- $o_t \in R^h$ - output gate's activation vector
- $h_t \in R^h$ - hidden state vector (output vector of the LSTM unit)
- $\widetilde{c_t} \in R^h$ - cell input activation vector
- $c_t \in R_h$ - cell state vector
- $W \in R^{hxd}$, $U \in R^{hxd}$, $b \in R^h$ - weight matrices and bias vector parameters which need to be learned during training

Activation functions are:

- σ_g - sigmoid function
- σ_c - hyperbolic tangent function
- σ_h - hyperbolic tangent function or, linear function

The LSTM training is made in a supervised mode by a set of algorithms like gradient descent, combined with backpropagation through time to compute the gradients needed during the optimization process, in order to change each weight of the LSTM network in proportion to the derivative of the error (at the output layer of the LSTM network) with respect to corresponding weight. With LSTM units, when error values are backpropagated from the output layer, the error remains in the LSTM unit's cell. This allows to avoid the problem with standard RNNs where error gradients vanish exponentially with the size of the time lag between important events. The system is trained using the equations (6).

4. Experimental results

The goal of the experiments was to evaluate the four methodologies and find the optimal configuration for each one. The experiments considered only three classes of sounds which could exhaust the specific sounds in the forest environment susceptible to illegal deforestation. They are chainsaw, vehicle, and genuine forest sounds. The identification process was closed set. Segments of 3s were considered and each segment was evaluated individually. We have assessed several lengths of subframes (analysis frames), based on an above remark (see figures 2-4). So, the analysis frames lengths considered are mainly 22ms, 44ms, 88ms. Concerning the frequency interval length, [f_{low}, f_{high}], we have investigated lengths of 3.7, 7.4, 10 and 12 kHz. We have conducted these experiments using the MATLAB framework.

The acoustic material contains 99 recordings of the three classes of sounds, in average about 15s each, 39 were used for training and 60 for testing. The testing set resulted in 685 segments of three seconds. The performance of each of the approaches we tested is presented subsequently. The performance was evaluated in terms of Identification rate, the ratio of numbers of correctly identified segments and the evaluated segments.

4.1. MFCC-GMM

We have applied GMM on the feature space consisting of Mel-cepstral features, accompanied or not by ZCR and Spectral

Flatness. We have tested several hierarchical clustering initialization methods, using different distance measures between branches, varied the number of Gaussian components, the number of the cepstral coefficients, the values of the frequency interval [f_{low}, f_{high}], and the length of the analysis frame. On the given acoustic material, the performance obtained with the PNN initialization, using the Euclidian Standardized distance, were slightly better than when using the other hierarchical methods. The MATLAB settings for analysis frame, 25ms, f_{low} =300 Hz, f_{high}=3.7 kHz are the most beneficial. Moreover, adding ZCR and SF improved the results. 12-13 GMM components and 13-14 MFCCs seemed to be the best configuration. Some results are presented in Table 1. We have chosen to assign an identical number of Gaussian components, as our acoustic material is currently quite scarce, and the investigated problem is less complex than, for instance, the task of an audio scene recognition. A more rigorous approach should consider the structure of the underlying acoustic feature space, as shown in [40], to assign the number of components to each category of sound.

Table 1: Performance of AESR using the MFCC-GMM approach, with additional ZCR and SF features, different values of f_{low} and f_{high} expressed in kHz, number of Gaussian components, number of Mel-cepstral coefficients

	f_{low}	f_{high}	Components	Coeficients	Identification Rate			
					Chainsaw	Forest	Vehicle	General
Euclidia	0.3	3.7	14	13	66.47	77.16	61.07	68.07
	0.3	4	13	13	61.85	76.72	59.92	66.27
	0	3.7	14	13	67.63	79.31	56.87	67.47
Standard	0.3	3.7	14	13	62.43	80.17	62.21	**68.52**
	0.3	3.7	15	13	60.69	80.17	62.21	68.07
	0.3	3.7	15	12	67.63	76.29	61.83	68.37

4.2. Spectra Alignment using Dynamic Time Warping

On applying DTW we have aligned Fourier Power Spectra. More precisely, we have aligned segments of these spectra defined on frequency intervals [f_{low}, f_{high}]. The frequency domain was restricted to some intervals included in [0, 7400] Hz. The length of the analysis window was set to 22ms. For analysis frames of 22ms the number of samples per frame, given the sampling frequency of 44.1 kHz, is 970, and the length of the Fourier Transform is the closest power of 2 greater than the number of signal samples, that means 1024. For shorter frequency intervals the corresponding Fourier subset involves less samples, less data to be processed, and a shorter execution time:

- [0, 7400] Hz - 171 samples;
- [0, 3700] Hz - 86 samples;
- [300, 3700] Hz - 80 samples;
- [300, 7400] Hz - 165 samples.

For this reason, the length of the analysis frame was set to 22ms. A 25ms frame would have meant a 2048 long Discrete Fourier Transform, and hence, power spectrum.

We have compared the performance of the DTW alignment for several lengths of the frequency interval [f_{low}, f_{high}], different lengths of the Sakoe- Chiba band, and the two distance measures in the calculus of the distortion measure, the Euclidian norm, and the 1-norm (sum of absolute values). The best results for analysis frames of 22ms, f_{low}=0, f_{high}=7.4 kHz, for the largest applied Sakoe- Chiba band, and the 1-norm. Some of the results can be viewed in the table 2.

Table 2: AESR Performance Using Dynamic Alignment of Power Spectra intervals of various lengths and applying Saloe-Chiba windows endowed with different widths

distance	Sakoe band	f_{low}	f_{high}	Identification Rate			
				Chainsaw	Forest	Vehicle	General
Norm1	15	0	7.4	69.57	74.14	67.94	70.44
		0	3.7	57.39	81.9	67.94	69.06
	20	0	11	50.43	76.72	67.94	65.19
		0	7.4	68.7	75.86	67.94	70.72
		0	3.7	57.39	83.62	67.94	69.61
	25	0	3.7	57.39	84.48	67.94	69.89
		0	7.4	68.7	76.72	67.94	70.99
	30	0	7.4	69.57	76.72	67.94	71.27
		0	3.7	57.39	86.21	67.94	70.44
Standardised	20	0	3.7	58.26	86.21	66.41	70.17
		0	7.4	68.7	76.72	67.18	70.72
	25	0	3.7	58.26	87.07	66.41	70.44
		0.3	7.4	66.09	77.59	67.18	70.17
		0	7.4	68.7	77.59	67.18	70.99
	30	0	7.4	68.7	77.59	67.18	70.99

4.3. Experiments using the FFNN

We have applied FFNN methodology in two hypostases: the first by feeding at input Mel-cepstral features (coming with or without Spectral Flatness, and/or ZCR) and the second, by feeding Fourier power spectrum features. In the first case we have extracted 12 to 20 Mel-cepstral coefficients, on an analysis window, using different frequency intervals [f_{low}, f_{high}], and different analysis window lengths. In the second case the number of coefficients depended on the length of the frequency interval.

At training we fed the information at the of sample level, each sample being associated with the expected outputs 1, 0 or -1, depending on the nature of the sound sample (chainsaw, genuine forest, vehicle engines). A sample in this case means a feature vector (of Mel-cepstral coefficients or Fourier spectrum coefficients, etc., calculated on an analysis window).

We applied the batch training and evaluated the BR, and LMA training algorithms. At classification, when training by feeding vectors of features, we evaluated each 3s segment by assessing each sample in the segment and finally the whole segment. A sample belongs to a certain class if its output *score* in (11) is closest to the respective class expected output, 1, 0 or -1. The overall decision on the 3s level is taken by applying one of the rules:

- Majority voting (the segment is associated with the class for which most of the samples of the segment belong to the respective class);
- Average output: the average output score of the samples on the segment is closest to the expected output of a certain class.

Concerning the network architecture, we have tested several configurations of FFNNs, 2 to 4 layers, with 6 to 10 neurons on each layer. As the performance of the test depends on the initialization of the training process, we provided 5 tests for each configuration. Because of the great choice of parameters, such as the length of the analysis window, or f_{low}, f_{high}, and configurations to be investigated, we could not exhaust all the possible combinations. The tables 3 and 4 present some relevant results.

Table 3: FFNN – MFCC -results of 5 tests using networks of 4 layers, with 9, 8, 7, 6 neurons respectively 88ms analysis windows, , [f_{low}, f_{high}] =[0, 7.4] kHz, 18 Mel-coefficients, and Spectral flatness. Training was accomplished by Bayesian Regularization, and classification using the majority voting rule

	Chainsaw	Forest	Vehicle	General
test1	52.36	93.10	62.69	70.13
test2	51.83	75.86	62.69	64.13
test3	48.69	89.66	60.00	66.91
test4	50.26	81.90	68.46	67.94
test5	58.64	93.10	61.92	71.60

Table 3 shows one the best performance, identification rates expressed in percent, obtained using Mel-cepstral features as input, using a 4 layers FFNN, with 9, 8, 7, 6 neurons on each layer, 88ms analysis windows, [0, 7.4] kHz, frequency interval for which the coefficients were computed. 18 Mel-coefficients were extracted, and Spectral flatness and ZCR added on each analysis frame. Training was accomplished by Bayesian Regularization, the default in MATLAB, and classification using the majority voting rule. The average performance was 68.14%. Similar results were obtained using other configurations, for instance an identification rate of 67.43% was achieved with a 3-layer network, 88ms analysis frame, [f_{low}, f_{high}] =[0, 10] kHz, 17 Mel cepstral coefficients, with SF added. However, all the tests provided a low identification rate for the "chainsaw" class. This performance is lower, or comparable to the ones obtained applying the classical GMM and DTW approaches.

Table 4 presents the results of 5 tests using FFNN of 4 layers, with 9, 8, 7, 6 neurons respectively, with Power Spectrum coefficients as input, 88ms analysis windows, [0, 7400] Hz frequency interval for spectral features. At training we applied the Bayesian Regularization algorithm and at classification the majority voting rule. The average performance was 78.82%.

Table 4: FFNN – Power Spectrum – results of 5 tests using networks of 4 layers, with 9, 8, 7, 6 neurons respectively, 88ms analysis windows, [f_{low}, f_{high}] =[0, 7.4] kHz, training models obtained by Bayesian Regularization and classification using the majority voting rule

	Chainsaw	Forest	Vehicle	General
test1	74.86	92.24	74.61	80.67
test2	65.96	91.37	78.46	79.35
test3	78.01	87.06	77.69	80.96
test4	67.53	80.17	78.46	75.98
test5	75.39	77.58	78.07	77.16

Tables 5 and 6 present some relevant results obtained using the LMA algorithm at training.

Table 5 presents the identification rate (in percent) of 5 tests using a 4-layernetwork, an analysis frame of 88ms, [f_{low}, f_{high}] =[0, 10] kHz, 18 Mel cepstral coefficients, without adding extra parameters. At classification, the majority voting rule was applied. The average achieved performance was 69.54%.

Table 5: FFNN – MFCC -results of 5 tests using networks of 4 layers, with 9, 8, 7, 6 neurons respectively 88ms analysis windows, , [f_{low}, f_{high}] =[0, 10] kHz, 18 Mel-coefficients. Training was accomplished using LMA, and classification using the majority voting rule

	Chainsaw	Forest	Vehicle	General
test1	58.12	90.52	70.77	73.94
test2	52.36	85.35	65.39	68.52
test3	44.50	87.07	75.00	70.57
test4	40.31	76.72	67.31	62.96
test5	49.74	87.93	72.31	71.30

Table 6 presents the results of 5 tests using networks of 3 layers, with 9, 8, 7 neurons respectively, 88ms analysis windows, and the frequency interval [0, 3700] Hz, applying LMA training and classification using the majority voting rule. The average recognition rate was 79.17%.

Table 6: FFNN applied on Power Spectra results of 5 tests using networks of 3 layers, with 9, 8, 7 neurons respectively, 88ms analysis windows, , [f_{low}, f_{high}] = [0, 3700] Hz, training models obtained by LMA and classification using the majority voting rule

	Chainsaw	Forest	Vehicle	General
test1	72.77	84.48	77.69	78.62
test2	75.39	79.31	85.76	80.67
test3	69.11	89.65	69.23	76.13
test4	69/63	90.08	79.61	80.38
test5	64.92	87.93	84.23	80.08

Table 7: FFNN applied on Power Spectrum results of 5 tests using networks of 4 layers, with 10, 9, 8, 7 neurons respectively, 88ms analysis windows, [f_{low}, f_{high}] = [0, 3700] Hz, training using LMA and classification and the average score on the 3s frames

	Chainsaw	Forest	Vehicle	General
test1	50.26	92.24	65.769	70.42
test2	61.25	95.69	70.769	76.57
test3	43.97	94.82	65.00	69.25
test4	53.40	93.10	70/38	73.35
test5	57.59	98.27	62.30	73.20

Table 8: Average Identification rates obtained using FFNN applied on Power Spectra, with 3 layers with 9, 8, 7 neurons, respectively, using different values for f_{high} and lengths of the analysis frame, training with Bayesian Regularization and majority vote classification.

		analysis frame lengths		
		22ms	44ms	88ms
Frequency interval (Hz)	[0, 3700]	71.04	76.31	78.83
	[0, 7400]	74.98	76.02	77.22
	[0, 10000]	69.64	74.76	76.02
	[0, 12000]	72.61	75.17	74.00

Table 7 contains the results of 5 tests applying FFNN to spectral coefficients, calculated on the frequency domain [0, 3.7] kHz, using a 88ms analysis windows, and applying the LMA training on 4-layer networks with 10, 9, 8, 7 neurons. The classification algorithm used the average score on 3s frames. The average performance was 72.56%.

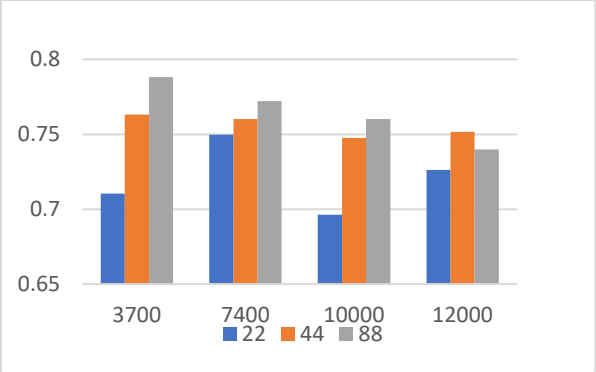


Figure 14: Average Identification rates for different values of f_{high} and length of the analysis frame

As an overall conclusion of the results, the Fourier spectrum as input to FFNN yielded very good results when applying the classification majority vote rule. The average score rule produced poorer results, with a low performance for the "chainsaw" class, but they are still better than using Mel-cepstral analysis or the GMM and DTW approaches. The BR and LMA produced comparable results, maybe LMA results were more balanced among the 5 tests (the standard deviation among the identification rates is lower). Concerning the network architecture for the Fourier spectrum variants of 2, 3 or 4 layers produce comparable results, especially when using the majority voting rule. The LMA training resulted in performance quite similar results as those obtained using the BR, for many configurations besides the one illustrated in Table 6, and the results are well balanced among the three classes of sounds. Perhaps the identification rates for the "chainsaw" class are a bit lower. In what concerns the average score classification, the 3 layers FFNN seemed to work better than 4-layer nets.

Concerning the analysis window, the results are better in all the cases for lengths of 44ms or 88ms. Table 8 and Figure 14 present the average overall identification rates for the FFNN applied on power spectra using the BR training, and majority voting at classification, several analysis window lengths and frequency intervals. The best average score is obtained for the spectrum restricted to [0, 3700]Hz, and an analysis window of 88ms, but in fact the results are very close among the frequency intervals. Among the 5 tests for each configuration there were many identification rates above 80%.

With regard to the results obtained using the Mel-cepstral coefficients as input, the conclusion concerning the optimum analysis window length is that window lengths greater than 44ms produced better performance. The frequency intervals [0, 7.4] kHz and [0, 10] kHz yielded better results. The general conclusion is that adding Spectral flatness and sometimes ZCR helped to increase the performance, although the example of Table 5 is an exception.

4.4. Experiments using the LSTM

In the experiments using LSTM we used the same input as in the FFNN experiments. The number of hidden units was set to 100 and each cell configured with 5 layers, the default MATLAB configuration. Table 9 presents the best results obtained so far by applying LSTM. We have fed as input 18 dimensional sheer Mel-cepstral vectors, calculated on 44ms analysis window and filtering the frequency domain to [0, 12] kHz. The average performance among the 5 tests is 64.85%. As can be seen the identification rates are unbalanced among the three classes. In any other configurations the results were even worse.

Table 9: Results of 5 tests using LSTM applied on an input set of Mel-cepstral 18-dimensional vectors, calculated on 44ms analysis windows, and the frequency interval of [0, 12] kHz

	Chainsaw	Forest	Vehicle	General
test1	48.69	86.20	47.69	61.05
test2	37.69	93.10	56.92	63.83
test3	42.40	97.41	45	62.07
test4	67.53	90.51	47.69	67.78
test5	62.82	93.10	53.46	69.54

Concerning the experiments using as input the Fourier spectrum we failed to obtain interesting results, as the network did not behave well at training Figures 15, 16 present the estimation of the achieved accuracy during the training process for LSTM applied to Mel-cepstral input and power spectra respectively. While the first process achieves maximum accuracy in less than 100 iterations the LSTM applied to power spectra achieves less than 80% in more than 300 iterations.

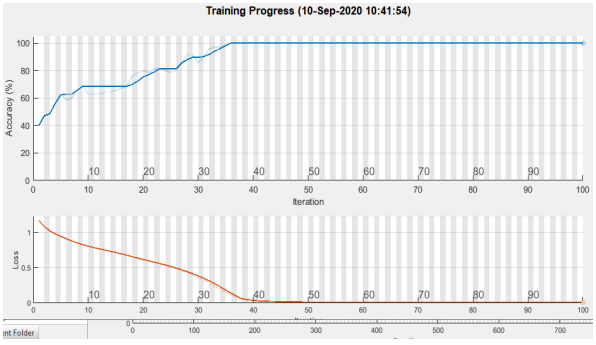


Figure 15: Accuracy estimation during training for a LSTM - MFCC process

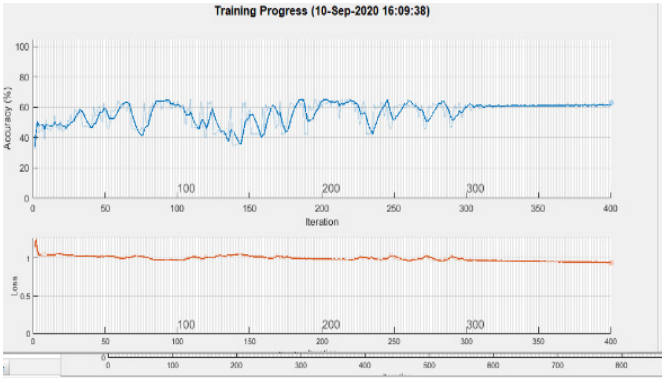


Figure 16: Accuracy estimation during training for a LSTM applied on Fourier power spectra.

Regarding the experiments using as input the Fourier spectrum we failed to obtain interesting results, as the network did not behave well at training Figures 15, 16 present the estimation of the achieved accuracy during the training process for LSTM applied to Mel-cepstral input and power spectra respectively. While the first process achieves maximum accuracy in less than 100 iterations the LSTM applied to power spectra achieves less than 80% in more than 300 iterations.

5. Conclusions and future work

The goal of this study was to test some state-of-the-art methodologies applied in AESR, Gaussian Mixtures Modelling, Dynamic Time Warping, and two types of Deep Neural Networks, in the context of forest acoustics. Another specific objective was to evaluate the behaviour of these techniques, in several configurations, such as different lengths of the analysis window, or find the frequency intervals on which the Fourier spectrum is more relevant for such type of applications.

We have succeeded to achieve significantly better performances using Feed Forward Neural Networks, in a certain setup, compared to the classical methods, GMM, and DTW. We used two types of networks (Deep Feedforward Neural Network and LSTM) and have fed as inputs two types of data, Mel-cepstral and Fourier power spectral coefficients. In this context we tested two training methods, the Levenberg-Marquardt Algorithm, and the Bayesian Regularization, and two different classification approaches.

Deep Feed Forward Neural Networks experiments output the best results when using the sheer spectral features, and especially when using the majority voting rule, with an average identification rate of over 78%, with about 10% higher than other methods performance. This fact suggests that FFNN, based on Fourier spectral features, using a less complex processing sequence, is able to produce more valuable features than the elaborate Mel cepstral analysis. A difference is in the number of features at input, while the Mel features are fewer than 20, the spectrum on [0, 7400] Hz frequency interval means about 170 coefficients.

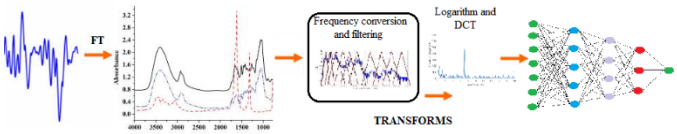


Figure 17: FFNN using Mel cepstral input applies a range of transforms (frequency conversion. spectrum filtering, logarithm, Discrete Cosinus Transform) on the Fourier spectrum and feeds the result to the network

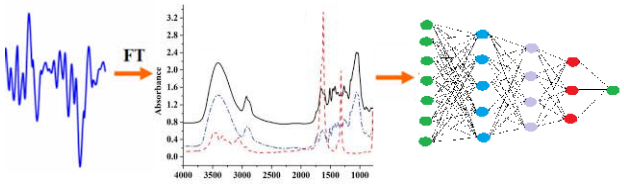


Figure 18: FFNN using Fourier Spectrum coefficients as has a simpler schema, and probably devises more valuable features through the layers of the network

Figures 17, 18 summarize this idea. Figure 17 presents the more complex row of operations to be accomplished on the power spectrum when the input to the FFNN involves Mel-cepstral analysis. Figure 18 presents the straightforward processing of spectrum by the FFNN, when just spectral coefficients are fed to the network.

The disappointing results using the LSTM network may have several reasons. One of them may be the unproper use of the LSTM MATLAB tool. A second reason may reside in the fact that this type of network might be not suited to the kind of problem we want to solve.

Another advantage of using FFNN is the fact that it is easy to implement in programming environments other than MATLAB. While the models can be generated in MATLAB, the classification part can be implemented in other programming languages, like C++, Java, etc. using the parameters established at training.

Concerning the length of the analysis window the experimental results have shown that its length must be set above 44ms or higher. We have chosen the length of the analysis window somehow empirically, therefore the use of an analytical approach, e.g., [41], to establish the proper length of the frame would be a future direction of research.

We could not draw a well-founded conclusion about the optimal frequency interval, as for 3.7 kHz to 10 kHz, the results do not vary too much.

Although the neural networks have apparently the advantage of training jointly several classes of data, this did not result in better results in comparison with the classical methods.

As future work we intend to extend our research by including the CNN framework.

Another important objective would be investigation of methods to merge decision of several sources, possibly by using a probabilistic logic.

Another important objective is to extend the field of research to other AESR applications, in the field of scientific environment monitoring (e.g., detect bird or species), or early detection of disasters such as land sliding or avalanches, where acoustic emissions are among the data used as input.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work has been supported in part by UEFISCDI Romania through projects SeaForest, CitiSim, SenSyStar and WINS@HI, and funded in part by European Union's Horizon 2020 research and innovation program under grant agreement No. 777996 (SealedGRID project) and No. 643963 (SWITCH project).

References

[1] S. Segarceanu, E. Olteanu, G. Suciu, "Forest Monitoring Using Forest Sound Identification," in 2020 International Conference on Telecommunications and Signal Processing (TSP-2020), 346-349 2020, doi: 10.1109/TSP49548.2020.9163433.

[2] S. Cowling, M. Cowling, R. Sitte, "Recognition of Environmental Sounds Using Speech Recognition Techniques," Advanced Signal Processing for Communication Systems, 31-46, 2006,. doi : 10.1007/0-306-47791-2_3.

[3] X. Guo, Y. Toyoda, H. Li, J. Huang, S. Ding, Y. Liu, "Environmental Sound Recognition Using Time-Frequency Intersection Patterns," Applied Computational Intelligence and Soft Computing, vol. 2012, 6 pages, 2012, https://doi.org/10.1155/2012/650818.

[4] S. Sivasankaran, K.M.M. Prabhu, "Robust features for Environmental Sound,"

in 2013 IEEE Conference on Electronics, Computing and Communication Technologies (CONECCT), 1-6, 2013, doi:10.1109/CONECCT.2013.6469297.

[5] S. Sigtia, A. Stark, S. Krstulovic, M. Plumbley, "Automatic environmental sound recognition: Performance versus computational cost," IEEE/ACM Transactions on Audio, Speech, and Language Processing, **24**(11). 2096-2107, 2016, doi: 10.1109/TASLP.2016.2592698.

[6] M. Babiš, M. Ďuríček, V. Harvanová, M. Vojtko, "Forest Guardian – Monitoring System for Detecting Logging Activities Based on Sound Recognition," in 2011 7th Student Research Conference in Informatics and Information Technologies, **2**, 354-359, 2011.

[7] J. Papán, M. Jurecka, J. Púchyová, "WSN for Forest Monitoring to Prevent Illegal Logging," in 2012 IEEE Proceedings of the Federated Conference on Computer Science and Information Systems, 809-812, 2012.

[8] R. Buhuș, L. Grama, C. Rusu, 2017, "Several Classifiers for Intruder Detection Applications," in 2017 9th International Conference on Speech Technology and Human-Computer Dialogue (SpeD), 2017, doi: 10.1109/SPED.2017.7990432.

[9] L. Grama, C.Rusu, "Audio Signal Classification Using Linear Predictive Coding And Random Forest," in 2017 9th International Conference on Speech Technology and Human-Computer Dialogue (SpeD), 1-9, 2017, doi: 10.1109/SPED.2017.7990431.

[10] S. Zhang, D. Gao, H. Lin, Q. Sun, "Wildfire Detection Using Sound Spectrum Analysis Based on the Internet of Things," in 2019 Sensors, **19**(23), 2019, doi: 10.3390/s19235093.

[11] L. A.Venier, M. J. Mazerolle, A. Rodgers, K. A. McIlwrick, S. Holmes, D. Thompson, "Comparison of semiautomated bird song recognition with manual detection of recorded bird song samples," Avian Conservation and Ecology, **12**(2), 2017, doi:10.5751/ ACE-01029-120202.

[12] D. Mitrovic, M. Zeppelzauer, C. Breiteneder, "Discrimination and Retrieval of Animal Sounds," in 2006 12th International Multi-Media Modelling Conference. IEEE, 2006, doi: 10.1109/MMMC.2006.1651344.

[13] S.Chandrakala, S. Jayalakshmi, "Environmental Audio Scene and Sound Event Recognition for Autonomous Surveillance: A Survey and Comparative Studies," ACM Computing Surveys, **52**(3), 1-34, 2019, doi: 10.1145/3322240.

[14] N. Dixon, A., Smith, J. A. Flint, R. Khanna, B. Clark, M. Andjelkovic, "An acoustic emission landslide early warning system for communities in low-income and middle-income countries," Landslides, **15**, 1631–1644, 2018, doi:10.1007/s10346-018-0977-1.

[15] Z. M. Hafizi, C.K.E. Nizwan, M.F.A. Reza & M.A.A. Johari, "High Frequency Acoustic Signal Analysis for Internal Surface Pipe Roughness Classification," Applied Mechanics and Materials, **83**, 249-254, 2011, doi:10.4028/www.scientific.net/AMM.83.249.

[16] V. Gibiat, E. Plaza, P. De Guibert, "Acoustic emission before avalanches in granular media," The Journal of the Acoustical Society of America, **123**(5), 3270, 2008,doi: 10.1121/1.2933600 .

[17] T. Thüring, M. Schoch, A. van Herwijnen, J. Schweizer, "Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays," Cold Reg. Sci. Technol., **111**, 60–66, 2015. doi: 10.1016/j.coldregions.2014.12.014

[18] L. G. Evers, D. Brown, K. D. Heaney, J. D. Assink, P. S.M. Smets, M.Snellen, "Infrasound from underwater sources," The Journal of the Acoustical Society of America, **137**(4), 2372-2372, 2015, doi: 10.1121/1.4920617.

[19] M. Cowling, R Sitte, T. Wysock, "Analysis of Speech Recognition Techniques for use in a Non-Speech Sound Recognition System," Advanced Signal Processing for Communication Systems. Springer, 2002.

[20] S. Chachada, C.-C. Jay Kuo , "Environmental Sound Recognition: A Survey," APSIPA Transactions on Signal and Information Processing, **3**, 2014, doi: 10.1109/APSIPA.2013.6694338.

[21] I. V. McLoughlin, Hao-min Zhang, Zhi-Peng Xie, Yan Song, Wei Xiao, "Robust Sound Event Classification using Deep Neural Networks. Audio, Speech, and Language Processing," IEEE/ACM Transactions on Audio, Speech and Language Processing, **23** (3). 540-552, 2015, doi:10.1109/TASLP.2015.2389618) (KAR id:51341.

[22] T. Lidy, A. Schindler, "CQT-based convolutional neural networks for audio scene classification," in 2016 Proceedings of the Detection and Classification of Acoustic Scenes and Events Workshop (DCASE'16), **90,** 1032-1048. doi: 10.1145/3322240.

[23] A. Rabaoui, M. Davy, S. Rossignol, N. Ellouze, "Using one-class SVMs and wavelets for audio surveillance," IEEE Trans. on Information Forensic and Security, **3(**4), 763–775, 2008, doi: 10.1109/TIFS.2008.2008216.

[24] D. Stowell, D. Clayton., "Acoustic event detection for multiple overlapping similar sources," in 2015 Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'15). 1–5, 2015, doi: 10.1109/WASPAA.2015.7336885.

[25] H. Phan, P. Koch, L. Hertel, M. Maass, R. Mazur, A. Mertins, "CNN-LTE: A class of 1-X pooling convolutional neural networks on label tree embeddings for audio scene classification," in 2017 Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP'17), 136-140, 2017, doi: 10.1109/ICASSP.2017.7952133.

[26] Karol J. Piczak, "Environmental sound classification with convolutional neural networks," in 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), 1–6, 2015, doi: 10.1109/MLSP.2015.7324337.

[27] Y. Han, K. Lee, "Convolutional neural network with multiple-width frequency-delta data augmentation for acoustic scene classification," in 2016 Proceedings of the IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events, 2017.

[28] R. Ahmed, T. I. Robin, A. A. Shafin, "Automatic Environmental Sound Recognition (AESR) Using Convolutional Neural Network," International Journal of Modern Education and Computer Science, **12**(5), 41-54, 2020, doi: 10.5815/ijmecs.2020.05.04.

[29] H. Eghbal-zadeh, B. Lehner, M. Dorfer, G. Widmer, "A hybrid approach with multichannel i-vectors and convolutional neural networks for acoustic scene classification," in 2017 25th European Signal Processing Conference (EUSIPCO), 2749-2753, 2017, doi: 10.23919/EUSIPCO.2017.8081711.

[30] D. Reynolds, "Gaussian mixture models," Encyclopedia of Biometrics, 659–663, 2009, doi:10.1007/978-0-387-73003-5_196.

[31] X. Huang, A. Acero, H. Hon, Spoken Language Processing:A Guide to Theory, Algorithm, and System Development, Prentice Hall, 2001.

[32] O.Virmajoki, P. Fränti, "Fast pairwise nearest neighbor based algorithm for multilevel thresholding," Journal of Electronic Imaging, **12(4)**, 648-659, 2003, doi: 10.1117/1.1604396.

[33] S. B. Davis, P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE Trans. on ASSP, **28**, 357–366, 1980, doi: 10.1109/TASSP.1980.1163420.

[34] S. Furui, "Cepstral analysis technique for automatic speaker verification," IEEE Transactions on Acoustics, Speech, and Signal Processing, **29**, 254–272, 1981, doi: 10.1109/TASSP.1981.1163530.

[35] P. Senin,"Dynamic Time Warping Algorithm Review," Technical Report Information and Computer Science Departament University of Hawaii, 2008.

[36] H. Sakoe, S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," Trans. Acoustics, Speech, and Signal Proc., **26**, 43-49, 1978, doi:10.1109/TASSP.1978.1163055.

[37] V. Zacharias, AI for Data Science Artificial Intelligence Frameworks and Functionality for Deep Learning, Optimization, and Beyond, Technics Publications, 2018.

[38] M. H. Beale, M. T.Hagan, H. B. Demuth, Neural Network Toolbox™ User's Guide, The MathWorks, Inc., Natick, Mass., 2010.

[39] H. Sak, A. Senior, F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in 2014, Fifteenth annual conference of the international speech communication association (INTERSPEECH-2014), 338-342, 2014.

[40] M. Kuropatwiński, "Estimation of Quantities Related to the Multinomial Distribution with Unknown Number of Categories," in 2019 Signal Processing Symposium (SPSympo), 277–281, 2019, doi:10.1109/sps.2019.8881992.

[41] E. Punskaya, C. Andrieu, A. Doucet, W. Fitzgerald, "Bayesian curve fitting using MCMC with applications to signal segmentation," IEEE Transactions on signal processing, **50**(3), 747-758, 2002, doi**:** 10.1109/78.984776.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 5, 16-23 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Learning the Influence between Partially Observable Processes using Score-based Structure Learning

Ritesh Ajoodha*, Benjamin Rosman

School of Computer Science and Applied Mathematics. The University of the Witwatersrand, Johannesburg. 2050. South Africa.

ARTICLE INFO

Article history:
Received: 03 July, 2020
Accepted: 11 August, 2020
Online: 08 September, 2020

Keywords:
Stochastic Processes
Dynamic Bayesian Networks
Structure Learning

ABSTRACT

The difficulty of learning the underlying structure between processes is a common task found throughout the sciences, however not much work is dedicated towards this problem. In this paper, we attempt to use the language of structure learning to address learning the dynamic influence network between partially observable processes represented as dynamic Bayesian networks. The significance of learning an influence network is to promote knowledge discovery and improve on density estimation in the temporal space. We learn the influence network, defined by this paper, by learning the optimal structure for each process first, and thereafter apply redefined structure learning algorithms for temporal models. Our procedure builds on the language of probabilistic graphical model representation and learning. This paper provides the following contributions: we (a) provide a definition of influence between stochastic processes represented by dynamic Bayesian networks; (b) expand on the conventional structure learning literature by providing a structure score and learning procedure for temporal models; and (c) introduce the notion of a structural assemble which is used to associate two stochastic processes represented by dynamic Bayesian networks.

1 Introduction

The problem of describing the interaction or influence between stochastic processes has received little scrutiny in the current literature, despite its growing importance. Solving this complex problem has large implications for density estimation and knowledge discovery. In particular, for making predictions about later aspects of the process, or even for learning how processes influence each other.

Usually, the individual structure of each stochastic process is ignored and all are merged into one big process which is modelled by some probabilistic temporal model. This approach undermines the explanatory importance of the relations between these processes. [1] has explored the problems with this approach. The core of the issue mentioned by [1], is that we lose the underlying structure of the relationships between the processes which is essential to learn how one process influences another.

In this paper, we provide a complete method for learning the dynamic influence network between processes. This paper also explores the case when we are learning the influence relationship between partially observable processes. This is a significantly harder problem since the likelihood of the temporal model to the data has multiple optima which is induced from the missing samples [1]. Unfortunately, given that learning parameters from missing data is also a NP-hard problem, heuristic approaches are then needed to solve for a suitable local optimum of the likelihood function of the parameters to the data [1].

We assess this problem by providing an algorithm to learn the influence relations between partially observable stochastic processes by building on the language of probabilistic graphical modelling. In particular, we consider structure learning which searches for an appropriate structure by using scoring metrics and provide evidence for the effectiveness of our approach over the standard benchmarks selected. We notice that our derived penalty-based score paired with a greedy structure search outperforms using a random structure or a tree structure built using the maximum weighted spanning tree algorithm.

The application of this research is broad. Influence networks for stochastic processes can capture the complex relationships of how processes impact others. For example, we can learn the influence of traffic in a network of roads to determine how the traffic condition of a road congestion will impact on another road. In educational data-mining we may want to determine the influence of participants in a lecture environment to encourage student success. We may wish to learn the influence between an IoT network [1, 2]; influence in music [3]; or influence between the skills of learners or their attrition [4, 5].

*Corresponding Author: Ritesh Ajoodha, The University of the Witwatersrand, Johannesburg, +27 74 418 3978 & ritesh.ajoodha@wits.ac.za

An overview of the proposed algorithm in this paper is given by the below instructions. This algorithm is expanded on later in the paper.

(i) The stochastic processes are given as input.

(ii) The parameters for a dynamic Bayesian network is learned for each of the stochastic processes (temporal structure remains static - time invariant and Markov).

(iii) A structure is imposed between the dynamic Bayesian networks (using a relation function called an assemble). This gives us a dynamic influence network (DIN). The parameters for the DIN are relearned.

(iv) The structure score for the DIN is computed.

(v) A structure search algorithm is initiated to find a DIN structure which is has an optimal likelihood to the observable data;

(vi) The optimal DIN is output.

The following contributions is made by this paper:

- The concept of dynamic influence networks (DINs) representing the influence (relationships) between partially observable stochastic processes.
- The derivation of a dynamic Bayesian information criterion (d-BIC score) for DINs.
- The concept of a structural assemble which is able to relate dynamic Bayesian networks.
- The greedy structure learning procedure for learning DINs.

The following structure is used by this paper: Section 2 provides the background and related work on DINs; Section 3 defined the representation of DINs between partially observable stochastic processes; Section 3.3 derives the notion of a dynamic structure score using the notion of an assemble; Section 3.5 provides a greedy structure learning learning algorithm for learning DINs; the results and discussion is illustrated by Section 4; and lastly, Section 5 provides the conclusion of the research and suggestions on future work.

2 Related Work

Many statistical procedures have been used to identify influence between variables [6]–[9]. These statistical procedures have been extended to the temporal environment to learn relationships the between processes (variables over time). A significant contribution is the use of dynamic Bayesian networks which is defined as a set of parameters and conditional independence assumptions which together make up an acyclic structure between variables defined using factors [10, 11]. The values in these factors are referred to as the parameters, and the list of conditional independence assumptions between variables are referred to as the structure of the dynamic Bayesian network.

Learning the independence assertions of a dynamic Bayesian network can be used to make conditional independence inferences over time (density estimation) or to simply learn the relationships between variables (knowledge discovery) [12]–[15]. On the one hand, a sparse graph structure may have more generalisability for density estimation, and on the other hand having a more dense graph can reveal unknown relationships for knowledge discovery. Care must be taken when considering for what purpose is the network required (more on this in the discussion) [11].

A successful approach to structure learning is using score-based structure learning [11, 16]. In score-based structure learning we develop a set of hypothesis structures which are evaluated using a score-based function that computes the likelihood of the data to the hypothesised structure. The likelihood is usually expressed as the information gain (mutual information) of the structure and parameters of the distribution to the data.

A search algorithm is then performed to identify the highest (possible) structure based on the structure score [17]–[20]. Viewing this problem as an optimisation problem allows us to adopt the already established literature on search methods in this super-exponential space to find the optimal structure given the data [21]–[24].

The structure of this section is as follows. In subsection 2.1 we introduce the well established BIC score which offers a way to trade-off the fit to data vs model complexity (the amount of independence assumptions between variables in the data). Finally, in subsection 2.2 we introduce a greedy search method to find the an optimal graph structure.

2.1 The BIC score

The BIC score models the structural fit to data verses the complexity of the conditional independence assumptions between variables, that is, the amount of independence assumptions made on the structure [25]. This makes it a popular choice for structure learning methods since the model complexity has a direct impact on the performance of inference tasks. This is because the amount of conditional independence assumptions on a particular variable increases the factor size of that variable exponentially. The mathematical expression of the BIC score comprises of two terms: the first term models the fit to data; and the second term penalises the fit to data based on the complexity of the structure considered. The complete BIC score is as follows:

$$score_{BIC} = \ell(\hat{\theta}_{\mathcal{G}} : \mathcal{D}) - \frac{\log M}{2} DIM[\mathcal{G}],$$

where the count of instances is denoted by M and the count of independent parameters is denoted by $DIM[\mathcal{G}]$ in the Bayesian network.

The intuition of the Bayesian score is that as the amount of samples increase (ie. M) the score is willing to consider more complicated structures if enough evidence (samples, ie. M) is considered [26, 27]. The BIC core is particularly effective since the likelihood score (one without a penalty to complexity) will always prefer the most complicated network. However, the most complex networks also impose the risk of fragmentation, which is the exponential increase to the size of the factors caused by the increase of the in-degree of a variable. Penalty-based structure scores allows us

to explore the opportunity to adopt more complicated structures if there is enough justification that the likelihood of the structure and parameters to the data is high-enough to compromise on the models speed to perform inference tasks caused by fragmentation.

There has been much contributions in the literature on the properties of the BIC score [25, 28, 29]. Key constitutions include a proof the it is consistent and is score equivalent which are necessary for efficient search procedures [30]–[32].

2.2 Learning General Graph-structured Networks

Since the search space for the optimal Bayesian structure is super-exponential, the difficulty of learning a graph structure for a Bayesian network is NP-hard. More specifically, for any $d \geq 2$, the problem of finding a structure with a maximum score with d parents is NP-hard [21]–[24]. See [33, 34] for a detailed proof.

Despite this, there have been many contributions to learning an optimal structure. A key contribution is using heuristic search procedure to find an optimal acyclic graph structure [35]. These heuristic search procedures make use of search operators (changes to the graph structure) and a search algorithm (e.g. greedy search, best first search or simulated annealing) [36]. The intuition of this approach is find an optimal acyclic structure by gradually improving the choice of the structure using the search operators [37]–[41].

3 Dynamic Influence Networks

We present the following algorithm to learn dynamic influence networks between a set of partially observable processes:

(i) Our stochastic processes are given as a set of partially observed data. This is the input.

(ii) From this data, we learn a dynamic Bayesian network for each partially observable stochastic processes. Expectation maximisation is used to learn the latent variables.

(iii) Build a network with the set of independence assumptions and relearn the parameters for that model.

(iv) Perform expectation maximisation once again to relearn the latent parameters of the resulting network.

(v) Evaluate the resulting dynamic influence network using a scoring function and structural assemble.

(vi) Determine if convergence has occurred or if we exceed the threshold for convergence.

(vii) Apply the structural operator to the model and reevaluate the dynamic influence network using a structure score. Repeat steps (iii - vii) until the structure score can not be improved or a threshold is reached.

(viii) Output the resulting network.

Figure 1 provides a flowchart of our method to learn dynamic influence networks between partially observable processes.

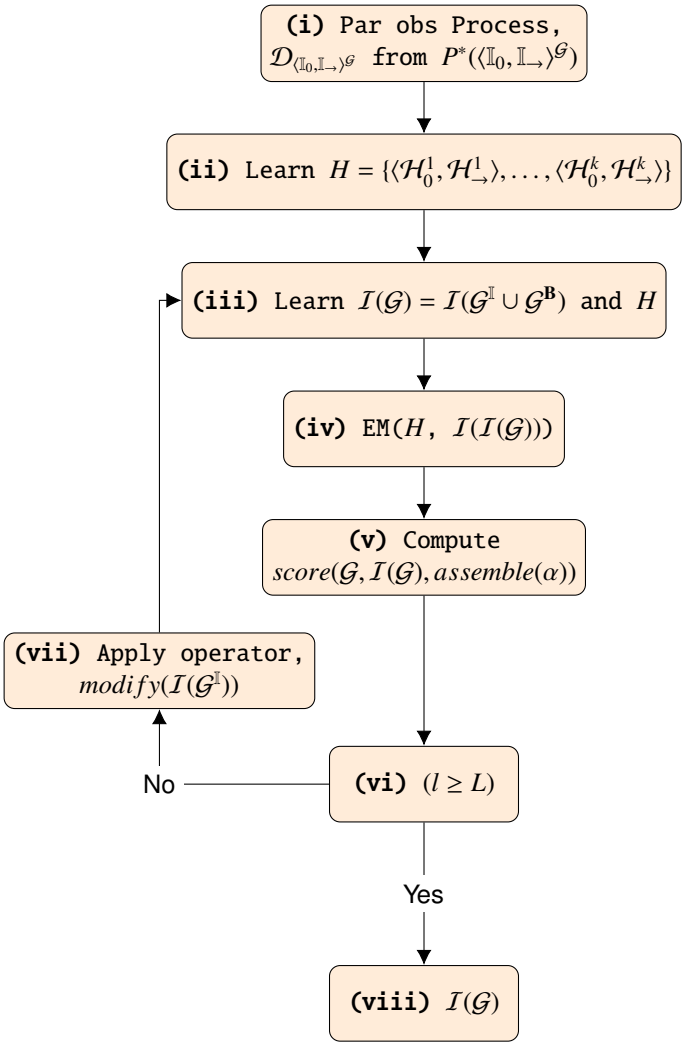


Figure 1: A flowchart of our method to learn dynamic influence networks between partially observable processes.

3.1 Assumptions

There are various assumptions we need to make about our dynamic influence network (DIN). For the below definitions, we denote $\mathcal{B}_i^{(t)}$ to be a shorthand for a Bayesian network $\mathcal{B}_i$ at time-point t.

Time Granularity Assumption We select a time-granularity, denoted $\triangle$, to split observable data into temporal time-slices at different intervals. We use the notation $t\triangle$ to represent the influence state with t time-slices.

The Markov Assumption We also adopted the Markov assumptions between consecutive states.

Definition 1 *The Markov assumption is satisfied for a DIN over the template Bayesian networks,* $\mathbf{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_R\}$*, if for all* $t \geq 0$*,* $(\mathbf{B}^{(t+1)} \perp\!\!\!\perp \mathbf{B}^{(0:(t-1))} \mid \mathbf{B}^{(t)})$.

The Time-Invariance Finally, we assume that the unrolled template structure of the DIN (which persists through time) does not change.

3.2 Dynamic Influence Networks

Given the assumptions above we define the dynamic influence network as follows:

Definition 2 *A dynamic influence network, denoted as DIN, is a couple* $\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle$*, where* $\mathbb{I}_0$ *is an influence network over the set of Bayesian networks,* $\mathbf{B}^{(0)} = \{\mathcal{B}_1, \ldots, \mathcal{B}_R\}$*, representing the starting distribution and* $\mathbb{I}_{\rightarrow}$ *is a 2-time-slice influence network for the rest of the influence distribution (*$P(\mathbf{B}' \mid \mathbf{B}_I) = \prod_{i=1}^{R} P(\mathcal{B}_i^{'} \mid Pa_{\mathcal{B}_i^{'}})$*). For any specified time-span* $T \geq 0$*, the joint distribution over* $\mathbf{B}^{(0:T)}$ *is defined as an unrolled influence network, where, for any* $i = 1, \ldots, n$*: the structure and conditional probability assumptions between variables of* $\mathcal{B}_i^{(0)}$ *are the same as those for* $\mathcal{B}_i$ *in* $\mathbb{I}_0$*; and the structure and conditional probability assumptions between variables of* $\mathcal{B}_i^{(t)}$ *for* $t > 0$ *are the same as those for* $\mathcal{B}_i^{'}$ *in* $\mathbb{I}_{\rightarrow}$*.*

3.3 The Structure Score

In this paper we adapt the celebrated Bayesian information criterion (BIC) to a dynamic Bayesian information criterion (d-BIC) for our dynamic influence networks. The d-BIC score make the same trade-off between model complexity and fit to the data, only the d-BIC can be applied to dynamic networks.

The d-BIC score is as follows:

$$\text{score}_{BIC}(\mathcal{H}_0 : \mathcal{D}) = M \sum_{k=1}^{K} \Big(\sum_{t=1}^{T} \Big(\sum_{i=1}^{N} \big(\mathbf{I}_{\hat{P}} \big(X_i^{\langle \mathcal{H}_0^k, \mathcal{H}_{\rightarrow}^k \rangle^{(t)}} ; \mathbf{Pa}^{\mathcal{G}}_{X_i^{\langle \mathcal{H}_0^k, \mathcal{H}_{\rightarrow}^k \rangle^{(t)}}} \big) \big) \Big) \Big)$$

$$- \frac{\log M}{c} DIM[\mathcal{G}],$$

where the amount of samples is given by M; the amount of dependency models is given by K; the amount of time-slices is given by T for any dependency model; the amount of variables in each time-slice is given by N; $\mathbf{I}_{\hat{P}}$ denotes the information gain in terms of the empirical distribution; and $DIM[\mathcal{G}]$ is the amount of independent parameters in the entire DIN.

The d-BIC score is designed to exchange the complexity of the dynamic influence network, $\frac{\log M}{c} DIM[\mathcal{G}]$, for the fit to the data, $\mathcal{D}$. As the amount of samples increases, the information gain term grows linearly, and the model complexity part logarithmically grows. The intuition of the d-BIC score is that we will be willing to consider more complicated structures, if we have more data that justifies the need for a more complex model (i.e. more conditional independence assumptions).

3.4 Structure Assembles

Choosing the set of parent variables in a DIN establishes the notion of a structural assemble. A structural assemble is a template which relates temporal models. The structural assemble defines the parent sets for variables to construct an dynamic influence network. More sepcifically, the assemble relation is defined as follows:

Definition 3 *Consider a family of dynamic Bayesian networks (D), where* $\langle D_0^0, D_{\rightarrow}^0 \rangle$ *represents the child with the parent set* $\mathbf{Pa}^{\mathcal{G}}_{\langle D_0^0, D_{\rightarrow}^0 \rangle} = \{\langle D_0^1, D_{\rightarrow}^1 \rangle, \ldots, \langle D_0^k, D_{\rightarrow}^k \rangle\}$*. Further assume that* $\mathcal{I}(\langle D_0^j, D_{\rightarrow}^j \rangle)$ *is the same for all* $j = 0, \ldots, k$*. Then the delayed dynamic influence network, denoted by* $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle$*, will satisfy all the independence assumptions in* $\mathcal{I}(\langle D_0^j, D_{\rightarrow}^j \rangle)$ $\forall\, i = 0, \ldots, k$*. In addition,* $\forall\, j$ *and* $\forall\, t$*,* $\langle \mathcal{A}_0, \mathcal{A}_{\rightarrow} \rangle^{(t)}$ *also satisfies the following independence assumptions for each hidden or latent variable denoted* L_i *and some* $t > \alpha \in \mathbb{Z}^+$*:*

$$\forall\, L_i^{\langle D_0^0, D_{\rightarrow}^0 \rangle^{(t)}} : \big(L_i^{\langle D_0^0, D_{\rightarrow}^0 \rangle^{(t)}} \perp\!\!\!\perp NonDescendants_{L_i^{\langle D_0^0, D_{\rightarrow}^0 \rangle^{(t)}}} \mid L_i^{\langle D_0^k, D_{\rightarrow}^k \rangle^{(t)}}, L_i^{\langle D_0^k, D_{\rightarrow}^k \rangle^{(t)-1}}, \ldots, L_i^{\langle D_0^k, D_{\rightarrow}^k \rangle^{(t)-\alpha}}, Pa_{L_i}^{\langle D_0^0, D_{\rightarrow}^0 \rangle^{(t)}} \big).$$

The assemble above an expressive representation to capture influence relationships that persist through time between temporal models. However, the choice of α is important since choosing a large α will render many dependencies on variables cause a fragmentation bottleneck, and therefore a larger computational burden for learning and inference tasks.

3.5 Structure Search

At this point we have the following well-defined optimisation problem:

1. A training set $\mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}} = \{\mathcal{D}_{\langle D_0^1, D_{\rightarrow}^1 \rangle}, \ldots, \mathcal{D}_{\langle D_0^k, D_{\rightarrow}^k \rangle}\}$, where $\mathcal{D}_{\langle D_0^i, D_{\rightarrow}^i \rangle} = \{\xi_1, \ldots, \xi_M\}$ is a set of M instances from underlying ground-truth DBN $\langle D_0^i, D_{\rightarrow}^i \rangle$;
2. a structure score: $\text{score}(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle : \mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}})$;
3. and, finally, we have an array of L distinct candidate structures, $\mathcal{G} = \{\mathcal{G}^1, \ldots, \mathcal{G}^L\}$, where each structure $\mathcal{G}^l$ represents a unique list of condition independence assertions $\mathcal{I}(\mathcal{G}) = \mathcal{I}(\mathcal{G}^{\mathbb{I}} \cup \mathcal{G}^{\mathbf{B}})$.

Our objective of this optimisation problem is output the DIN which produces the maximum score. We present the following influence structure search algorithm in Algorithm 1, where $\mathcal{S} = \{\mathcal{S}_{\rightarrow}^1, \ldots, \mathcal{S}_{\rightarrow}^P\}$ represents the set of stochastic processes; *assemble*, is the option of the parameters for an assemble relation; and *score*, the selected scoring function used for the search procedure.

Algorithm 1: Influence structure search

Input: $\mathcal{S} = \{\mathcal{S}_{\rightarrow}^1, \ldots, \mathcal{S}_{\rightarrow}^P\}$, *assemble*, *score*
Output: $\mathcal{G}_i$
For each process we learn a temporal model ($H = \{\langle D_0^1, D_{\rightarrow}^1 \rangle, \ldots, \langle D_0^P, D_{\rightarrow}^P \rangle\}$);
Using the models in H we generate a search space (ie. $\mathbf{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$);
Find the structure $\mathcal{G}_i$ which produces the highest *score* (w.r.t. *assemble*) in $\mathbf{G}$;
return $\mathcal{G}_i$

The dynamic influence network, $\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}$, holds a distribution between a set of DBNs, denoted $\langle D_0^1, D_{\rightarrow}^1 \rangle$, ... , $\langle D_0^k, D_{\rightarrow}^k \rangle$, with the conditional independence assumptions listed by $\mathcal{I}(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$. We further assume that $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$ is induced by another model, $\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$, we will refer to this model as the underlying ground-truth model. The model is evaluated by recovering the set of local

independence assertions in $\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}})$, denoted $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}}))$, by only observing $\mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}}}$. This structure learning procedure is referred to in this paper as greedy structure search (GESS).

3.6 Computational Complexity and Savings

The overall computational complexity of the above structure search algorithm is given by [42]. In order to allow for notable computational savings we suggest using a cache to store sufficient statistics and the of max priority queues (implemented using heaps) to arrange contending structure using their scores as keys. Random restarts and Tabu lists are also used.

4 Experimental Results

This sections presents the performance of modelling influence between partially observable stochastic processes using dynamic influence networks (DINs). We evaluate the performance of model aside several benchmarks.

The experimental setup is as follows. We constructed a ground-truth DIN which was used to sample sequential data. To simulate a partially observed process, several variables were removed from the sequential data sample. The Algorithm provided in section 3 was used to learn candidate networks. Several variations of the algorithm was also used, such as using the d-AIC score instead of the d-BIC; using prior knowledge of the ground-truth structure such as the maximum in-degree used in the generative distribution; using tree structure for sparse generalisability; a even using no structure.

More specifically, the empirical evaluation of our method was set against the following benchmarks:

1. a random DIN structure;
2. a DIN with no structure (i.e. all DBNs are mutually independent);
3. a DIN with a tree like structure (each DBN has one and only one parent DBN);
4. a structure that incorporates prior knowledge of the true structure;
5. a learned structure with the d-AIC score instead of the d-BIC score, which is the dynamic extension of the AIC score;
6. using the full knowledge of the DIN ground-truth structure.

The parameters for the ground-truth DIN distribution is summarised by Table 1.

Figure 2 illustrates a logarithmic scale plot indicating the relative entropy (also known as KL-divergence) to the ground-truth DIN over the amount of samples alongside the aforementioned benchmarks. The vertical axis represents the logarithmic scale of the relative entropy to the ground-truth generative model (Table 1) and the horizontal axis represents the amount of samples. 10 trials were run for each experiment and the mean of the result was plotted with the standard deviation as error bars (shaded regions). All of the model parameters for each experiment is provided in Table 2 for reproducibility.

In Figure 2 we record that providing no structure, a random structure, tree structures, and finally, learning with knowledge of the maximum order in-degree executes in a same way with reference to their relative entropy to the ground-truth DIN. However, knowledge about the maximum order in-degree executes better on average than the other procedures for a large amount of instances (greater than one thousand). The d-AIC and d-BIC scores execute on average better than the other learning procedures (not counting learning using the true structure). However, the d-BIC and d-AIC penalty scores execute similarly.

Figure 3 illustrates a logarithmic scale plot indicating the execution times (in milliseconds) over the amount of samples alongside the aforementioned benchmarks. The vertical axis represents the logarithmic scale of the execution time of each experiment in milliseconds and the horizontal axis represents the amount of samples. 10 trials were run for each experiment and the mean of the result was plotted with the standard deviation as error bars (shaded regions).

In Figure 3 we provide the results of the execution times for the learning procedures considered. With respect to their execution times, the d-AIC, d-BIC scores, and learning with knowledge of the ground-truth maximum in-degree yield the best run-time. Learning tree-structures, generating a random structures, using no structure, or being given the true structure can be achieved in constant time. It is also noticed that learning with the maximum in-degree from the ground truth can be done faster than using penalty score procedures, which have roughly the same execution time.

In the experimental learning scenario the three penalty-based learning procedures outperformed the benchmarks. Notably these penalty-based learning procedures provide significant improved performance than using no or a random structure for the DIN.

The results also indicate that learning a tree structure for the DIN still significantly outperformed the use of a random structure. This result is particularly useful from a computational saving perspective as tree structures are sparse since they capture less complex dependence relations between variables. Tree structure summarise effective independence assertions and thus offer better generalisability. Another notable result from Figure 2 is that when we have fewer samples (> 250) we may be better suited to use no structure since imposing a structure with little training data weakens the inferences we can draw from the model.

5 Conclusion

In this paper we empirically demonstrated the a score-based structure learning procedure to learn a DIN to represent the influence relationships between partially observable stochastic processes. Why we would want to learn a DIN depends on what the structure will be used for.

On the one hand, if we are trying to identify the original DIN structure for knowledge discovery, then we will need to identify each of the original conditional independence assumptions of the ground-truth network. This means we will need to find the set $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}}))$. This is not a promising task since there are many perfect maps for $P^*(\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}})$ that can be derived from $\mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}}}$.

Recognising $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}}))$ from the set of structures from $\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_\rightarrow \rangle^{\mathcal{G}})$ will yield the same fit to the data. Therefore iden-

Table 1: A table summarising the parameters for the ground-truth DIN distribution.

Ground-truth DIN Distribution	
No. DBNs	10
Random variable values	3
No. time-slices	5
No. layers	2
No. CPDs between DBNs	15
max in-degree	2
No. Obs	5 p.t.
No. Latent	3 p.t.

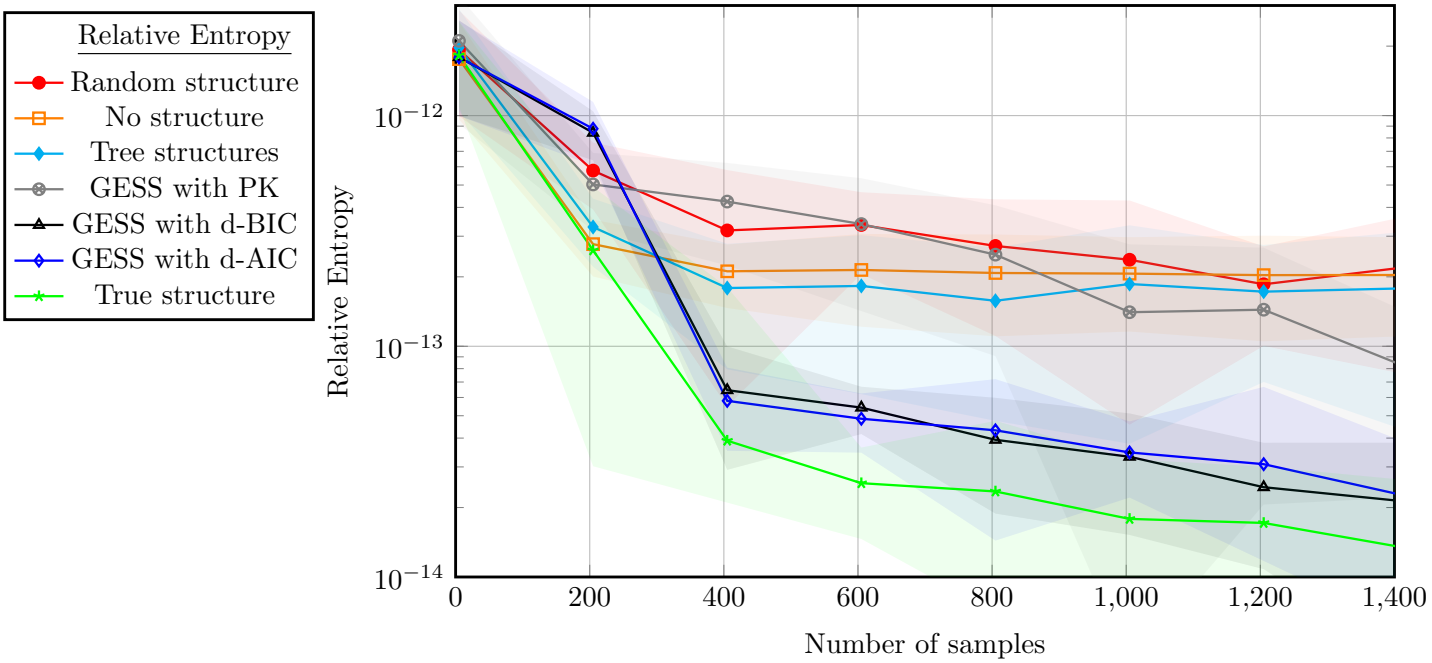


Figure 2: The relative entropy (also known as KL-divergence), to a ground truth DIN, for seven learning tasks to construct a dynamic influence network between dynamic Bayesian networks with respect to the amount of training samples.

Table 2: A table showing all of the parameters used in the structure learning methods in this paper.

	Rand	**No struc**	**Tree**	**GESS with PK**	**GESS with BIC**	**GESS with AIC**	**True**
α	2	2	2	2	2	2	2
No. edges	-	-	-	15	-	-	15
Max in-degree	3	-	-	3	-	-	-
No. observable var	5	5	5	5	5	5	5
Dirichlet prior	5	5	5	5	5	5	5
Parameter threshold	-	-	-	-	5000	5000	-
EM iterations	20	20	20	20	20	20	20
EM accuracy $(\mu\%, \sigma)$					$(76\%, 10)$		
Likelihood score	-	-	Yes	Yes	Yes	Yes	-
Penalty score	-	-	-	-	BIC	AIC	-
Search iterations	-	-	-	50	50	50	-
No. random restarts	-	-	-	5	5	5	-
Tabu-list length	-	-	-	10	10	10	-
α	2	2	2	2	2	2	2

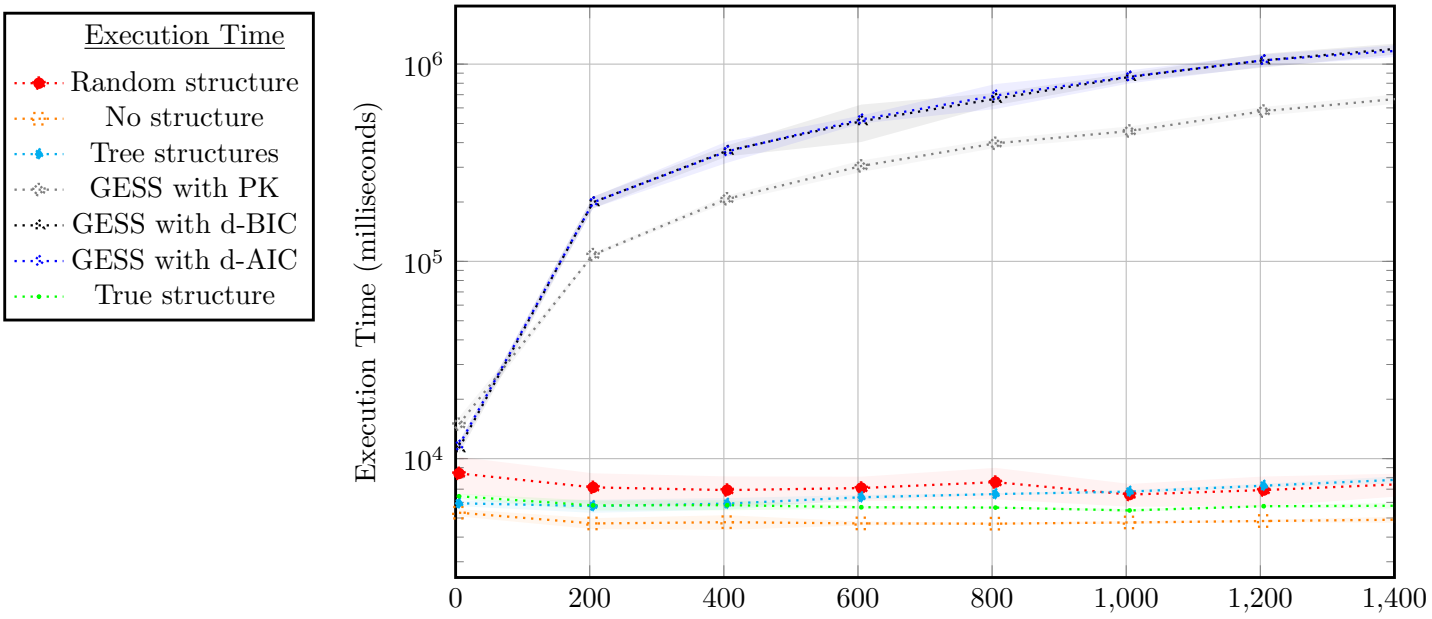


Figure 3: The execution time in milliseconds for seven learning tasks to construct a dynamic influence network between dynamic Bayesian networks with respect to the amount of training samples.

tifying the original ground-truth structure is not identifiable from $\mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}}$. This is because the structures in the I-equivalent structure set all produces the same numeric likelihood (mutual information) for $\mathcal{D}_{\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}}$. Therefore, we should rather try to learn a set of structures that are I-equivalent to $\mathcal{G}^*$.

On the other hand, if instead we are trying to learn a DIN structure for density estimation (i.e. to draw probabilistic inferences), then we are interested in capturing the distribution $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}})$. If we can successfully constructure such a distribution then we can reason about new data instances and also sample new one.

There are two implications when learning a structure or density estimation: Firstly, Although capturing more independence assertions than specified in $\mathcal{I}(\mathcal{G}^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}))$ may still allow us to capture $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}$, our selection of more independence assumptions could result in *data fragmentation*. Secondly, selecting too sparse structures can restrict us to never being able to learn the true distribution $P^*(\langle \mathbb{I}_0, \mathbb{I}_{\rightarrow} \rangle^{\mathcal{G}}$ no matter how we change the parameters. However, often sparse DIN structures can be used to promote computational complexity savings [11].

Conflict of Interest The authors declare no conflict of interest.

Acknowledgement This work is based on the research supported in part by the National Research Foundation of South Africa (Grant number: 121835).

References

[1] R. Ajoodha, B. Rosman, "Learning the influence structure between partially observed stochastic processes using IoT sensor data," in Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence, 2018.

[2] R. Ajoodha, B. Rosman, "Tracking influence between naïve Bayes models using score-based structure learning," in 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), 122–127, IEEE, 2017.

[3] A. Anshel, D. A. Kipper, "The influence of group singing on trust and cooperation," Journal of Music Therapy, **25**(3), 145–155, 1988.

[4] R. Ajoodha, A. Jadhav, S. Dukhan, "Forecasting Learner Attrition for Student Success at a South African University," in In Conference of the South African Institute of Computer Scientists and Information Technologists 2020 (SAICSIT '20), September 14-16, 2020, Cape Town, South Africa. ACM, New York, NY, USA, 10 pages., ACM, 2020, doi:https://doi.org/10.1145/3410886.3410973.

[5] T. Abed, R. Ajoodha, A. Jadhav, "A Prediction Model to Improve Student Placement at a South African Higher Education Institution," in 2020 International SAUPEC/RobMech/PRASA Conference, 1–6, IEEE, 2020.

[6] J. Hatfield, G. J. Faunce, R. Job, "Avoiding confusion surrounding the phrase "correlation does not imply causation"," Teaching of Psychology, **33**(1), 49–51, 2006.

[7] R. Opgen-Rhein, K. Strimmer, "From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data," BMC systems biology, **1**(1), 37, 2007.

[8] T. Grinthal, N. Berkeley Heights, "Correlation vs. Causation," AMERICAN SCIENTIST, **103**(2), 84–84, 2015.

[9] D. Commenges, A. Gégout-Petit, "A general dynamical statistical model with causal interpretation," Journal of the Royal Statistical Society: Series B (Statistical Methodology), **71**(3), 719–736, 2009.

[10] M. Bunge, Causality and modern science, Routledge, 2017.

[11] D. Koller, N. Friedman, Probabilistic graphical models: principles and techniques. (Chapter 16; 17; 18; and 19), MIT press, 2009.

[12] D. Heckerman, D. Geiger, D. M. Chickering, "Learning Bayesian networks: The combination of knowledge and statistical data," Machine learning, **20**(3), 197–243, 1995.

[13] A. Mohammadi, E. C. Wit, "Bayesian structure learning in sparse Gaussian graphical models," Bayesian Analysis, **10**(1), 109–138, 2015.

[14] A. L. Madsen, F. Jensen, A. Salmerón, H. Langseth, T. D. Nielsen, "A parallel algorithm for Bayesian network structure learning from large data sets," Knowledge-Based Systems, **117**, 46–55, 2017.

[15] X. Fan, C. Yuan, B. M. Malone, "Tightening Bounds for Bayesian Network Structure Learning." in AAAI, 2439–2445, 2014.

[16] C. P. d. Campos, Q. Ji, "Efficient structure learning of Bayesian networks using constraints," Journal of Machine Learning Research, **12**(Mar), 663–689, 2011.

[17] S. Kok, P. Domingos, "Learning the structure of Markov logic networks," in Proceedings of the 22nd international conference on Machine learning, 441–448, ACM, 2005.

[18] J. B. Tenenbaum, C. Kemp, T. L. Griffiths, N. D. Goodman, "How to grow a mind: Statistics, structure, and abstraction," science, **331**(6022), 1279–1285, 2011.

[19] I. Tsamardinos, L. E. Brown, C. F. Aliferis, "The max-min hill-climbing Bayesian network structure learning algorithm," Machine learning, **65**(1), 31–78, 2006.

[20] S.-I. Lee, V. Ganapathi, D. Koller, "Efficient structure learning of markov networks using *l*_1-regularization," in Advances in neural Information processing systems, 817–824, 2007.

[21] D. M. Chickering, D. Geiger, D. Heckerman, "Learning Bayesian networks is NP-hard," Technical report, Technical Report MSR-TR-94-17, Microsoft Research, 1994.

[22] D. M. Chickering, "Learning Bayesian networks is NP-complete," Learning from data: Artificial intelligence and statistics V, **112**, 121–130, 1996.

[23] D. M. Chickering, D. Heckerman, C. Meek, "Large-sample learning of Bayesian networks is NP-hard," Journal of Machine Learning Research, **5**(Oct), 1287–1330, 2004.

[24] J. Suzuki, "An Efficient Bayesian Network Structure Learning Strategy," New Generation Computing, **35**(1), 105–124, 2017.

[25] G. Schwarz, et al., "Estimating the dimension of a model," The annals of statistics, **6**(2), 461–464, 1978.

[26] S. Chen, P. Gopalakrishnan, "Speaker, environment and channel change detection and clustering via the bayesian information criterion," in Proc. darpa broadcast news transcription and understanding workshop, volume 8, 127–132, Virginia, USA, 1998.

[27] Y. Tamura, T. Sato, M. Ooe, M. Ishiguro, "A procedure for tidal analysis with a Bayesian information criterion," Geophysical Journal International, **104**(3), 507–516, 1991.

[28] J. Rissanen, "Stochastic complexity," Journal of the Royal Statistical Society. Series B (Methodological), 223–239, 1987.

[29] A. Barron, J. Rissanen, B. Yu, "The minimum description length principle in coding and modeling," IEEE Transactions on Information Theory, **44**(6), 2743–2760, 1998.

[30] D. Geiger, D. Heckerman, H. King, C. Meek, "Stratified exponential families: graphical models and model selection," Annals of statistics, 505–529, 2001.

[31] D. Rusakov, D. Geiger, "Asymptotic model selection for naive Bayesian networks," Journal of Machine Learning Research, **6**(Jan), 1–35, 2005.

[32] R. Settimi, J. Q. Smith, "On the geometry of Bayesian graphical models with hidden variables," in Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, 472–479, Morgan Kaufmann Publishers Inc., 1998.

[33] M. Koivisto, K. Sood, "Exact Bayesian structure discovery in Bayesian networks," Journal of Machine Learning Research, **5**(May), 549–573, 2004.

[34] T. Silander, P. Myllymaki, "A simple approach for finding the globally optimal Bayesian network structure," arXiv preprint arXiv:1206.6875, 2012.

[35] D. Chickering, D. Geiger, D. Heckerman, "Learning Bayesian networks: Search methods and experimental results," in proceedings of fifth conference on artificial intelligence and statistics, 112–128, 1995.

[36] W. Buntine, "Theory refinement on Bayesian networks," in Proceedings of the Seventh conference on Uncertainty in Artificial Intelligence, 52–60, Morgan Kaufmann Publishers Inc., 1991.

[37] A. Moore, M. S. Lee, "Cached sufficient statistics for efficient machine learning with large datasets," Journal of Artificial Intelligence Research, **8**(3), 67–91, 1998.

[38] K. Deng, A. W. Moore, "Multiresolution instance-based learning," in IJCAI, volume 95, 1233–1239, 1995.

[39] A. W. Moore, "The anchors hierarchy: Using the triangle inequality to survive high dimensional data," in Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence, 397–405, Morgan Kaufmann Publishers Inc., 2000.

[40] P. Komarek, A. W. Moore, "A Dynamic Adaptation of AD-trees for Efficient Machine Learning on Large Data Sets." in ICML, 495–502, 2000.

[41] P. Indyk, "Nearest neighbors in high-dimensional spaces," Citeseer, 2004.

[42] R. Ajoodha, Influence modelling and learning between dynamic Bayesian networks using score-based structure learning, Wirespace, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 269-273 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Pre-University Students' Learning Styles and Attitude towards Mathematics Achievements

Nurhilyana Anuar, Nurashikin Abdullah*, Sharifah Norasikin Syed Hod

Universiti Teknologi MARA, Cawangan Selangor, Kampus Dengkil Dengkil,Selangor, 43800, Malaysia

ARTICLE INFO

Article history:
Received: 08 September, 2020
Accepted: 01 November, 2020
Online: 10 November, 2020

Keywords:
Learning styles
Attitude
Academic achievement

ABSTRACT

Learning styles studies have been discussed widely in academic field. It can be considered as a factor that contributes to the achievement of the students in their learning because everyone has a unique learning style when learning. In this paper, we aimed to examine the association of learning styles and attitude towards academic achievement. By identifying the student's learning styles preference, it will offer benefits to the student and instructor by improving learning achievement. Sample data were collected among 328 pre-university science and engineering students using cross-sectional survey. Attitudes Towards Mathematics Inventory (ATMI) and Index of Learning Style© (ILS) were used as the instruments to collect data from the respondents. Data were analysed using Stata statistical tool to examine the result. Result showed that there was no association of learning styles and attitude towards academic achievement in this study.

1. Introduction

Mathematics is one of the subjects in Sijil Pelajaran Malaysia (SPM), a high school national examination taken by secondary school students in Malaysia, as it is one of the subject requirements to continue study in pre-university program. However, the percentage of students who passed with distinction and credit in this subject has not improved from 2017 to 2019. Students who passed with distinction were 31.5% in 2017, 31.2% in 2018 and 31.2% in 2019 while those who passed with credit were 25.2%, 24.9% and 24.1% for 2017, 2018 and 2019 respectively [1, 2]. It is quite worrisome to see these statistics as mathematics does not only exist as individual subject, but it is also used as a tool in other subjects such as physics, chemistry, accountancy and economics to name a few. There are many aspects needed to be studied and improved in order to overcome this problem and the aspects of the students' learning styles and attitudes toward mathematics are among the aspects worth looked into.

In [3], the author mention that the learning style theory was originated from psychology research domain. However, this theory appears to have wide acceptance in many educational researches in exploring the issue of student achievement through learning style. There are several disciplines that have applied this theory in their studies such as nursing, engineering, mathematics [4-6] and online learning mode course [7].

In literatures, some studies on learning styles and academic achievement have been discussed extensively. For example, studies in [8], the studies found that there was no significant difference between learning style and academic achievement among Polytechnic undergraduate business students in Malaysia. The study concluded that by promoting student self-awareness and learning strategies, it will affect their learning style. In [9], the author studied the relationship of learning styles and academic achievement in mathematics among Grade 8 students. The study concluded that learning style is significantly related to academic achievement.

Based on the literature, there is a lack of study that discussed on pre-university students' learning style and attitude towards mathematic achievements. Moreover, the mixed results found in literature have motivated this study to be conducted. In addition, this study is an extension of the original research [10] that examined pre-university student's attitude towards mathematics achievement. Despite the fact that pre-university students from science program have always gained better achievement compared to students from engineering program, the original research showed that mathematical capability of the students was not influenced by the program they enrol in a local public university. Thus, the result showed that there was no significant findings on the attributes associated to academic achievement [10].

Therefore, it is significant to broaden this study on learning styles and attitudes towards mathematics achievement for pre-university students. The objectives of the studies are:

a. To study the student's preference learning style towards mathematics.

*Corresponding author: Nurashikin Abdullah., nurash2856@uitm.edu.my

b. To study the differences of student attitudes towards mathematics.

c. To measure the association of learning style and attitudes towards mathematics achievement.

In the next section, it discusses the related work on learning style. It is followed by research methodology while the final section is result and conclusion.

2. Related Works

What makes individual learning style distinct? Individual learning style is driven by some attributes such as characteristics and behaviour of the learner. According to literature, learning style has three basic elements which are cognitive, motivational and physiological. These elements indicate how the learner perceive, respond, and cooperate with the learning environment [11]. Every student has different level of cognitive, motivation and psychology responsive to the learning atmosphere [12]. Moreover, learning styles give an impact to the academic performance [13].

In an analysis done in [14], result showed there was a positive relationship between students' learning styles and the attitude towards mathematics. Many researches agreed that the effect of learning styles and positive attitude contributed to better achievement in mathematics. In [15], the author suggested that learning environment designed according to learning styles may increase students' success while the researcher in [16] said that individual learning style correlated the highest with mathematics performance. Likewise, studies conducted in [17], the author stated that positive attitudes towards mathematics and science are also associated with higher achievement in each subject.

There are several types of learning model discussed in literature on learning style which includes Kolb's Learning Style instrument (LSI), Learning Style Questionnaire (LSQ), Canfield Learning Style Inventory (CLSI), Cognitive Styles Analysis (CSA), Index of Learning Survey (ILS) and VARK (Visual, Aural, Reading or Write and Kinaesthetic) model [18].

Table 1: ILS dimension

Domains	Description
Sensing vs. Intuiting	Sensing – concrete and practical learner. Prefer facts, sound, sight, data, and physical sensation.
	Intuiting – abstract and theoretical learner. Prefer concept, theories, and insights.
Visual vs. Verbal	Visual – learn from what they have seen.
	Verbal – learning through the words whether from spoken or written.
Active vs. Reflective	Active – learn by performing practical activity and collaboration with others.
	Reflective – learn through thinking or pondering and working alone.
Sequential vs. Global	Sequential – linear thinking process. Prefer to do things step by step.
	Global – holistic thinking process. Prefer to study the whole picture to understand entire subject.

The most common models used in literatures are VARK model and ILS model [5-6],[19]. Different models of learning style have their own strengths and weaknesses. VARK model examines on sensory modality that includes visual, aural, read/write and kinaesthetic of learning styles. This model is used to identify learner preferred learning mode in which the learner may have multimodal learning styles [19]. Another model used among many researchers is Richard M. Felder and Linda K. Silverman's Learning and Teaching Styles based on its instrument of Index of Learning Styles (ILS). Initially, the model consists of five dimensions of Felder and Silverman's Learning Styles; Sensory-Intuitive, Visual-Auditory, Inductive-Deductive, Active-Reflective and Sequential-Global [12] but it has been updated to four dimensions which are Sensing-Intuiting, Visual-Verbal, Active-Reflective and Sequential-Global. In ILS model, information was perceived and processed by four domains [20, 21].

Furthermore, literature on learning style preferences among engineering student showed that students prefer active, sensing and visual learning styles in their problem-based learning (PBL) curriculum course [22]. Moreover, other studies also found that engineering students favoured sensing, visual, active and sequential learning styles [6]. In another perspective, learning styles preferences may differ between genders in which females have reflective, sensing, visual and sequential learning styles while males prefer active, sensing, visual and sequential [23]. In 2015, the author found that there was a significant relationship between learning styles and educational achievement such as in mathematical field [7]. In addition, there is a positive significant relationship between students learning style who used Active-Reflective and Visual-Verbal learning style and their score means.

ILS is more reliable, valid, and suitable for identifying learning styles [20]. It is a well-established tool and widely used to assess learner's learning styles. Moreover, various studies have been done on learning style using ILS [23]. Thus, this study adapted ILS model to fulfil the objectives of the study. Instruments used in this study are Attitudes Towards Mathematics Inventory (ATMI) and Index of Learning Style© (ILS). How does ATMI measure the attitude of the learner in mathematics? In ATMI instrument, it consists of four attributes; i) Self-confidence: learner's confidence and self-concept in comprehend mathematics; ii) Value: belief of usefulness, relevance and worth; iii) Enjoyment: learner has fun working with mathematics; and iv) Motivation: learner's interest and desire to study in mathematics [24].

3. Methodology

This study used cross-sectional survey among students attending pre-university studies in a local university and all students were invited to participate in this survey. A total of 328 students participated in this survey, giving the response rate of 13%.

A set of questionnaire consisted a range of information including socio-demographic as well as questions measuring their attitudes towards mathematics using Attitudes Towards Mathematics Inventory (ATMI) [25] and Index of Learning Style© (ILS) [26] was distributed among the participants. There were four self-reported domains focused in ATMI which included to measure self-confidence, value, enjoyment and motivation. ATMI consisted of 40 items on 5-point Likert scale ranging from strongly disagree to strongly agree. Its Cronbach's alpha indicated good reliability (α= 0.97). To find the result on the attitude of the student, each domain score was summed up. The higher scores showed positive attitudes towards mathematics.

The ILS consisted of 44-items which were divided into four subscales (active/reflective, sensing/intuitive, visual/verbal, and sequential/global). The Cronbach's alpha for ILS was considered

good with 0.74 [21]. The scores in ILS were computed by taking the difference on the number of items selected for each of the two styles in a scale. The total scoring was between -11 to 11. The scoring was categorised as strong (+/- 11-9), moderate (+/--5) and mild (+/-3-1) preferences.

Data analysis was done using Stata statistical software [27]. Prior to analysis, data were cleaned and verified for any errors, outliers or missing values. Continuous data were reported as means and standard deviations while categorical data were reported as frequencies and percentages. Normality test was conducted using Shapiro-Wilk's test. Bartlett's test for equal variances was conducted prior to mean comparisons between groups. P-value of more than 0.05 from Bartlertt's test confirmed the assumption of equality of variance. Bivariate analysis between continuous variables was run using t-test, while ANOVA was used for more than 2 comparison groups. Pairwise correlation was measured using Pearson correlation. The association between mathematics attitude and learning style on mathematics achievement was measured by linear regression. All statistical tests were conducted at 5% significant level.

4. Results

4.1. Participants

Table 2 shows a total of 328 pre-university students who participated in this study. It reflected the same participants who were used in both studies. The respondents were about 53% female and 47% male students. About 60% of the participants were from science program and 40% of the participants were from engineering program. There was a statistically significant difference between program and gender (Chi-square=49.233, p-value <0.001). More than half of the respondents reported that they attended boarding school previously (57%). More than 70% from both male and female students scored A+ in mathematics during their Higher School Certificate (HSC) exam. However, result showed that there was no statistically significant association between school type and gender (Chi-square=0.658, p-value=0.417).

Table 2: Distribution of participants in the survey

	All (N)	Male (n)	Female (n)	Chi-square test	p-value
Program					
Science	195	61	134	49.233	<0.001
Engineering	133	94	39		
Type of school					
Boarding school	187	92	95	0.658	0.417
Non-boarding school	141	63	78		
HSC Mathematics					
A+	235	115	120	0.939	0.333
A	93	40	53		

4.2. Finding from the original research

Result in Table 3 was taken from the original research which was presented in ICEED, 2019. The original research was more focused on the students' attitude towards mathematics final exam achievement for pre-university students. ANOVA test was conducted and the result showed there was no statistical difference between attitude and mathematics achievement (F=074, p-value = 0.658) of the students [10].

Table 3: Attitude and mathematics achievement

Attitude	Sum of squares	Df	Mean square	F	Sig.
Between groups	606.636	8	75.829	0.74	0.658
Within groups	32784.668	319	102.773		
Total	33391.304	327			

4.3. Learning styles preferences

Figure 1 shows the distribution of domain in learning style preferences. Female students were found to score more than 50% in most of the domains. Those who attended boarding school scored higher in all domains compared to non-boarding school attendees. Students who got A+ in their HSC mathematics exam were found to dominate all domains (more than 65%) while those who got between B+ and A for their HSC Additional Mathematics exam were found to score between 20% and 35% in all domains. Majority of the students from science program were found to prefer reflective (65%), sensing (67%), visual (60%) and global (62%) learning styles.

	Learning Style Preferences (Domain)							
	Active	Reflective	Sensing	Intuitive	Visual	Verbal	Sequential	Global
Gender								
Male	(47.1%)	(47.6%)	(40.4%)	(56.4%)	(47.7%)	(41.7%)	(47.2%)	(47.3%)
Female	(52.9%)	(52.4&)	(59.6%)	(43.6%)	(52.3%)	(58.3%)	(52.8%)	(52.7%)
School type								
Boarding	(60.3%)	(51.6%)	(59.0%)	(54.3%)	(57.6%)	(50.0%)	(53.9%)	(57.0%)
Non-boarding	(39.7%)	(48.4%)	(41.0%)	(45.7%)	(42.4%)	(50.0%)	(46.1%)	(43.0%)
HSC Mathematics								
A+	(70.6%)	(73.4%)	(67.6%)	(77.1%)	(72.0%)	(66.7%)	(70.6%)	(73.0%)
A	(29.4%)	(26.6%)	(32.4%)	(22.9%)	(28.0%)	(33.3%)	(29.4%)	(27.0%)
HSC Additional Mathematics								
A+	(4.4%)	(3.2%)	(2.7%)	(5.7%)	(4.3%)	(0%)	(5.0%)	(2.7%)
A	(30.4%)	(31.4%)	(27.7%)	(35.0%)	(30.6%)	(33.3%)	(31.1%)	(30.4%)
A-	(29.4%)	(29.8%)	(28.2%)	(31.4%)	(28.9%)	(37.5%)	(27.2%)	(32.4%)
B+	(23.0%)	(18.5%)	(24.5%)	(17.1%)	(21.7%)	(16.7%)	(21.1%)	(21.6%)
B	(12.2%)	(16.9%)	(16.5%)	(10.7%)	(14.1%)	(12.5%)	(15.0%)	(14.0%)
B-	(0.5%)	(0%)	(0.5%)	(0%)	(0.3%)	(0%)	(0.6%)	(0%)

Chi-square test: *<0.05, **<0.01, ***<0.001

Figure 1: Percentages of Learning Style Preferences (Domain)

Figure 2 shows the distribution of domain in learning style preferences for science and engineering students. Majority of the students from science program were found to prefer reflective (65%), sensing (67%), visual (61%) and global (62%) learning styles.

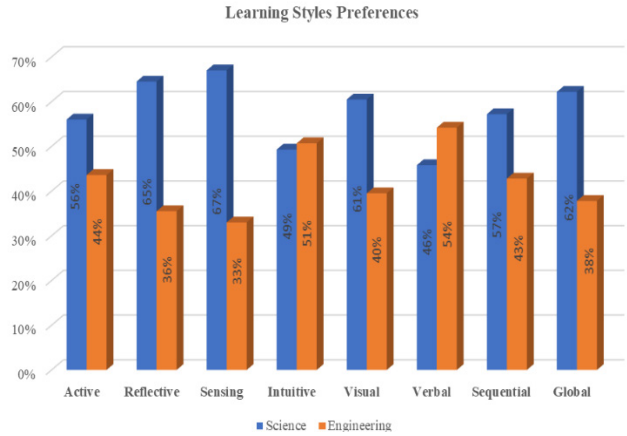


Figure 2: Learning Style Preferences (Domain) by program

4.4. Learning Style Preferences (Domain) during pre-university studies

Students who did very well (grade A) for both pre-university mathematics examinations during their studies had preference in all domains of learning styles. For semester 1 paper, 50.6% of those who scored A grade preferred sequential learning style while for semester 2 paper, students who scored A preferred both verbal (70.8%) and sequential (70%) learning style, respectively. In comparison, students who scored B grades (B- to B+) were found to prefer global learning style for both semester 1 and semester 2 papers as shown in Figure 3.

Learning Style Preferences (Domain)								
Grade	**Active**	**Reflective**	**Sensing**	**Intuitive**	**Visual**	**Verbal**	**Sequential**	**Global**
Semester 1								
A	48.0%	40.3%	44.7%	45.7%	45.1%	45.8%	50.6%	38.5%
A-	22.1	26.6%	22.3%	25.7%	23.0%	33.3%	23.3%	24.3%
B+	12.7%	12.9%	13.8%	11.4%	12.8%	12.5%	14.4%	10.8%
B	7.8%	12.9%	9.0%	10.7%	10.5%	0%	6.7%	13.5%
B-	5.9%	4.0%	6.4%	3.6%	5.3%	4.2%	2.2%	8.8%
C+	2.9%	1.6%	2.7%	2.1%	2.3%	4.2%	1.7%	3.4%
C	0.5%	0. %	1.1%	0%	0.7%	0%	0.6%	0.7%
D	0%	0.8%	0%	0.7%	0.3%	0%	0.6%	0%
Semester 2								
A	67.2%	60.5%	63.3%	66.4%	64.4%	70.8%	70.0%	58.1%
A-	10.3%	12.9%	11.7%	10.7%	10.5%	20.8%	9.4%	13.5%
B+	7.8%	11.3%	8.5%	10.0%	9.5%	4.2%	8.3%	10.1%
B	8.3%	7.3%	8.5%	7.1%	8.5%	0%	7.2%	8.8%
B-	3.4%	2.4%	3.2%	2.9%	3.0%	4.2%	1.1%	5.4%
C+	1.5%	2.4%	2.7%	0.7%	2.0%	0%	2.2%	1.3%
C	1.5%	1.6%	2.1%	0.7%	1.5%	0%	1.1%	2.0%
D+	0%	0.8%	0%)	0.7%	0.3%	0%	0%	0.7%
D	0%	0.8%	0%)	0.7%	0.3%	0%	0.6%	0%

Figure 3: Learning Style Preferences (Domain) in pre-university program

4.5. Students' attitude towards mathematics

The summary from ATMI survey that measured the students' attitude towards mathematics is presented in Table 4. The results indicate that male students have more positive attitudes compared to female students. Students from both types of school showed interest or positive attitudes towards mathematics. Those who did very well in both mathematics papers during HSC examinations have much higher positive attitudes compared to others. The same was true for those who were from engineering program. There were variations of attitude scores found across grades in mathematics subjects during pre-university studies. For semester 2 paper, students who scored B- and below were found to have higher positive attitudes' scores compared to the others. However, for semester 1 paper, their attitudes towards mathematics seemed to be equally distributed across all grades.

Table 4: Score on ATMI by gender, school type, HSC Mathematics and HSC Additional Mathematics

Characteristics	**N**	**Mean**	**SD**
Gender			
Male	155	137.941	10.237
Female	173	136.584	9.971
School			
Boarding	187	137.064	10.093
Non-boarding	141	137.410	10.153
HSC Mathematics			
A+	235	137.745	10.334
A	93	135.914	9.429
HSC Additional Mathematics			
A+	13	140.538	9.812
A	101	139.129	8.975
A-	97	137.505	10.586
B+	70	135.514	9.970
B	46	134.283	10.946
B-	1	130.0	0
Program			
Science	195	136.533	10.470
Engineering	133	138.241	9.492
Pre-university Mathematics Semester 1 grade			
A	148	138.412	9.926
A-	78	136.667	9.979
B+	42	134.048	12.791
B	32	137.375	8.454
B-	17	136.353	9.006
C+	8	138.500	6.655
C	2	139.500	2.121
D	1	134.000	0
Pre-university Mathematics Semester 2 grade			
A	212	137.439	10.209
A-	37	136.567	9.907
B+	30	136.867	10.274
B	26	134.154	10.376
B-	10	141.000	9.333
C+	6	138.667	10.875
C	5	140.400	5.771
D+	1	148.000	0
D	1	134.000	0

4.6. Association between learning style and attitudes towards mathematics achievement

Multiple linear analysis was conducted to measure the association between learning style and attitudes towards mathematics on the mathematic score while controlling for the effect of gender, type of school and HSC grade in mathematics. A significant regression model was observed ($F(5,322)=7.44$, Prob>F= <0.001 and Adjusted $R^2=0.089$). In this model, we found that learning style and attitude were not statistically associated with mathematics scores ($B_{learning\ style}=0.446$, $p=0.054$ and $B_{attitude}=0.019$, $p=0.708$, respectively) as presented in Table 5.

Table 5: Association between learning style and attitudes towards mathematics on the mathematics score in pre-university (semester 2 paper)

Predictors	**B**	**t**	**p-value**
Learning style scores	0.446	1.93	0.054
Attitude scores	0.019	0.37	0.708
Female	4.047	3.75	<0.001
Non-boarding school	3.682	3.39	0.001
HSC Mathematics, (A+)	3.221	2.690	<0.001

F (5,322)=7.44, Prob>F= <0.001 and Adjusted R2=0.089

5. Conclusion

This study showed that preference learning style was dominated by the female students from science program who had attended boarding school. Those students who got A+ in HSC mathematics dominated more than 65%. These students' preference learning styles were reflective, sensing, visual and global learning styles. Meanwhile, the result showed that when the students were in pre-university studies, those who scored

grade A during semester 1 preferred sequential learning style while students in semester 2 who scored A preferred verbal and sequential learning styles. Thus, it shows that there is no specific learning style that the students prefer to use because it depends on the environment of the studies. In terms of the student's attitude, we can conclude that there was a variation of attitude score found across grades in mathematics subject during pre-university studies. Those who had positive attitude scored a good grade in mathematics both in HSC and pre-university studies. Besides that, we have studied the factors that contributed towards academic achievement. The result showed no association was found between learning style and attitudes towards mathematics based on the mathematics scores. Thus, this study can help the instructors to plan their teaching through student's learning style to improve student achievement in mathematics. In addition, this study was conducted on pre-university students from science and engineering background which only focused on mathematics academic achievement. Therefore, this research can be expanded by investigating pre-university learning style for all subjects taken by the students and their academic achievement.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank Institute of Research Management and Innovation, Universiti Teknologi MARA for the financial support in this research under the grant 600- IRMI/Dana KCM 5/3/LESTARI (227/2017).

References

[1] Lembaga Peperiksaan Malaysia, "Pengumuman Analisis Keputusan Sijil Pelajaran Malaysia (SPM) Tahun 2018," 15, 2020.

[2] Lembaga Peperiksaan Malaysia, "Laporan Analisis Keputusan Sijil Pelajaran Malaysia (SPM) Tahun 2019," 17, 2020.

[3] F. Shahzad, S.A. Saif-ur-Rehman, "Learning Styles: History, Conceptualization and Continuum," Social Sciences Review (SSR), **2**(2), 15–31, 2014.

[4] N. Ojeh, N. Sobers-Grannum, U. Gaur, A. Udupa, M.A.A. Majumder, "Learning style preferences: A study of pre-clinical medical students in Barbados.," Journal of Advances in Medical Education & Professionalism, **5**(4), 185–194, 2017.

[5] C.K. Lee, M.S. Sidhu, "Engineering students learning preferences in UNITEN: Comparative study and patterns of learning styles," Educational Technology and Society, **18**(3), 266–281, 2015.

[6] M. Rezaeinejad, A. Azizifar, H. Gowhary, "The Study of Learning Styles and its Relationship with Educational Achievement Among Iranian High School Students," Procedia - Social and Behavioral Sciences, **199**, 218–224, 2015, doi:10.1016/j.sbspro.2015.07.509.

[7] M.M. El-Bishouty, A. Aldraiweesh, U. Alturki, R. Tortorella, J. Yang, T.W. Chang, S. Graf, Kinshuk, "Use of Felder and Silverman learning style model for online course design," Educational Technology Research and Development, **67**(1), 161–177, 2019, doi:10.1007/s11423-018-9634-6.

[8] H. Awang, N. Abd Samad, N.S. Mohd Faiz, R. Roddin, J.D. Kankia, "Relationship between the Learning Styles Preferences and Academic Achievement," IOP Conference Series: Materials Science and Engineering, **226**(1), 2017, doi:10.1088/1757-899X/226/1/012193.

[9] A.-M.L. Villajuan, "Relationship between Learning Styles & Academic Achievement in Mathematics of Grade 8 Students," International Journal of English Literature and Social Sciences, **4**(4), 1052–1055, 2019, doi:10.22161/ijels.4419.

[10] N. Abdullah, N. Anuar, N.F. Zolkipli, R.H. Shaharudin, S.N.S. Hod, "Assessing Pre-University Students' Attitude Towards Mathematics," Proceedings of the 2019 IEEE 11th International Conference on Engineering Education, ICEED 2019, 251–254, 2019, doi:10.1109/ICEED47294.2019.8994930.

[11] J. Keefe, "Assessment of Learning Style Variables: The NASSP Task Force Model, " Theory Into Practice, **24**(2), 138-144, 1985.

[12] R. Felder, L. Silverman, "Learning and Teaching Styles in Engineering Education.," Engineering Education, **78**(7), 674–81, 1988.

[13] M. İlçin, N. Tomruk, S.S. Yeşilyaprak, D. Karadibak, S. Savcı, "The relationship between learning styles and academic performance in Turkish physiotherapy students," Physiotherapy, **102**, e84–e85, 2016, doi:10.1016/j.physio.2016.10.084.

[14] N. Sirmaci, "The relationship between the attitudes towards mathematics and learning styles," Procedia - Social and Behavioral Sciences, **9**, 644–648, 2010, doi:10.1016/j.sbspro.2010.12.211.

[15] A. Ozerem, B. Akkoyunlu, "Öğrenme Stillerine Gore Tasarlanan Öğrenme Ortamları Ve Matematik Başarısına Etkisi," Egitim Arastirmalari - Eurasian Journal of Educational Research, **15**(61), 61–80, 2015, doi:10.14689/ejer.2015.61.4.

[16] A. Bosman, S. Schulze, "Learning style preferences and mathematics achievement of secondary school learners," South African Journal of Education, **38**(1), 1–8, 2018, doi:10.15700/saje.v38n1a1440.

[17] N. Berger, E. Mackenzie, K. Holmes, "Positive attiitudes towards mathematics and science are mutually beneficial for student achievement: a latent profile analysis of TIMSS 2015," The Australian Educational Reseacher, **47**, 409-444, 2020. doi: 10.1007/s13384-020-00379-8

[18] G. Balasubramaniam, I. K, "A Study of Learning Style Preferences among First Year Undergraduate Medical Students Using VARK Model," Education in Medicine Journal, **8**(4), 15–21, 2016, doi:10.5959/eimj.v8i4.440.

[19] H.R. Mozaffari, M. Janatolmakan, R. Sharifi, F. Ghandinejad, B. Andayeshgar, A. Khatony, "The Relationship Between the VARK Learning Styles and Academic Achievement in Dental Students," Advances in Medical Education and Practice, **11**, 15–19, 2020, doi:10.2147/amep.s235002.

[20] K. Middleton, E. Ricks, P. Wright, S. Grant, "Examining the Relationship Between Learning Style Preferences and Attitudes Toward Mathematics Among Students in Higher Education," Institute of Learning Styles Journal, 1(January), 1–15, 2013.

[21] S.R. Viola, S. Graf, Kinshuk, T. Leo, "Analysis of felder-silverman index of learning styles by a data-driven statistical approach," ISM 2006 - 8th IEEE International Symposium on Multimedia, 959–964, 2006, doi:10.1109/ISM.2006.30.

[22] A. Kolmos, J.E. Holgaard, "Learning styles of science and engineering students in problem and project based education," Proceedings of 36th European Society for Engineering Education, SEFI Conference on Quality Assessment, Employability and Innovation, 2008.

[23] C. SENELER, H. PETRIE, "Adaptation Of The Felder-Soloman Index Of Learning Styles (ILS) Into Turkish And An Assessment Of Its Measurement Quality," Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (December), 711–736, 2018, doi:10.16953/deusosbil.332763.

[24] M. Tapia, G.E. Marsh, An Instrument to Measure Mathematics Attitudes, Academic Exchange Quarterly, **8**(2), 16–22, 2004.

[25] M. Tapia, "Attitudes toward Mathematics Inventory Redux," Academic Exchange Quarterly, **9**(3), 272, 2005.

[26] T.A. Litzinger, S.H. Lee, J.C. Wise, R.M. Felder, "A psychometric study of the index of learning styles©," Journal of Engineering Education, **96**(4), 309–319, 2007, doi:10.1002/j.2168-9830.2007.tb00941.x.

[27] StataCorp 2013. Stata Statistical Software: Release 13, 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 184-188 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Comparison of Learning Style for Engineering and Non-Engineering Students

Mimi Mohaffyza[1,*], Jailani Md Yunos[1], Yee Mei Heong[1], Junita[1], Fahmi Rizal[2], Badaruddin Ibrahim[1]

[1]*Faculty of Technical and Vocational Education, Universiti Tun Hussein Onn Malaysia, Parit Raja Batu Pahat, 86400, Malaysia*

[2]*Fakultas Teknik, Universitas Negeri Paandg, Jln Prof Hamka Air Tawar Paandg, Sumatera Barat Paandg, 25131, Indonesia*

ARTICLE INFO

Article history:
Received: 31 December, 2020
Accepted: 05 April, 2021
Online: 20 July, 2021

Keywords:
Learning styles
Accommodator
Converger
Diverger

ABSTRACT

Educators should be considered the learning style of students so that the best practice approach can be applied in learning activities. As students understand their learning style, they will be able to integrate it into their learning process. Kolb Learning Style was the learning style that was widely used based on the theory of learning experiences. Therefore, this study aimed to describe engineering and non-engineering students' learning style. The survey research design with a quantitative approach was applied in this study. A total of 300 respondents were selected randomly from all faculties in Universiti Tun Hussein Onn Malaysia. The survey questionnaire consisted of two main sections representing Learning Goals, Learning Style, and Learning Activities. The result explains that both engineering and non-engineering students are more dominant to adopt the Accommodator learning style, followed by the Converger learning style, and then Assimilator learning style and Diverger learning style. It is concluded that the engineering and non-engineering students are more incline to be a kinesthetic learner. These learning preferences and learning styles will contribute to their engagement in the concept of learning and for educators to plan teaching strategies.

1. Introduction

Learning about students' learning styles can be very beneficial for both teachers and students. Involving students in the active learning phase necessitates recognizing and comprehending learners' learning styles and teachers' teaching styles. Types of learning play a considerable role in learners' lives. Students may incorporate their learning style into their learning process as they become more aware of it. Students learn in various ways, and teachers must design their courses according to different types of learning. Learning skills, creativity and life and career skills are evidence that students master the process of capability and development, integration and knowledge assessment from different subjects and sources of understanding [1]. Identifying students' learning styles will help educators plan their teaching methods and activities effectively to achieve their learning outcome [2]. The learning style of students is f the supporting forms of active learning [3]. The style of learning plays an important role in ensuring that the learning process is performed effectively.

*Corresponding Author: Mimi IMohaffyza, mimi@uthm.edu.my

Students should have 21st-century skills, especially soft skills, to enhance their employability and values. [4]. Universities must make vital elements of education to conduct learning by introducing effective student learning processes in the growth of international education in the formulation of skills in the twenty-first century. [5]. To ensure that all students receive knowledge from the learning process, educators must observe and consider the discrepancies and similarities between students and use the knowledge to prepare for the learning process [6] to design learning regardless of the learning style of the students [7]. To compare learning style preferences between engineering and non-engineering students in Malaysia, this study used a measurement method called the Kolb Learning Style Inventory (LSI) since LSI is able to provide a simple validation of the Experiential Learning Theory.

1.1. Kolb Learning Style

The Experiential Learning Theory of Kolb forms the basis of the paradigm of Kolb's learning style. Experiential learning, which is distinct from other cognitive learning theories, notes the increase in learning process interactions [8]. The Kolb Learning Style Inventory (LSI) is one method for measuring learners'

preferential teaching style. Kolb's learning style, or more generally known as Experiential Learning Theory (ELT), describes learning as a process in which information is created by transforming experience [9]. Learning is a process, according to Kolb, and knowledge is the transformation of experience [10]. Kolb also indicates that, to have a complete learning experience, students must go through all four phases of the learning cycle, as depicted in Figure 1. These four stages not only allow students to explore a subject through various activities and viewpoints thoroughly, but they also accommodate different learning styles.

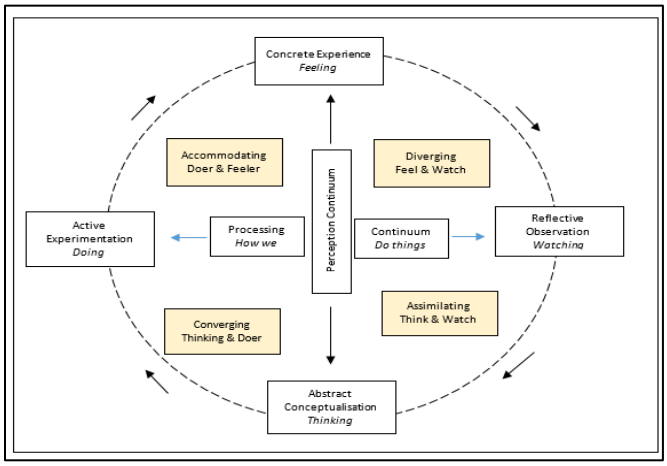


Figure 1: The Learning Model of Experiential Learning at Kolb's Learning Styles

In the Kolb view, learning styles refer to processes in which the person organizes the ideas, rules and principles that address them in dealing with new situations. In practice, one of the most powerful methods in the learning analysis of the individual is the theory of the Kolb learning style. The learning styles as a collection of values, interests, and habits that people attempt to learn about a particular situation by using it [11]. Previous research put it another way: the learner first conducts an action (concrete experience), then tries to think about it (reflective observation), then develops a hypothesis (abstract conceptualization), and finally attempts to exempt it (active experimentation) [12]. According to Kolb, experiential learning can be used in both engineering and non-engineering educational environments [13]. It enables students to participate actively in the learning process to develop awareness, skills, values, and attitudes through direct experience. The learning stages will promote knowledge transfer by providing direct practice tailored to student expertise's scope [14]. This learning method enables students to create their awareness and experience and the acquisition of new skills and knowledge. It stresses that learners learn to use their expertise and experience to solve their problems. This study aimed to compare the preferential learning styles of engineering and non-engineering students in Malaysia by using a Kolb Learning Style Inventory (LSI) model as a reference model because it can provide a basic foundation for validating the theory of experiential learning.

2. Learning Style in Technical and Vocational Education

Technical and Vocational Education is an important road to vocational education and the growth of skills. To meet Malaysia's economic needs, the country's TVET enrollment must be increased by 2.5 times by 2025. Transformation Programme [15]. The human resources to meet this demand, however are inadequate. Right at the time. Moreover, TVET is regarded to be less appealing than traditional university education. This has led to a shortage of, especially highly skilled, TVET students. Malaysia must therefore move from the commonly accepted assumption that the only career path for Malaysian youth is traditional university education, and also emphasize TVET as a valid higher education choice.

Technical and Vocational education students are exposed to an educational system aimed at getting a job. (1) A component of an educational activity aimed at providing the necessary knowledge and skills to carry out a specific job, occupation or professional activity in the labour market can be technical and vocational education. At the same time, other types of education, by training people not only as workers but also as citizens, act as an additional form; (2) an activity associated with the technology transition, innovation, and growth processes Knowledge and skills must be transferred since they form the foundation of technical progress and growth [16]. In technical and vocational teaching, as in many fields of knowledge, it is important to identify and understand students differences to adopt the institute's needs to best suit the students' learning conditions and skills. A fact in the classroom, which can be seen in actual scenarios or in virtual techniques, is the need to adapt teaching methods to student learning styles and interests.

If learning styles are not identified, they may influence the teaching and learning process [17]. A lack of knowledge of the modes of learning can also be problematic. In the implementation among students of the acceptable and successful learning styles [18]. Academic success will be impaired as a result [19]. Unfortunately, teacher-centred learning sessions are held by most educators, allowing fewer students to engage in the process and activities of learning [17]. Therefore, for the performance of students, learning style is an important matter. The style of learning will ensure that a learner learns well [19]. Students need to identify their learning styles to build on their learning skills and expand their learning skills. By posing a challenge or using various education methods, educators are also expected to encourage their students to identify their learning style. [20].

3. Material and Method

The survey research design with a quantitative approach was applied in this research. A set of questions was designed based on the collected learning style and activities found in literature based on the Kolb Learning Style Inventory. A total of 300 respondents were randomly selected from all faculties in Universiti Tun Hussein Onn Malaysia, UTHM (i.e. Faculty of Civil Engineering and Built Environment, Faculty of Technology Management and Business, Faculty of Technical and Vocational Education, Faculty of Electrical and Electronic Engineering, Faculty of Computer Science and Information Technology, Faculty of Applied Sciences and Technology and Faculty of Engineering Technology). The survey questionnaire consisted of two main sections representing the Learning Goals, Learning Style and Learning Activities. This questionnaire was deployed online from the university's online forum and platform. Respondents

were able to complete the questionnaire in approximately 10-15 minutes.

4. Finding and Discussion

The findings discussed are based on the data of the Learning Goals, Learning Style and Learning Activities items that were constructed. Data that had been collected were used to analyze in the context of Learning Style characteristics, and T-test was conducted to determine whether there are any variations between the two groups of fields, as well as descriptive statistics such as frequency and percentage, to evaluate and interpret the results in this report. The interpretation in the research instrument was used to explain the frequencies and percentages. The agreement level was used to assess the students' perceptions in both areas, which were either Yes or No.

4.1. The Learning Style Between Engineering and Non-Engineering Students (Descriptive Results)

Based on a survey conducted, the different learning styles of engineering and non-engineering students were gathered and divided into four forms of learning style defined by Kolb, following the learning style Diverger, Assimilator, Converger and Accommodator. To better understand both of these learning styles, it should be understood that the Assimilator learning style (think and watch) is a variation of Reflective Observation (RO) and Abstract Conceptualization (AC).Converger learning (think and do) is a synthesis of Abstract Conceptualization (AC) and Active Exploration (AE) (AE). Accommodation learning style (feel and do) is a combination of Active Experimentation (AE) and Concrete Experience (CE) and Diverger learning style (feel and watch) is a combination of Concrete Experience (CE) and Reflective Observation (RO) [21]. The percentage of students' data distribution on each Kolb learning style determined by The Kolb Learning Style Inventory is shown in Table 1 and illustrated in Figure 2 below.

Table 1: Results of Learning Style in Vocational Education between Malaysian engineering and non-engineering students

Field	Kolb Learning Style									
	Converger		Assimilator		Accomodator		Diverger		Total	
	f	%	*f*	%	*f*	%	*f*	%	*f*	%
Engineering	42	28	29	19.3	51	34	28	18.7	150	100
Non-Engineering	51	34	21	14	60	40	18	12	150	100

The results show that Accommodator learning style in engineering students is the highest percentage than others learning style with value (*f* = 51, 34%) followed by Converger (*f* = 42, 28%) and Assimilator (*f* = 29, 19.3%). While Diverger learning styles shows the lowest percentage within engineering students with values (*f* = 28, 18.7%). Other than that, a similar condition was shown by non-engineering students where the Accommodator learning style shows the highest worth percentage (*f* = 60, 40%) followed by Converger (*f* = 51, 34%) and Assimilator (*f* = 21, 14%). While the lowest value of percentage is Diverger which is (*f* = 18, 12%).

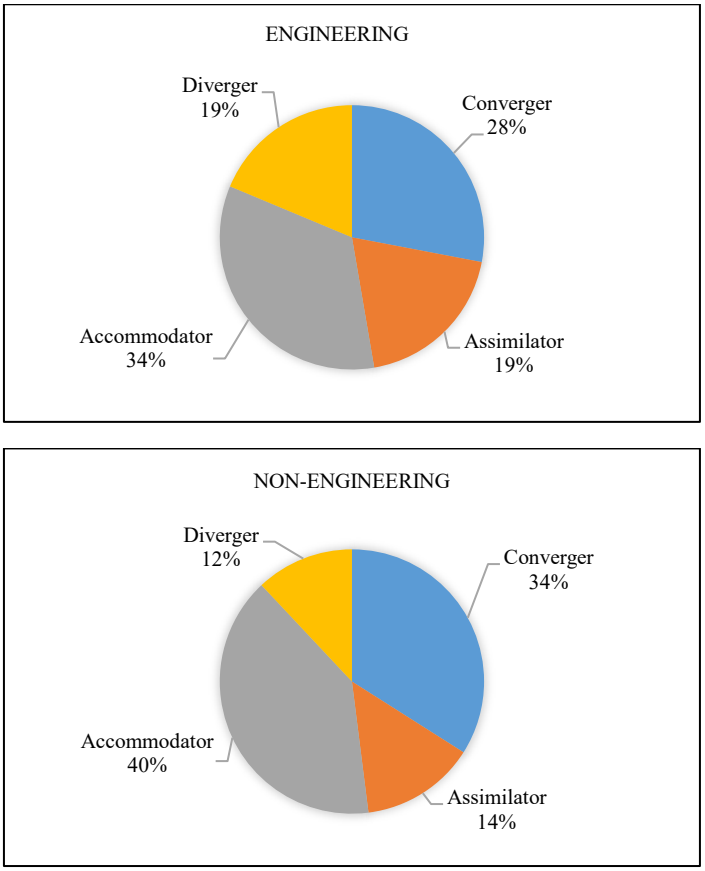


Figure 2: Results of Learning Style in Vocational Education between Malaysian engineering and non-engineering students

The findings of the study can be seen more clearly by referring to figure 3 below where you can see the significant differences between the four learning styles. Although there is a percentage difference between the two fields, it shows that most of the engineering and non-engineering students can be described as an accommodator, which indicates they are most potent in Concrete Experience and Active Experimentation. Instead of logic, they rely on intuition which prefers learning from personal experience, relies on given knowledge rather than carrying out his/her research, requires a clear explanation before starting work [22]. It also shows that both engineering and non-engineering students have strengths that lie in their desire to execute plans and tasks to take part in new events [23]. This result is in line with the Kolb Learning Styles trend, which states that students who use the Doer and Feeler learning styles are best suited for teaching, technician, and engineering jobs and have a background in education, technical studies, and engineering [24].

Other than that, the overall result shows Converger is the second-highest percentage for both fields. In contrast to engineering students, non-engineering students prefer technical tasks and better interpret complex concepts and hypotheses. They also enjoy experimenting. This type of learning style's strengths

lie in their ability to set goals, solve problems, and make decisions [23].

Apart from that, Assimilator shows the third-highest percentage for both engineering and non-engineering fields. The results show that engineering students have a higher percentage than non-engineering students. This means that engineering student who learns in this style has a wide range of knowledge and arrange it in the most logical way [22] compared to non-engineering students. It also indicates that these students prefer rational, factual, and well-thought-through knowledge [24]. The strengths of this learning style lie in their ability to schedule, coordinate, evaluate and engage in inductive reasoning systematically. The results of this study are confirmed by a study conducted by [25] in which engineering students need more diverse knowledge gathered from different sources since they must observe how to execute the task before beginning to perform it. The knowledge is presented from different angles and concluded in a logical, simple, and concise manner. Finally, the type of learning that shows the fourth-highest percentage for both engineering and non-engineering is Diverger. The findings showed that there was a higher proportion of engineering students compare to non-engineering students. It indicates that engineering students with a particular style of learning observe a situation and then look at the situation later from multiple viewpoints, learning from each one [22]. Besides, it also shows that engineering students have more effective at seeing a particular situation from different perspectives than non-engineering students [26].

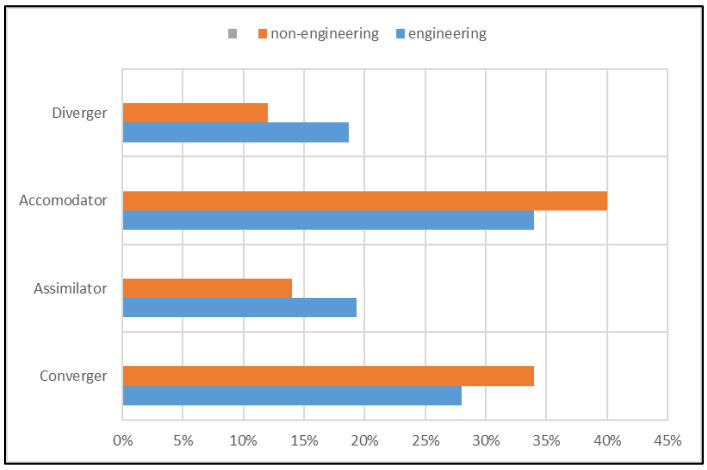


Figure 3:Comparison of Learning Style in Vocational Education between Malaysian engineering and non-engineering students

Because of the interest in designing the learning process, educators have to consider the style of learning of different students. Other than that, to maximize learning effectiveness, the learning method that relates to each learning style is more important because the learning method has a learning style related to it [27]. The difference between the way information was obtained and interpreted was more related to the style of learning that students had. One of the key reasons for gathering learning efficacy is the type of learning [21]. The suggest that knowledge of the learning styles of learners can be important for curriculum and teaching improvement. Similarly, [28] state that If the learning styles of students are evaluated, it is possible to systematically plan learning activities that further strengthen strengths or develop weaker phases to maximize thinking and problem-solving skills."

4.2. Comparison of Learning Style Between Engineering and Non-Engineering Students (Inferential Results)

As for the comparison between Engineering and Non-Engineering students, the inference analysis had shown a non-significant value between both groups in practising learning style in their learning process with mean and significant value (Engineering = 0.538, Non-Engineering = 0.562, p=0.543). Although there is a difference in percentage and frequency values, the inference value indicates no significant value for engineering and non-engineering students learning style. This shows that both groups of students have approximately the same learning style between engineering and non-engineering students.

Table 2: The Differences of Learning Style Between Engineering and Non-Engineering Students

Field	Mean	Std Deviation	Significant
Engineering	0.538	0.133	0.543
Non-Engineering	0.562	0.130	

5. Conclusion

The study showed that learning styles are necessary for a course to achieve total value from learning. While sharing certain characteristics, may show major differences in other aspects that affect learning. Educators who are mindful of these differences and can articulate these characteristics have a better chance of creating good instruction for a wide range of learners. In knowing their strengths and interests and utilizing the learning cycle, all learning styles will become stronger for students exposed to learning style models and Kolb's Learning Style Inventory (LSI), which will enable them to become more active learners. This research can be very beneficial for educators who want to increase the effectiveness of the learning process. Recognizing and reacting to individual learning styles may improve students' ability to accept and retrain content and help to avoid possible learning difficulties by selecting the appropriate teaching method. This may also aid in selecting the most suitable materials and activities for the individual students.

Acknowledgement

We would like to thank the team of Project Matching Grant K135 who participated in this study consisting of experts from two universities, Tun Hussein Onn University, namely Maizam Alias, Tee Tze Kiong, Lee Ming Foong and Faizal Amin Nur Yunus and Universitas Negeri Paandg, namely Ganefri, Nizwardi Jalinus, Syahril, Sukardi, Risfendra and Rahmat Azis Nabawi for their contribution. Finally, we would like to thank the Tun Hussein Onn University of Malaysia for the financial support under UTHM Grant Vot U940.

References

[1] M.C. Sahin, "Instructional design principles for 21 st century learning skills," **1**(1), 1464–1468, 2009, https://doi.org/10.1016/j.sbspro.2009.01.258

[2] S.Kalbasi, M.Naseri, G.H.Sharifzadeh, A.Poursafar, "Medical students' learning styles in Birjand University of medical sciences strides in development of medical education," Journal of Medical Education Development Center of Kerman University of Medical Sciences, 2008, **1**(5): 10-16. doi: 10.5681/rdme.2013.017

[3] Kolb, Y,Alice, & D.A. Kolb, "The Kolb Learning Style Inventory — Version 3," 2005, Technical Specifications. Experience Based Learning Systems, Inc.

[4] N. Azid, R. Rawian, Shaik-Abdullah & T.K. Tee, " The Development of Interactive Case-Based Smart Thinking and Industrial Problem-Solving Stimulator to Enhance TVET students' Thinking Skills," Journal of Engineering Science and Technology, 2019, **14**(5), 2643-2656.

[5] A. Halstead & L. Martin, "Learning styles: A tool for selecting students for group work," International Journal of Electrical Engineering Education, **39**(3), 245–2522, 2002. https://doi.org/10.7227/ijeee.39.3.8

[6] R. Laura, "What is differentiated instruction? New York: Rowan University,"2007. https://doi.org/10.1093/ww/9780199540884.013.u289492

[7] C. Weselby, "What is differentiated instruction? Examples of how differentiated instruction in the classroom," Oregon: Concodia University, 2017. https://doi.org/10.4135/9781483346243.n103

[8] S. Cassidy, "Learning styles: An overview of theories, models, and measures," Educational Psychology, 2004. https://doi.org/10.1080/0144341042000228834

[9] Y.M. Yousafzai, N.Baseer, S.Fatima, A.Ali & I.Shah, I. "Investigating the Relationship between Learning Styles and ESP Reading Strategies in Academic Setting," International Journal of Applied Linguistics & English Literature, **7**(3). 11-115, 2018. https://doi.org/10.7575/aiac.ijalel.v.7n.3p.156

[10] C.Babadoğan, C. "Öğrenme Stilleriyle İlgili Araştırmaların Taranması (Survey of Researches Related to Learning Styles) Eğitim Bilimleri Dergisi," Ankara Üniversitesi, **24**(2), 603-606, 1991.

[11] T. De Jong, "Cognitive load theory, educational research, and instructional design: some food for thought," Instructional science. **38**(2): 34105, 2010. https://doi.org/10.1007/s11251-009-9110-0

[12] A. Z. Moghadam, M.M. Fard "Surveying the Effect of Metacognitive Education on the on the Mathematics Achievement of 1st Grade High Junior School Female Students in Educational District 5, Tehran City, 2009-10 Educational Year, Procedia-Social and Behavioral Sciences, **29**(15):31-40, 2011. https://doi.org/10.1016/j.sbspro.2011.11.394

[13] C. Tseng, "Connecting Self-directed Learning with Entrepreneurial Learning to Entrepreneurial Performance," International Journal of Entrepreneurial Behavior and Research, **19**(4), 425- 446, 2013. https://doi.org/10.1108/ijebr-08-2011-0086

[14] Dumiyati, "Pendekatan Experiential learning dalam Perkuliahan Kewirausahaan di Perguruan Tinggi untuk Menghadapi Asean Economic Community [Suatu Kajian Teoretis]," Prosiding Seminar Nasional Profesionalisme Pendidik dalam Dinamika Kurikulum Pendidikan di Indonesia pada Era MEA, 87- 97, 2015

[15] Ministry of Education Executive Summary Malaysia Education Blueprint 2015-2025, 2015.(Higher Education).

[16] O.O.Oskay, E.Erdem, B.Akkoyunlu, A.Yilmaz, "Prospective chemistry teachers' learning styles and learning preferences," Procedia Social and Behavioral Sciences, **2**, 1362-1367, 2010. https://doi.org/10.1016/j.sbspro.2010.03.201

[17] M.H. Yee, J. Yunos, W.Othman, R. Hassan, T.K.Tee, & M. Mohamad, "Disparity of Learning Styles and Higher Order Thinking Skills among Technical Students," Procedia - Social and Behavioral Sciences, **204**(2014), 143–152, 2015. https://doi.org/10.1016/j.sbspro.2015.08.127

[18] H.Pashler, M.McAndiel, D.Rohrer, & R.Bjork, "Learning styles: Concepts and evidence," Psychological Science in the Public Interest, 2009.https://doi.org/10.1111/j.1539-6053.2009.01038.x

[19] I.J. McCoog, "21st Century teaching and learning. Education Resource Center," 2018.

[20] E. Marin, " Experiential learning: empowering students to take control of their learning by engaging them in an interactive course simulation environment," Procedia - Social and Behavioral Sciences **180**, 854 – 859, 2015. https://doi.org/10.1016/j.sbspro.2015.02.224

[21] A. Y. Kolb & D.A. Kolb, " Experimental Learning Theory Bibliography," Cleveland: OH : Experience Based Learning System Inc, 2003 https://doi.org/10.1007/978-1-4419-1428-6_227

[22] M. MiMuro, A, Terry, "A matter of style: Applying Kolb's learning style model to college mathematics teaching practices," Journal of College Reading and Learning, **38** (1), 53-60, 2007. https://doi.org/10.1080/10790195.2007.10850204

[23] E.F. Turesky & D. Gallagher, " thyself: Coaching for leadership using Kolb's experiential learning theory," Coaching Psychologist, **7**(1), 5–14, 2015.

[24] D.A. Kolb, "Learning Style Inventory: Self Scoring Inventory and Interpretation Booklet," McBer and Company, Boston, 2005. https://doi.org/10.1007/978-3-8350-9212-9_4

[25] M.N. Ghufron & R. Risnawita, "Gaya Belajar: Kajian Teoritik," Yogyakarta: Pustaka Pelajar.

[26] S. Nasution, "Berbagai Pendekatan dalam Proses Belajar and Mengajar," Jakarta: Bumi Aksara. 2009.

[27] Ö. Şimşek, "Marmara Öğrenme Stilleri Ölçeği"nin Geliştirilmesi Ve 9–11 Yaş Çocuklarının Öğrenme Stillerinin İncelenmesi," Marmara Üniversitesi Eğitim Bilimler Enstitüsü. İlköğretim Anabilim Dalı. Yayınlanmamış Doktora Tezi. 2019. İstanbul

[28] J.A.Gyeong & S.Y. Myung, " Critical thinking and learning styles of nursing students at the baccalaureate nursing program in Korea," Contemporary Nurse, **29**(1), 100–109, 2008. https://doi.org/10.5172/conu.673.29.1.100

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 17-35 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

The Role of RFID in Green IoT: A Survey on Technologies, Challenges and A Way Forward

Zainatul Yushaniza Mohamed Yusoff [1,2], Mohamad Khairi Ishak[*,1], Kamal Ali Alezabi[3]

[1]*School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Nibong Tebal, Pulau Pinang, 14300, Malaysia*

[3]*Majlis Amanah Rakyat, Kuala Lumpur, 56000, Malaysia*

[2]*Institute of Computer Science and Digital Innovation (ICSDI), UCSI University, Cheras, Kuala Lumpur, 56000, Malaysia*

ARTICLE INFO

Article history:
Received: 11 October, 2020
Accepted: 28 November, 2020
Online: 10 January, 2021

Keywords:
Green Power
Green IoT
RFID
Application
Challenges
Impacts

ABSTRACT

The Internet of Things (IoT) is a technology that enables communication between everyday life using different sensor actuators that work together to identify, capture, and distribute critical data from the planet. Massive machines and devices are therefore linked and communicate with them. The use of resources in this area presents new challenges for this technology. The goal was to find a green IoT that focuses on energy efficiency and IoT efficiency. Green IoT is an energy-efficient way to reduce or eliminate the greenhouse effect of current applications. Radio Frequency Identification (RFID) is one of the Green IoT and Master IoT components that identifies a person or entity in a high-frequency electromagnetic spectrum when combining electromagnetic or electrostatic couplings. If the predictions are also correct, energy use issues arise as active battery-powered RFID detection needs to be addressed by incorporating new solutions for Green IoT technology. Past studies and assessments have attempted to evaluate RFID technology and its functions. Unfortunately, however, they concentrated on a single RFID view of technique and technology. This paper examines holistically and systematically the impact of RFID applications on green IoT, focusing on three categories: the challenges, environmental consequences, and the benefits of green IoT RFID applications. The impacts, performance and safety of RFID IoT applications have been carefully described. The benefits and examples of RFID applications, including their key advantages and disadvantages, are also discussed. Overall, this paper highlights the potential efforts of RFID to address existing Green IoT issues.

1 Introduction

Green power comes from natural sources such as sunlight, wind, rain, tides, plants, heat and algae as shown in Figure 1. These energy resources are renewable and naturally refilled to offer the most significant environmental benefit [1]. Conventional energy can be a valuable source of air pollution and emissions of greenhouse gases. The use of renewable energy helps to perform the following [2]:

1. Reduce the carbon footprint with purchased electricity.
2. Future price increases and volatility of electricity (individual products).

The purpose of green power is to find energy conservation measure in order to reduce energy consumption and increase the energy efficiency. Energy conservation means making use of less energy to accomplish work, namely reducing waste energy. Energy efficiency has many advantages: reducing greenhouse gas emissions, declining energy demand and reducing our household and economic costs. Although the use of renewable energy technologies also improves this, increased energy usage is the simplest and often most efficient way to reduce fossil fuel use. There are enormous opportunities to enhance productivity in every sector of the economy, whether it be infrastructure, transportation, manufacturing or energy generation [3]. In reality, without Internet of Things (IoT), it isn't easy to imagine the future of renewable energies. Today sensor-based technology and data science enable wind farms and solar fields to be productive and automated. Green IoT technology programs should be taken into account because of the growing perception of environmental issues worldwide. The Greening IoT definition refers to innovations that make the IoT ecosystem healthier and environmentally pleasant through the use of facilities and storage that enables subscribers to capture, store, access, and handle different information.

[*]Corresponding Author: Mohamad Khairi Ishak, Universiti Sains Malaysia, Nibong Tebal 14300, Malaysia Email:khairiishak@usm.my

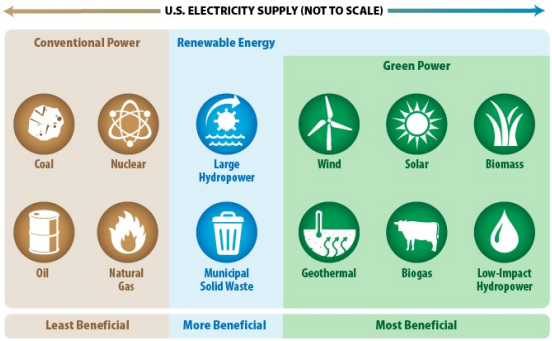


Figure 1: Green Power Generation [1]

The technology that enables green IoTs is called information and communication technologies (ICTs) [4]. ICTs will cause climate change in the world because the rising application of ICT has used much more energy [5, 6]. The emphasis on sustainability for ICTs has been on data center optimization through infrastructure sharing techniques that improve energy efficiency, minimize CO_2 emissions and e-waste disposal [7]. Greening ICT enables green IoT technology which includes green wireless sensor networks (GWSNs), green machine-to-machine communication (GM2M), green RFID (GRFID), green data center (GDC), green cloud computing (GCC), green Internet (G-Internet), and green communication network (GCN) as shown in Figure 2. Greening ICT offers green IoT technologies. Green ICT innovations play a crucial role in green IoT and offer many benefits for society, including the reduction in energy used to develop, produce, and distribute ICT devices and equipment [8].

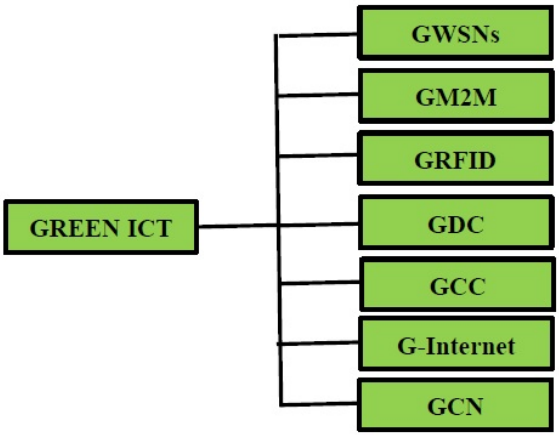


Figure 2: Green ICT technologies [8]

Green IoT consists of aspects of nature and exploitation. Green IoT focuses on reducing the use of IoT resources and CO_2 emissions, which are essential for creating an innovative and prosperous world of intelligent everything. As Figure 3 shows, the green IoT design elements apply to computer creation, energy conservation, communication protocols, and architectures for networking [9]. It is possible to use the IoT portion to remove CO_2 emissions, minimize pollution, and increase energy efficiency [10]. In networking, Green IoT aims to decide where the relay is located and how many nodes meet budget constraints and energy saving [11]–[12]. For achieving green IoT, there are three principles, namely, design technologies, leverage technologies, and enabling technologies. Build technologies that apply to computers, communication protocols, network interfaces, and interconnections for energy efficiency. Leverage technology relates to reducing emissions of carbon and rising energy efficiency [13]. By reducing electricity, toxic emissions, use of resources, and pollution, green IoT has become more effective due to green ICT technologies. As a result, green IoT ensures the protection of natural resources, minimizes the effect of technology on the environment and human health, and dramatically reduces costs. In reality, therefore, green IoT focuses on green development, green usage, green design, and green disposal [14].



Figure 3: Green IoTs Environments [9]

RFID is a complementary solution to green IoT technology to save energy. RFID uses radio frequency to transmit, identify, classify and track data between readers and moving objects. It is fast and does not require a physical view or connection between the reader/scanner and the object to be inspected. The operating cost factor is deficient low and can use in many applications such as healthcare, manufacturing, inventory management, shipping, retail sales and home use [15]. RFID technology internal divided into two parts, namely RF communication; reader/tag protocols and middleware architecture, where RFID tag can be attached to almost anything such as items, cases or product palettes, high-value goods, vehicles, assets, livestock or personnel. RFID tags divided into two sits; which is a passive tag and an active tag. RFID middleware is RFID's software that exists between readers and enterprise/business applications. Middleware has several functions and plays an essential role in the operation and management of RFID systems. Middleware handles RFID readers and printers and communicates between these devices and business applications, as well as manage, filter, aggregate, and understand data derived from RFID tags [15]. RFID application used in manufacturing and processing for supervision, production of warehouse processing and fulfillment orders. In supply chain management, it should face in Inventory Tracking and Logistics management system. It also used in local controls for customer inventory and views; and automated checks with reverse logistics which benefit the environment. Whereas securities applied on access control, counterfeiting and theft control/prevention [16]. However, there is little discussion in the existing literature on RFID management in the environment. Is seen that ICT harms the environment [17], and RFID, as part of ICT, has gained a reputation

[18] as one of the leading computing technologies, using sufficient energy to impact the environment negatively [19].

The contributions of this paper are summarized as follows:

1. A detailed review of current and recent RFID literature on IoT, notably the Green IoT, in order to show that RFID can be part of the green IoT, not only for the benefit of the economy but also for the benefit of the organization, but also to improve responsible IT practice.

2. As this paper reviews the application of the Green IoT study proposed by RFID in recent years, it also outlines some of the early research related to this field.

Section 2 provides a summary of some IoT device RFID technologies. Section 3 Discussion about the security of RFID and addresses RFID implementation issues and environmental consequences in Section 4. In section 5, the benefit of greening IT. The future work of RFID in Green IoT is discussed in Section 6, and Section 7 outlines the study's findings.

2 Studying RFID Technology in IoT Application

The Internet of Things (IoT) intended to describe a range of technologies and research disciplines that enable global connectivity through the world's physical objects [20]. Activation technologies such as RFID, sensor networks, biometrics, and nanotechnologies are now widely used [21]. The IoT implemented in real implementations for different applications such as smart grid, e-health, and intelligent transport. RFID is a stand-alone technology used to enable various systems to perform functions based on object recognition, data recording, and different control objects [22]. The iconic RFID system consists of two parts: a tag (transmitter/receiver) and a reader (transmitter/receiver). The tag contains a chip attached to an object for detection and serves as an object identifier. Instead, RFID readers regularly exchange information with RFID tags using radio waves [23]. The main advantages of RFID technology are automatic object recognition, data acquisition and cost reduction in various activities [24]. The distributed RFID readers around the world are connecting to the Internet according to an appropriate routing protocol. Readers can then exchange information to identify, track, and monitor objects based on the tags they carry. Instead, IoT refers to virtual presentations in the IoT architecture, where identifiable objects are called "things". RFID is a prerequisite for the proper operation of an IoT system [25]. IoT is a global infrastructure that connects and ignores the distance between virtual and physical objects based on data usage and routing capabilities. The IoT architecture consists of three layers, namely the service or application layer, the network layer, and the perceptual layer, where the level of perception includes real-time information and the Internet of Things [26]. This layer uses sensors, wireless sensor networks, tags, RFID systems, cameras, GPS, smart terminals, EDIs, and creators to collect and display information about objects in the physical world. RFID tags, readers and application systems are composed of RFID systems. RFID systems are fast and effective for the electronic identification, control, and identification of different artifacts. RFID uses wireless transmission to send a signal for an RFID tag that responded to a data collection reader linked to the host computer's data for analysis by a single ID code [27] RFID application functionality divided into three categories: control, tracking, and supervision. Monitoring concerns the current device status. For medical circumstances to be detected, irregular behavior and alarm must be an identifier. Tracking sees moving objects and provides a modeling sequence for every location of the data. Control refers to the evaluation and monitoring carried out without the awareness of the object [28]. RFID is used to prevent misplacement or to locate items quickly. One approach is to track the movement of objects in the database. For example, an RFID reader may attach to a door in a building between rooms. Thus, the location of the object can be determined by analyzing RFID reader recordings. It mainly used in supply chain management systems, item monitoring processes, and industrial process control. 3 Great advantages of identification of RFID supply chain properties as seen in the following Table 1 [29]. The use of RFID is also developing in manufacturing, agricultural management, health care management and tele-medicine, maritime operations, management of payment transactions, and others.

Table 1: 3 Great advantages of identification of RFID supply chain [29]

Item	Advantages
Increase the visibility of assets	RFID asset monitoring to increase inventory usability and visibility. This ensures that RFID tags can be read from all over the facility so that workers can easily display each item's location and value. Supplier chains can overcome issues created by bad management, misplaced components, tracking and manual process failures by increased transparency. Naturally, it can easily move things and make decisions about fly changes.
Improve employee productivity	RFID technology workers do not have to search for any items. No wasting time on tedious processes like scanning individual barcodes, finding the wrong tools and devices, and calculating inventory. This technology will increase the speed of sales of jobs and projects, lower labor and transportation costs, and improve the plant's overall productivity. Manufacturing and shipping products in unmanaged facilities are challenging and slow. By eliminating manual steps, human error and monitoring can avoid. Corrections can take a long time to improve inventory and asset visibility.
Reducing risk, theft and loss	RFID asset tracking to quickly and easily access inventory information and locations to track during items storing, distributing, or replenishing. With this type of visibility, there is little room for mistakes or mischief. Track each asset's movement in the facility to consider all items and pallets, minimizing costly risks and reducing misalignment's and losses. Struggle against counterfeiting and theft; Promote callbacks. It is imperative to avoid such errors as much as possible, since fixing these errors can lead to litigation, high fines, or loss of customers, as well as time and money. Full transparency and real-time information on inventory and assets increase profitability for significant organizational and productivity improvements.

2.1 Review Papers of RFID applications

This section discusses recent studies related to RFID applications in the IoT world.

1. In [30], building the Green Campus in IoT Architecture, RFID technology is used to detect students entering the lab. The system will read, and store data related to students or any-one entering the lab. Middleware acts as a connector between temperature sensors, air conditioners, RFID, and computers [31]. Each of these tools has its information. This information is used to determine the number of labs that should open to students, when the air conditioner needs to turn on, and finally, to monitor computer usage correctly [32]. The effectiveness of this system is when:

    - Laboratory management runs smoothly according to consumer demand.
    - The computer management mechanism can adequately monitor to save energy on laboratory equipment.
    - Energy-saving can do excellently when the air conditioner was only functioning when the laboratory temperature reached.

2. In [33], the System of global positioning and geographical information, RFID, general packet radio service and webcam combined into one device. The RFID reader in the camion reads both the customer and the bin details [34, 35]. This system improves waste management and decreases waste emissions resulting in ozone pollution [36].

3. Researchers in [37] introduced a new authentication strategy for home health applications using low-cost RFID tags. For systems where RFID tags need not be measured, only data stored for their storage sent. A tablet container has stored with the RFID tag inside the patient. The RFID reader receives the details when the patient recovers the [38] bag. In the RFID tag, the information is encrypted. Only valid ID tags and ID readers can decrypt the data decryption key – valid RFID tag ID and RFID reader, which generate the use key. Authentication between RFID tags and readers of RFID will combat several attacks, including tag-spoofing, reading readers. One of the benefits is that it reduces the likelihood of human error in data collection and that doctors can also provide reliable information on patients [39].

4. The environment, including the distribution system, has changed considerably in the past few years in [40]. With consumer demand changing, distribution structures are becoming increasingly centralized and globalized, and public changes such as deregulation, innovation and IT dissemination are moving forward. Simultaneously, there is a substantial demand for reducing CO_2 in logistics processes to solve global environmental problems. As a result, RFID (Supply Chain Management) technology was designed to use CO_2 reduction levels, deemed an impact identification object, to be identified and thought that improvement would be a criterion. To deploy an SCM with high efficiency, the information would be necessary, namely in real-time, which is a turning point for success or loss. In this respect, the RFID tag can be a useful tool for automatically defining status by merely placing at the point of departure or arrival items, collapse containers and pallet vehicles using RFID technology. This process would help to understand conditions in real-time, such as the transportation of stocks and shipments [41]. They, therefore, expect RFID to be a practical tool rather than an existing barcode system and argue that the introduction of RFID is a new way to view the information in the SCM. The CO_2 reduction ratio is estimated by researchers as a function of transport expenses, with decreased labor costs, thus, reducing transport expenses. The amount of energy used, such as a roadway, is commensurate with transport weight (t) and distance (km). It means that transport costs and energy consumption are positively correlated. Since CO_2 is proportional to the amount of energy consumed, emissions of CO_2 may also decrease if power is reduced. The decreasing costs of transport can thus reduce CO_2 emissions, with a definite link between energy reduction and CO_2 emissions [42]. Consequently, the transport CO_2 emissions reduction rate was 3.2% as an effect of RFID introduction.

5. In [43], while metal fractures in the automotive industry have been a topic of productive study, plastic recycling in vehicles is poorly understood. It is now key to effective recycling of vehicles, due to legal requirements. Besides, the key problems faced are recognizing plastics and the technological standard of logistics activities, as well as scarce resources related to plastics. Therefore, authors in [44] introduce the RFID framework that automatically recognizes, register, distribute, and search for information during the product life-cycle. The research produces an efficient decision support solution (DSS) in a case study on automotive plastic recycling. DSS has applied during the life-cycle of the drug. It can be divided into three phases depending on the product life cycle management (PLM), initial life expectancy (BOL), life expectancy (MOL) and end-of-life (EOL) of the contemporary product [45]. BOL functions as product design, a manufacturer's preparation, and process management. MOL is used to assist with product repair, condition diagnosis, and prognosis. At the other hand, the classification of EOL has many meanings, including the duty to comply with economic benefits and legal requirements. The DSS method is as follows [46, 47]:

    - Throughout the BOL and MOL phases of the product life cycle, the flow of knowledge completely facilitated by knowledgeable systems such as CAD / CAM, Product Data Management (PDM) and other Knowledge Management systems used successfully in the Automotive Plastic Recycling industry.
    - Often the knowledge that comes to the consumer about the product is incomplete during the transfer process from MOL to EOL.
    - Data sources from MOL tell BOL that drug producers disrupted after product sales. The findings are typically based on BOL data, or limited but imperfect practical experience, based on full knowledge of past transactions.

Then, experiments used to produce missing information, but the cost of EOL operations increased significantly.

- RFID technology is then used to fill the knowledge gap and make product information reliable, accessible, timely, complete and available whenever needed during its life cycle. Different information can be collected and transmitted using RFID technology via MOL and EOL phases.
- Using the unique identifier stored in the RFID tag, EOL will obtain the information required to evaluate the recovery path from the product life-cycle information service. At the end of the support phase, the data collection unit is integrated to collect information from the stores using RFID technology.

6. Researchers in [48] provide the RecycleBank with the resources it requires to organize recycling activities for households, using RFID systems to allow municipalities to calculate the number of recycled materials per household. They have created a community-friendly approach with advanced, creative technology to tackle the rising cost of waste collection. Weight of recycled plastic turned into incentives for RecycleBank. Hundreds of local and national participants can award it. In collaboration with Texas Instruments, RecycleBank offers an RFID that adds low-frequency RFID tags with exact recycling containers and includes integrated RFID readers fitted to a counterbalancing system [49]. RFID tags developed and implemented for RecycleBank, stores Serial Number (referring to the name and address of the container owner), reads and data transmitted to the internal computer of the vehicle. The data collected are then transferred to the data collection and processing system of RecycleBank. Personal data (name, address, etc.) not scheduled for the tag. Only the recorded serial numbers are kept in the secure database of RecycleBank. The proprietary RFID tag stores provide unique advantages for waste and recycling applications with low-frequency technology as a weather-resistant, all-weather, secure solution that works in snowy, wet, or harsh environments throughout the world. Labels must be physically cute while preserving their technical characteristics over the life of the [50] container. The results are as follows:
   - Within one week, the recycling rate in New Jersey, Cherry Hill City, increased by 135%. The average household recycled 12 pounds of recyclables before RecycleBank. Homes now earn an average of over 26 pounds each week.
   - The city of Wilmington, Delaware, has been recycling in just six months from 0% to 37% using the RecycleBank programme.
   - Elk City, New Jersey, averages 16 tons per week before RecycleBank of recyclable material. The municipality generates an average of 42 tons of recyclable material per week after the usage of RecycleBank, a rise of 136%.
7. The case study in [51] demonstrates a highly advanced waste electrical and electronic equipment reuse and recycling program (WEEE). This center is named the Multi Life Cycle Center (MLC), using state-of-the-art advances in automation and IT to provide cost-effective and environmentally sustainable processes. WEEE is back to a valuable commodity [52]. IT and communicators, household appliances, electrical and mobile devices, toys, and recreational equipment are product categories subject to reuse. All other classes of items are stored and recycled. RFID technology is used for communication and organization, together with reverse logistics systems. This technology can be used as a monitoring device that transparentizes and prevents misuse of the WEEE stream. Furthermore, container level measurements are currently being tested, so that it can improve the collection route of the truck and prevent theft. A special reading station is located at the collection point to allow continuous monitoring to collect RFID information [53]. IoT is the link to the Internet of objects utilizing several information recognitions tools such as RFID, infrared sensors, global positioning systems, laser scanners and the identification of certain system information according to the accepted protocols. The computer is a large network. IoT can be used to define wisely, locate, map, control and manage to share information and communications [54]–[55]. Main IoT technologies, software, research and development technologies, external layers, network layers and device layers.Table 2 displays main perception layer technologies.

Table 2: Key technologies in the perceptive layer [53]

Key Technology	Performance Description	Application
RFID	Recognition and identification of objectives. Its purpose is to describe the object.	Application stage
Sensor	Sensitive organs, providing raw data.	The exploratory stage
Intelligent embedded technology	Intelligence is the target of IoT.	Experimental stage
Nanotechnology	Enables the interaction and connection of small objects and reduces system power consumption	Research stage

8. IoT requires finding, scanning, monitoring, and handling, and not just ”connecting objects.” As RFID is a ”speaker” technology, RFID technology plays a vital role in critical technologies in the IoT layer of perception [56, 57]. RFID technology is, therefore, the first choice to establish a full system for IoT identification. EPC encoding, RFID EPC systems and EPC network systems are critical applications of RFID technology with RFID tags in the ”IoT” EPC environment. In addition to handling individual objects, IoT encoding discovers objects at any time and enables coded requests, statistical tests, unique e-commerce categories and open information for objects. ”One single-code entity.” EPC encoding technology [58, 59] can best serve this need. The EPC RFID framework is a module for collecting codes automatically and shapes the EPC information network systems for local internet networks.

In the meantime, the global internet network is a platform for handling and flowing knowledge [60].

9. The goal of the analysis in [61] is to establish an overview of current RFID research from an IoT overview. The analysis includes passive UHF (860-960 MHz) devices (e.g. low battery) that can provide an effective sensor network to track human health and control the efficiency of the controlled environment [62, 63], with sufficient measurements and readings. Environmental sensors that calculate physical parameters such as temperature, humidity, and toxic substances may calculate the health of the environment and evaluate their human health relations [64]. Wearable and implantable labels offer more contextual knowledge about a person's in-room appearance, activity, contact with the individual, and innovative applications based around the body. Data on the health status of artificial organs can be produced [65]. When operating with special compounds or inter-linked with chips which include detection functions, passive RFID tags can detect changes in environmental chemical / physical parameters. Autonomous RFID tags are an important instrument for designing body-centric healthcare systems that are fully transparent to the consumer [66, 67] for touching and approaching the body. Most specifically, wearable passive UHF tags have long been technical tabus because of the significant power outages caused by human tissues. Previous body center network studies are mainly focused on active computers [68, 69]. Over the past five years, the development of commercially available low-power RFID (COTS) chips has changed gaming entirely, creating and testing of several prototypes of body-centered, sensitive, distance-readable tags compatible with remote-controlled monitoring [70, 71]. Since high electromagnetic losses characterize the human body, antenna power output is meager and the required reading distance for passive devices. The performance of the antenna power is deficient. This combination of wearable and ambient tags helps to track the situation of children, people with disabilities, and the elderly, identify human behaviors, and build passive RFID systems in a fully-passive manner [72, 73].

10. Researchers in [74], which explicitly addresses the safety criteria for RFID validation schemes, have submitted an authentication scheme based on elliptical curve cryptography (ECC) suitable for their protection and health care climate. The aircraft device Classify Friend or Foe (IFF) was the beginning of the RFID technology used during World War II. The benefits of RFID are that it is easy to use RFID tags, has read/write power, does not need long-distance RFID readers communication, and can simultaneously read multiple RFID tags [75]. RFID technology in the healthcare field is applied to the verification [76, 77], birth and patient recognition [78], medical treatment detection and confirmation [79], healthcare centers [80] and surgery sites [81]. A typical healthcare network with RFID technology [82], as shown in Figure 4. RFID authentication is one of the most critical measures to ensure secure communication inside the RFID system. However, there is a wide range of security features in the process of sending messages between RFID tags and RFID readers. Some researchers have established the safety criteria for secure RFID communication and reliable and effective authentication schemes when it is applied [83]. Mutual authentication between RFID tags, RFID readers, and servers are among the security requirements. Besides, before transmission, the confidentiality of sensitive information contained in the RFID tag must be encoded. The identity of the tag must also be protected for the protection of the owner [84]. The next requirement of protection is to validate the RFID authentication scheme during the life cycle of the RFID tag that gives alerts updating the confidential details that have shared during the implementation of the authentication scheme. It meant that the verification scheme would be invalid if the modified coordinate is lost. To ensure secure transfer [85], Transfer Protection requires an RFID authentication scheme. For certain RFID authentication systems, if the enemy successfully retrieves sensitive information from the RFID tag, it can detect the last position in the tag. It will severely breach the privacy of the user. As the RFID tag validation process is required, scalability, RFID authentication scheme must scale, an RFID device server must find the corresponding records in their database [86].

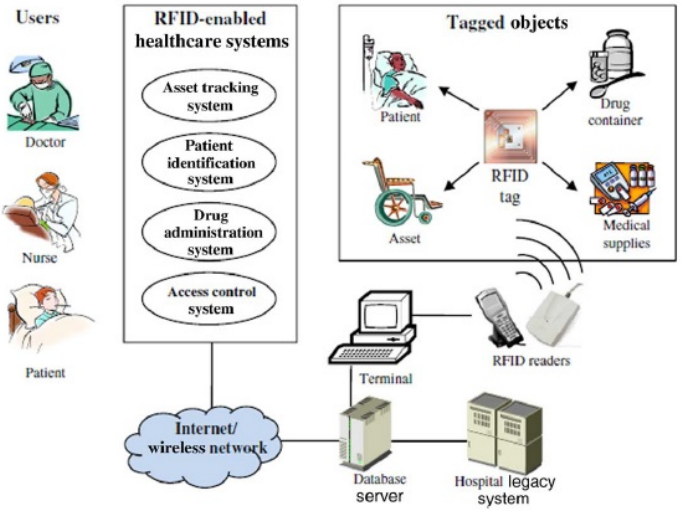


Figure 4: Typical RFID-based healthcare system [82]

As the number of RFID tags increases, the device does not scale if the workload of the search algorithm is significantly increased. RFID authentication processes must defend against multiple attacks, duplicate attacks, tag defense attacks, server spoofing attacks, mid-level attacks, cloning points, and others. Resistance to attacks to ensure the security of RFID communication [87]. The researchers found that the ability and memory of tag calculations were minimal when comparing the performance and safety aspects of the ECC-based RFID authentication schemes to determine whether ECC-based RFID validation schemes are adequate for practical applications. Cost measurements, communication costs, and storage requirements for practical applications are, therefore, important. The researchers investigated and found that most of the current ECC-based authentication scheme proposed

cannot comply with all safety requirements. To ensure secure contact, researchers will first build acceptable safety models for ECC-based RFID systems by using ECC-based techniques in RFID systems. They then developed an ECC-based RFID authentication system that protected security models securely.

11. IoT consists of an increasing number of interconnected smart devices that reduce hardware costs and architectural complexity, where communications occur at any time or anywhere. The reading of ultra-high frequency RFID tags is a cost-effective operation because of its technology and the reading system for these frequencies. The researcher proposes an architecture that reduces the cost of ultra-high frequency RFID reading equipment operating with cloud computing and micro-services in an infrastructure. Due to the scalability and management of large quantities of data that can be produced using ultra higher frequency RFID tags and the nature of the architecture linked to the theme of the research, the use of cloud and micro-computing was required [88]. To collect data and analyze the findings, the architecture proposed to address the research questions intended for the case study of this work has been created. After the RFID read tag has been detected by SparkFun Simultaneous RFID Reader — M6E Nanoboard, the Asset Micro-service must interact with the Read microservice to see if a similar physical object remains. This test takes place via the RFID tag area Electronic Product Code (EPC) present in the Asset microservice relational database and the Read microservice NoSQL database. Since the RFID tag EPC is a unique identifier, this is the only field needed to check whether an RFID tag is in the Asset Microservice database, identified by the SparkFun Simultaneous RFID Reader — M6E Nanoboard and registered in the Read microwave database. The monitoring microservice is used to view the collected data after telemetry information has been collected from the SparkFun Simultaneous RFID Reader — the M6E Nano Temperature, Position and Intermittent Connectivity Information through Microsoft Azure, the IoT Hub service. The result obtained from the case study as below:

    - The board was able to read RFID tags with a distance of approximately 1 m and could read approximately 100 RFID tags per second only with its internal antenna. Having an external UHF antenna together with the board will improve the board's reading power. If the reading distance has a minimum value of the required project execution, the use of an external UHF antenna is needed.
    - There were no charges for the software provided by Microsoft Azure. The non-cost is due to the use of all the free layer services. The cost will slowly increase as additional resource resources such as further computing, memory, or monitoring of other IoT boards increases.

3 Security of RFID

Some of the advantages of using an RFID device is that it performs tasks to classify, achieve efficiency and simplify its implementation automatically. Even these added benefits do not guarantee personal data protection. Therefore, when designing an RFID device, several criteria must be addressed to provide the highest degree of security. They are exchanged for the user or included in tags sent to the backend server. Data is private, accessible, and not modifiable (CIA). The safe maintenance and operation of the entire network A fully secure network continue to be a challenging challenge with numerous attacks and threats which that happen in various scenarios at the same time. RFID forgery: RFID can split into three groups by processing ability, which are simple tags, symmetrical key tags, and public-key tags. Basic tags are not encrypted and can be fooled easily. Supply chain and the operation of the travel industry are well established. To access or test product reliability [89], an attacker might write information to a primary black tag or change the data in a simple write tag. Many articles listed in the solution identified typical attacks which can occur on RFID systems during the investigation. Some of the most common attacks are shown below. In the literature, many steps and features have been suggested, including pseudorandom solutions, anonymous authentication schemes, symmetric and asymmetric cryptographic algorithms, and other hash-based systems. Several review studies related to the following security attacks are discussed in subsection 3.1.

1. Eavesdropping
   These attacks include eavesdropping on the network and recording examples of all interactions between the tag and the reader. The primary purpose of this attack is to collect data and use the information collected.

2. Denial of service
   Denial of service is the creation of a device that cannot use in an RFID system. An example of a use case for this attack is spam, which is sent to readers requesting a specific tag frequency to deactivate the reader.

3. Cloning
   The primary purpose of a clone attack is to spread RFID tags. Reverse engineering technology allows it to change and duplicate the label by extracting all the functions of the vehicle, precisely the secret key.

4. Tracking
   When detecting attacks, the tags voluntarily tied to a person, and their presence and movement are detected while the attacker's leader is inside the tag.

3.1 Security Review Studies

1. The IoT is the opportunity to have a means of distinguishing everyday objects and another way of interacting with each other. There is an extensive range of IoT technology domains like smart houses, smart cities, wearables, e-health, and others. Thus, tens and even hundreds of billions of devices are related. These machines would have intelligent capabilities for collecting, evaluating, and even making decisions without any human contact. Under these situations, protection is a supreme necessity, and authentication under particular

is of great importance given the harm that might result from a malicious, unauthenticated device in an IoT network [90]. This paper provides a general overview of the security issues and requirements in a layer-based approach in the IoT context. Instead, it offers an updated survey of the various IoT authentication schemes. It contrasts and analyzes the current authentication protocols through a multi-criteria framework, showing their advantages and disadvantages. The researchers found out the following:

- The authors proposed a lightweight RFID tag authentication protocol based on Physical Unclonable Functions (PUF) functions. There are three transactions in the protocol: tag identification, authentication, and update. The tag reader identifies the tag in the first transaction. The second transaction is a test where the reader and the tag verify each other's validity. The latest used key for the next check should be retained in the last transaction (Update).
- The authors have allowed the authentication and perceptibility of the IoT devices to secure the supply chain of connected devices through an RFID-based solution. The process of authentication consists of two steps: the verification of the connection between the tag and the IoT system and the tag perceptibility.
- The RFID reader is connected to the Internet in an IoT-RFID network to form an IoT end unit. On the other hand, it is linked via RFID communication protocols to the tagged objects. The tagged object is portable and transfers from a reader to another, so authentication is necessary to verify another identity. The device is vulnerable to security attacks, such as impersonation or clone attack, because of the absence of cryptographic features in RFID. The authors proposed an authentication protocol to be used with lightweight encryption algorithms in the case of IoT-RFID use.
- The authors suggested an offline authentication for PUF-based RFID-tags to prevent cloning attacks on the RFID tags. It combined security protocols for identification and digital signature. The tag generates a hidden key in authentication by questioning the PUF and gathering the answer. Such a reaction with the support data will create a certificate that will be stored inside the tag's ROM. First, the tag is authenticated by the verifier verifying the certificate validity. The authors proposed a PUF-based authentication scheme for the classic RFID tags to provide anonymous authentication for RFID systems. Then, they created an enhanced PUF noisy atmosphere scheme. The main downside of this scheme is not to re-feed the server with new Challenge-Response Pair (CRP) when the current pool is empty.
- The authors suggested a shared authentication scheme for IoT RFID-based applications in fifth generation of mobile networks (5G) by providing a cache to the user, storing the keys (used for authentication) for the tags recently visited, speeding up authentication, minimizing computation costs and the storage security. On the other hand, the authors in [91] has been proposed authentication methods for heterogeneous networks which can be applied to RFID systems as well.

2. The detection of ultra-high frequency radio frequencies (UHF RFID) is becoming a central technology in the Internet of Things. It allows batteryless and wireless nodes to be deployed like sensors and actuators. Safety is a critical issue because of its possible transmission range is greater than 10 m, and the ability to hold sensitive information. For this reason, in 2014 the International Organization for Standardization released several UHF RFID crypto packages under the ISO-29167 standard. Nevertheless, no UHF RFID readers are currently available. In order to cope with the rapid growth of software-defined radios, "simple" and versatile readers are required. This paper illustrates a way to implement the decryption of ISO 29167-19 RAMON on a prototyping reader platform based on National Instruments PXIe chassis. PXIe programming is based on LABView. The first LABView implementation revealed that decryption only requires 51.3 s. It would mean being able to process just one tag per minute. Therefore, a prototyping platform intended to conduct practical experiments in the creation of new technology in UHF RFID systems is not suitable. Therefore, they prefer a different approach. The LabVIEW also acts as a platform and monitors the transceivers which are integrated. Nevertheless, a TCP/IP connection to a Java application is also provided. This Java code is decrypted. A TCP/IP link is used to return the decrypted plain message to LabVIEW and can be further processed [92]. The results are as follows:
    - The average time spent with LabVIEW in tandem with the Java app is 2.2 ms, which improves the time-optimized approach to LabVIEW by 23,318 times. Within LabVIEW, the principal difficulty is that large numbers are difficult to deal with since they are widely used in encryption. The problem is not limited to the implementation of the RAMON method by outsourcing the calculation in a Java program. The approach can be used for any encryption method and for other processes in LabVIEW which handle large numbers. In this way, the RAMON encryption can be incorporated into the existing reader protocol and a fully working, easy to adapt reader can be obtained. Particularly for UHF RFID systems, it creates entirely new possibilities for rapid growth.
3. The medium of contact between the tag and the reader is vulnerable to multiple threats in conventional RFID systems, including denial of service, spoofing, and desynchronization. Therefore, the confidentiality and validity of the data transmitted can not be guaranteed. In this paper, the researcher proposes a new RFID authentication protocol based on a lightweight block cypher algorithm, SKINNY (short for LR-SAS), in order to solve these security problems. The 128-bit block size and key size n SKINNY encryption algorithm are being used to test the implementation in passive 96-bit

EPC-encoded radio frequency identification systems. The tag and the reader need to be mutually authenticated when implementing the LRSAS protocol, and the SKINNY algorithm encrypts the information used by the tag and the reader for mutual authentication. The key is already stored during the initialization process. This paper selects a SKINNY lightweight block cypher that has the advantages of low hardware power consumption and low computational complexity based on ensuring safe encryption. It can be used in Things terminal equipment's low-cost Internet [93]. The results are as follows:

- The protocol conceived in this paper uses one of the algorithms for SKINNY encryption and can accommodate 96-bit EPC encoding. In the encryption process, the calculation time for the round function used by the SKINNY encryption algorithm is shorter than the calculation for Hash, ECC, and Present encryption. However, the overhead measure also applies to low-cost RFID tags. Additionally, the tag's overhead storage is 3 L which significantly reduces the tag's storage capacity compared to other protocols, and reduces the complexity of the storage structure's logic gate design. Besides, the protocol has five information interactions in the mutual authentication of the tag and the user, and the total amount of data collected and transmitted is 6 L, which is relatively high, thereby ensuring the efficiency of the information interaction.
- Concerning the number of logic gates equivalent, multiple versions of SKINNY have varying numbers of logic gates equivalent. This protocol uses version SKINNY-128-128; the equivalent number of logical gates is 2391, less than 3K. It can also be used in low-cost tags. Furthermore, the number of equivalent logic gates for other protocols often allows specific security attacks to become vulnerable.

4 Challenges and environmental implication of RFID application

RFID is an exciting technology, but it does pose many technological problems and implications.

4.1 Challenges

The challenges of RFID technology applications in IoT were divided into seven: collision problems, privacy and security concerns, interference issues, miscellaneous challenges, technical issues, interoperability issues and energy efficiency.

4.1.1 Collision Problems

Electromagnetic interference can cause tags and readers to fail, leading to wrong choices. For tags and readers using the same radio frequency channel, multiple transmissions between the reader and the tag can cause problems in a collision. To recognize multi-tag communication to promote the implementation of large-scale RFID systems, a secure and active collision avoidance protocol is required—a database tree protocol, a binary tree protocol, a split system protocol and others. Nevertheless, not all protocols produce the requisite efficiency as their identification efficiency is less than 50% commercial. For a new, better protocol will look for the best features of the detection protocol [94, 95].

4.1.2 Privacy and Security concerns

RFID authentication tags and security issues can affect people and organizations significantly. Tags are not properly secured and they considered an easy target to eavesdropping, DoS strike, analyze traffic and others. Besides, unauthorized readers can view multiple codes, which can breach privacy, without single access control. Predictable perceptions of tags can accompany secure tags. Traffic analysis attacks may jeopardize the privacy of the user. Appropriate confidentiality and protection protocols are needed for RFID systems given the associated costs. Extensive work is underway to define and incorporate cost-effective protection and privacy protections to improve RFID systems' reliability and usability. Several light-weight protocols have been proposed, but still have high costs and low protection and do not completely solve security issues. An excellent work section is available to develop ultra-light encryption protocols for low-cost RFID systems [96].

4.1.3 Interference Issues

Recently, a very high-frequency passive RFID (UHF) system has received much coverage. UHF RFID systems are widely agreed that many commercial applications, such as supply chain management will revolutionize UHF RFID systems. In the 860-960 MHz band more popular supply chain companies like Wal-Mart and Tesco will promote the use of UHF RFID systems in supply chains. However, with this distribution, dozens or hundreds of readers will work closely together and trigger to severe interference [97]. The UHF RFID interference has three types: tag interference, reader-to-multi-tag interference and reader-to-reader interference. The specifics of the intrusion issues are shown below.

1. Tag interference
   It happens when the reader simultaneously supplies power to several tags, and each signal reflects the reader. Owing to the dispersion of dispersed waves, readers cannot differentiate between individual IDs and marks. Anti-collision mechanisms, called binary trees and aloha, are therefore essential.

2. Reader-to-multi-tag interference
   It occurs when the mark is at the intersection of two or more reader requests, and the reader tries to interact with the tag simultaneously causing the tag to behave and interact unwantedly.

3. Reader-to-reader interference
   These are activated when a signal from one reader hits another and can also take place if the readers' surveys are not intersected. Interference between readers can cause serious

problems when using UHF RFID systems, as remote reader transmitted signals are so powerful that precise decoding of messages transmitted by neighbouring tags is prevented.

Active, semi-passive and passive RFID tags add comfort and prominence to RFID technology in our environment. Such tags are cheaper to produce and can be small enough to suit virtually any product. The type of RFID device works is shown below. Readers and tags focused on wireless two-way communication. The reader transmits a signal in passive RFID tags which induces a current in the tag to supply power to the transmitter of the tag. On the contrary, an active RFID device does not require an induced current to transmit its presence. Wireless communication may create a difference between RFID and other wireless technologies that are used in the workplace or data center since other RFID technologies often operate in the atmosphere and cause significant disruption [97]. Two classes ought to consider for new action. First, interruptions prevent data transmission properly and reduce the efficiency of one or more devices. Second, the danger of interpreting one signal from one device is that data from another is valid. The key cause of this disturbance is the environmental implications of use. This problem occurred in the past when the RFID tags placed on metal or liquid containers. The tag is inactive in both cases and refuses to respond to the reader. Recent advances allow these problems with proper antenna and tunnel design to tackled during the planning process. The UHF system may cause interference with the radiation of the power signal. Passive systems, on the other hand, are less vulnerable to intrusion than active systems with markers continually shifting. Crosstalk between RFID and WLAN or WPAN such as Bluetooth is another interference scenario when devices have a joint or an adjacent frequency band. Several cases of interference between wireless microphones and wireless endoscopy recorded in the medical field [98, 99].

1. Active and semi-passive RFID tags
   It uses an internal battery for the circuit. Active tags use an array to transmit radio waves to the reader, while semi-passive tags depend on the reader for transmitting power. Active and semi-passive tags reserved for expensive items capable of reading long distances, transmitting high frequencies from 850 to 950 MHz and reading more than 30.5 metres. If the label needs to read from a distance, an additional battery can increase the name range to more than 100 metres.
   It occurs when the mark is at the intersection of two or more reader requests, and the reader tries to interact with the tag simultaneously causing the tag to behave and interact unwantedly.

2. Passive RFID tags
   It entirely depends on the reader as strength. These tags are read and produced for use up to 20 feet (6 meters) and with disposable consumer goods.

4.1.4 Miscellaneous Challenges

Several other issues prevent RFID devices from spreading too. The key issue here is cost. The RFID tags are more expensive than the [100, 101] written stickers. A further issue is the complexity of the RFID system. They are owing to their high-reliability readers, and labels are not on the market yet. Moreover, there is no guarantee of accurate recognition. The incorporation of RFID systems into the current network is another big problem needing a solution for using [102], large RFID systems. The designers need to know the approach to incorporate the new RFID system into the current network in a successful development phase.

4.1.5 Technological Issues

Implementation issues are most apparent and should be dealt with immediately. Some papers dealing with technological issues such as durability, precision, and contact with metals are still in the pilot process. For DHL, read speed accuracy is not 100%, because Tag interaction readers do not comply with defined specifications [97]. Even, it is not feasible to prohibit maximum use of short contact tags. The concept in citeb101 is to make sure that the people do not forget the process of activating RFID devices.

4.1.6 Interoperability Issues

If the RFID device is not working alone, but in conjunction with other electronic devices, the issue of information becomes apparent. For example, an RFID Smart Vareflyt system can operate independently, but will not work if it combined with an existing data system [100]. This question occurs when the real data is different from the reported data, does not change, or involves correction from other sources. On average, the accuracy rate in Nestle Italy's system is only 80-90% [99], due to the order output depending not only on the role of the RFID portion but also on the RFID system and the entire enterprise policy [99]. However, one of the hardest problems facing RFID technology is privacy and security. As corporations and public agencies continue to search for collaborations, sharing resources poses obstacles that involve sophisticated organizational solutions. The most challenging problem in the case of Smart Vareflyt is not just the issue of data incorporation at the knowledge level, but also the exchange of information and resources and the application of competition law to the level of the company [100].

4.1.7 Energy efficiency

The energy-efficient passive tag in the RFID device is a preferred option because it used without the need for an on-board energy source. The selection of passive tags is, however, limited compared to the active tag, which used on a broader scale. Nevertheless, the inactive energy efficiency tag is still an open research field that needs to be addressed [103]. As power usage decreased, energy efficiency improved. Several studies on power usage when using the RFID technology are discussed below .

1. The first large-scale backscatter communication system was developed by researchers in [104] for reflecting signal from an RF source to synthesize data packets, which are then decoded by a receiver. It has two significant drawbacks, and the backscatter must first organize the information so that the receiver can transmit it at a signal force less than -135 dBm and work efficiently in the event of robust external interference. Secondly, backscatter signals developed using

simple, affordable standard hardware, which makes it simpler to implement and improve the design needed instead of dedicated receivers for costly backscatter signals (e.g. RFID readers).

- The result shows that wireless power consumption, including Wi-Fi, BLE, ZigBee, Lora, and SigFox, consumes 10-50 mW of electricity, which is 3 to 4 magnitude orders higher than the system's energy consumption. Backscattering technologies, including RFID, passive Wi-Fi, and LoRa backscattering technologies, however, consume just $9.25\mu W$ for continuous frequency-modulated chirps generation and harmonic cancellation.

2. Researchers in [105] introduce an on-board digital temperature sensor with perovskite PV-RFID. This sensor can transmit temperature data wirelessly under sufficient lighting within a distance of up to 4 m. It is converted from light with perovskite PV into ICs and RF signals that used as contact links. The user acts as a portal to the Internet between these hundreds of nodes.

- According to the simulation performance, the overall energy consumption of the sensor depends on the sensor load (the number of sensor measurements per hour). For instance, a 20,000 / hour measurement requires approximately $20\mu W$ of power (calculated with a constant current of $6/muA$ in stand-by mode and 30μA in 8 mm a 1,5V input voltage each measurement).

3. Photovoltaics (PV) for use by researchers in [106] are the use of photovoltaics (PVs) to increase the energy available during the day, enhance read range and detection capability and add to integrated circuits (ICs) with minimum additional electronics.

- The researchers found that the IC's additional power not only enhances the ability to disperse the RF energy needed for additional tag areas but also enhances the ability to measure and transmit sensor data. This versatile sensor is also suitable for long-term detection embedding in 3D structures if the battery can not replace it. Thus, using PV-RFID can build wireless sensors that are autonomous, inexpensive, and reliable over a long-range. The frequency of PV-RFID measurements is about eight times that of traditional passive tags. By reading a wide area of PV-RFID tags with a limited number of readers and antennas, it lowers the network costs. The PV-RFID sensor additionally records 0.14 million measurements in 24 hours. It shows how sensors can monitor and transmit critical environmental data at low cost, without batteries.

4. PLoRa, an environmental backcast architecture that enables a remote wireless link for IoT devices without a battery, is available from the researchers. PLoRa uses the current LoRa transmission as a new signal, modulates the new signal into the standard LoRa chirping signal, transfers the data to another LoRa channel received by a remote gateway, and adjusts this original signal [107].

- The results of experiments show that the prototype of the PCB-PLoRa tag backscatters an environmental broadcast of LoRa that is transmitted from a local LoRa node (20 cm) to an entrance up to 1.1 km away, with data accessible indoors or 284 bytes every 14 minutes. Simulate a lower power FPGA 28 nm prototype whose digital baseband processor delivers 220μW in power consumption.

4.2 Environmental Implication

RFID tags are used for managing product movement from factories to store in cartons and pallets. RFID has a great many useful applications for the world. RFID also has environmental drawbacks at the same time. The next topic will concentrate on ways and concerns that help create useful solutions and reduce possible impacts on the environment.

4.2.1 Reduced traffic using RFID

In order to pay for tolls, RFID congestion, and area pricing systems now deliver major environmental advantages. Automatic charging reduces downtime as well as braking and acceleration at the mail stop. Through the use of RFID repeaters on the New Jersey Turnpike tolls, 1.2 million gallons can save annually, air pollutant emissions decrease, and travel times decrease dramatically [101]. London has lowered the congestion level by 26%, nitrogen oxide, and particulates (PM-10) by 15%, and oil usage inside the congestion billing zone by 20% [102]. With the surface charging network introduced in Singapore since 1975, circulation has dropped by 45%. In 2006, the Stockholm scheme had reduced carbon dioxide emissions by 10 to 14% [108]. While it is very effective to reduce traffic, air pollution, and energy usage, plans for congestion charges are not common. The framework became popular after implementation [107].

4.2.2 RFID transponders on curbside waste and recycling bins

The waste truck fitted with an RFID reading panel. The device recognizes the bin when a truck gathers it. In 2001, a German public waste management firm surveyed RFID schemes, unimplemented containers, and productive waste collection routes through data that were exposed and evaluated. The RFID system implemented reported an average increase of 2.85%. In cities with invoices focused on waste numbers, non-recyclable waste decreased by 35% on average, and total waste decreased by 17% cent on average. Contains recycling and non-recycling stuff [109]. As a result, almost half of the overall waste reduction accounts for a 35% percent drop in the garbage disposal, and the other half is for recycling. This finding is consistent with the US surrender program, which reduces municipal waste by 25-35% for disposal [110].

4.2.3 RFID for information on the consumer environment

Item level RFID tags may provide information about different goods as well as generic products. That sort of knowledge can use to identify tainted beef, poultry, or fast food. Regulators, the foodservice industry, and regular customers are helpful with this [111]. At the same time, users will test how far they have transported, greenhouse gas emissions, allergens, and much more. Manufacturers and retailers have commodity carbon labels where a UK Tesco supermarket network sets carbon labels on each product sold, calculating the amount of greenhouse gas released during the manufacture, transport, and consumption of each product [112].

4.2.4 RFID for refurbishment and reuse

The RFID tag increases the efficacy of recycling as the UPC process enhances companies' productivity in the supply chain. For example, battery labels may make the sorting of cell types cheaper and more effective for recycling. Electronic equipment labels may connect to websites that show how to disassemble the device. Tags on hazardous goods, such as household chemicals, will define the material, how, and where to dispose of these [113]. Attributes on consumer goods make it easier to sell products online and are an extension of how UPCs used in the one-way supply chain where the majority of items sold today have UPCs. This code is intended for revenue estimation as well as for inventory control [114], distribution, consolidation, and managing. For RFID, instead of being discarded with it, it will leave the object's security code. The product code will allow many kinds of recycling opportunities. It may create systems that encourage recycling or punish improper disposal by recognizing when recyclable (or harmful or valuable) products placed into the garbage (or garbage can or trash can). Recycling opportunities, such as bottling, are not new, but an automated product-based solution will minimize costs and promote for innovation [115]. RFID systems provide a modest incentive for recyclable low-value goods and a significant incentive for high-value and high-risk goods. One sidewalk recycling company will handle specific items electronically for different goods, different markets, and various geographic areas [116].

Product codes can be applied both as UPC and RFID barcodes. When a large amount of device is picked up in a curb or code and put into a truck quickly, RFID could be preferred over bar code, as RFID can read from other devices at the same time. The benefits of recycling not deemed adequate to justify the expense of the product-level RFID tags. However, if RFID correlated with an item for other purposes, the side benefits of RFID processing can be significant [117]. The use of RFID for electronics replacement, roadside recycling, smart household waste disposal, the management of product deposits and returns in stores and recycling centers, and the collection of waste and recyclables in incinerators, landfills and resource recovery plants addressed in many different applications.

- RFID for Electronics Refurbishment
  Electronic recycling companies typically recognize multiple labels, computer models, printers, or mobile devices related to the most crucial element of the faulty parts market in their product instruction codes. The recycling industry, the product supplier, or the whole industry union RFID labels in products may use these instructions to connect with databases where both parties may learn how to recycle papers. Will major component can have its RFID tag so that repairs can easily recognize parts and updates [113].

- Item-Level Curbside Recycling
  The use of RFID transponders has addressed in waste management and recycling containers. Through extending this kind of method, residual or recyclable material can detect at the item level. It opens up a whole variety of reward schemes. Incentives (rebates and fees) sent to the customer based on what put on the bin. Municipalities can also pay recyclable collectors extra fees to transport dangerous and challenging things to handle, such as mercury-containing fluorescent lamps, and recycled collectors can quickly sell their collections if they know what they have in-depth. This approach raises problems of a technological nature. The objects in the bin are one or two feet of the antenna for recycling vehicles, and probably other side and rear products which can absorb radio signals. Although RFID tags can be detected theoretically in products in recycling bins and empty bins can be ejected in lorries, this has not commercially demonstrated.

- Smart Trash Cans
  The RFID reader can read each object in the container for home recycling tanks. Drum items may either be sold online or published by traders. Consumers and businesses will automatically check for waste material and schedule recycling services. Online retailers check for the recycling containers, buy recyclable and reselling items, give back some of the consumer's income, and leave vital recycling services (paper, plastics, artistry, metal). Home intelligent trash will play mostly the same role as the item level curb set. Intelligent home trash can be technically pure but probably a more expensive alternative than the identification of congestion objects.

- In-Store Recyclables Collection
  RFID allows the storage in shop or recycling centers of recyclable items. Some retailers use recycled items already. Recycling Corporation Recycling Corporation (RBRC) provides electronics and electronic stores with batteries throughout the United States and Canada. For different computer producers Costco has a standard marketing and recycling plan [114]. Mobile telephone providers accept used telephones as well. Retailers often use UPC bar codes to handle returns. The use of product codes to provide customers with discounts using RFID tags is economical and straightforward. Like every other voucher or swap system, retailers may benefit from a discount plan. Selling machines can also accept recyclable goods and offer cash refunds or credits based on details on the product label.

- RFID at Materials Recovery Facilities
  To organize things in warehouses, incinerators, and dumps, using RFID readers. Modern materials recovery facilities include magnetic steel seams, aluminum vortex separators, and plastic seam infrared detectors. RFID devices can track small electronics and batteries. Previous studies have shown

that RFID-labeled items such as batteries, small electronics, and mobile devices can obtain at a much cheaper collection facility than recycled [113].

4.2.5 Environmental drawbacks of RFID

Many consumer goods include RFID tags, including food packaging. Over time, the number of tags produced and discarded increases. Usage of the global tag will reach 2 trillion a year [115]. Such marks end up in urban solid waste or systems for recycling paper, carton, metal, glass, and plastic. RFID tags typically include copper, aluminum, or silver antennas as well as integrated silicon circuits, adhesives, plastics [116]. In short, the use of RFID materials and removable tag properties allows the avoidance of problems that cannot be degraded or recycled as RFID tags can prevent the recycling of various products and other materials [117]. When carbon-based conductive ink used as a substrate antenna and bio-based plastic, most RFID tags can be biodegradable or ecologically friendly. However, silicone chips may have environmental effects, which at the development and not the stage of life are challenging to deal with reliably. The processing of silicon chips requires considerable energy and water and significant emissions of pollutants. A specific problem was the pollution of glass, steel, aluminum, paper, and plastics recycling, discussed in turn below.

- Glass Recycling
  Since silicon melts at speed other than glass, silicon RFID chips can pose problems for RFID tags attached to glass containers. If the silicon remains in the glass of the entire recycling line and the glass furnace, silicon beads can create in a new package. It is a potential weakness, especially for the media, and it shows that RFID tags used in the glass must be removed entirely or placed inside the lid and not on the glass itself [117].

- Steel Recycling
  RFID tags can contain approximately 20 mg of copper. On a regular 20 g steel box, this corresponds to a working copper concentration of 0.1%, and the steel copper content accumulates with continuous recycling. It is rising steel consistency. Therefore, RFID tags for steel goods should be entirely removable or made with copper-free antennas. RFID tags mounted on aluminum do not impact the recycling of steel. Inadvertent aluminum falls from tags removed during the exothermic reaction at very high steel temperatures [117].

- Aluminium Recycling
  Aluminum can be recycled with high-value many times. Flows of feed do not permit pollution. Pollutants shall be picked or removed by other means. RFID tags for aluminum items are either produced or made of eco-friendly materials.

- Paper Recycling
  RFID tags will join an old carboard container in the recycling facility. The National Aeronautical Research Committee (NCA) tests have shown that copper foil antennas are kept intact and can be easily captured and removed during the entire hydro pumping process. The silver ink antennas, however, contain 2 to 3 microns of silver particles. Pilot experiments using labels containing around 16 mg of silver found that much of the silver in the dough remained. In certain cases, however, there was insufficient proof of rates exceeding the limits and regulations. Found in reverse water.

- Plastic Recycling
  The use of High-density polyethylene (HDPE) for milk bottles, washing machines, and other applications. Application of polyethylene terephthalate (PET) in bottles and other applications for soda and wine. Some of the most often recycled plastics are HDPE and PET. In general, RFID labels applied to substrates of PET. Thanks to the denser RFID tags than HDPE and their isolation during the first stage of the HDPE recycling process, RFID labels are anticipated to differentiate easily from HDPE bottles during the recycling process. However, it is a struggle to remove RFID tags from PET bottles. PET is more than one (1.2–1.4 g / cm3) in specific gravity, and both RFID and PET fragments immersed in water. As RFID tags typically performed in PET substrates, the mass difference between PET and RFID tags may not be sufficient using cyclonic separation. Every RFID tag used on PET bottles should also be designed to be easily removed, either by non-PET substrates or other methods.

- Towards RFID biodegradable systems
  However, most RFID tags can not be biodegraded. These include petrochemicals-based metal parts, plastics, or other materials. Smaller tags also make it hard to isolate and recycle different label components. Recent researchers have developed completely biodegradable tags, in particular for medical and feed supply chain applications [118].

4.2.6 Environmental Sustainability

While most case studies at various rates have achieved environmental sustainability, Walmart's case shows that the bigger the size, the higher the effect. A breakthrough in RFID implementation by Walmart that replaced the barcode system by Walmart decreased carbon footprint by 3.2% by allowing less product movement because of improved product visibility and handling and improved inventory management [116]. The TruckTag project is another new environmental protection program. RFID tags allow for automatic high-speed safety checks, reducing traffic jams, and improving air quality by reducing the lorry idle time [116].

4.2.7 Business Value

The average organization aims to reduce costs when introducing a green RFID scheme. The City of London School for Girls, with its RFID temperature sensor networks, will reduce the cost up to 80% compared to its wired networks and gain greater control over room temperature via RFID [117]. Using RFID to monitor temperature, Nestle can save up to 10% on energy-related costs, reduce harm, and boost efficiency. Although direct cost savings are legendary, indirect cost savings should also be acknowledged (such as loss-of-life insurance and increased quality). Many businesses use RFID for revenue generation. Recycle Bank, and Recycling Rewards offer recycling coupons to customers to launch a new business model and

promote more recycling families. Bank Recycle has raised its sales by collecting more recyclable products. The use of RFID saved over 318,000 trees and 21 million gallons of petroleum in the Recycle Tank [44].

4.2.8 Social Responsibility

Green RFID initiatives have an impact not only on the safety and sustainability of the ecosystem of business but also on environmental conservation, which adds value to customer-to-company relationships. It is because consumers know what they want in investment in the environment. Nestlé Italia uses an RFID-compliant management framework to support green ads and to boost brand identity successfully. Under public demand scrutiny, major businesses urgently have to meet their social obligations [99]. Through its smart truck initiative, DHL has minimized CO_2 emissions and fuel consumption by automating package scans to speed up driver handling of the parcels and idle periods to shorten the truck engines during delivery. In more environmentally aware businesses, the culture in which they work may have a positive effect. People and economies with their bank recycle point have raised more than one million dollars [97].

5 Benefit of Greening IT

Green technology is a concept that includes almost every type of practical use of technology to benefit the environment. Despite industrial technology advancing so far, its green predecessors are supposed to do the same. Many green technology examples include items that come under the given use of the subject, such as green chemistry, environmental monitoring, and others. Anything that fits the accepted concept of green technology will do one or two things. It is either to preserve or to help the ecosystem in any way. The words ”environmental technology” or ”clean technology” can also be used for this technology. Green tech’s primary goal is to protect nature and help reverse the adverse effects humans have had on the planet [118]. Green technology enables us to turn our contaminants and waste from unnecessary by-products into a sustainable workforce that benefits humans and the environment. Green technology has some of the best recycling methods: turning waste into fertilizer. One of the most significant benefits of green technology is that it helps people recycle things like water and air. Green technology leads to emissions reduction and air purification for future generations. Above all, green technology has permitted people to save vast quantities of resources, dramatically reducing the remaining carbon footprint. Alternatives to systems that use a high volume of electricity and fuel are provided to the public at this very moment. The use of items such as electric vehicles and solar panels has become natural, daily, and the Planet could not be the best. It is highly recommended that people use environmentally friendly tools and appliances to help facilitate this energy saving. The ability to rejuvenate ecosystems that have suffered a great deal because of the effects of human carbon emissions and other factors is also a tremendous use of green technology. The use of this technology replants trees, replaces plants, and replaces entire habitats to their original glory. Besides, waste is collected and recycled, so that animals, which once existed, will now return to these destroyed habitats. It means a new beginning for many species worldwide [119].

Environmental concerns [120] again impact the IT industry’s competitive landscape, and businesses with the expertise and vision to offer goods and services that tackle environmental problems profit from a competitive advantage. For example, when making transactions, leases, or contracting decisions, many consumers consider service provider registration and environmental initiatives. Companies face rising energy prices and could be subject to additional government taxes if they do not recognize the effect of environmental activities. Investors and customers are starting to seek more information about carbon emissions and environmental measures and business outcomes. Because of these issues, businesses start proving environmental efficiency, such as the Carbon Disclosure Project (https://www.cdp.net/en) an initiative that allows multinational companies to report their CO_2 emissions. The implementation of green IT practices gives companies and individuals economic and other benefits. IT operations increase energy efficiency, especially when electricity is costly and energy prices are increasing, through economically beneficial green initiatives. In a Sun Microsystems Australia survey, which included 1,500 responses from 758 Australian and New Zealand organizations, interviewed respondents said energy conservation and cost management were focused on environmentally sustainable environmental policies and environmental impacts. The primary reason for using the method shown in Figure 5 is that less is better said.

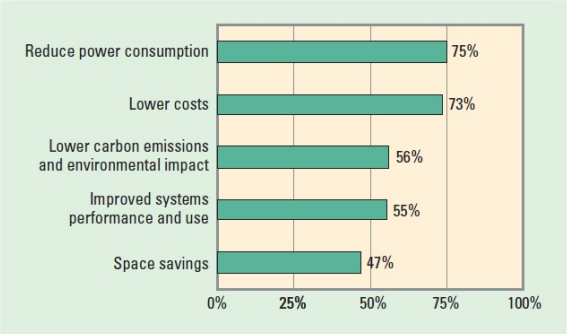


Figure 5: Reasons and benefits for using green IT practices[120]

Many businesses will concentrate on environmental conservation, energy usage, and cost control issues. The organization focuses on environmental protection through the frameworks for resolving business demands, legislation, and climate change. Business and institutional customers ask manufacturers to ”relive” their goods and processes. For instance, businesses like Dell and Wal-Mart take actions that compel suppliers to follow environmental-friendly practices. People start evaluating the essence of the IT environment, and green IT will become a regular function in the next five years. The company provides a range of innovative green goods and services, generating new markets for companies.

6 Future Work of RFID in Green IoT

Green IoT’s promising future will make our future world safer and greener, very high quality, socially and environmentally sustainable

and economically viable. Today, green communication and networking, green architecture, green IT and application infrastructure, energy-saving techniques, integrated RFIDs and networks, mobility and network management, the co-operation of homogenous and diverse networks, intelligent artefacts and green location all concentrate on exciting areas [121]. From two points of view, green IoT architecture can be applied to achieve outstanding performance and high quality. Identifying suitable techniques to improve QoS parameters (i.e. bandwidth, delay, and throughput) will make an effective and useful contribution to greener IoT. It would need to use less energy when it comes to greening IoT, searching for new resources, mitigating the adverse effects of IoT on human health, and disruption to the environment. Green IoT will then make a significant contribution to sustainable, intelligent, and green societies. Radio-frequency energy harvesting should be taken into account in order to achieve an energy balance to help green communication between IoT devices. Energy saving and CO_2 emission reduction is a major challenge for smart and green environment.

Logistics, Aero-ID, counterfeiting, supply chain control, spaces and much more will accelerate the growth of modern society and technology in the coming decades. Mobility, climate, safety, protection and energy, including medical, pharmaceutical and military applications, will become the leading technological drivers of the future, further growing the need for sustainability—wireless and electronic modules that are reliable, versatile, robust, low power and durable. Work on electronics and RFID seeks to develop versatile systems, based on organic/inorganic substrates and materials, to solve the limitations of ceramics and silicone in order to address these significant challenges. A challenge for potential applications is the convergence of elegant, lightweight, heterogeneous, versatile, environmentally sustainable, newer materials and cost-effective processes in a wide area. Continuous development and innovation involve a wide range of electrical components that can directly produce and incorporate from coils to affordable methods, mainly using organic electronics, along with work into a variety of electrical conductivity and semiconductor materials. Organic materials can be printed and modeled using several methods, each with its advantages and disadvantages, such as flexographic, width, offset, sorting, and injection. Based on these materials and techniques, primarily low-cost RFID transponders, different types of sensors, storage devices, photovoltaic cells, passive devices such as displays or batteries have tested active components such as diodes and transistors with the production of them. Though seen only in the low UHF (RF) range, the trend moves to higher speeds, resolutions and frequencies without compromising versatility, eco-friendliness and low costs. To address these problems includes a new approach to RFID and the design of electronic systems in terms of alternative materials, processes and functions. The researchers in [122] are proposing an innovative approach focused on developing a new low-cost inkjet printing platform with efficient diode integration for applications with optimized metallization traces and enhanced recognition capabilities. The method theoretically proposed can form the basis for the first generation of precise "smart" RFID devices.

More fields of research need to be explored to establish optimal and efficient Greening IoT solutions:

- Crewless aerial vehicles (UAVs), particularly in agriculture, traffic, and surveillance, are required to replace many IoT devices, which will reduce power consumption and emissions. UAV is a promising technology that leads to low cost and efficient green IoT.

- Sensor data to the mobile cloud must be more useful. Sensorcloud combines a network of wireless sensors and a mobile cloud. It is a hot and promising IoT greening technology. The device, service, WSN, and cloud management's energy efficacy can be examined through a green social network such as service (SNaaS).

- M2M connectivity plays a crucial role in reducing energy consumption and harmful emissions. In order to allow automated systems, smart machines must be smarter. In the event of traffic, computer automation delays must be reduced and intervention required and immediate.

7 Conclusion

IoT has an enormous influence on transforming our everyday lifestyle. IoT technology connects a large number of devices and items linked to the Internet around the world. The main concern is how data from these devices can be collected and analyzed effectively. RFID is one of the leading IoT tools, used for the identification and processing of information from IoT devices due to its cost and energy consumption characteristics. Studying this technology is therefore essential and paying attention to mitigating its adverse effects. This paper discusses and summarizes the latest RFID applications, including studies that have improved RFID applications in Green IoT efficiency. This paper also addressed the challenges and implications of green, IoT RFID applications. However, there are still some challenges to this field, such as sophistication, protection and lack of trust. These problems form a relevant research field for future study.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors would like to thank Majlis Amanah Rakyat (MARA) and Universiti Sains Malaysia, for providing the financial support under Research Grant RUI(1001/PELECT/8014049) to carry out this research.

References

[1] "What Is Green Power," https://www.epa.gov/greenpower/what-green-power, accessed: 15-July-2020.

[2] "Benefits of Using Green Power," https://www.epa.gov/greenpower/benefits-using-green-power, accessed: 15-July-2020.

[3] "Energy Efficiency," https://www.eesi.org/topics/energy-efficiency/description, accessed: 16-July-2020.

[4] C. Zhu, V. C. Leung, L. Shu, E. C.-H. Ngai, "Green internet of things for smart world," IEEE access, **3**, 2151–2162, 2015, doi:10.1109/ACCESS.2015.2497312.

[5] H. Eakin, P. M. Wightman, D. Hsu, V. R. Gil Ramón, E. Fuentes-Contreras, M. P. Cox, T.-A. N. Hyman, C. Pacas, F. Borraz, C. González-Brambila, et al., "Information and communication technologies and climate change adaptation in Latin America and the Caribbean: a framework for action," Climate and Development, **7**(3), 208–222, 2015, doi:doi.org/10.1080/17565529.2014.951021.

[6] A. P. Upadhyay, A. Bijalwan, "Climate change adaptation: services and role of information communication technology (ICT) in India," American Journal of Environmental Protection, **4**(1), 70–74, 2015, doi:10.11648/j.ajep.20150401.20.

[7] O. Balaban, J. A. P. de Oliveira, "Sustainable buildings for healthier cities: assessing the co-benefits of green buildings in Japan," Journal of Cleaner Production, **163**, S68–S78, 2017, doi:doi.org/10.1016/j.jclepro.2016.01.086.

[8] S. Rani, R. Talwar, J. Malhotra, S. H. Ahmed, M. Sarkar, H. Song, "A novel scheme for an energy efficient Internet of Things based on wireless sensor networks," Sensors, **15**(11), 28603–28626, 2015, doi:doi.org/10.3390/s151128603.

[9] A. H. Sodhro, S. Pirbhulal, Z. Luo, V. H. C. de Albuquerque, "Towards an optimal resource management for IoT based Green and sustainable smart cities," Journal of Cleaner Production, **220**, 1167–1179, 2019, doi:doi.org/10.1016/j.jclepro.2019.01.188.

[10] R. Arshad, S. Zahoor, M. A. Shah, A. Wahid, H. Yu, "Green IoT: An investigation on energy saving practices for 2020 and beyond," IEEE Access, **5**, 15667–15681, 2017, doi:10.1109/ACCESS.2017.2686092.

[11] M. M. K. M. S. Akshaya, M. M. N. Priya, "Green Internet of Things for Smart World," .

[12] F. K. Shaikh, S. Zeadally, E. Exposito, "Enabling technologies for green internet of things," IEEE Systems Journal, **11**(2), 983–994, 2015, doi:10.1109/JSYST.2015.2415194.

[13] C. Zhu, V. C. Leung, L. Shu, E. C.-H. Ngai, "Green internet of things for smart world," IEEE access, **3**, 2151–2162, 2015, doi:10.1109/ACCESS.2015.2497312.

[14] C. S. Nandyala, H.-K. Kim, "Green IoT agriculture and healthcare application (GAHA)," International Journal of Smart Home, **10**(4), 289–300, 2016, doi:dx.doi.org/10.14257/ijsh.2016.10.4.26.

[15] B. Nath, F. Reynolds, R. Want, "RFID technology and applications," IEEE Pervasive computing, **5**(1), 22–24, 2006, doi:10.1109/MPRV.2006.13.

[16] I. Bose, S. Yan, "The green potential of RFID projects: A case-based analysis," It Professional, **13**(1), 41–47, 2011, doi:10.1109/MITP.2011.15.

[17] Y. Karakasa, H. Suwa, T. Ohta, "Evaluating effects of RFID introduction based on CO2 reduction," in Proceedings of the 51st Annual Meeting of the ISSS-2007, Tokyo, Japan, 2007.

[18] A. Ilic, T. Staake, E. Fleisch, "Using sensor information to reduce the carbon footprint of perishable goods," IEEE Pervasive Computing, **8**(1), 22–29, 2008, doi:10.1109/MPRV.2009.20.

[19] S. Murugesan, "Harnessing green IT: Principles and practices," IT professional, **10**(1), 24–33, 2008, doi:10.1109/MITP.2008.10.

[20] F. Al-Turjman, A. Kamal, M. Husain Rehmani, A. Radwan, A.-S. Khan Pathan, "The Green Internet of Things (G-IoT)," 2019, doi:doi.org/10.1155/2019/6059343.

[21] S. Park, "A development of UHF RFID device for mobile IoT service," in 2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 1–7, IEEE, 2016, doi:10.1109/UEMCON.2016.7777830.

[22] R. Colella, L. Catarinucci, L. Tarricone, "Improved RFID tag characterization system: Use case in the IoT arena," in 2016 IEEE International Conference on RFID Technology and Applications (RFID-TA), 172–176, IEEE, 2016, doi:10.1109/RFID-TA.2016.7750760.

[23] R. Colella, L. Catarinucci, L. Tarricone, "Evaluating the suitability of specific RFID tags for IoT applications through a new characterization platform," 2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech), 1–3, 2016, doi:10.1109/SpliTech.2016.7555934.

[24] F. Nekoogar, F. Dowla, "Passive RFID for IOT using UWB/UHF hybrid signaling," 2016 IEEE/ACES International Conference on Wireless Information Technology and Systems (ICWITS) and Applied Computational Electromagnetics (ACES), 1–2, 2016, doi:10.1109/ROPACES.2016.7465304.

[25] "The 3 Biggest Benefits of RFID Asset Tracking in the Supply Chain," https://lowrysolutions.com/blog/the-3-biggest-benefits-of-rfid-asset-tracking-in-the-supply-chain/, accessed: 15-February-2020.

[26] Kubo, "The Research of IoT Based on RFID Technology," 2014 7th International Conference on Intelligent Computation Technology and Automation, 832–835, 2014, doi:10.1109/ICICTA.2014.199.

[27] H. Wang, "Constructing the Green Campus within the Internet of Things Architecture," International Journal of Distributed Sensor Networks, **10**, 2014, doi:10.1155/2014/804627.

[28] "Green Campus at the University of Copenhagen," https://greencampus.ku.dk/greencampus20/, accessed: 15-July-2020.

[29] B. Singh, D. K. Lobiyal, "A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks," Human-centric Computing and Information Sciences, **2**, 1–18, 2011, doi:10.1186/2192-1962-2-13.

[30] B. S. Malapur, V. R. Pattanshetti, "IoT based waste management: An application to smart city," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2476–2486, 2017, doi:10.1109/ICECDS.2017.8389897.

[31] T. Anagnostopoulos, A. Zaslavsky, "Effective Waste Collection with Shortest Path Semi-Static and Dynamic Routing," in NEW2AN, 2014, doi:10.1007/978-3-319-10353-2_9.

[32] P. Jajoo, A. Mishra, S. Mehta, V. Solvande, "Smart Garbage Management System," 2018 International Conference on Smart City and Emerging Technology (ICSCET), 1–6, 2018, doi:10.1109/ICSCET.2018.8537390.

[33] T. Anagnostopoulos, A. Zaslavsky, A. Medvedev, "Robust waste collection exploiting cost efficiency of IoT potentiality in Smart Cities," 2015 International Conference on Recent Advances in Internet of Things (RIoT), 1–6, 2015, doi:10.1109/RIOT.2015.7104901.

[34] S. Jisha, M. Philip, "Rfid based security platform for internet of things in health care environment," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), 1–3, 2016, doi:10.1109/GET.2016.7916693.

[35] D. He, S. Zeadally, "An Analysis of RFID Authentication Schemes for Internet of Things in Healthcare Environment Using Elliptic Curve Cryptography," IEEE Internet of Things Journal, **2**, 72–83, 2015, doi:10.1109/JIOT.2014.2360121.

[36] C. Occhiuzzi, G. Contri, G. Marrocco, "Design of Implanted RFID Tags for Passive Sensing of Human Body: The STENTag," IEEE Transactions on Antennas and Propagation, **60**, 3146–3154, 2012, doi:10.1109/TAP.2012.2198189.

[37] G. Calcagnini, F. Censi, M. Maffia, L. Mainetti, E. Mattei, L. Patrono, E. Urso, "Evaluation of thermal and nonthermal effects of UHF RFID exposure on biological drugs," IEEE Transactions on Information Technology in Biomedicine, **16**(6), 1051–1057, 2012, doi:10.1109/TITB.2012.2204895.

[38] W. Zhou, S. Piramuthu, "Framework, strategy and evaluation of health care processes with RFID," Decision Support Systems, **50**(1), 222–233, 2010, doi:doi.org/10.1016/j.dss.2010.08.003.

[39] P. Schmitt, F. Michahelles, "Economic impact of RFID report," Retrieved November, **1**, 2009, 2008.

[40] Z. Lu, H. Cao, P. Folan, D. Potter, J. Browne, "RFID-based information management in the automotive plastic recycling industry," in Information Technologies in Environmental Engineering, 397–408, Springer, 2007, doi: 10.1007/978-3-540-71335-7_41.

[41] D. Duval, H. Maclean, "The role of product information in automotive plastics recycling: a financial and life cycle assessment," Journal of Cleaner Production, **15**, 1158–1168, 2007, doi:10.1016/J.JCLEPRO.2006.05.030.

[42] T. C. Poon, K. L. Choy, H. K. H. Chow, H. Lau, F. Chan, K. Ho, "A RFID case-based logistics resource management system for managing order-picking operations in warehouses," Expert Syst. Appl., **36**, 8277–8301, 2009, doi: 10.1016/j.eswa.2008.10.011.

[43] N. Wu, M. A. Nystrom, T. R. Lin, H. C. Yu, "Challenges to Global RFID Adoption," 2006 Technology Management for the Global Future - PICMET 2006 Conference, **2**, 618–623, 2006, doi:10.1109/PICMET.2006.296595.

[44] J.-S. Park, "Evaluating Green IT Initiatives Using the Sustainability Balanced Scorecard," Journal of the Korea Safety Management and Science, **19**(3), 81–87, 2017, doi:doi.org/10.12812/ksms.2017.19.3.81.

[45] A. Nasir, U. Kasimu, B. Ige, A. Mohammed, "Characterization and management of solid waste generated in Nasarawa LGA in Nasarawa state, Nigeria," Middle-East J Sci Res, **24**, 1128–1134, 2016, doi:10.5829/idosi.mejsr.2016.24.04.23231.

[46] R. Knoth, B. Kopacek, P. Kopacek, "Case study: multi life cycle center for electronic products," in Proceedings of the 2005 IEEE International Symposium on Electronics and the Environment, 2005., 194–198, IEEE, 2005, doi:10.1109/ISEE.2005.1437022.

[47] M. Favot, R. Veit, A. Massarutto, "The evolution of the Italian EPR system for the management of household Waste Electrical and Electronic Equipment (WEEE). Technical and economic performance in the spotlight," Waste management, **56**, 431–437, 2016, doi:doi.org/10.1016/j.wasman.2016.06.005.

[48] H. R. Culver, J. R. Clegg, N. Peppas, "Analyte-Responsive Hydrogels: Intelligent Materials for Biosensing and Drug Delivery." Accounts of chemical research, **50 2**, 170–178, 2017, doi:10.1021/acs.accounts.6b00533.

[49] N. Deng, "RFID technology and network construction in the internet of things," in 2012 International Conference on Computer Science and Service System, 979–982, IEEE, 2012, doi:10.1109/CSSS.2012.248.

[50] R. H. Weber, "Internet of Things–New security and privacy challenges," Computer law & security review, **26**(1), 23–30, 2010, doi:10.1016/J.CLSR.2009.11.008.

[51] R. H. Weber, "Internet of things–Need for a new legal environment?" Computer law & security review, **25**(6), 522–527, 2009, doi:doi.org/10.1016/j.clsr.2009.09.002.

[52] M. Balaji, S. K. Roy, "Value co-creation with Internet of things technology in the retail industry," Journal of Marketing Management, **33**(1-2), 7–31, 2017, doi:10.1080/0267257X.2016.1217914.

[53] X. Zhu, S. K. Mukhopadhyay, H. Kurata, "A review of RFID technology and its managerial applications in different industries," Journal of Engineering and Technology, **29**, 152–167, 2012, doi:10.1016/J.JENGTECMAN.2011.09.011.

[54] T. Taleb, M. Corici, C. Parada, A. Jamakovic, S. Ruffino, G. Karagiannis, T. Magedanz, "EASE: EPC as a service to ease mobile core network deployment over cloud," IEEE Network, **29**, 78–88, 2015, doi:10.1109/MNET.2015.7064907.

[55] B. Fabian, O. Günther, "Security challenges of the EPCglobal network," Commun. ACM, **52**, 121–125, 2009, doi:10.1145/1538788.1538816.

[56] J. Cooper, A. James, "Challenges for database management in the internet of things," IETE Technical Review, **26**(5), 320–329, 2009, doi: 10.4103/0256-4602.55275.

[57] M. Z. Hasan, F. Al-Turjman, "Optimizing multipath routing with guaranteed fault tolerance in Internet of Things," IEEE Sensors Journal, **17**(19), 6463–6473, 2017, doi:10.1109/JSEN.2017.2739188.

[58] S. F. Khan, "Health care monitoring system in Internet of Things (IoT) by using RFID," in 2017 6th International Conference on Industrial Technology and Management (ICITM), 198–204, IEEE, 2017, doi:10.1109/ICITM.2017.7917920.

[59] S. Amendola, R. Lodato, S. Manzari, C. Occhiuzzi, G. Marrocco, "RFID technology for IoT-based personal healthcare in smart spaces," IEEE Internet of things journal, **1**(2), 144–152, 2014, doi:10.1109/JIOT.2014.2313981.

[60] S. Manzari, G. Marrocco, "Modeling and applications of a chemical-loaded UHF RFID sensing antenna with tuning capability," IEEE transactions on antennas and propagation, **62**(1), 94–101, 2013, doi:10.1109/TAP.2013.2287008.

[61] S. Manzari, A. Catini, G. Pomarico, C. Di Natale, G. Marrocco, "Development of an UHF RFID chemical sensor array for battery-less ambient sensing," IEEE Sensors Journal, **14**(10), 3616–3623, 2014, doi:10.1109/JSEN.2014.2329268.

[62] S. Amendola, L. Bianchi, G. Marrocco, "Combined passive radiofrequency identification and machine learning technique to recognize human motion," in 2014 44th European Microwave Conference, 1044–1047, IEEE, 2014, doi:10.1109/EuMC.2014.6986617.

[63] J. Zhang, G. Y. Tian, A. M. Marindra, A. I. Sunny, A. B. Zhao, "A review of passive RFID tag antenna-based sensors and systems for structural health monitoring applications," Sensors, **17**(2), 265, 2017, doi:doi.org/10.3390/s17020265.

[64] L. K. Fiddes, N. Yan, "RFID tags for wireless electrochemical detection of volatile chemicals," Sensors and Actuators B: Chemical, **186**, 817–823, 2013, doi:doi.org/10.1016/j.snb.2013.05.008.

[65] E. Moradi, K. Koski, L. Ukkonen, Y. Rahmat-Samii, T. Björninen, L. Sydänheimo, "Embroidered RFID tags in body-centric communication," in 2013 International Workshop on Antenna Technology (iWAT), 367–370, IEEE, 2013, doi:10.1109/IWAT.2013.6518367.

[66] R. Krigslund, S. Dosen, P. Popovski, J. L. Dideriksen, G. F. Pedersen, D. Farina, "A novel technology for motion capture using passive UHF RFID tags," IEEE Transactions on Biomedical Engineering, **60**(5), 1453–1457, 2012, doi:10.1109/TBME.2012.2209649.

[67] S. Manzari, C. Occhiuzzi, S. Nawale, A. Catini, C. Di Natale, G. Marrocco, "Humidity sensing by polymer-loaded UHF RFID antennas," IEEE Sensors Journal, **12**(9), 2851–2858, 2012, doi:10.1109/JSEN.2012.2202897.

[68] A. A. Babar, S. Manzari, L. Sydanheimo, A. Z. Elsherbeni, L. Ukkonen, "Passive UHF RFID tag for heat sensing applications," IEEE Transactions on Antennas and propagation, **60**(9), 4056–4064, 2012, doi:10.1109/TAP.2012.2207045.

[69] G. Li, Y. Huang, G. Gao, X. Wei, Z. Tian, L.-A. Bian, "A handbag zipper antenna for the applications of body-centric wireless communications and Internet of Things," IEEE Transactions on Antennas and Propagation, **65**(10), 5137–5146, 2017, doi:10.1109/TAP.2017.2743046.

[70] T. Kellomaki, "On-body performance of a wearable single-layer RFID tag," IEEE Antennas and Wireless Propagation Letters, **11**, 73–76, 2012, doi: 10.1109/LAWP.2012.2183112.

[71] S. Manzari, C. Occhiuzzi, G. Marrocco, "Feasibility of body-centric passive RFID systems by using textile tags," IEEE Antennas Propag. Mag, **54**(4), 49–62, 2012, doi:10.1109/MAP.2012.6309156.

[72] J. Kim, A. Banks, H. Cheng, Z. Xie, S. Xu, K.-I. Jang, J. W. Lee, Z. Liu, P. Gutruf, X. Huang, et al., "Epidermal electronics with advanced capabilities in near-field communication," small, **11**(8), 906–912, 2015, doi: doi.org/10.1002/smll.201402495.

[73] S. Wang, S. Liu, D. Chen, "Analysis and construction of efficient RFID authentication protocol with backward privacy," in China Conference on Wireless Sensor Networks, 458–466, Springer, 2012, doi:10.1007/978-3-642-36252-1_43.

[74] Y.-P. Liao, C.-M. Hsiao, "A secure ECC-based RFID authentication scheme integrated with ID-verifier transfer protocol," Ad hoc networks, **18**, 133–146, 2014, doi:10.1016/j.adhoc.2013.02.004.

[75] J. Chou, "An efficient mutual authentication RFID scheme based on elliptic curve cryptography," The Journal of Supercomputing, **70**, 75–94, 2013, doi:10.1007/s11227-013-1073-x.

[76] C. Jin, C. Xu, X. Zhang, J. Zhao, "A secure RFID mutual authentication protocol for healthcare environments using elliptic curve cryptography," Journal of medical systems, **39**(3), 24, 2015, doi:10.1007/s10916-015-0213-7.

[77] Z. Zhang, Q. Qi, "An efficient RFID authentication protocol to enhance patient medication safety using elliptic curve cryptography," Journal of medical systems, **38**(5), 47, 2014, doi:10.1007/s10916-014-0047-8.

[78] A. K. Singh, B. Patro, "Elliptic Curve Signcryption Based Security Protocol for RFID," Ksii Transactions on Internet and Information Systems, **14**, 344–365, 2020, doi:10.3837/tiis.2020.01.019.

[79] C. Lai, H. Li, X. Liang, R. Lu, K. Zhang, X. Shen, "CPAL: A conditional privacy-preserving authentication with access linkability for roaming service," IEEE Internet of Things Journal, **1**(1), 46–57, 2014, doi:10.1109/JIOT.2014.2306673.

[80] Z. Ali, A. Ghani, I. Khan, S. A. Chaudhry, S. H. Islam, D. Giri, "A robust authentication and access control protocol for securing wireless healthcare sensor networks," J. Inf. Secur. Appl., **52**, 102502, 2020, doi:10.1016/j.jisa.2020.102502.

[81] Y.-l. Liu, X.-l. Qin, C. Wang, B.-h. Li, "A lightweight RFID authentication protocol based on elliptic curve cryptography," Journal of computers, **8**(11), 2880–2887, 2013, doi:10.4304/jcp.8.11.2880-2887.

[82] Y.-P. Liao, C.-M. Hsiao, "A secure ECC-based RFID authentication scheme using hybrid protocols," in Advances in Intelligent Systems and Applications-Volume 2, 1–13, Springer, 2013, doi:10.1007/978-3-642-35473-1_1.

[83] J. Bongaarts, "Global fertility and population trends," in Seminars in reproductive medicine, volume 33, 005–010, Thieme Medical Publishers, 2015, doi:10.1055/s-0034-1395272.

[84] L. Batina, S. Seys, D. Singelée, I. Verbauwhede, "Hierarchical ECC-based RFID authentication protocol," in International Workshop on Radio Frequency Identification: Security and Privacy Issues, 183–201, Springer, 2011, doi:10.1007/978-3-642-25286-0_12.

[85] C. Lv, H. Li, J. Ma, Y. Zhang, "Vulnerability analysis of elliptic curve cryptography-based RFID authentication protocols," Transactions on Emerging Telecommunications Technologies, **23**(7), 618–624, 2012, doi:10.1002/ett.2514.

[86] B. Chander, K. Gopalakrishnan, "Security Vulnerabilities and Issues of Traditional Wireless Sensors Networks in IoT," 2020, doi:10.1007/978-3-030-33596-0_21.

[87] K. Fan, Y. Gong, Z. Du, H. Li, Y. Yang, "RFID Secure Application Revocation for IoT in 5G," 2015 IEEE Trustcom/BigDataSE/ISPA, **1**, 175–181, 2015, doi:10.1109/Trustcom.2015.372.

[88] Y. L. Santos, E. Canedo, "On the Design and Implementation of an IoT based Architecture for Reading Ultra High Frequency Tags," Inf., **10**, 41, 2019, doi:10.3390/info10020041.

[89] "Top 10 RFID Security Concerns and Threats," https://securitywing.com/top-10-rfid-security-concerns-threats/, accessed: 10-September-2020.

[90] M. El-hajj, A. Fadlallah, M. Chamoun, A. Serhrouchni, "A Survey of Internet of Things (IoT) Authentication Schemes †," Sensors (Basel, Switzerland), **19**, 2019, doi:10.3390/s19051141.

[91] G. Saxl, M. Ferdik, M. Fischer, M. Maderboeck, T. Ussmueller, "UHF RFID Prototyping Platform for ISO 29167 Decryption Based on an SDR," Sensors (Basel, Switzerland), **19**, 2019, doi:10.3390/s19102220.

[92] L. Xiao, H. Xu, F. Zhu, R. Wang, P. Li, "SKINNY-Based RFID Lightweight Authentication Protocol," Sensors (Basel, Switzerland), **20**, 2020, doi:10.3390/s20051366.

[93] K. Fan, C. Liang, H. Li, Y. Yang, "LRMAPC: A Lightweight RFID Mutual Authentication Protocol with Cache in the Reader for IoT," 2014 IEEE International Conference on Computer and Information Technology, 276–280, 2014, doi:10.1109/CIT.2014.80.

[94] O. Hongzhi, G. Jin-yun, L. Ke-sheng, "An UHF RFID reader in IOT," Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), 672–675, 2013, doi:10.1109/MEC.2013.6885148.

[95] C. P. Schuch, G. Balbinot, M. Jeffers, M. W. McDonald, D. Corbett, "An RFID-based activity tracking system to monitor individual rodent behavior in environmental enrichment: Implications for post-stroke cognitive recovery," Journal of Neuroscience Methods, **324**, 2019, doi:10.1016/j.jneumeth.2019.05.015.

[96] "DHL Says Its SmartTrucks Save Money, Time and CO2," Accessed: 15-September-2020.

[97] "Communities Turn to RFID to Boost Recycling," http://www.rfidjournal.com/article/articleview/4259/1/1, accessed: 03-June-2020.

[98] "Nestlé italy finds rfid brings roi for ice cream," Accessed: 03-June-2020.

[99] X. Chen, X. Wang, M. Zhou, "Firms' green R&D cooperation behaviour in a supply chain: Technological spillover, power and coordination," International Journal of Production Economics, 2019, doi:10.1016/j.ijpe.2019.04.033.

[100] X. Zhu, Y. Zhang, Z. Zhao, J. Zuo, "Radio frequency sensing based environmental monitoring technology," in International Workshop on Pattern Recognition, 2019, doi:10.1117/12.2541098.

[101] T. Litman, London congestion pricing: Implications for other cities, Victoria Transport Policy Institute, 2012.

[102] L. Ecola, T. Light, "Making congestion pricing equitable," Transportation Research Record, **2187**(1), 53–59, 2010, doi:doi.org/10.3141/2187-08.

[103] V. Talla, M. Hessar, B. Kellogg, A. Najafi, J. R. Smith, S. Gollakota, "LoRa Backscatter: Enabling The Vision of Ubiquitous Connectivity," arXiv: Networking and Internet Architecture, 2017, doi:doi.org/10.1145/3130970.

[104] S. N. R. Kantareddy, I. Peters, I. Mathews, S. Sun, M. Layurova, J. Thapa, J.-P. Correa-Baena, R. Bhattacharyya, T. Buonassisi, S. Sarma, "Perovskite PV-Powered RFID: Enabling Low-Cost Self-Powered IoT Sensors," IEEE Sensors Journal, **20**, 471–478, 2020, doi:10.1109/JSEN.2019.2939293.

[105] S. N. R. Kantareddy, I. Mathews, R. Bhattacharyya, I. Peters, T. Buonassisi, S. Sarma, "Long Range Battery-Less PV-Powered RFID Tag Sensors," IEEE Internet of Things Journal, **6**, 6989–6996, 2019, doi:10.1109/JIOT.2019.2913403.

[106] Y. Jabareen, "An assessment framework for cities coping with climate change: The case of New York City and its PlaNYC 2030," Sustainability, **6**(9), 5898–5919, 2014, doi:doi.org/10.3390/su6095898.

[107] V. M. Thomas, "Environmental implications of RFID," in 2008 IEEE International Symposium on Electronics and the Environment, 1–5, IEEE, 2008, doi:10.1109/ISEE.2008.4562916.

[108] J. Su, Z. Sheng, V. C. M. Leung, Y. Chen, "Energy Efficient Tag Identification Algorithms For RFID: Survey, Motivation And New Design," IEEE Wireless Communications, **26**, 118–124, 2019, doi:10.1109/MWC.2019.1800249.

[109] R. Gradus, G. C. Homsy, L. Liao, M. E. Warner, "Which US municipalities adopt Pay-As-You-Throw and curbside recycling?" Resources, Conservation and Recycling, **143**, 178–183, 2019, doi:doi.org/10.1016/j.resconrec.2018.12.012.

[110] N. E. Weckman, N. Ermann, R. Gutierrez, K. Chen, J. Graham, R. Tivony, A. Heron, U. F. Keyser, "Multiplexed DNA identification using site specific dCas9 barcodes and nanopore sensing," ACS sensors, **4**(8), 2065–2072, 2019, doi:doi.org/10.1021/acssensors.9b00686.

[111] V. M. Thomas, "A universal code for environmental management of products," Resources, Conservation and Recycling, **53**(7), 400–408, 2009, doi: doi.org/10.1016/j.resconrec.2009.03.004.

[112] S. Saar, M. Stutz, V. M. Thomas, "Towards intelligent recycling: a proposal to link bar codes to recycling information," Resources, Conservation and Recycling, **41**(1), 15–22, 2004, doi:doi.org/10.1016/j.resconrec.2003.08.006.

[113] A. Condemi, F. Cucchiella, D. Schettini, "Circular Economy and E-Waste: An Opportunity from RFID TAGs," Applied Sciences, **9**(16), 3422, 2019, doi:doi.org/10.3390/app9163422.

[114] P. Krauchi, P. Wager, M. Eugster, G. Grossmann, L. Hilty, "End-of-life impacts of pervasive computing," IEEE Technology and Society Magazine, **24**(1), 45–53, 2005, doi:10.1109/MTAS.2005.1407747.

[115] M. Brink, F. Berkemeyer, J. Ohlendorf, G. Dumstorff, K. Thoben, W. Lang, "Challenges and Opportunities of RFID Sensortags Integration by Fibre-Reinforced Plastic Components Production," Procedia Manufacturing, **24**, 54–59, 2018, doi:10.1016/J.PROMFG.2018.06.008.

[116] K. Kunz, N. Lui, "Process Chemistry and Copper Catalysis," Copper-Mediated Cross-Coupling Reactions, 725–743, 2013, doi:10.1002/9781118690659.

[117] L. Guest, W. Guibène, S. Guruacharya, "Wireless Communications and Mobile Computing Green Internet of Things (IoT) : Enabling Technologies , Architectures , Performance , and Design Issues," 2018, doi: https://doi.org/10.1155/2018/3747562.

[118] Y. Duroc, D. Kaddour, "RFID potential impacts and future evolution for green projects," Energy Procedia, **18**, 91–98, 2012, doi:doi.org/10.1016/j.egypro.2012.05.021.

[119] S. H. Alsamhi, O. Ma, M. Ansari, Q. Meng, "Greening Internet of Things for Smart Everythings with A Green-Environment Life: A Survey and Future Prospects," ArXiv, **abs/1805.00844**, 2018.

[120] G. Orecchini, L. Yang, A. Rida, F. Alimenti, M. Tentzeris, L. Roselli, "Green Technologies and RFID : Present and Future," 2010.

[121] A. Ahmed, "Benefits and Challenges of Internet of Things for Telecommunication Networks," 2019, doi:10.5772/INTECHOPEN.81891.

[122] K. A. Alezabi, F. Hashim, S. J. Hashim, B. M. Ali, A. Jamalipour, "Efficient authentication and re-authentication protocols for 4G/5G heterogeneous networks," EURASIP Journal on Wireless Communications and Networking, **2020**, 1–34, 2020, doi:10.1186/s13638-020-01702-8.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 27-34 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Electroencephalogram Based Medical Biometrics using Machine Learning: Assessment of Different Color Stimuli

Md Mahmudul Hasan[1,*], Nafiul Hasan[2], Dil Afroz[3], Ferdaus Anam Jibon[4], Md. Arman Hossen[2], Md. Shahrier Parvage[2], Jakaria Sulaiman Aongkon[5]

[1]*Centre for Accident Research and Road Safety- Queensland, Queensland University of Technology, Brisbane, QLD 4059, Australia*

[2]*Department of Electrical and Electronic Engineering, Bangladesh Army University of Engineering and Technology, Qadirabad, 6431, Bangladesh*

[3]*Department of Computer Science and Engineering, Bangladesh Army University of Engineering and Technology, Qadirabad, 6431, Bangladesh*

[4]*Department of Computer Science and Engineering, Sheikh Fazilatunnesa Mujib University, Jamalpur, 2000, Bangladesh*

[5]*Department of Electrical and Electronic Engineering, Shahjalal University of Science and Technology, Sylhet, 3100, Bangladesh*

ARTICLE INFO

Article history:
Received: 17 December, 2020
Accepted: 29 March, 2021
Online: 05 May, 2021

Keywords:
Electroencephalogram
Biometrics
Machine learning
Color stimuli
Performance measure

ABSTRACT

A methodology of medical signal-based biometrics has been proposed in this paper for implementing a human identification system controlled by electroencephalogram in respect of different color stimuli. The advantage of biosignal based biometrics is that they provide more efficient operation in simple experimental condition to ensure accurate identification. Red, Green, Blue (primary colors) and Yellow (secondary color) were chosen as the color stimuli for making more comfortable EEG regenerating environment. Four supervised classification models, namely, Logistic Regression (LR), K- Nearest Neighbor (KNN), Support Vector Machine (SVM) and Random Forest Classifier (RFC) were trained and tested for assessing the performance of the EEG based biometric identification, with five-fold cross-validation. Four different measures (sensitivity, specificity, accuracy and area under the receiver operating characteristic curve) were used to evaluate the overall performance. The results suggested that Blue color stimuli perform the best among all the color stimulus with an accuracy ranging from (77.2-88.9%). The classifiers identify each of the subjects with any color having an accuracy ranged from (70.9-88.9%), and the RFC shows the best accuracy which is 88.9% in the case of blue color stimuli.

1. Introduction

This paper is an extension of work originally presented in 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 [1]. The presented paper utilized electroencephalogram (EEG) for medical biometrics using color stimuli using only one classifier (artificial neural networks) where the current article is expanded further to validate the EEG based biometrics using multiple machine learning models. Also, this paper examines the utility of the different color stimulus on the EEG based human identification system.

Biometrics refers to the process of identifying and authenticating a person based on a unique identifier. By utilizing a person's unique feature, the human identification system is built for different security applications. Existing biometric systems use knowledge or possession-based features like passcode, PIN, fingerprint, voice which are extensively being used for device security and other security purposes [2]. There are some limitations in the existing authentication technologies, for example, the fingerprints can be replicated, and facial detection can be fooled [3]. The Boston Marathon bombing incident has shown the failure of the so-called advanced facial identification system [3]. Furthermore, with the development in hacking

*Corresponding Author: Md Mahmudul Hasan, 2 Rochester Terrace, Brisbane, QLD 4059, Australia; Email: mahmudul.hasan.eee.kuet@gmail.com

techniques such as replication of fingerprints has resulted in reduced reliability of the conventional biometric identification systems [3, 4].

In the given circumstances, unique electroencephalogram remains the only choice due to its non-biasing nature [3-6]. For this reason, physiological signal based biometrics such as EEG uses neurons activity is becoming research of interest due to its person to person variability characteristics [7]. To utilize the medical signals for biometric identification, a study was performed [3], where the authors have tried to develop an EEG-based approach in order to make an efficient human identification. In [3], the authors found the beta band as the most influential rhythm of EEG for human identification.

However, generating similar EEG signal frequently is tough and can't be declared as a standard method. The main issue with the EEG is that it is variable and needs a specific environment to reproduce the similar patterned brainwave [3, 4]. To simplify this experimental condition, the use of color stimuli is the right choice. Every color has a unique effect on the brainwaves, which was observed by statistical analysis of the signal as per the previous studies [8]. In [4], the authors explained that frequency domain shows better performance than time domain, and the value of power spectrum density varies a lot among individuals while varies a little within an individual. It is also observed that the use of both time and frequency domain feature comes out with the best performance for human identification system [4]. A study developed an EEG based architecture for identifying the individuals based on the brainwaves using color stimulus based experiment, which was performed on three participants in a laboratory-based work [9]. The authors used three fundamental colors (Red, Green and Blue) [10] and one secondary color Yellow color stimulus for human identification purpose. The results suggested that the blue color is the most sensitive to the human identification, whereas the secondary color Yellow gave the worst performance in identification. The study showed promising results but utilized only one classifier (ANN) for the classification approach and only one performance measure (mean square error) for the assessment of the system [1]. However, considering the sensitivity and specificity metrics are most important for an EEG based detection system. As higher sensitivity with lower specificity leads to the higher false alarm, and the opposite trend causes the missing of a lot of positive states, a compromise between the two metrics is crucial. Though there are very limited works on the field of biometric identification using brainwaves, these systems are worth for the IoT devices and cyber security with the application of AI.

In this study, the data was recorded using the BIOPAC® data acquisition unit, the pre-processing and feature extraction was done using the Acqknowledge 4.1® software [11]. Most importantly, four supervised classification models, namely, Logistic Regression, K- Nearest Neighbor (KNN), Support Vector Machine (SVM) and Random Forest Classifier (RFC) were trained and tested for evaluating the performance of each of the EEG rhythm, with five-fold cross-validation. Moreover, four different performance measures (sensitivity, specificity, accuracy and area under the receiver operating characteristic curve-AUROC) were utilized to examine the performance of the human identification system.

The following part of this paper is organized as follows- a brief methodology, including experimental design and tools, then the result section with the findings. Last, the paper was concluded with as short summary, followed by a discussion on the outcomes, research implications and future works.

2. Methodology

2.1 Experimental Design

The experiment was conducted in the laboratory-based environment. The different steps for developing the electroencephalogram based human identification system by color stimuli is shown in Figure 1. With the given experimental conditions, EEG were obtained by the BIOPAC® system from the selected participants. The next step is removing the noise and artifact due to eye blinking and body movement. Then FIR band pass filter was utilized to separate bands in Acqknowledge 4.1 software. Afterwards, eight features were extracted for each band, and the selected features were supplied towards the machine learning tools. Four different supervised learning techniques, namely K-nearest neighbors (KNN), support vector machines (SVM), logistic regression (LR) and random forest classifier (RFC) were developed in python 3.6.9 platform, and models were applied for human identification. The best classifier was evaluated by comparing their performance metrics.

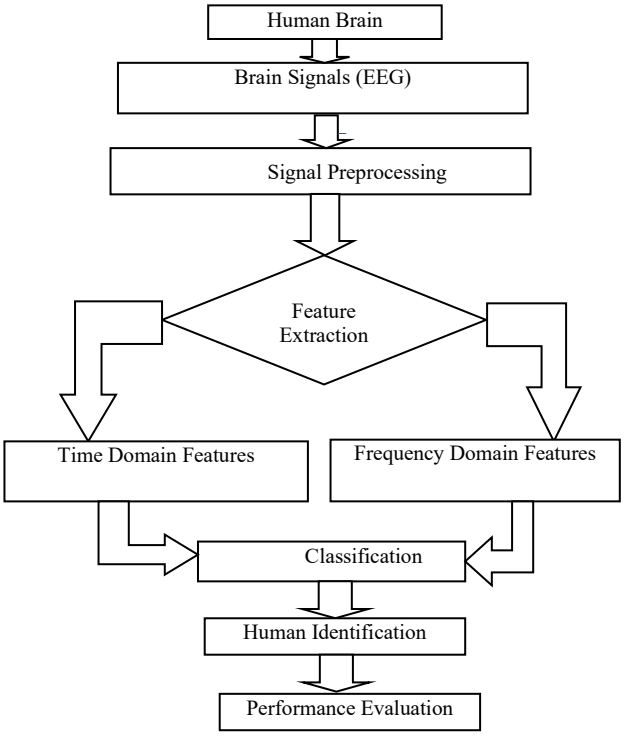


Figure 1: Block diagram of the proposed EEG based human identification system

2.2. Experimental Equipment

2.2.1. Hardware tool

For the signal acquisition, a BIOPAC® MP 36 system [11] was used at the Biomedical Engineering Lab, Khulna University of Engineering and Technology (KUET), Bangladesh. This is a wired data acquisition system which has a signal acquisition and a processing unit to interface with the computer.

2.2.2. Pre-processing software tool

As a software tool, BIOPAC student Lab Pro® was used to record the physiological signals. Acqknowledge 4.1® software was used for the feature extraction purpose [11,12]. Machine learning based classification models were developed and applied in the python 3.6.9 version in Google Colab platform, which is research based online environment affiliated by Google .

2.3. Participants

For participation in this experiment, recruitment was done with online advertisement on Facebook, Twitter and LinkedIn. In total, three subjects participated in the experiment who were male, healthy and not suffering from any color blindness or psychological illness. The color blindness was tested using the Ishihara 38 Plates CVD Test [13] upfront after they arrive in the laboratory to check their vision and to ensure they are not suffering from difficulties in choosing colors, especially the deuteranopic vision (red-green color blindness). Then, color stimulus were shown in a computer monitor (21.5" with a 1920 ×1080 resolution) while they were instructed to focus on indefinite color for 15 minutes long, with their normal blinking. In total, twenty trials were taken for each color (red, green, blue and yellow). The electrodes were placed on the right central (C4), and the right occipital (O2) position.



Figure 2: Experimental environment in BME, KUET

2.4. Experimental Procedure

2.4.1. Signal Preprocessing

The recorded signal contains artefacts due to muscle movement, eye blinking, hand movement and the background effect behind the color stimulus, the primarily obtained EEG signals could have contained noise. Additionally, the line frequency was 50 Hz, which also adds noise to the data. To pre-process raw EEG was gone through band-pass finite impulse response (FIR) filter with a range of 0.5 to 44 Hz, as it removes the non-linear trends of the signals. Later, the signals was further smoothed, taking a moving average over a short period of the signal. The pre-processing makes the signal viable for extracting different time and frequency domain features.

2.4.2. Feature Extraction

Feature extraction is one of the major steps of biosignal processing and analysis, which contains valuable information from the time-series signal. Several features were extracted in time and frequency domain, including maximum value (Emax), standard deviation (STDDEV), skew (sk), kurtosis (k), power spectrum density (PSD) mean, PSD max, Fast Fourier Transform (FFT) mean, FFT max (total eight feature) were extracted for different subjects using the Acknowledge 4.1® software. The feature were tabulated in the excel sheet for the future stages.

2.4.3. Feature Scaling

Machine learning models work on different strategies, and thus the range of the feature values is an essential factor. The features extracted from time and frequency domains have a different range in their magnitude. Since different machine learning models works with different features putting them in a same matrix, it is necessary to put all the features in a same range, which is referred to as feature scaling. Two common types of feature scaling is done in preliminary data: standardization and normalization [14]. As a part of the normalization process, MinMaxScaling was performed in this study in python platform. Using the MinMaxscaler() function from sklearn library. Here the data is shrunk within a range between [-1,1].

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

If x_{new} is symbolized as the the normalized value of a feature point x, within a range x_{min} and x_{max}, then the normalization formula can be given by the equation (1) [14].

2.5. Classification

The ultimate goal of the study is to identify the individual by means of the EEG features, applying the machine learning techniques. Machine learning is a hot topic nowadays, which is referred to as the application of artificial intelligence, which provides a system capable of learning nature of a given dataset. Basically, there are three categories of Machine Learning models and application, supervised learning, unsupervised learning and reinforcement learning. Supervised learning is extensively used for the classification and regression problem [15]. Previous studies worked with EEG have used supervised learnings, especially KNN [16], SVM [17], RFC [18] and LR [19]. Based on the previous studies, these four classifiers were chosen for the data classification in this research. Moreover, these models works on different algorithms of learning from the given data points, which might be worthful to see which one perform the best for the human identification purpose.

2.5.1. Logistic Regression

The simplest way to classify data points was linear regression, and the disadvantages of the linear regression models were overcome by the logistic regression models. Logistic regression is a supervised learning model, which works based on the linear method, and the predictions are made using a logistic or sigmoid function σ(t). The sigmoid function is an 'S' patterned curve that takes a real number and maps within a range between 0 and 1, which is given by equation (2).

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}} \quad (2)$$

Contemplating the two types of variables, dependent and independent, Logistic regression predict dependant variable basing on the independent variable. The 'C parameter was tuned here in the Logistic Regression model to lessen the overfitting problem, which could produce over-optimistic results [20].

2.5.2. K-nearest neighbours (KNN)

One of the simplest supervised learning models is KNN, which is a non-parametric method where k nearest training examples in the feature space is taken as input and neighbours vote do the classification generally used for classification and regression. At the very starting point, KNN read the value of K, type of distance D and test data; then it finds the K nearest neighbours D to the test data and thus sets the maximum label class of K to test data. The same processes are gone through an iterative process named looping. In details, its algorithm initializes the value of K from 1 (setting as initial iteration value). After loading data, iteration from initial K =1 (generally) to the total number of training data point. Then, distances specifically Euclidean distance between test data and each row of training data is measured and sorted in ascending order to get topmost K rows from the sorted array and the most frequent class is returned as the predicted class [21]. The value of K was tuned, and the K for best efficiency was chosen in the classifier model in this research to reduce overfitting. With the chosen 'K' value the model was further developed, trained and tested with the given data.

2.5.3. Support Vector Machines (SVM)

One of the most popular supervised learning approaches, SVM aims to obtain a hyperplane which classifies the data point (data points can be at any side of hyperplane) in feature dimensional space while depending on both linear and non-linear regression. Data points distance across to hyperplane are called support vector whose detection can exchange hyper plane's location [21]. The model used a Gaussian kernel for SVM classifier in this research due to the non-linear trend of the dataset. Two parameters- 'C' and 'gamma' was adjusted within a set of values using the grid search algorithm to reduce the overfitting problem, which could cause a non-generalized model.

2.5.4. Random Forest Classifier (RFC)

In addition to the binary logic or decision tree-based classifiers, ensemble-based models are now getting popularity due to their robustness. Random forests are made of individual decision trees with a logic of group of weak learners to finally make a strong learner while the decision trees operate as divided or conquer. A class is predicted from every decision tree and a final class is predicted by model depending on their vote [21]. Two parameters were tuned in the RFC models, namely, 'n_estimate', which implies the number of trees in the forest and 'max-depth' which signifies the depth of each tree. With the tuned parameters, the model was further developed, trained and tested to find out the performance measures.

2.6. Performance Measures

2.6.1. Sensitivity or True Positive Rate (TPR)

True positive rate or Sensitivity is the proportion of the true positives (desired factor), which is correctly identified from the given test set [22]. The definition of sensitivity can be provided by equation (3), where, TP = True Positive and FN = False Negative. In this study, sensitivity is the measure of the proportion of successfully identifying a specific person.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+ FN}} \quad (3)$$

2.6.2. Specificity or True Negative Rate (TNR)

True negative rate or Specificity is the proportion of true negative (undesired factor) in which was correctly excluded from the given test sets [22]. The definition of specificity can be provided by equation (4), where, TN = True Negative and FP = False Positive. In the case of this study, specificity is the measure of not correctly identifying a specific person.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+ FP}} \quad (4)$$

2.6.3. Accuracy

The accuracy is the proportion of true results, in an experiment, being either true positive or true negative [22]. The definition of accuracy can be provided by equation (5), given that TP = True positive, TN= True Negative, FP= False Positive and FN = False Negative. In this study, accuracy is the proportion of the successful identification, either a specific person or not being that person.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (5)$$

2.6.4. Area under the receiver operating characteristic (ROC) curve (AUC)

A system having a higher discrepancy between sensitivity can cause false alarm or missing positive states (in this case, identification of a specific individual). Therefore, it is essential to find out the best compromise between them. As a part of this step, ROC is performed, which is a plot of the sensitivity (true positive rate) against the (1- specificity) or false positive rate. Here all the possible combination of TPR and FPR are plotted, showing the trade-off between them [23].

Validation of the performance of the models is an important step towards evaluation of a model. Five-fold cross-validation was done in this study while evaluating the performance measures. The mean value and the standard deviation (SD) were noted, considering the five experimental validations. As the classification is a four-class problem, one vs. all method was used in all the classification approach, splitting the four-class problem in binary class. Thus, the mean sensitivity, specificity and AUC was calculated from the obtained confusion matrix, which was used for further analysis.

3. Results

3.1. Data visualization

Visualizing the data is one of the main steps to understand the data points, and thus it helps to take the further decisions in the machine learning approach. The data points found from the selected features were plotted in box and violin plots to observe the range of each of the features. The following Figure 3 shows that the time and frequency domain features are having a versatile variation in the range. Range of the difference features varies

among themselves either in the time domain or in the frequency domain. Thus, feature scaling was done on the given dataset, and the scale features were supplied to the learning models.

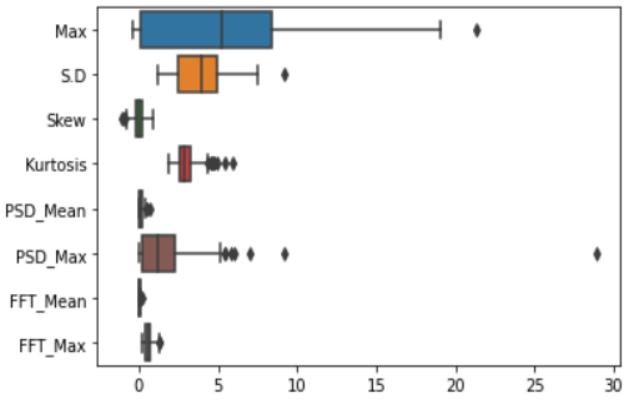


Figure 3: Boxplot of part of the EEG data showing the varying magnitude of different features

3.2. Classification Performance

Finalizing the feature scaling, the scaled features were supplied towards the machine learning models after necessary parameter tuning. Four different performance measures were evaluated, namely, sensitivity or true positive rate (TPR), specificity or True negative rate (TNR), accuracy and area under the receiver operating characteristic (ROC) curve (AUC). The obtained results are listed below in Table 1, which visualizes the performance metrics with respect to the classification models.

Table 1: Performance measures (mean value) for EEG biometrics using four different classifiers, using a five-fold cross-validation

		LR	KNN	SVM	RFC
Red Stimuli	Sensitivity	75.4	70.6	77.1	74.8
	Specificity	72.5	70.7	76.6	81
	Accuracy	79.1	74.1	78.3	79.8
	AUC	74.1	80.3	69	83.7
Blue Stimuli	Sensitivity	89.8	90.9	84.1	93.3
	Specificity	89.8	80.8	81.9	90.6
	Accuracy	**82.8**	**81.7**	**77.2**	**88.9**
	AUC	75.6	84.1	77.3	90
Green Stimuli	Sensitivity	76.6	72.9	68.6	83.3
	Specificity	75.8	77.4	71.8	76.3
	Accuracy	75	70.9	81	83.6
	AUC	78.6	80.6	70.4	83.8
Yellow Stimuli	Sensitivity	76.4	76.11	73.3	84.1
	Specificity	81.6	75.63	73.5	84.4
	Accuracy	75.5	74.69	76.5	86
	AUC	76.9	72.83	79.5	84

3.2.1. Scenario-1: HID using Red Color stimuli

While using the red color stimuli for human identification, the performance measures (mean ± SD) obtained from the human identification from four different classifiers, namely logistic regression (LR), K-nearest Neighbours (KNN), Support Vector Machine (SVM) and Random Forest Classifier (RFC) are shown in the Figure 4. The plots show that the gap between sensitivity and specificity is highest in RFC (6.1%) and lowest in the case of KNN (0.1%). RFC shows the highest gap between sensitivity and specificity (6.1%). Overall, considering the accuracy and ROC, RFC gives the best performance with an accuracy of 79.8%.

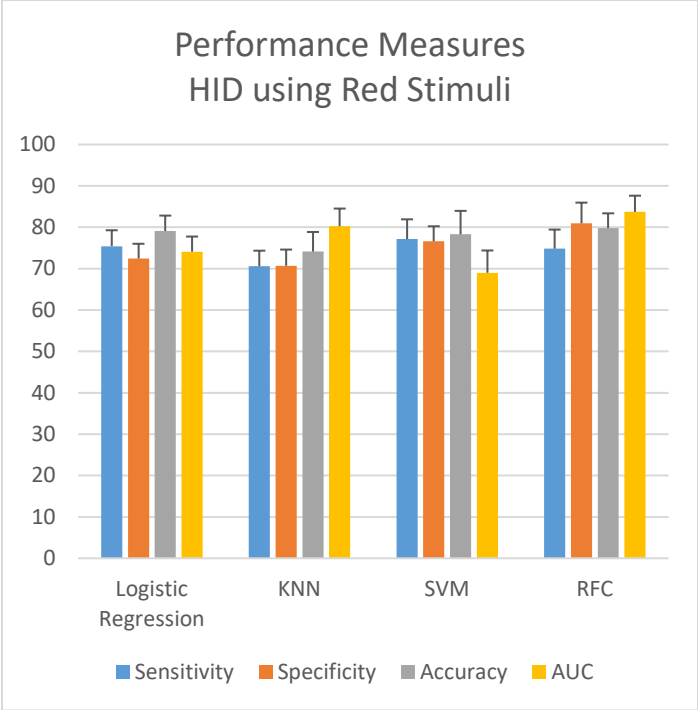


Figure 4: Performance measurement of red stimuli

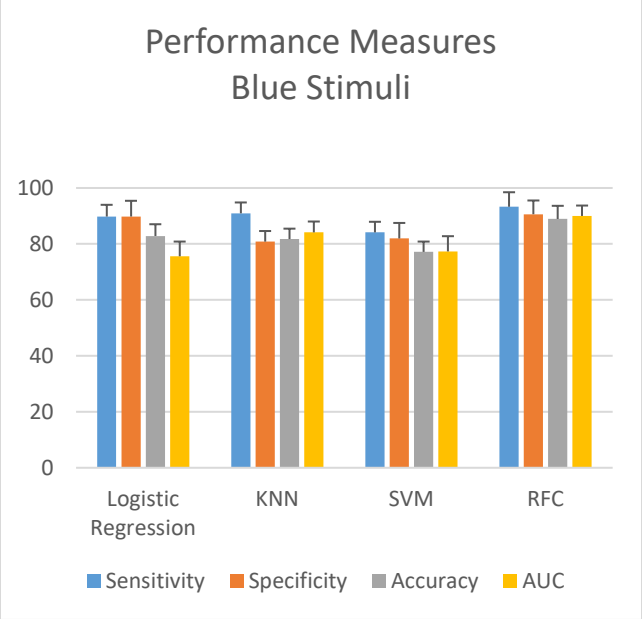


Figure 5: Performance measurement of blue stimuli

3.2.2. Scenario-2: HID using Blue Color stimuli

While using the blue color stimuli for human identification, the performance measures (mean ± SD) obtained from the human identification from four different classifiers are shown in the Figure 5. Here, KNN shows the highest gap (10.1%) between sensitivity and specificity, and LR shows a zero gap between the two metrics. Overall, RFC gives an accuracy of 88.9%, which performs the best.

3.2.3. Scenario-3: HID using Green Color stimuli

While using the green color stimuli for human identification, the performance measures (mean ± SD) obtained from the human identification from four different classifiers are shown in the Figure 6. The plots show that the gap between sensitivity and specificity is highest in RFC (7.03%) and lowest in the case of LR (0.85%). Overall, RFC gives an accuracy of 83.6%, which performs the best.

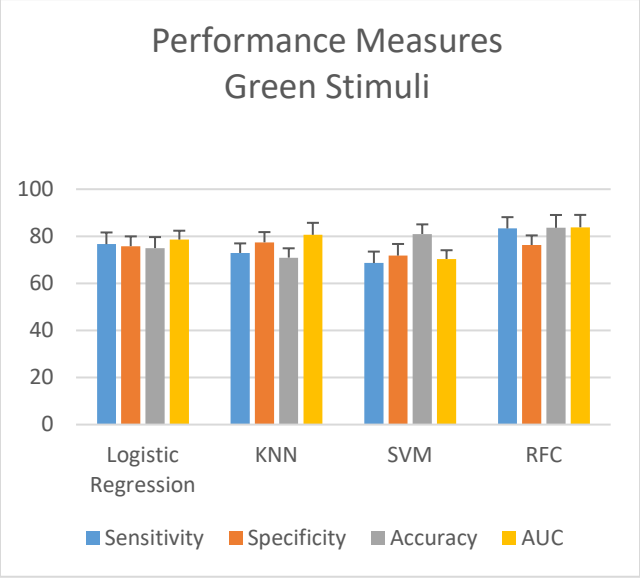


Figure 6: Performance measurement of green stimuli

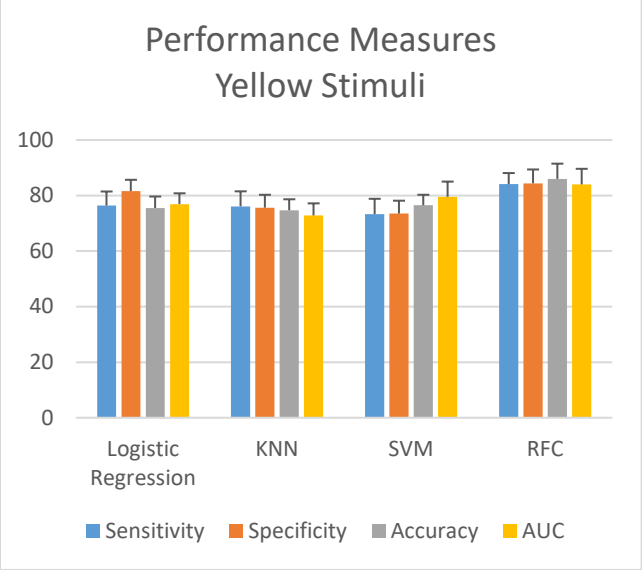


Figure 7: Performance measurement of yellow stimuli

3.2.4. Scenario-4: HID using Yellow Color stimuli

While using the yellow color stimuli for human identification, the performance measures (mean ± SD) obtained from the human identification from four different classifiers are shown in the Figure 7. The plots show that the gap between sensitivity and specificity is highest in LR (5.2%) and lowest in the case of SVM (0.2%). LR shows the highest gap between sensitivity and specificity. Overall, RFC gives an accuracy of 86%, which performs the best.

3.3. Choosing the best scenario/best performance in subject identification

In order to find out the best color stimuli for human identification, the accuracy and AUC measures was selected as two reference metrics as it is difficult to compare different classifiers using several factors. The plots of the accuracy for four different classifiers corresponding to the four color are shown in Figure 8 below. From the figure, it is evident that the accuracy for blue stimuli is better than any other colors for all the four classifiers. Overall, it is evident that all the classifiers identify subjects by using blue color more accurately. So, in the rest of the papers, the performance for the blue color will be considered.

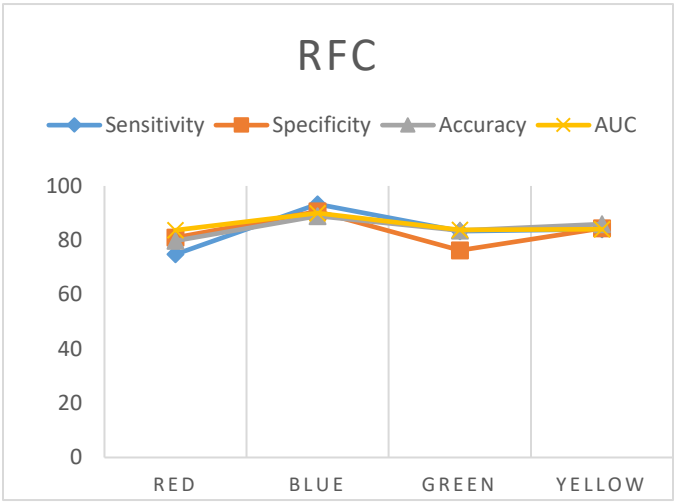


Figure 8: performance measurement of RFC

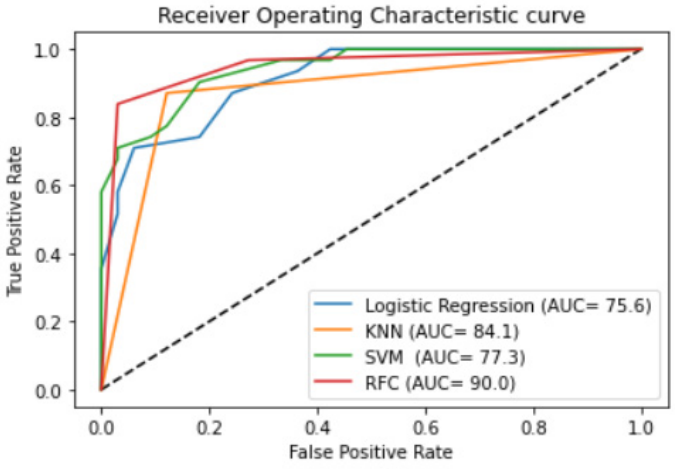


Figure 9: Comparison of AUC for four different classifiers for blue color based HID

3.4. Choosing the best Classifier

Considering the blue color for human identification, the plots for the area under the ROC curve for the classifiers are shown in the Figure 9. The figure illustrates that the RFC classifier show the best compromise between sensitivity and specificity, with covering the highest area under the ROC curve (AUC= 0.90%). Thus, the next part the paper will compare the performance of the color stmuli considering RFC. Overall, all the plots show that RFC performs the best on EEG based human identification for blue color stimuli.

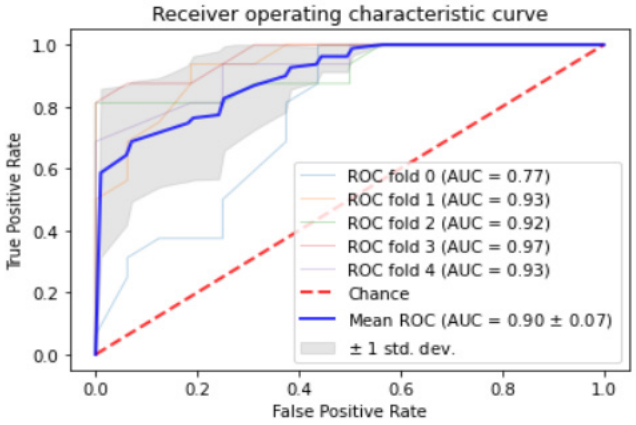


Figure 10: ROC Plots for Blue Stimuli based RFC model with 5 fold CV

As the RFC performs the best, the Area under the AUC curve plots for this model with blue color stimuli in 5 different experiments is given in Figure 10 for five-fold cross-validation. The AUC for the blue color stimuli ranged from (0.77-0.93), with a mean of 0.90 and 0.07 standard deviation. This signifies that, the blue color shows an excellent performance than the other colors while using RFC classifier in order to make EEG based medical biometric system.

4. Discussion

Four different color stimulus were used in this study for assessing the performance of the EEG rhythms for medical biometrics. The results revealed that Blue stimuli perform the best among the other colors. It also revealed that the maximum performance was obtained using the RFC Classifier, with a sensitivity, specificity and accuracy of 93.3%, 90.6% and 88.9%. Moreover, RFC based model with blue stimuli based dataset shows promising AUC (0.90), which is a good compromise between sensitivity and specificity. The finding of this study is consistent with the previous study [1], where the authors found the Blue stimuli as the best performing rhythm, though they have used only one classifier (ANN) and one performance metrics (mean square error). The possible reason behind the best performance of ANN in that study could be the backpropagatipon algorithm, which is strong enough to learn the inherent features and complex structure of the data. Nonetheless, the random forest algorithm works on the majority voting of the multiple decision trees, thus it provides very precise performance, and it is less prone to overfitting. Thus, achieving the similar outcome validates the use of the blue color stimuli for medical biometrics. On the other hand, while using the random forest classifier, the red stimuli perform the worst (RFC accuracy= 79.8%).

However, the scope of the paper is not out of limitations. The background effect is one of the main challenges while using the color stimulus. Inter-individual difference among participants is another factor, which is needed to be considered. As the paper represented a novel methodology of EEG based medical biometrics system using four different colours in a laboratory-based condition, more research is required to find out the feasibility in real-world condition as well.

5. Conclusion

In order to develop an EEG based medical biometrics system using this proposed model, an analysis was done in this study to find out the feasibility of the time and frequency domain EEG features, with respect to different color stimuli. Here efficiency is obtained after applying several steps- feature scaling, tuning of classifiers and finally with five-fold cross-validation of the developed models. The analysis of the results show that the blue color based biometrics system shows the best accuracy than the other color stimuli and the accuracy profiles are promising, i.e. LR (82.8.3%), KNN (81.7%), SVM (77.2%), RFC (88.9%). The further research investigation found that the blue color stimuli with RFC classifier showed the best accuracy while the red color stimuli showed the most insufficient accuracy. However, the experiment could be done on more number of participants to validate the model based on leave one participant out approach. Also, efficiency will increase with the addition of more EEG channels which can be considered for future work. Using a deep-learning framework would be more appropriate, which will reduce the time for handcrafted feature extraction.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors want to acknowledge the signal recording facilities provided from the Biomedical Engineering Laboratory, KUET.

References

[1] M.M. Hasan, N. Hasan, M.M. Islam, M. Suhi, "Human identification based on color stimuli," in 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE: 1–4, 2019. DOI: 10.1109/ICASERT.2019.8934893

[2] F.F.I.E. Council, "Authentication in an internet banking environment," FFIEC Gencies (August 2001 Guidance), 2005.

[3] M.M. Hasan, M.H.A. Sohag, M.E. Ali, M. Ahmad, "Estimation of the most effective rhythm for human identification using EEG signal," in 2016 9th International Conference on Electrical and Computer Engineering (ICECE), IEEE: 90–93, 2016, doi:10.1109/ICECE.2016.7853863.

[4] M.M. Hasan, M.H.A. Sohag, M. Ahmad, "EEG biometrics based on small intra-individual and large inter-individual difference of extracted features," in 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), IEEE: 1–4, 2016. DOI: 10.1109/ICECTE.2016.8934893

[5] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, G. Ruffini, "Unobtrusive biometric system based on electroencephalogram analysis," EURASIP Journal on Advances in Signal Processing, **2008**(1), 143728, 2007.

[6] C.R. Hema, M.P. Paulraj, H. Kaur, "Brain signatures: A modality for biometric authentication," in 2008 International Conference on Electronic Design, IEEE: 1–4, 2008. DOI: 10.1109/ICED.2008.8934893

[7] Q. Gui, Z. Jin, W. Xu, "Exploring EEG-based biometrics for user identification and authentication," in 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), IEEE: 1–6, 2014.

[8] D. Zakzewski, I. Jouny, Y.C. Yu, "Statistical features of EEG responses to

color stimuli," in 2014 40th Annual Northeast Bioengineering Conference (NEBEC), IEEE: 1–2, 2014.

[9] M.M. Hasan, N. Hasan, A. Rahman, M.M. Rahman, "Effect of Smoking in EEG Pattern and Time-Frequency Domain Analysis for Smoker and Non-Smoker," in 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), IEEE: 1–4, 2019. DOI: 10.1109/IC4ME247184.2019.9036492

[10] D.Q. Utama, T.L.R. Mengko, R. Mengko, M.N. Aulia, "Color blind test quantification using RGB primary color cluster," in 2016 International Conference on Information Technology Systems and Innovation (ICITSI), IEEE: 1–4, 2016.

[11] D. Acquisition, Analysis with BIOPAC MP Systems, AcqKnowledge 4 Software Guide. pdf, Biopac Systems, Inc. ISO 9001: 2008.

[12] M.M. Hasan, M. AshfaqulIslam, S.A. Imtiyaz, M. MahbubHasan, "Presumption method for detecting and analyzing human mental behavior by employing EEG signal," in 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), IEEE: 519–523, 2018. DOI: 10.1109/ICME247184.2018.9036492

[13] D. Flück, "Ishihara 38 Plates CVD Test| Colblindor," Colblindor. Wordpress. Nd Web, 7–11, 2016.

[14] C. Colantuoni, G. Henry, S. Zeger, J. Pevsner, "SNOMAD (Standardization and NOrmalization of MicroArray Data): web-accessible gene expression data analysis," Bioinformatics, **18**(11), 1540–1541, 2002, doi:https://doi.org/10.1093/bioinformatics/18.11.1540.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, "Scikit-learn: Machine learning in Python," The Journal of Machine Learning Research, **12**, 2825–2830, 2011.

[16] S. Bhattacharyya, A. Khasnobish, S. Chatterjee, A. Konar, D.N. Tibarewala, "Performance analysis of LDA, QDA and KNN algorithms in left-right limb movement classification from EEG data," in 2010 International conference on systems in medicine and biology, IEEE: 126–131, 2010.

[17] H. Lee, S. Choi, "Pca+ hmm+ svm for eeg pattern classification," in Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings., IEEE: 541–544, 2003.

[18] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," Computer Methods and Programs in Biomedicine, **108**(1), 10–19, 2012, doi:https://doi.org/10.1016/j.cmpb.2011.11.005.

[19] J. Kim, J. Lee, C. Lee, E. Park, J. Kim, H. Kim, J. Lee, H. Jeong, "Optimal feature selection for pedestrian detection based on logistic regression analysis," in 2013 IEEE International Conference on Systems, Man, and Cybernetics, IEEE: 239–242, 2013.

[20] H. Rajaguru, S.K. Prabhakar, "Non linear ICA and logistic regression for classification of epilepsy from EEG signals," in 2017 international conference of electronics, communication and aerospace technology (ICECA), IEEE: 577–580, 2017.

[21] K. AlSharabi, S. Ibrahim, R. Djemal, A. Alsuwailem, "A DWT-entropy-ANN based architecture for epilepsy diagnosis using EEG signals," in 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), IEEE: 288–291, 2016, doi:10.1109/ATSIP.2016.7523093.

[22] W. Zhu, N. Zeng, N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," NESUG Proceedings: Health Care and Life Sciences, Baltimore, Maryland, **19**, 67, 2010.

[23] D. Justin, R.S. Concepcion, A.A. Bandala, E.P. Dadios, "Performance Comparison of Classification Algorithms for Diagnosing Chronic Kidney Disease," in 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE: 1–7, doi:10.1109/HNICEM48295.2019.9073568.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 29-35 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Curved Pyramidal Metamaterial Absorber: From Theory to an Ultra-Broadband Application in the [0.3 - 30] GHz Frequency Band

Zeinab Fneish[1,2,3,*], Hussam Ayad[1], Moncef Kadi[2], Jalal Jomaa[4], Ghaleb Faour[3]

[1]*Lebanese University, Physics Departement, Faculty of Science, Beirut, 2038, Lebanon*

[2]*Normandie University, Ecole superieur d'ingenieur ESIGELEC, Research Institute, Rouen, 76800, France*

[3]*Scientific Research National Center (CNRS), Beirut, 2038, Lebanon*

[4]*Maaref University, Beirut, 2038, Lebanon*

ARTICLE INFO

Article history:
Received: 20 September, 2020
Accepted: 04 February, 2021
Online: 10 March, 2021

Keywords:
Metamaterial absorber
Relative Absorptive Bandwidth
Absorption
Resonance
Curved Altitude
Pyramidal Absorber
Broadband Response

ABSTRACT

For its importance nowadays in a wide range of applications such as the anechoic chamber, we introduce a microwave ultra-broadband polarization-independent metamaterial absorber (MA) in the Ultra High Frequency (UHF)/ Super High Frequency (SHF) frequency bands. Through this work, we improved the Relative Absorptive Bandwidth (RAB) of the conventional pyramidal absorber (CPA) by modifying its altitude to a curved shape. As a result, the RAB increased from 25.9 % to 71.82 % with an absorptive level greater than 90% paving the way to an optimized structure for a broader band of absorption. As a second target, we looked for widening the broadband absorption of the CPA in the low-frequency region. To achieve this aim, we introduced two new prototypes. The first with a total thickness of 12.7 cm, consisting of 35 curved resonant layers where numerical simulations show an enhanced design with an absorption band from 0.3 GHz to 30 GHz referring ta a RAB of 182%. The second prototype consists of a cell containing different pyramidal absorbers grouped in-plane in a unit cell; such structures operate in complementary bands. This prototype is dedicated to combining these bands of absorption. After that, an enhancement is presented of this latest to reach a well-combined band with a RAB of 128.69%. We used for simulation, testing, and collecting results the High-Frequency Structure Simulator (HFSS) tool.

1. Introduction

This paper is an extension of work originally presented at the *7th Mediterranean Congress of Telecommunications (CMT) conference* [1] where we presented an enhanced prototype of a broadband electromagnetic Metamaterial Absorber (MA).

Because of its matched impedance due to the electric and magnetic resonance, the MA is able to omit the reflection by strongly absorbing the incident wave in the dielectric [2]. Based on the latter proposition and in different frequency bands [3-6], other designs of MAs have been introduced. Because its principle is based on resonance, the absorption bandwidth of MAs relies on narrow resonant frequencies. However, broadband absorption is an important factor in many applications one of them is the anechoic chamber. To overcome this issue, the literature proposes to increase the number of simple resonators with size variation either in longitudinal directions [6] as the pyramidal Absorber (PA) structure by stacking patches resonators or in transverse directions [7]. One of the novelties addressed in this article is a novel prototype that combines these two ways in one model in such a way that the PA structure is used as a part of a unit cell that has different PAs with size variation in the transverse direction.

Because its negligible incident angle dependence [8] and its negligible polarization dependence due to its symmetrical design geometry, the PAs structures achieve a great importance. PA is composed of a periodic array of multilayered patches forming a quadrangular pyramid where these pyramids possess resonant

*Corresponding Author: Zeinab Fneish, fneichzeinab@gmail.com

absorption modes at multi-frequencies in each patch layer. Due to the stacked multi-dimensional patches, we obtain a resonant response on successive multi-frequencies, which the overlapping conduct to a total absorption of the incident wave over a broad band of frequency.

Researches are rarely attempting to resolve the broadband PA below Ultra High Frequency (UHF)/ Super High Frequency (SHF) bands even though many articles have been published [9-12] focusing on the former issue in the microwave, Terahertz (THz), infrared and visible regions.

A metamaterial absorber at this region suffers from a big dimension view of its wavelength dependence. In literature, most of the miniaturized MA that is operative in low frequencies has a narrow band response [13]. Besides, from the industrial point of view, regarding the absorber operation, there exists always a problem of immensity in the frequency band [0.3-1 GHz]. Also, in the latest anechoic chamber, the absorption at low frequency is achieved by two main ways: through heavy ferrite material or by a 1-meter thickness of pyramidal Urethane foam. This specific band gains a high importance because it includes the region of the telecommunication waves. That is why in this work, we design an ultra-broadband PA working in UHF and above regions. We will use a high permittivity dielectric material as a substrate to decrease the dimension of resonators. Moreover, we will enhance the absorption response of a pyramidal design by adding a new factor in optimization. Then the enhanced pyramidal will be used as a unit in two novel prototypes that can achieve an ultra-broadband of absorption.

The first section contains the theoretical model analysis of a conventional PA that is considered as the basic's unit in our broadband design. The study is promoted by simulation results. The second section highlights the effect of using the curved altitude on enhancing the absorption response of the CPA by means of numerical simulations highlighting the improvement of the RAB. After that, two novel prototypes based on PA units are presented aimed to make the absorption response broader in the UHF/SHF band.

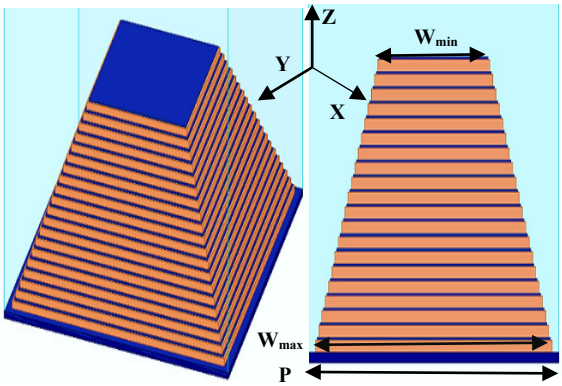
All calculations were performed on a High-Performance Cluster (HPC) of 24 cores with a systems memory of 192 GB RAM. Discrete frequency analysis mode was adjusted to make simulation results with very high precision.

2. Conventional Pyramidal Structure

Based on the design equation of the simple patch resonator (1), the resonant absorption frequency f_r is related to the side W of the patch and relative permittivity of the dielectric substrate (ε_r) [14]. Then, we can conclude that its response is limited to a narrow band of absorption. For that, stacking multi-dimensional patches is proposed in the literature by PA structures as an efficient way that can widen the band of absorption.

$$f_r = \frac{c}{2W\sqrt{\varepsilon_r}} \tag{1}$$

PA design provides broadband absorption. It consists of a periodic array of pyramids with a quadrangular cross shape and a homogeneous metal ground in the bottom, the latter blocking any transmission of the incidence wave ($S_{21} = 0$).



(a) (b)

Figure 1: Design of an ultra-broadband PMA, (a) 3-D illustration of the simulation MA, (b) Side view of the PMA unit cell [1].

In this model, the thickness of the metallization layer and the dielectric layers in each patch layer are optimized to be; t_m=35 µm, t_d = 140 µm respectively. The metal used is copper with electric conductivity of σ =5.8×10^7 S/m. FR4 is used for the dielectric substrate with the relative permittivity of 4.4 and loss tangent equal to 0.02. In this primary model, 20 resonator layers are chosen.

In the simulation, the periodic boundary conditions are assigned along the *x* and *y*-directions. A wave port is launched along the *z*-direction with *E* field polarized along the *y*-direction.

RAB defined in (2) is the factor that describes the absorption bandwidth performance of an absorber.

$$RAB = \frac{f_{\max} - f_{\min}}{f_c} \tag{2}$$

where f_c is the central frequency of the absorptive band presented in (3) is given by:

$$f_c = \frac{f_{\max} + f_{\min}}{2} \tag{3}$$

Where f_{max} and f_{min} are the upper and lower limits of a frequency range with a specific absorption level, respectively.

The absorption coefficients as a function of frequency from 10 GHz to 28 GHz were shown in Figure 2. There are 20

absorption peaks with an absorbance of approximately above 80%, corresponding to the number of dielectric patches between two neighboring metal interfaces.

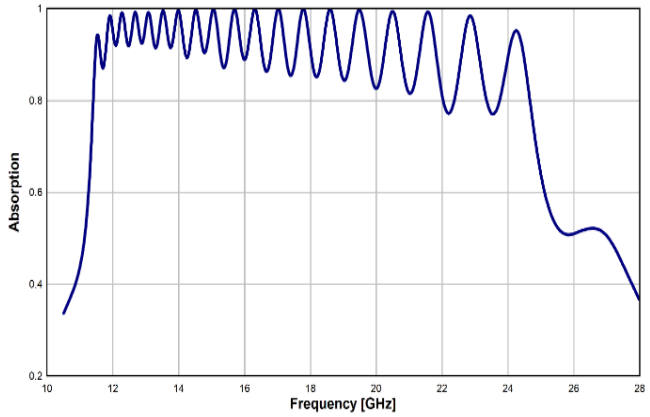


Figure 2: The simulation absorption response of the optimized pyramidal structure with dimensions of: Wmin =2.975 mm, Wmax=6.3 mm, P= 6.65mm, tm= 35µm, td= 140 µm [1]

As shown in Figure 3, Electric and Magnetic field distributions are plotted at three frequencies (11.5, 17, and 20.45 GHz).

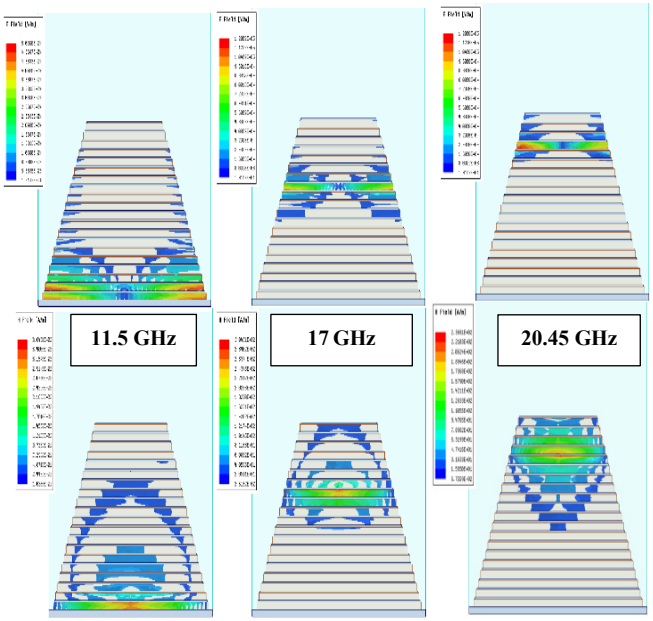


Figure 3: Obtained results of the Magnetic and Electric Magnitude distributions taken at three different resonance frequencies [1]

Results show that each two adjacent metal layers spaced with a dielectric layer can localize an electric and magnetic field at special resonant frequencies. Such resonances lead to zero-reflection by impedance matching to the free space at resonance.

From figure 3, we can observe that, at lower frequencies, the electromagnetic field is localized at the bottom layers of the pyramid. As the frequency is increased, the electromagnetic fields are localized in the upper layers of the pyramid. Since the layers of the pyramid are gradually decreased in width from the bottom to the top, they resonate at different successive frequencies that ensure an ultra-broadband absorption response. Based on this theoretical study, as the number of layer increases, the absorption window broadens gradually due to additional resonant frequencies

3. Curved Altitude Enhancing Factor

In this section, we include the numerical simulations that show the improvement in the RAB by considering the curved altitude modification on the pyramidal structure. The latter modification adds a new degree of freedom that can improve any pyramidal structure. The unit cell of the pyramidal structure in Figure 1 is considered as a criterion for evaluation of the impact of applying a curved altitude to the pyramidal structure.

Figure 4 shows three different forms of altitudes applied to the same structural parameters. We take the sawtooth altitude (Figure 4 (a)), the linear altitude (Figure 4 (b)), and the curved altitude (Figure 4 (c)).

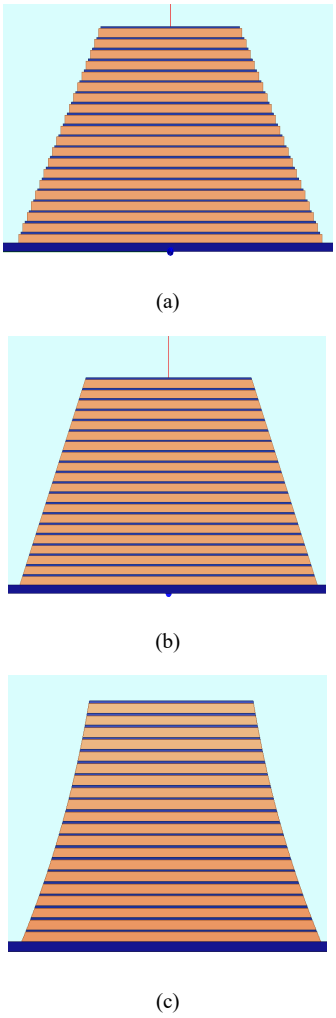


Figure 4: Diverse altitude design shapes of a Pyramidal structure. (a) Sawtooth Altitude, (b) Linear Altitude, and (c) Curved Altitude [1].

The curve altitude is drawn by a 3-point arc and optimized for the sake of achieving the best outcomes. In Figure **5**, we plot the frequency-dependent absorption response for the designs shown in Figure 4.

The simulation results show that the level of absorption in linear and sawtooth altitudes is below 90% at a higher frequency, and it can cover a broad band of absorption but with an oscillator level of absorption (Figure 5 (a) and (b)). This oscillation phenomenon is explained by the need to decrease the difference in dimension between the highest patches on top to ensure nearer resonant frequencies. This result guided us to think about a novel altitude shape. With the curved altitude design, the resonance peaks, at the higher frequency, became nearer to each other and well combined that explains the increase in the RAB. Therefore, the Curved PA reaches 71.82 % RAB with an absorption level above 90% showing a larger absorption compared to that obtained by other altitude forms (see Table 1). This result paves the way to found a new parameter in optimizing such a design.

Table 1: Relative Absorption Bandwidth for different altitude form obtained by simulation results.

	RAB with Absorption level above 80%	RAB with Absorption level above 90%
Sawtooth Altitude	**63.3%**	**25.9%**
Linear Altitude	**46.3%**	**33.3%**
Curved Altitude	**73.4%**	**71.82%**

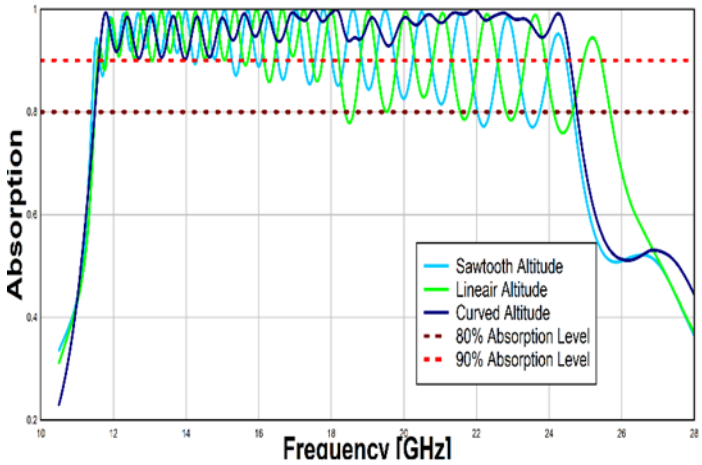


Figure 5: Absorption response for diverse altitude design shapes: Sawtooth Altitude, Linear Altitude, and Curved Altitude.

4. Absorber Prototypes with an Ultra-Broadband Absorption Spectrum

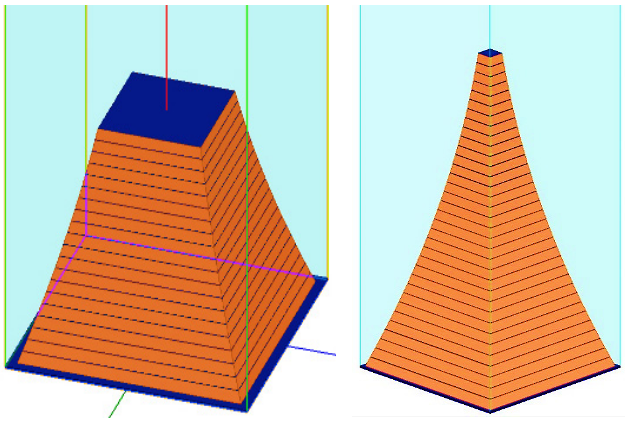
In this section, we will present two novel structures that can achieve broader absorption results by increasing the number of resonators either in transverse or longitudinal directions as addressed in the subsections below.

As we cited before, our study is directed to get a broad absorption band at lower frequencies. For that, according to the design equation (1) to get a low resonant frequency response, we shall increment the dimension of the considered patch in addition to the relative permittivity of the substrate material.

After the full parametric study, we reached a state in which we can tune the operating frequency to drive the new design to become absorptive in the UHF band by using dielectric material with ε_r = 55 and modifying the fundamental patch to be 72 mm.

4.1. The prototype of Adding Supplementary Patches by Altitude

Due to the role of the conic foam form in the anechoic chamber and for the goal of enlarging the band of absorption, we increased the number of supplementary patches in altitude by 15 using the same stitching method to the pyramid scaled to low frequency consisting of 20 layers as shown in Figure 6 (a). Built up from 35 stacked patches (see Figure 6(b)), the new prototype makes it possible to get a sharp peak at the top; this step can help in decreasing the interface between an electromagnetic wave and the design leading to a higher absorption compared to reflection.



(a) (b)

Figure 6: Optimized Pyramidal Design. (a) Prototype of optimized pyramid in UHF Band. (b) Prototype of the Pyramid Absorber after adding 15 patches, the biggest patch is 72 mm in length and the smaller one at the top is 11.7mm [1].

First, we adjust an optimized model of PA with an optimized curved altitude to be operative in the UHF region (Fig.6(a)). Figure 7 displays the spectrum of absorption of such model, it shows an absorptivity exceeding 80% in the frequency band [300-625] MHz. In this structure, we inserted the 15 supplementary layers by height (Fig. 6(b)). Figure 8 aims to verify that the proposed new design reached an ultra-broadband of absorption covering the UHF/SHF frequency regions.

Theorically, the new design should gives 35 resonant peaks of absorption referring to the basic operation of PA that declare that every patch must resonates in a unique peak of frequency. Luckily, the achieved outcomes crossed over the expectations with an additional factor of absorption at higher frequencies.

In the range [0.3 GHz - 1.92 GHz] of absorption (Phase I), we can explain the absorptive results to the successive resonant absorption modes that occurred at multi-frequencies as expected (fig.9). The reason behind the plots of the electric and magnetic field distributions in Figure 9 is to understand physically the absorption took place in Phase I. In resonant mode, both magnetic and electric resonance exist where they lead to a matching between the absorber impedance and the free space impedance then the wave is absorbed in the dielectric.

Phase II [1.92 GHz- 30 GHz] , with a level greater than 90% (Fig. 8 (b)) the ultra broad-band continues. In this region, the

reason for absorption can be linked to a strong coupling between spaced patches due to the formation of some absorbing modes (see Figure 10). Two main factors clarify the observation: the high permittivity of the considered material that causes a high absorption and the conical geometry with a peaked top that may be responsible for the decrease in the reflection, and To make sure about this result, it has to be experimentally validated.

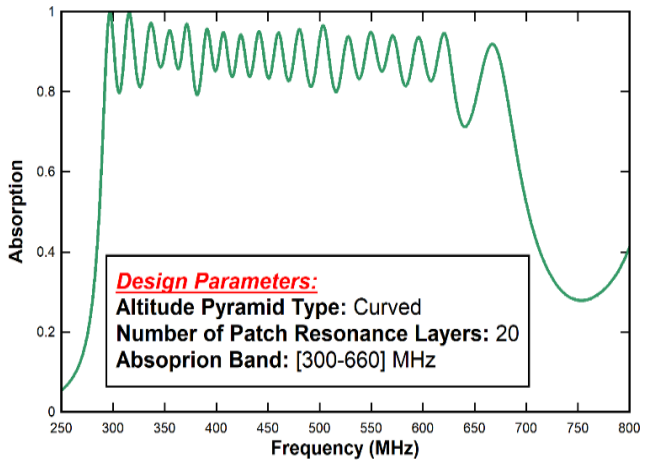
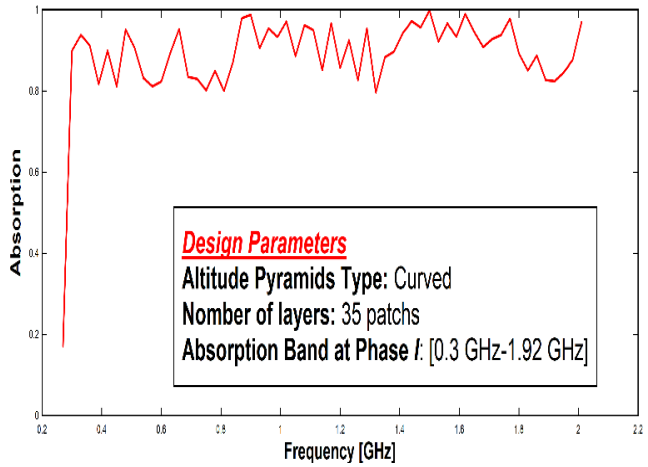
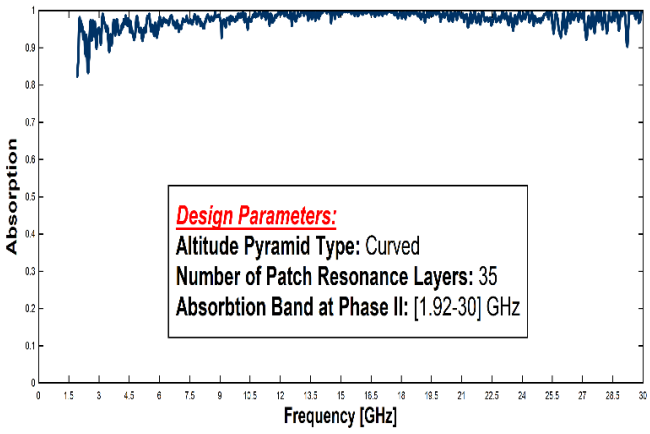


Figure 7: The absortption response of Figure 3(a) in the UHF band [1]



(a)



(b)

Figure 8: Absorption response of the EP design of Figure 3(b) in the UHF/SHF band after adding 15 layers. (a) Phase I of absorption. (b) Phase II of absorption [1]

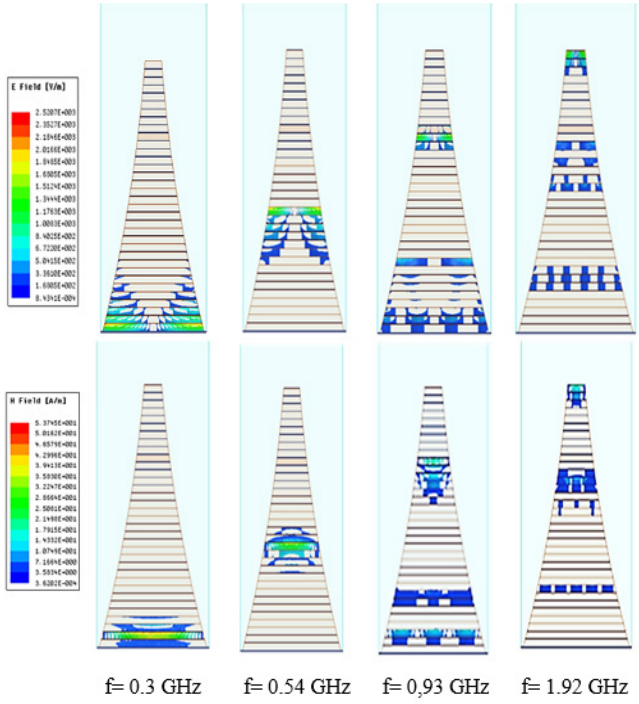


Figure 9: Simulated Electric magnitude distributions at some frequencies at Phase I of absorption [0.3 GHz - 1.92 GHz] [1]

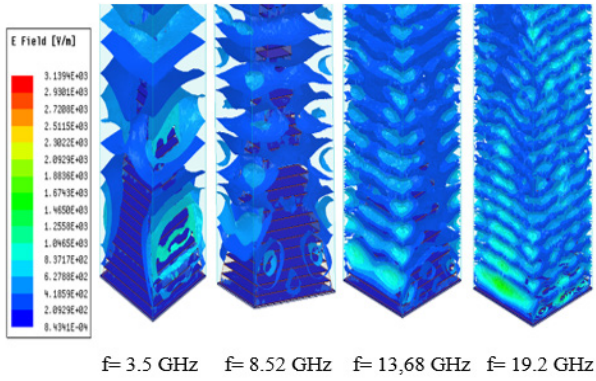


Figure 10: Simulated Electric magnitude distributions at some frequencies of the second phase of absorption [1]

4.2. The prototype of combining pyramids with complementary bands planar in one unit cell

In this section, we test the efficiency of adding supplementary resonators in the transverse direction to have the broadest absorption response. For that, we apply a novel idea that consists of grouping two kinds of curved and optimized PAs that operate on complementary frequency bands in a one-unit cell. This combined unit cell should also respect the geometrical symmetry that permits to make the model insensitive to the wave polarization as shown in Figure 11.

For that, we adjust two PAs with the same material properties to be absorptive in complementary bands (Figure 12).

As shown in Figure 12, structure 1 operates in the lower band covering [0.29-0.68 GHz] and structure 2 operates in the complementary band covering [0.68-1.48 GHz] with an absorption level greater than 80% in these two structures.

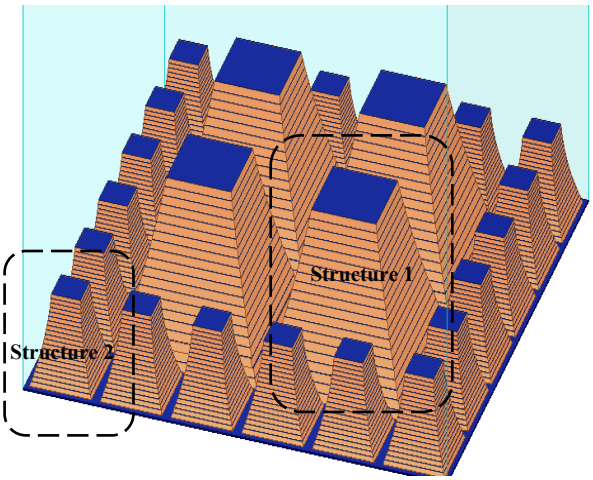


Figure 11: Structure of the combined structure operating in complementary bands

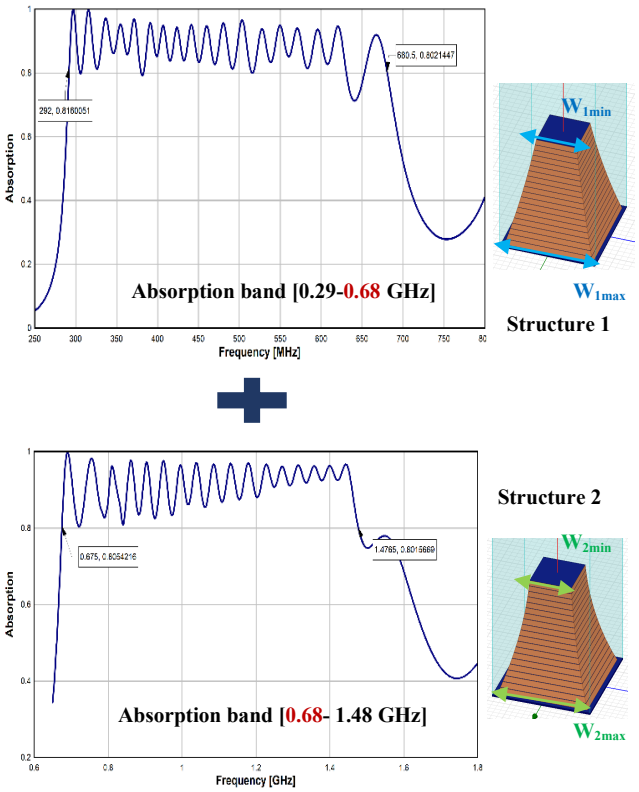


Figure 12: The operating absorption band of structures composing the unit cell with geometrical dimensions: W_{1max}=72mm; W_{1min}=33mm; W_{2max}=31.5 mm; W_{2min}= 14.7mm.

This design imposes a geometrical spacing between pyramids with the same dimensions that can degrade the grouping performance of complementary bands. For that, the periodicity between pyramids is taken to be as minimum as possible in such a way to get the pyramids closed to each other and prevent any additional spacing between the same pyramid's dimensions.

To resolve this geometrical disadvantage an enhanced design is presented inserting a third pyramid as shown in Figure 13. The latest pyramid is structured and optimized to operate in a medial band covering [0.42 -0.99 GHz].

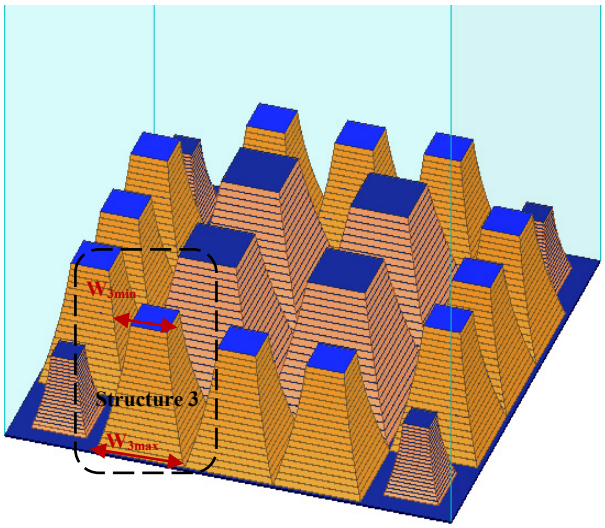


Figure 13: Enhanced Model of the combined structure, considering these geometrical dimensions for structure 3: W_{3max} =50.22 mm; W_{3min} = 21.06 mm.

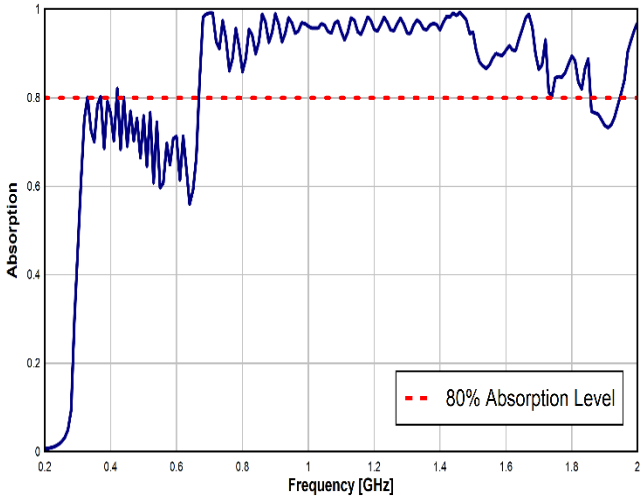


Figure 14: Absorption response of the first combined design presented in Figure 11.

From Figure 14, it is shown that the bands are not combined properly, what is particularly striking is the loss of absorption in the lower frequency band due to the increase in the distance between patches of the same size. Increasing the distance between patches will affect the damping coefficient of the structure that directly determines the level of absorption.

The enhanced structure comes to resolve such a problem by adding a third structure operating in a middle band. Results show that the bands are now combined better but a part of the lower band is still below 80% of absorption (Figure 15). By this result, we obtained an absorption band covering [0.41-1.89 GHz] with an absorption level greater than 80%. Such band refers to a RAB of 128.69 %. By this enhanced prototype, the combination is successfully obtained from [0.41 GHz-1.89 GHz]. Such a result shows that there is an extra band of absorption from [1.48 GHz-1.89 GHz]. Figure 10 shows the absorbing modes that cause a strong coupling between spaced patch; such phenomenon represents the cause of the observed extra band of absorption.

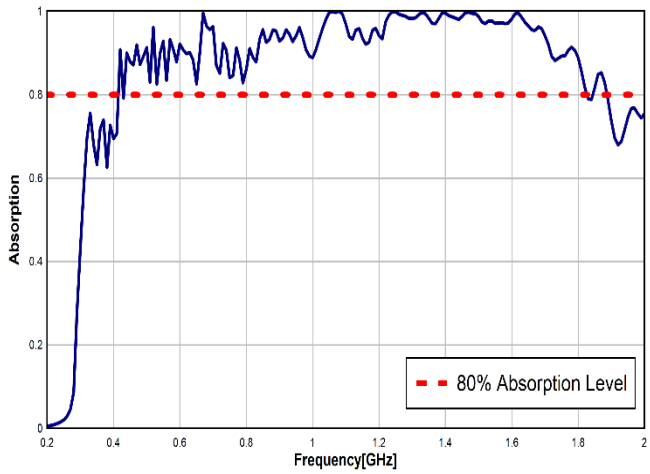


Figure 15: The absorption level of the design presented in Figure 13.

5. Conclusion

Due to the remarkable characteristics of the pyramidal design, we covered a full theoretical study in this paper. We take this structure as a basic model in our study by different geometrical shapes. The curved altitude is found as an enhancing factor where the relative absorption bandwidth increase from 63.3 % to 73.4 % with an absorption level greater than 80 %. Added to these results, we presented two novel prototypes based on the PA that can achieve a broader absorption response.

Generally, the broadband objective can be achieved by increasing the number of resonators with size variation either in transverse or longitudinal directions. That is was applied in this article based on pyramidal units instead of simple patches where they were combined in-plane and in altitude. Each idea was showed its efficiency in broadening the absorption band. We reached a RAB of 182.83 % by applying the curved pyramid of 35 layers and 128.69 % when applying the transverse combining method. The novelty in the ways of patches structuring applied in this article, admits to a RAB values unreachable in many broadband absorber essays in literature.

For achieving important results from the considered prototypes, there is a need for experimental testing after fabrication.

Acknowledgment

The Lebanese University (LU) and the National Center of Scientific Research (CNRS) and the support this work.

References

[1] Z. Fneish, H. Ayad, M. Kadi, F. Mazeh, J. Jomaah, G. Faour, "Ultra Broadband Curved Pyramidal Absorber Metamaterial in the UHF/SHF Region," in 7th Mediterranean Congress of Telecommunications 2019, CMT 2019, IEEE: 1–4, 2019, doi:10.1109/CMT.2019.8931391.

[2] N.I. Landy, S. Sajuyigbe, J.J. Mock, D.R. Smith, W.J. Padilla, "Perfect metamaterial absorber," Physical Review Letters, **100**(20), 207402, 2008, doi:10.1103/PhysRevLett.100.207402.

[3] H. Tao, N.I. Landy, C.M. Bingham, X. Zhang, R.D. Averitt, W.J. Padilla, "A metamaterial absorber for the terahertz regime: design, fabrication and characterization," Optics Express, **16**(10), 7181–7188, 2008, doi:10.1364/OE.16.007181.

[4] A.-X. Wang, S.-B. Qu, H. Dai, G. Zhang, W.-J. Wang, M.-B. Yan, "Design of Ultrathin Five-band Polarization Insensitive Metamaterial Absorbers," in 2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT), IEEE: 658–661, 2019, doi:10.1109/ICEICT.2019.8846323.

[5] M.A. Ghaderi, P. Enoksson, R.F. Wolffenbuttel, "CMOS Compatible Fabrication of Mid Infrared Microspectrometers Based on an Array of Metamaterial Absorbers," in 2019 20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII), IEEE: 1580–1583, 2019, doi:10.1109/TRANSDUCERS.2019.8808691.

[6] C.-H. Lin, R.-L. Chern, H.-Y. Lin, "Polarization-independent broad-band nearly perfect absorbers in the visible regime," Optics Express, **19**(2), 415–424, 2011, doi:10.1364/OE.19.000415.

[7] A. Sellier, Absorbants à métamatériaux: étude théorique et expérimentale, 2014.

[8] F. Ding, Y. Cui, X. Ge, Y. Jin, S. He, "Ultra-broadband microwave metamaterial absorber," Applied Physics Letters, **100**(10), 103506, 2012, doi: 10.1063/1.3692178.

[9] M. Lobet, M. Lard, M. Sarrazin, O. Deparis, L. Henrard, "Plasmon hybridization in pyramidal metamaterials: a route towards ultra-broadband absorption," Optics Express, **22**(10), 12678–12690, 2014, doi: 10.1364/OE.22.012678.

[10] Q. Liang, W. Yu, W. Zhao, T. Wang, J. Zhao, H. Zhang, S. Tao, "Numerical study of the meta-nanopyramid array as efficient solar energy absorber," Optical Materials Express, **3**(8), 1187–1196, 2013, doi: 10.1364/OME.3.001187.

[11] Q. Liang, T. Wang, Z. Lu, Q. Sun, Y. Fu, W. Yu, "Metamaterial-based two dimensional plasmonic subwavelength structures offer the broadest waveband light harvesting," Advanced Optical Materials, **1**(1), 43–49, 2013, doi: 10.1002/adom.201200009.

[12] J. Zhu, Z. Ma, W. Sun, F. Ding, Q. He, L. Zhou, Y. Ma, "Ultra-broadband terahertz metamaterial absorber," Applied Physics Letters, **105**(2), 21102, 2014, doi: 10.1109/JPHOT.2018.2888971.

[13] S.T. Bui, Y. Yoo, K.W. Kim, D.L. Vu, Y. Lee, "Small-size metamaterial perfect absorber operating at low frequency," Advances in Natural Sciences: Nanoscience and Nanotechnology, **5**(4), 45008, 2014, doi: 10.1088/2043-6262/5/4/045008.

[14] H. Ayad, Antenna Performance Control using Metamaterials, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 35-39 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Mechanical Characterization of Recycled Aggregates Concrete Based on its Compressive Strength

Khaoula Naouaoui*, Toufik Cherradi

Mohammadia School of engineers, Mohamed V University Agdal, Rabat, 10000, Morocco

ARTICLE INFO

Article history:
Received: 23 December, 2020
Accepted: 12 April, 2021
Online: 05 May, 2021

Keywords:
Concrete
Mechanical properties
Natural aggregates
Recycled aggregates
Replacement percentage

ABSTRACT

The construction industry has undergone several changes in recent years linked to new laws and international conventions aimed at protecting the environment and combating pollution. Construction industry alone produces tons of waste annually due to debris produced either during construction or during deconstruction. To combat this, companies are forced to control their debris either by reusing it on site or by sending it to specialized landfills. Thus, new materials appear constantly based on the recycled materials. Recycled aggregate concrete was thus born. It is a concrete based on the use of recycled aggregates retrieved from the demolished structures to replace natural aggregates. Characteristics of this type of concrete depends of the chosen replacement percentage of natural aggregates specially the mechanical properties. This article is part of my research studies done in the civil laboratory of the Mohammadia School of engineers. The study is based on the identification of the recycled aggregates, the determination of physicals and mechanical characteristics of the aggregates, the determination of the effect of the use of recycled aggregates on the concrete characteristics and finally the improve of the quality of the concrete to prove so that it can replace ordinary concrete. The test results described in this article show that the increase of the replacement by recycled aggregates decreases the mechanical properties especially when it is up to 25-30%. The results also prove that the ad of additives especially plasticizer with 1% replacement of cement improves the compressive strength of concrete and allows us to use up to 50% of replacement.

1. Introduction

The field of construction knows a radical undergoing and innovative changes in the designs proposed, the types of materials used and the domains of their use. Concrete, as the main construction material, is now at the center of several studies, researches, innovations and improvements. Several types of new concrete are proposed for their quality, compliance with environmental and / or technical requirements, design and special needs. One of the concrete types proposed for their environmental value, their support for sustainable development is the recycled aggregates concrete.

The comparison of its properties with those of concrete based on naturel aggregates. A summary of the comparison between the two concretes is detailed below:

Related to the compressive strength, the majority of studies confirm that up to [20-30] percentage replacement of natural aggregates by recycled ones, the compressive strength is almost the same but beyond this replacement level, it decreases considerably. Main recent studies results are as folow.

Many experimental studies match our synthesis that up to 25% the compressive strength is the same and that the decrease's max is 20% for 100 % replacement level [1]. Others conducted a synthesis of several researches and works in his article. He concluded that in case compressive strength values of the original concrete, of which recycled aggregate is manufactured, and targeted values were approximately equal. The strength values of recycled aggregate concrete was 5 to 10% lower than those of the comparable naturel aggregate concrete were [2].

For the modulus of elasticity, the experimental studies prove that its values have almost the same trend as compressive strength results. The increase in insignificant for percentage level up to 75 % (9 %) and it reaches 18% for 100% replacement of natural aggregates by recycled ones [1]. In another study, they found a reduction of 34.8% when the replacement level is 100 % compared to the control concrete [3]. Meanwhile, other researchers found 45 % of reduction for the same replacement level. [4]

*Corresponding Author: Khaoula Naouaoui, naouaoui.khaoula@gmail.com

Studies done on durability properties as open porosity, sorptivity, chloride permeability confirm that they decrease with the increase of RCA content. The study also concluded that the rate of deterioration is higher compared to mechanical characteristics.

For example, for 100% replacement level, the resistance to chloride permeability decrease by 9.5% [5], the water absorption is lower by 7.37% [3] and the open porosity by 20% [6].

2. Experimental Protocol

This section is related to the description of the experimental protocol done to determinate the mechanical properties of concrete based on the restitution of the natural aggregates by recycled one.

2.1. Concrete formulation to determinate the constituents

The main of our study is to determine the possibility of replacing naturel concrete by recycled aggregates concrete in the majority of types of building, thus the concrete chosen to study have the following characteristics:

- Basic concrete, ordinary, used in all types of based constructions.
- F_{c28} = 25MPA.
- Slump: Average of 7cm.

To formulate this concrete, the constituents chosen are:

For aggregates, we have chosen two aggregate, the first are natural ones and the second recycled aggregates based on a crushed building. For the two of them we used two dimensions: [5-12.5] and [12.5-31.5]

The result of the granulometric test of the two aggregates is presented in the figure 1:

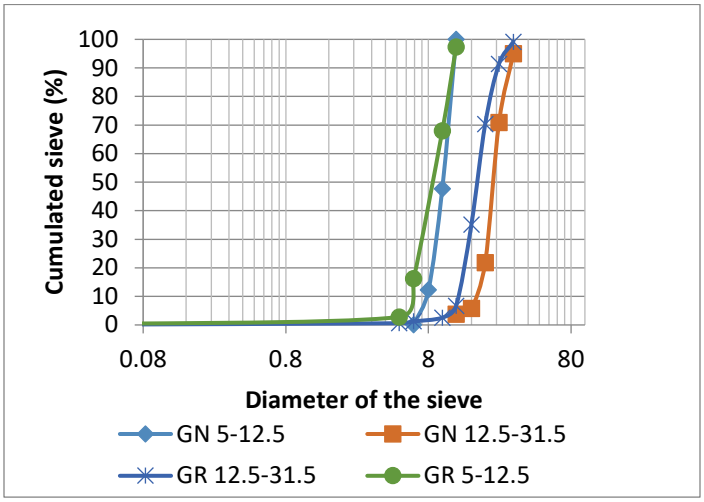


Figure 1: Granometric curve for natural and recycled aggregates used in experimental study (GN : Naturel gravel / GR : recycled gravel).

The quantities of materials used based on the dreux gorisse method [7] are, as follows:

- Water : 0.19 m^3 / 190 kg
- Cement : 0.29 m^3 / 350 kg
- Sand : 0.29 m^3 / 472.85 kg
- Gravillon G1 : 0.15 m^3 / 376.14 kg
- Gravillon G2 : 0.2 m^3 / 537.3 kg

2.2. Concrete mixtures

The aim of the study is to detect the effect of the replacement percentage of natural aggregates by recycled aggregates in the compressive strength of concrete and the effect of different types of additives in improving the mechanical characterisation of theses concretes.

The replacement ratio studied are 0% - 20% -50% -75% - 100%.

For the mixture with additives, we used three product from SIKA- MAROC:

- Plasticizer/ water reducer: BV40
- Superplasticizer/ high water reducer: ViscoCrete Tempo 10 M
- New generation superplasticizer: SikaFluid R.

The first step of experimental study was as explained before, the effect of the replacement of naturel aggregates by recycled ones. The quantity of component used for every mixture in detailed in the table.

Table 1: Quantity used in the compressive strength test (kg)

Component		0 % GR	20 % GR	30 % GR	50 % GR	75 % GR	100 % GR
Cement		350					
Water		193					
Sand		473					
GN	5-12.5	402	321,6	281,4	201	100,5	0
	12.5-31.5	512	409,6	358,4	256	128	0
GR	5-12.5	0	80,4	120,6	201	301,5	402
	12.5-31.5	0	102,4	153,6	256	384	512

The second step of study was to determinate the effect of additives in improving the concrete quality. To do this, we used three types of adjuvants with different percentages. The experimental protocol is as details below:

The test aims to view the effect of the ad of superplasticizer on the comparison between different replacement levels by calculating the compressive strength for various replacement levels with the superplasticizer. The component is similar to the previous one; it is based on using different percentages of replacement of natural aggregates by recycled one. The only difference is the use of superplasticizer.

Second part of the study is to determinate the Compressive strength with different adjuvants. For this comparative test, we

used a concrete based on 50% of recycled aggregates and 50% of natural aggregates. We, then, added the three types of adjuvant with 1% of replacement of cement with means 3.5 kg/m^3 of concrete.

The last step of comparison is the Compressive strength with different percentages of the adjuvant. For this comparative test, we used different percentage of the Tempo 10M (0.5%, 1% and 1.5%) as a partial replacement of cement. The concrete used is based on 50% of recycled aggregates and 50% of natural aggregates.

3. Experimental results and discussions

3.1. Physical characterization of recycled aggregates

The physical characterization of aggregates from the demolition of an old building can be summarized as detailed below.

Table 2: Summarize of the physical characterization tests

	Recycled aggregates 5-12.5	Recycled aggregates 12.5-31.5
Absolute Density	2480 kg/m^3	
Water content	1.49%	4.32%
Water absorption	11.11%	9.4%

The absolute density is 2.48t/m^3 approximatively 2.5t/m^3 wich is the absolute density of reinforced concrete.

The comparison between two recycled aggregates of two different classes GR1 (5-12.5) and GR2 (12.5-31.5) showed that the water content is greater for large aggregates. This is justified by the presence of several pores which disappear during fine crushing of the aggregates.

The comparison has also shown that the water absorption rate is higher for small aggregates. This is justified by the superiority of the contact surface.

3.2. Compressive strength without additives

The aim was to determinate the effect of the replacement percentage on the compressive strength, so we made mixtures with 0, 20, 30, 50 75 and 100 % replacement percentage.

The results of compressive strength in 7, 21, 28 and 90 days are detailed in Table 3.

Table 3: Compressive strength results for concrete without additives (MPA)

	0% RA	20% RA	30% RA	50% RA	75% RA	100% RA
7j	16,5	14,65	16,75	13,55	14,05	14,5
21j	23,4	19,9	14,8	14,3	19	14,7
28j	24,8	23,55	23,5	18,05	17,7	15
90j	26,3	25,35	23,55	20,25	19,4	17,8

In order to make the comparison easier and more remarkable, a j-day compression graph for all replacement percentages was established. (figure2)

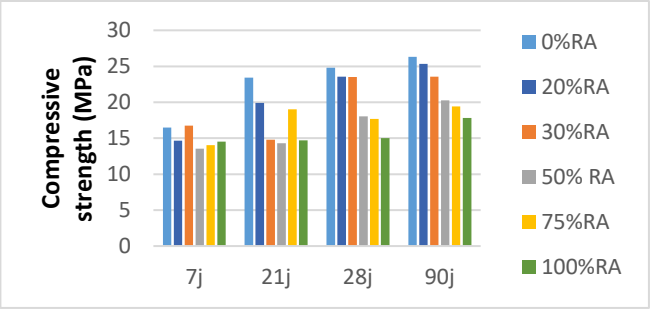


Figure 2: Compressive strength test results representation

At early age, 7 and 28 days, the resistance to compression cannot be decisive because the inter-component reactions are still in progress, so we can notice that the evolution of the resistance for mixtures with 0 or 20% restitution is greater than the others. For concrete containing 30% natural aggregates, it is noted that the compression at 7 days is greater than that of 21 days. This is probably due to poor sampling and that not reflects a real decrease in compression.

28 days is considered to be the reference for resistance in the short and medium term because the compression reaches 90% of its maximum value. The results at 28 days in our study are similar to the majority of researchers. The compressive strength decreases with increasing replacement rate.

After compression tests on recycled aggregate concrete with different percentages, it has been found that beyond 50%, an improvement in quality is required to adopt this type of concrete.

The first proposal to improve this is to add an additive that reduces the water requirement of the aggregates and thus improves the quality of the concrete.

In order to determine the best combination (type of adjuvant and percentage of addition), a series of mixtures are proposed:

- Concrete with 50% recycled aggregates and 1% added admixture with different types of admixture.

- Concrete with 50% recycled aggregates and one type of adjuvant with different percentage of adjuvant added.

- Concrete with several recycled aggregate's percentage and one type and percentage chosen for adjuvant.

3.3. Compressive strength with different adjuvants

The purpose of this step is to compare the effect of several types of adjuvants with a percentage of 1%.

These adjuvants are of several types:

- Plasticizer: BV 40

- Superplasticizer: Tempo 10 m

- New generation superplasticizer: SFR

The choosen percentage of recycled aggregates for this test is 50% as replacement of natural aggregates.

Compression tests are done at 7 and 28 days. The results are detailed in table 6.

Table 4: Compressive strength results for concrete with various additives (MPa)

Type of adjuvant	Compressive strength at 7 days	Compressive strength at 28 days
BV40	19,87	22,63
SFR	11,50	16,65
Tempo 10 M	17,47	21,45

From the results of these tests, it is noted that:

- The new generation superplasticizer gives the lowest result. This can be interpreted as a bad interaction between the aggregates and their cement paste attached with the additive.
- The plasticizer gives the best resistance
- The superplasticizer also gives a good result, better than without additives.

3.4. Compressive strength with different percentages of the adjuvant

These tests aim at the optimal choice of the percentage of the adjuvant in the mixture.

The results of these tests demonstrate that beyond 1%, compression decreases considerably.

Then, for our next mixes, the chosen percentage is 1%.

Table 5: Compressive strength results for concrete with different percentage of additive (MPa)

Type of adjuvant	Percentage of adjuvant	Compressive strength at 7 days	Compressive strength at 28 days
Tempo 10 M	0,5%	14,35	19,90
	1,0%	17,47	21,45
	1,5%	12,87	16,03

3.5. Compressive strength for various replacement levels with the superplasticizer

The purpose of this part of the tests is to see the result of the combination of the adjuvant tempo10M - percentage 1% with various percentages of recycled aggregates.

Table 6: Compressive strength results for concrete with additive (MPa)

percentage of RA	Compressive strength at 7 days	Compressive strength at 28 days
25	18,33	24,3
50	17,47	21,45
75	14,5	20,47
100	14,3	19,65

We conclude that these tests adhere to the results of concrete without additives where compression decreases with increasing percentage of recycled aggregates.

3.6. Conclusion

The plasticizer (BV40) as well as this superplasticizer (Tempo 10 M) with a percentage of 1% improves the compressive strength: 18.05 for concrete with 50% recycled aggregates without addition, 22.63 with addition of BV40 and 21.45 with addition of Tempo 10 M.

4. Conclusion and Perspectives

Our article summarizes the results of experimental works conducted as part of a PhD research on the characterization of recycled aggregates.

The tests focus on the mechanical characterization and essentially the compression test, considered essential in any concrete study.

It was split in four steps: Test on ordinary concrete with various percentages of natural aggregates with and without adjuvant, tests on concrete with 50% recycled aggregates with different additives and with different percentages of admixtures.

The results show that:

- Compression decreases with increasing percentage of recycled.
- Up to 30% replacement of natural by recycled aggregates, the concrete is acceptable from a mechanical characterization point of view
- The addition of plasticizer and superplasticizer has improved the quality of the concrete. This can be explained by the fact that recycled aggregates require more water and these additives reduce this need.
- The optimum percentage of addition of the plasticizer is 1%

A review of the literature shows that one of the main methods to improve the quality of concrete is to improve the quality of aggregates by removing the adhered mortar, causing higher water absorption and chemical interaction with other concrete components as used with a customized low-cost simple treatment method [8]

To conclude, a concrete based on recycled aggregates could be used with a maximum release rate of 25-30% without addition and 50% with the addition of superplasticizer additives.

This study focused on mechanical characterization, the goal being an ordinary building plan under current conditions.

For more complicated projects and more developed concretes, more in-depth studies must be established based on:

- The quality of the aggregates to be used
- The aggregate / cement interaction
- The results of the mechanical tests

- The results of durability tests; these tests are essential for buildings with a lifespan of more than 50 years or buildings in environments liable to attack concrete.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M.R. Riaz, R.Hameed, M.Ilyas, A.Akram, Z.A. Siddiqi , "Mechanical Characterization of Recycled Aggregate Concrete" , Pakistan. Journal of Engineering & Applied Sciences. **16**, 25–32,2015.

[2] M. Malešev, V.Radonjanin, G.Broćeta, "properties of recycled aggregate concrete, Contemporary Materials", **2** , 239-249 ,2014, doi: 10.7251/COMEN1402239M .

[3] M. Chakradhara Rao, S.K. Bhattacharyya, S.V. Barai , "Influence of field recycled coarse aggregate on properties of concrete", Materials and Structures, **44**, 205-220, 2011, doi:10.1617/s11527-010-9620-x.

[4] J. Xiao,W. Li , Y.Fan, X. Huang , A, "overview of study on recycled aggregate concrete in China (1996-2011)", Construction and Building Materials, 31, 364-383 , 2012 , doi:10.1016/j.conbuildmat.2011.12.074.

[5] S.C. Kou, C.S.Poon, "Long-term mechanical and durability properties of recycled aggregate concrete prepared with the incorporation of fly ash", Cement and Concrete Composites, **37,** 12-19, 2013, doi:10.1016/j.cemconcomp.2012.12.011.

[6] C. Thomas, J.Setién, J.A.Polanco, P.Alaejos, M.Sanchez De Juan, "Durability of recycled aggregate concrete", Construction Building Materials, **40**, 1054-1065 ,2013, doi: 10.1016/j.conbuildmat.2012.11.106.

[7] H. Chbani, B. Saadouki, M. Boudlal, M. Barakat, "Formulation of Ordinary Concrete using the Dreux-Gorisse Method", Procedia Structural Integrity, **28**, 430-439, ,2020 doi: 10.1016/j.prostr.2020.10.050

[8] G. Dimitrou, P.Savva, M.F.Petrou , "Enhancing mechanical and durability properties of recycled aggregate concrete", Construction Building Materials, **158**, 228-235 , 2018, doi: 10.1016/j.conbuildmat.2017.09.13 .

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 36-42 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Anonymity Preserving Framework for Associating Personally Identifying Information with a Digital Wallet

Qazi Mudassar Ilyas*, Muhammad Mehboob Yasin

College of Computer Sciences and Information Technology, King Faisal University, Al Ahsa, 31982, Saudi Arabia

ARTICLE INFO

Article history:
Received: 23 October, 2020
Accepted: 19 December, 2020
Online: 10 January, 2021

Keywords:
Anonymity
Cryptocurrency
CryptoRegistry
Digital Wallet Ownership
Lost Digital Wallet

ABSTRACT

Growing adoption of cryptocurrencies by institutional investors coupled with recurring incidents of loss of access to digital wallets have resulted in emergence of custodial services for high net worth individuals and institutions. However, such services are not economically feasible for the wider community of crypto owners who are left with no recourse for recovery of lost funds. This paper proposes a framework, AFAP, for associating certain personally identifying information with a digital wallet so that a user may prove ownership of the said wallet. Thus, the owner is able to establish his/her claim over funds associated with such a wallet in case of loss of access to it. We show that the proposed scheme has no adverse effect on anonymity, privacy and forgetfulness of personal information of the wallet owner.

1. Introduction

Cryptocurrencies are gaining much traction in the mainstream economic activity because of their global and fully decentralized nature, promise of freedom from political influence, and peer to peer nature of transactions [1]. The blockchain does not carry any information pointing to the owner of a particular digital wallet. Further, the use of Hierarchical Deterministic (HD) wallets [2] makes it harder to even group individual transactions carried out by one individual. The pseudo-anonymity thus provided by blockchain is one of the major factors contributing to the popularity of digital currencies. However, a drawback of this characteristic is the inability to associate a digital wallet address with an individual in order to prove ownership of a lost digital wallet.

The value of Bitcoin, leader of the pack, recently climbed to an all-time high of USD 24,058 on CoinDesk Bitcoin Price Index beating the earlier high of USD 20,000 reached on December 2017 [3]. Several businesses have already embraced digital currencies, while governmental entities are also preparing to bring them into mainstream financial systems [4]. Hence, cryptocurrencies are on track to become an important part of all financial transactions in future. However, digital currencies are in their infancy and several issues need to be resolved before they can become part and parcel of a common individual's life [5].

*Corresponding Author: Qazi Mudassar Ilyas, College of CS&IT, King Faisal University, PO Box 400, Al Ahsa, 31982, Kingdom of Saudi Arabia, qilyas@kfu.edu.sa

The Achilles' heel of Cryptocurrencies is the fact that access to all funds owned by an individual is tied to a complex long number that is almost impossible to remember. Therefore, this complex number, known as private key of a digital wallet, needs to be stored safely. Also, the private key needs to be kept in a secure manner because anyone with knowledge of this number can use the funds stored in the respective wallet.

A private key can be lost owing to hardware or software failure, personal carelessness in safekeeping the keys, or in some cases even death of the owner. Several users have reported losing their private keys thus making their funds inaccessible to them. A survey reported that a significant fraction of respondents had experienced loss of Bitcoin keys at least once [6]. According to careful estimates, about 4 million Bitcoins had been lost till 2017 [7]. More recently, more than 100,000 individuals lost Bitcoins worth nearly USD 190 million owing to the death of the owner of a cryptocurrency exchange QuadrigaCX [8]. Since there is no inherent mechanism in the Bitcoin framework to recover lost keys, a lost private key simply means all Bitcoins associated with that key are lost forever. Hence, there is need for a mechanism to restore access to funds in case respective private keys become permanently inaccessible.

Since a mnemonic passphrase is easier to memorize or record as compared to a complex number, it is widely used by digital wallets to create the private key [9, 10]. Though, this improves the user experience, yet this does not absolve the user from

responsibility of storing her/his passphrase in a safe and secure manner. As in the case of loss of private keys, loss of a passphrase implies loss of access to the respective digital wallet. In the following, we address the problem of loss of access to a digital wallet which may be because of loss of a private key or loss of the corresponding mnemonic passphrase.

We propose a two-step solution to this problem. As the first step, the owner is afforded the ability to prove ownership of a lost wallet. Subsequently, either the funds stored in the lost wallet may be returned to the rightful owner or he/she may be suitably compensated. In this paper, a framework for proving ownership of a wallet is proposed, while the second step of solution is subject of future work. The proposed framework, "Anonymity Preserving Framework for Associating Personally Identifying Information (PII) with a Digital Wallet (AFAP)", can be used to associate the PII of owner with a digital wallet in a secure fashion while maintaining the same level of anonymity as provided by the current cryptocurrency ecosystem.

The proposed framework comprises three entities, namely CryptoRegistry, PII Associating Transaction (PAT) and PII Wallet. CryptoRegistry is a trusted long-lived entity that registers a given PII against a digital wallet by digitally signing this association and helps in validating the same when required to do so. A PII Associating Transaction (PAT) stores the aforementioned association on the blockchain by embedding the said CryptoRegistry signature in an outgoing transaction. A PII wallet is a specialized digital wallet that facilitates the registration and validation processes.

Rest of the paper is organized as follows. The next section summarizes the related works. The proposed framework is presented in Section 3 detailing key components and schemes for registration of a digital wallet and claiming its ownership. A case study is provided in Section 4 that shows feasibility of the proposed framework. The paper is concluded with the discussion and future work.

2. Related Work

The issue of loss of access to a digital wallet has gained attention of researchers owing to an increasing number of incidents resulting in loss of digital currency worth millions of dollars belonging to thousands of individuals [11,12]. The solutions proposed till date focus on safekeeping of digital wallets to ensure continued availability of private keys as and when needed. These solution approaches can be divided into two main categories, namely user-centric approaches that facilitate one in safekeeping of one's digital assets and custodial approaches that transfer the responsibility of safekeeping to a trusted third party.

User-centric approaches include cold storage, offline wallets and specialized hardware wallets. Cold storage refers to safely storing the private key of a digital wallet on an offline medium ranging from secondary storage to plain paper [13,14]. An offline wallet is an air-gapped device that can generate transactions for export to a hot wallet, thus providing an improved user experience. A specialized hardware wallet incorporates Hardware Security Modules (HSM) to store the private keys in a protected area of microcontroller storage which cannot be accessed from outside the device [15]. As suggested in [16], one can derive two values from the private key and store one of these value on a remote server. Protection against brute force password attacks is provided by keeping the second value on the local machine. A shortcoming of this scheme is that one needs to remember password to retrieve the value from the server. Further, damage to the local machine can result in permanent loss of access to the key.

Though all aforementioned approaches provide varying levels of protection against loss of private keys, yet the possibility of damage to or loss of access to physical medium still exists, which will result in loss of access to the respective digital wallet.

Another class of solutions relies on Shamir's secret sharing scheme, as adapted by [17], based on semi-trusted social network comprising friends of an individual. However, such schemes are more suitable for corporate clients since social networks of individuals are quite fluid and cooperation of "friends" cannot be guaranteed for unforeseeable future.

In custodial approaches, such as a browser-based wallet, an online account is maintained with a service provider that is used for storing and retrieving digital assets; examples include CoinBase*, Kraken† and BitStamp‡. However, several instances have been reported where many clients lost access to their digital assets owing to the service provider going out of business. For an exposè, the reader may refer to [18,19]. To overcome this issue, clients are usually facilitated to backup their private keys as an assurance against the possibility of the service provider ceasing to exist. Although these online services relieve the owner from the responsibility of safekeeping digital assets, they are exposed to a multitude of vulnerabilities such as internal or external theft and cyber-attacks. Recently, specialized custodial service providers have also emerged that claim improved security against aforementioned vulnerabilities by utilizing air-gapped devices and employing higher standards of security. Two of the well-known examples of such service providers are Fidelity Digital Assets§ and Ledger Vault** . However, these services have high associated costs and thus are suitable only for institutional clients or high net-worth individuals. It needs to be emphasized that privacy and anonymity are obviously compromised in all custodial services as digital assets in custody need to be associated with identity of respective owners.

One may conclude that none of the solutions proposed so far gives absolute guarantee against loss of access to a digital wallet while maintaining privacy and anonymity. Another perspective is to accept the occurrence of such eventuality as an inherent weakness of the blockchain technology and devise mechanisms to handle such situations. One such mechanism may be to prove ownership of a lost digital wallet and claim compensation for unspent amount stored in that wallet.

Various initiatives have been proposed for using blockchain to prove existence or ownership of a digital asset at a certain point in

* https://www.coinbase.com/
† https://www.kraken.com/
‡ https://www.bitstamp.net/
§ https://www.fidelitydigitalassets.com/
** https://www.ledger.com

time. For example, the online service named "Proof of Existence" [20] stores cryptographic hash of a digital document on blockchain that can be used to prove that the said document existed at a particular point in time. Further, integrity of the document is also established. Additionally, various schemes, referred to as Colored Coins [21], exploit the immutable characteristic of blockchain technology to store real world asset manipulation information on the blockchain [22, 23]. In [24] a DNS like naming system is proposed that associates human-readable names with respective internet resources by storing mapping information on the blockchain. In all such schemes, non-financial data is made part of the blockchain by embedding it into script of an OP_RETURN output; a special output of a blockchain transaction [25]. We also exploit the said scripting feature of OP_RETURN output to associate PII with a digital wallet by storing requisite metadata on the blockchain. The details of the proposed framework follow.

3. Proposed Framework

In order to overcome the wallet ownership issue, one may associate some Personally Identifying Information (PII), such as national ID, driving license or passport number with the digital wallet using the aforementioned output script of a transaction. In the following, we first introduce some obvious extensions of the use of aforementioned scripting feature for storing PII on the blockchain and point out that anonymity and privacy of the wallet owner may be compromised by using such simple schemes. We then present our framework to associate PII with a digital wallet and show that it preserves anonymity and privacy of the wallet owner.

3.1. Simple schemes for storing PII on blockchain

In this section, we present three ways of storing PII on blockchain to associate ownership of a digital wallet in the order of increasing complexity.

3.1.1. Plaintext PII

In this scheme, the owner of a digital wallet constructs a transaction and embeds his/her PII in one of the output scripts. This transaction is then sent on the network for processing. Since output script of a transaction can only be created by the originating digital wallet and cannot be tampered with after the transaction has been broadcast, the said information is irrevocably tied to the corresponding digital wallet. The wallet owner may record ID of the said transaction (TxID) for future reference. Once this transaction is confirmed, it becomes part of the blockchain, thus permanently associating PII with the issuing wallet.

In case of losing access to the private key of a digital wallet, henceforth referred to as a lost wallet, blockchain will be scanned to extract the aforementioned transaction using TxID and identify the rightful owner of the issuing wallet from the embedded PII. This scheme may be used to serve as a basic naming system for digital wallets akin to BitAlias [26] and OneName [27] etc. However, this scheme violates one of the basic principles of blockchain technologies, namely, ensuring anonymity of parties to a transaction. Therefore, this way of associating PII with a wallet may not be of general appeal.

3.1.2. Encrypted PII

To preserve anonymity of the wallet owner, one may create a public-private key pair and encrypt the PII with this public key before embedding it into the script. Now, to prove ownership of a digital wallet, one needs to locate the corresponding transaction on the blockchain and decrypt the PII using the respective private key. This requires the corresponding private key to be readily accessible when needed. This problem is akin to safely storing the private key, of the digital wallet, itself. Hence, this scheme does not reduce the responsibility of the wallet owner in safeguarding a private key.

An alternate way to relieve the owner of this added responsibility is to use pubic key of a well-known entity, such as a coin exchange, for encrypting the PII. Now, one needs to approach the said entity for decrypting the PII in order to establish one's claim of ownership. Though a wallet owner is absolved from responsibility of safekeeping any private keys, this scheme suffers from two drawbacks.

Firstly, the well-known entity may not be aware of such use of its public key. Also, it is not bound to maintain all its public private key pairs for foreseeable future. Therefore, it may not be in possession of the required private key when approached. Hence, owner claim may not be substantiated based on the encrypted PII. Secondly, since public key is supposed to be widely available, the encrypted PII is prone to dictionary attack.

3.1.3. Hashed PII

This method requires PII to be hashed using a well-known hashing algorithm and the resulting hash to be embedded in the script of a transaction in blockchain. To establish the claim of ownership, one may create hash of one's PII and validate it against the hash recovered from the respective transaction. In order to avoid dictionary attack, a sufficiently large random, commonly called salt, may be added to the PII. However, PII validation process now depends upon availability of the salt used. Therefore, the wallet owner is made to bear the responsibility of safekeeping the salt. Though, the salt may be in the form of hash of a memorable passphrase, the situation is reduced to a problem that is akin to safely storing the private key of a digital wallet.

3.2. An Anonymity Preserving Framework for Associating PII with a Digital Wallet

By now, the need has been established for a long-lived trusted third party with the responsibility of safekeeping the information required for proving association of a given PII with a digital wallet in an anonymous way. We call this entity a CryptoRegistry and propose a framework that preserves owner's anonymity and privacy as long as the CryptoRegistry is deemed trustworthy. We call this framework AFAP - An Anonymity Preserving Framework for Associating Personally Identifiable Information (PII) with a Digital Wallet.

The proposed AFAP framework comprises three components, namely a CryptoRegistry, a PII Associating Transaction (PAT) and a PII wallet. In the following, we give the details of each one of these components.

3.2.1. CryptoRegistry

We introduce the concept of a CryptoRegistry that is a trusted long-lived entity. A CryptoRegistry is required to perform the following functions:

- To provide an Application Programmer Interface (API) for registering the PII provided by a wallet owner
- To provide an API for making available the following information:
  - A public key $K_{Pub,CR}$
  - Registration fee, F_{CR}
  - A wallet address, W_{CR}
  - A sufficiently large random number, also known as salt, for defense against dictionary attack as discussed below, S_{CR}
  - A hashing algorithm along with its parameters, $\mathcal{H}$
  - Hash of (PII || S_{CR}), H_{ID}
  - Digital signature on PII, S_{CR}, F_{CR}, W_{CR}, $\mathcal{H}$ and H_{ID} using $K_{Priv,CR}$
  - The digital certificate to the effect that the said CryptoRegistry is owner of the respective public key, $K_{Pub,CR}$

It may be noted that the proposed scheme introduces a random salt to preserve anonymity since use of hash of PII alone may be linked to the original PII by brute force attack.

- To ensure availability of the PII, S_{CR}, $TxID_{PAT}$, and respective public-private key pair, once a PAT constructed by a PII wallet paying the fee F_{CR} to the wallet W_{CR} is confirmed on the blockchain. By virtue of its digital signature, the CryptoRegistry is bound to the PAT. As a trustee, the CryptoRegistry is also required to ensure confidentiality and security of information entrusted with it throughout its life time.
- To provide a list of PATs and corresponding digital wallet addresses associated with a given PII. It is to be noted that, in the interest of anonymity and privacy, the aforementioned information will only be provided to the rightful owner of PII or any authorized legal entity.
- Optionally, a CryptoRegistry may advertise special prefix to be used with OP_RETURN to facilitate searching a particular PAT on the blockchain.

3.2.2. PII Associating Transaction (PAT)

Generally a PII Associating Transaction (PAT) consists of three outputs. Firstly, an output transferring Registration fee F_{CR} to CryptoRegistry wallet W_{CR}. Secondly, an output transferring change to the wallet owner's change address. Thirdly, a zero valued output with OP code OP_RETURN and output script comprising the following information:

- Special prefix provided by the CryptoRegistry (if any)
- $\mathcal{H}$(PII || S_{CR}), H_{ID}
- Digital signature on PII, S_{CR} ,F_{CR}, W_{CR} and H_{ID} using $K_{Pub,CR}$

It is to be noted that since every Bitcoin transaction is tied to the originating wallet and cannot be tampered with after it has been confirmed, thus the above information, by virtue of being embedded in the transaction, becomes irrevocably tied to the corresponding digital wallet.

3.2.3. PII Wallet

PII wallet is a specialized digital wallet capable of constructing and verifying PAT transactions for associating itself with PII. For this purpose, it needs to communicate with a CryptoRegistry and perform a number of functions as detailed below:

- Accepting and storing PII of the wallet owner in a secure way
- Providing a list of available CryptoRegistries to the wallet owner
- Initiating the process for registration of given PII with the selected CryptoRegistry
- Accepting response from CtyptoRegistry comprising S_{CR}, F_{CR}, W_{CR}, H_{ID} , a hashing algorithm $\mathcal{H}$ and a digital certificate
- Verifying digital signature and hash value, H_{ID}
- Constructing PAT as described above and transmitting to the network
- Sending H_{ID} and transaction ID of PAT to CryptoRegistry after confirmation of PAT
- Periodically querying the said CryptoRegistry for confirmation of H_{ID} registration in its database
- When required, querying all CryptoRegistries for registration status of provided PII
- Displaying to the user a list of CryptoRegistries with whom provided PII is registered

3.2.4. Associating PII with a digital wallet

The sequence of steps for the purpose of registering a PII of the owner of a digital wallet is depicted in Figure 1.

The details of each step are as follows:

1. The user provides his/her PII to the wallet.
2. The wallet sends the provided PII for registration with the selected CryptoRegistry.
3. The CryptoRegistry selects a wallet (W_{CR}) for receiving registration fee (F_{CR}), chooses a random salt (S_{CR}) and calculates H_{ID} which is hash of PII concatenated with S_{CR}. It saves this information in its database for future reference with registration status as 'Not Confirmed'.
4. The CryptoRegistry sends PII, S_{CR}, F_{CR}, W_{CR} and H_{ID} , its signature on this information and its corresponding digital certificate to the PII wallet.
5. The wallet creates a PII Associating Transaction (PAT) using the received information and transmits it to the Bitcoin network.
6. The wallet periodically queries the blockchain for confirmation of PAT.

7. Once PAT is confirmed, the wallet notifies the CryptoRegistry of the confirmation by sending transaction ID of PAT ($TxID_{PAT}$) and H_{ID}
8. The CryptoRegistry queries the blockchain for confirmation of $TxID_{PAT}$
9. Upon confirmation, The CryptoRegistry updates PII registration status as 'Confirmed' and records corresponding $TxID_{PAT}$
10. The wallet periodically queries the CryptoRegistry for PII registration status
11. Upon confirmation, the wallet notifies the user that the provided PII has been successfully registered with the selected CryptoRegistry and records $TxID_{PAT}$ for future reference.

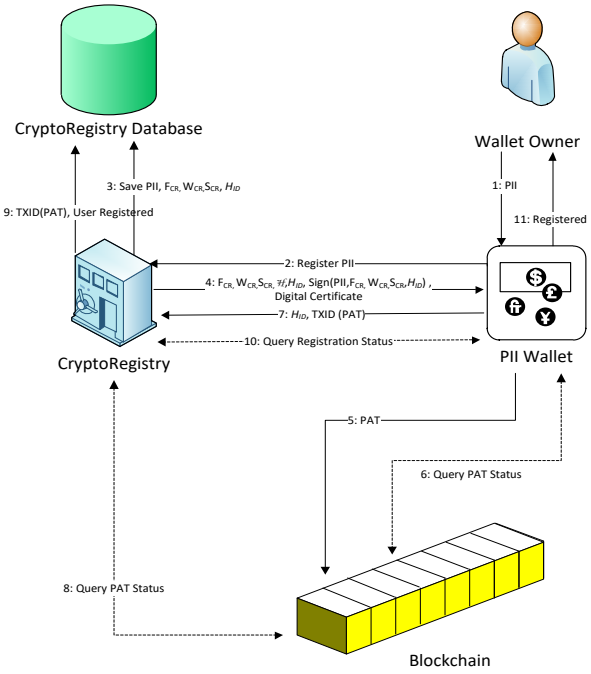


Figure 1: PII Registration in AFAP Framework

3.2.5. *Claiming ownership of a lost PII Wallet*

In the event of loss of a PII wallet, a PII wallet is installed and its recovery mode is initiated. The following sequence of events may follow (see Figure 2).

- The user enters her/his PII used for querying the registration status of the lost wallet.
- The wallet queries all known CryptoRegistries one by one using the provided PII.
- PII wallet receives response from the CryptoRegistries and displays the list of CryptoRegistries with affirmative response.

In the interest of owner's anonymity, the CryptoRegistry shall not reveal any information other than whether the subject PII is registered with it or not. The said CryptoRegistry may now be approached by the owner to obtain the transaction ID of the corresponding PAT along with the salt.

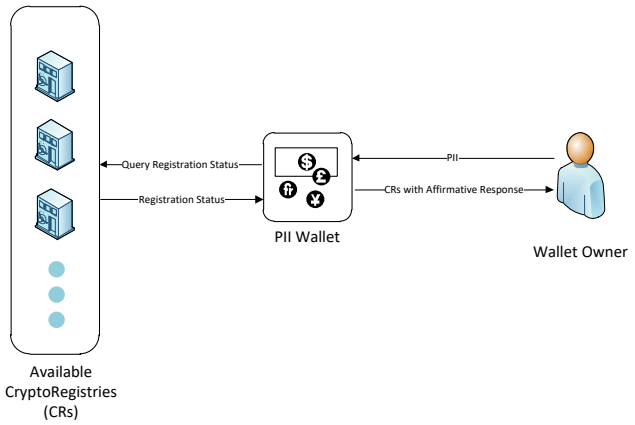


Figure 2: Claiming ownership of a lost wallet

The hash of given PII concatenated with salt is computed and matched against the one extracted from PAT. A successful match proves that the said PII was used to construct the PAT and that the wallet from which this PAT was broadcast is the wallet associated with the said PII. This way the ownership of the wallet, and any unspent funds associated with it, is established.

3.2.6. *Right to be forgotten*

One of the important principles incorporated into General Data Protection Regulation (GDPR) is a user's right to be forgotten (RtbF) allowing for retroactive erasure of all personal data. Considering the immutable nature of blockchain, it may appear that storing association of PII with a digital wallet may be in violation of the RtbF. However, the said right is safeguarded on two counts. Firstly, PII itself is not stored on the blockchain in any way, i.e., neither in plain text nor in encrypted form. Secondly, proving association of a particular PII with a digital wallet requires knowledge of the corresponding salt, which is a cryptographically strong random number. Hence, if the salt is forgotten, this association becomes unprovable. The right to forget this salt and any associated personal data can be exercised by the user with a request to CryptoRegistry. In [28], it has been shown that integration of such requirements into the current computing infrastructures is practically feasible.

4. Case study

In order to validate the proposed framework, a sample PAT was constructed by simulating interaction between a client and a typical CryptoRegistry as detailed below. The resulting transaction was successfully broadcast on Bitcoin Testnet. This was followed by simulating the situation of loss of this wallet. The lost wallet address was successfully retrieved from the blockchain by presenting only the PII. The following is the detail of interactions among the various components of the proposed AFAP framework.

4.1. *PII Association with a digital wallet*

The example client wishes to associate his PII "ABC1234567" with his digital wallet. The CryptoRegistry generates a salt, calculates hash of PII concatenated with this salt and returns the following information to the client along with his signature on the same:

CryptoRegistry Prefix = TESTPII

Salt: S_{CR} = 9586311452

SHA1 (PII || Salt): H_{ID} = d49cdc68c804114b402bbaa108d324a80a19e354

Registration Fee: F_{CR}= 5mBTC

Wallet address for receiving registration fee:

W_{CR} = mpvkKZ6ysUYEGZuXR1FYRCKzykvMmtovyv

Signature:

Sign(PII,F_{CR}, W_{CR},S_{CR},H_{ID}) = AAwPRuNFlPvKCaic12aFRLYXf9ZbytXjBowXeKZm0kGNETcplSyEE268401jgS8=

The client constructs the PAT with three outputs. Firstly, paying registration fee to W_{CR}, secondly the change to be returned to itself and finally a null data transaction with the script comprising prefix, hash and signature of the CryptoRegistry. The resulting script is given below:

Script: TESTPIId49cdc68c804114b402bbaa108d324a80a19e354AAwPRuNFlPvKCaic12aFRLYXf9ZbytXjBowXeKZm0kGNETcplSyEE268401jgS8=

This example transaction was broadcast on Bitcoin Testnet and can be found with the transaction id "985b15d99baf20facc67f1f925e51b7dd49b34dcbdc0a64bc0607adc9100e0d7". Figure 3 shows the details of the said transaction as depicted by Bitcoin Block Explorer.

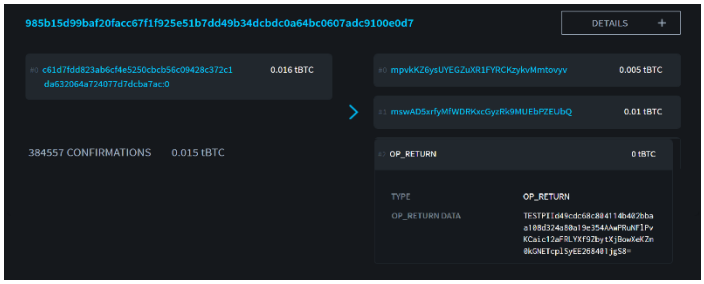


Figure 3: Details of PAT as shown by Bitcoin Block Explorer

It may be noted that the CryptoRegistry is required to store the transaction ID of PAT and other associated information in its database, as depicted in Figure 1.

4.2. Claiming ownership of a digital wallet

Assuming the owner knows the transaction ID of PAT and the salt used to construct the script of PAT, firstly PAT may be retrieved from blockchain using transaction ID. Then, the PII association may be proven by calculating hash of PII concatenated with salt and validating it against the script of null data output of PAT.

Since the main purpose of the proposed AFAP framework is to relieve the wallet owner from maintaining access to the private key of a digital wallet in a secure, safe and anonymity preserving manner, the owner is not assumed to maintain transaction ID of PAT and the respective salt. In such a case, the owner may present her/his PII to the CryptoRegistry to confirm association of the said PII with a digital wallet.

5. Discussion

The proposed AFAP framework uses blockchain itself to store the association of the Personally Identifying Information (PII) of the owner with her/his digital wallet. The proposed framework requires a persistent entity, called CryptoRegistry, which is mandated to maintain record of transaction IDs of PII Associating Transactions (PATs) and associated information in a secure way. The proof of ownership of a digital wallet lies in hash of PII and a random number, called salt, stored in a PAT.

In case of loss of a digital wallet, the owner may query available CryptoRegistries to confirm registration of her/his PII against a PAT from their databases. Once a confirmation is received, the owner may prove her/his identity to the satisfaction of the respective CryptoRegistry in order to retrieve transaction ID of PAT and respective salt. The hash extracted from PAT validates the PII and ownership of legal claimant of the PII is established. The proposed scheme, therefore, depends on long term availability of the CryptoRegistry used to register a digital wallet. In the interest of anonymity and privacy of the wallet owner, the CryptoRegistry may require due legal process in identifying the claimant before releasing transaction ID of PAT and respective salt.

From the aforementioned requirements of a CryptoRegistry, one may draw parallel with crypto exchanges or custodial services as trusted third parties for safekeeping of users' crypto assets. However, it may be noted that in the latter case, clients' funds are stored in the wallets owned by the trusted third party functioning as a custodian. Thus a user may lose all his/her funds in the event of hacking or bankruptcy of the custodian. In contrast, in the event of bankruptcy of a CryptoRegistry, the hash values stored in respective PATs cannot be verified owing to unavailability of corresponding salts, whereas the user funds remain intact. A user may safeguard against this eventuality by choosing to register with multiple CryptoRegistries. Though, the cost of registration may discourage registration with multiple CryptoRegisteries, it is expected that existing crypto exchanges will offer CryptoRegistry services and they may waive off these fees to increase their customer base.

The proposed framework brings the cryptocurrency ecosystem a step closer to the conventional banking system where an account holder may prove ownership of funds after losing access to corresponding checkbook or card. It may be noted that the protection afforded to an individual by AFAP is stronger than that of traditional banks since the funds and proof of ownership are stored in an immutable blockchain and there is no risk akin to bankruptcy.

6. Future Work

Although a user may claim ownership of her/his lost funds in the above manner, the process for recovery/compensation may involve legal procedures. One possible way of compensating the owner of a lost wallet is to create a compensation fund for such eventualities. Alternatively, a special transaction to credit the required funds to claimant's new wallet akin to the mechanism proposed for managing deflation of cryptocurrencies by [29]. Further, similar to the expiration flag proposed by them, the old wallet needs to be flagged as invalid in a way that blocks any future access to it or funds associated with it.

References

[1] Capgemini, BNP Paribas, "World Payments Report 2018," 55, 2018.
[2] V. Buterin, "Deterministic Wallets, Their Advantages and their Understated Flaws," Bitcoin Magazine, 2013.
[3] D.Z. Morris, "Bitcoin Hits a New Record High, But Stops Short of $20,000

| Fortune," Fortune, 2017.
[4] R. Krygier, "Venezuela launches the 'petro' its cryptocurrency," The Washington Post, 2018.
[5] M. Raskin, D. Yermack, Digital currencies, decentralized ledgers and the future of central banking, 2018, doi:10.4337/9781784719227.00028.
[6] K. Krombholz, A. Judmayer, M. Gusenbauer, E. Weippl, "The other side of the coin: User experiences with bitcoin security and privacy," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017, doi:10.1007/978-3-662-54970-4_33.
[7] J.J. Roberts, N. Rapp, "Exclusive: Nearly 4 Million Bitcoins Lost Forever, New Study Says," Fortune, 2017.
[8] T. Bult, "Thesis - Security analysis of blockchain technology Security analysis of blockchain technology."
[9] M. Palatinus, P. Rusnak, A. Voisine, S. Bowe, (bip-0039) Mnemonic code for generating deterministic keys, GitHub, 2013.
[10] Y. Liu, R. Li, X. Liu, J. Wang, L. Zhang, C. Tang, H. Kang, "An efficient method to enhance Bitcoin wallet security," in Proceedings of the International Conference on Anti-Counterfeiting, Security and Identification, ASID, 201-207, 2018, doi:10.1109/ICASID.2017.8285737.
[11] M. Conti, K.E. Sandeep, C. Lal, S. Ruj, "A survey on security and privacy issues of bitcoin," IEEE Communications Surveys and Tutorials, 2018, doi:10.1109/COMST.2018.2842460.
[12] U. Mukhopadhyay, A. Skjellum, O. Hambolu, J. Oakley, L. Yu, R. Brooks, "A brief survey of Cryptocurrency systems," in 2016 14th Annual Conference on Privacy, Security and Trust, PST 2016, 1-5, 2016, doi:10.1109/PST.2016.7906988.
[13] S. Eskandari, D. Barrera, E. Stobert, J. Clark, A first look at the usability of bitcoin key management, ArXiv, 2018, doi:10.14722/usec.2015.23015.
[14] R.N. Akram, K. Markantonakis, D. Sauveron, "Recovering from a lost digital wallet: A smart cards perspective extended abstract," Pervasive and Mobile Computing, 2016, doi:10.1016/j.pmcj.2015.06.018.
[15] O. Boireau, "Securing the blockchain against hackers," Network Security, 2018, doi:10.1016/S1353-4858(18)30006-0.
[16] E.M.D.W. Brickell, SECURE STORAGE OF PRIVATE KEYS, United States, 2005.
[17] S. He, Q. Wu, X. Luo, Z. Liang, D. Li, H. Feng, H. Zheng, Y. Li, "A Social-Network-Based Cryptocurrency Wallet-Management Scheme," IEEE Access, 2018, doi:10.1109/ACCESS.2018.2799385.
[18] L.J. Trautman, "Virtual Currencies: Bitcoin & What Now after Liberty Reserve and Silk Road?," SSRN Electronic Journal, 2014, doi:10.2139/ssrn.2393537.
[19] U. Chohan, "The Problems of Cryptocurrency Thefts and Exchange Shutdowns," SSRN Electronic Journal, 2018, doi:10.2139/ssrn.3131702.
[20] M. Crosby, Nachiappan, P. Pattanayak, S. Verma, V. Kalyanaraman, "Blockchain Technology - BEYOND BITCOIN," Berkley Engineering, 2016.
[21] M. Rosenfeld, "Overview of colored coins," 13, 2012.
[22] L. Bell, W.J. Buchanan, J. Cameron, O. Lo, "Applications of Blockchain Within Healthcare," Blockchain in Healthcare Today, 1-7, 2018, doi:10.30953/bhty.v1.8.
[23] H. Min, "Blockchain technology for enhancing supply chain resilience," Business Horizons, 2019, doi:10.1016/j.bushor.2018.08.012.
[24] M. Ali, J. Nelson, R. Shea, M.J. Freedman, "Bootstrapping Trust in Distributed Systems with Blockchains," USENIX ;Login:, 2016.
[25] M. Bartoletti, L. Pompianu, "An analysis of bitcoin OP_RETURN metadata," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017, doi:10.1007/978-3-319-70278-0_14.
[26] Y.G. Malahov, "BitAlias I Aka Usernames for Bitcoin," Medium, 2015.
[27] H. Kalodner, M. Carlsten, P. Ellenbogen, J. Bonneau, A. Narayanan, "An empirical study of Namecoin and lessons for decentralized namespace design," 14th Annual Workshop on the Economics of Information Security (WEIS), 2015.
[28] E. Politou, E. Alepis, C. Patsakis, Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions, Journal of Cybersecurity, **4**(1), tyy001, 2018, doi:10.1093/cybsec/tyy001.
[29] H. Gjermundrød, I. Dionysiou, "Recirculating lost coins in cryptocurrency systems," Lecture Notes in Business Information Processing, 229-240, 2014, doi:10.1007/978-3-319-11460-6_20.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 36-42 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Super Resolution Based Underwater Image Enhancement by Illumination Adjustment and Color Correction with Fusion Technique

Md. Ashfaqul Islam[*,1], Maisha Hasnin[2], Nayeem Iftakhar[3], Md. Mushfiqur Rahman[4]

[1]*Department of Electrical and Electronic Engineering, Bangladesh Army University of Engineering and Technology (BAUET), Natore, 6431, Bangladesh*

[2]*Department of Information and Communication Technology, Bangladesh University of Professional (BUP), Dhaka, 1216, Bangladesh*

[3]*Department of Computer Science and Engineering, Bangladesh Army University of Engineering and Technology (BAUET), Natore, 6431, Bangladesh*

[4]*Department of Computer Science and Engineering, Rajshahi University of Engineering and Technology (RUET), Rajshahi, 6204, Bangladesh*

ARTICLE INFO

Article history:
Received: 08 November, 2020
Accepted: 17 February, 2021
Online: 10 March, 2021

Keywords:
Underwater Image
Illumination Adjustment
Color Correction
Fusion
Neural Network
Super Resolution

ABSTRACT

In underwater photographs are look like low-quality images, the main reason is behind that due to attenuation of the propagated light, absorption and scattering effect. The absorption significantly reduces the light energy, while the dispersion causes changes in the light propagation path. They result in foggy appearance and degradation of contrast, causing misty distant objects. So, for getting the most effective result from that type of image, there must be an enhancement technique that has to be applied. We propose an efficient technique to enhance the images captured underwater by applying a fusion-based technique using super-resolution. For enhancing images, we have followed two steps. The first one illumination adjustment and another one is color correction. Then fusion technique is applied to the resultant image from illumination adjustment and color correction as two inputs and combined them with their maximum coefficient value and received output from there. After that, on the fused output image, we used the Super-Resolution method. In the Super-resolution procedure, low resolution and high-resolution images are used then a bi-cubic interpolation algorithm and finally, VDSR (very-deep super-resolution) neural network has been used to get the most effective result from an obscure underwater image. For getting the most effective result from an obscure image, a new high-quality and efficient image enhancement method has been proposed in this paper.

1. Introduction

Under water photographs have been commonly used in marine life under water studies in recent years. Underwater photography has jointly been a primary objective of concern in various technology and science branches, such as scrutiny of underwater infrastructures and cables, object tracking, underwater vessel control, marine biology review, and archeology [1]. Underwater images are primarily depicted by their poor visibility as a result of the light is exponentially weakened within the underwater atmosphere [2]. The entities at a distance more than ten meters are almost undetectable in common seawater pictures; the colors get faded [1]. Water absorbs light, reducing light rays' energy, resulting in image under-exposure [3]. The reasoning for the color cast is that there is a differing absorption rate for the varying wavelengths of light [4]. The key challenges, including color casting, fuzz, and under-exposure, can be faced by underwater imaging processing. It ought to be handled to recoup the comparatively true color and normal aberrance to make the collected photographs more appropriate for observation [5]. This can be done by either image restoration or image enhancement

[*]Corresponding Author Name: Md. Ashfaqul Islam, Bangladesh Army University of Engineering and Technology (BAUET), Natore, 6431, Bangladesh, +8801723859090 & niloy064@gmail.com

procedures [6]. For enhancing the underwater images, it is the most important part to develop some algorithms like illumination adjust, color correction, fusion technique & restoration process to get the most effective result from a nun-uniform image. We do super-resolution by revising the arbitrary visual appearance of the image by generating higher images from low-resolution pictures [4].

2. Previous Research Work

To solve the obscurity of a photograph shot in an underwater setting, a lot of work has already been completed. Three big problems occur in underwater photography, including color casting, under-exposure, and fuzz. In [1], the author proposes a new method formed by two components: correction of color and illumination modification. In order to overcome the color cast, they used an appropriate color enhancement technique and then followed the Retinex model to make the illumination adjustment by successively applying a gamma correction on and removing the illumination map. In order to eliminate the haze in underwater photographs that do not require advanced hardware or information about the underwater environments or scene composition. Their approach relies on the fusion of multiple inputs, but by contrast correcting and enhancing the sharpness of a white-balanced version of a single native input image, they derive the two inputs to merge [7]. Eliminating the color cast by white balancing to create a realistic appearance of the underwater pictures, and the multi-scale application of the method of fusion results in an artifact-free blending. To dramatically upgrade the color of underwater images, the author used contrast stretching and HSI color space [8]. Again, author of [1] implemented a fusion approach that used the global min-max windowing strategy for maximizing contrast. In [6], the author order to distinguish less contrast, researchers have used histogram expanding technique and CLAHE for color correction and contrast enhancement techniques. The unsupervised color correction method is applied by the author of [9] to increase the resolution of a low-quality picture.

For increasing the accuracy of the images of underwater, we are doing this because it has been used in many research sectors in recent years. By reviewing some previous works, we can see that they have separately worked with different techniques, but in our work, we have combined the different techniques - illumination adjustment and color correction with fusion technique and super resolution for the better output than others.

In our article, we have worked on the various procedures for enhancing an underwater image. There we have worked on illumination adjustment, color correction first to enhance the image. Then the fusion technique is applied to the modified image we get by balancing the illumination and correcting the color, and finally, we have used the super-resolution process on fused output image to get the best effective result from the input image.

3. Proposed Architecture

Basically, there are four sections in the proposed architecture. In the first section, we have described the illumination adjustment technique. In section two we have described the color correction part. In section three, the fusion technique has been described and finally, we have briefly described the super-resolution process in section four.

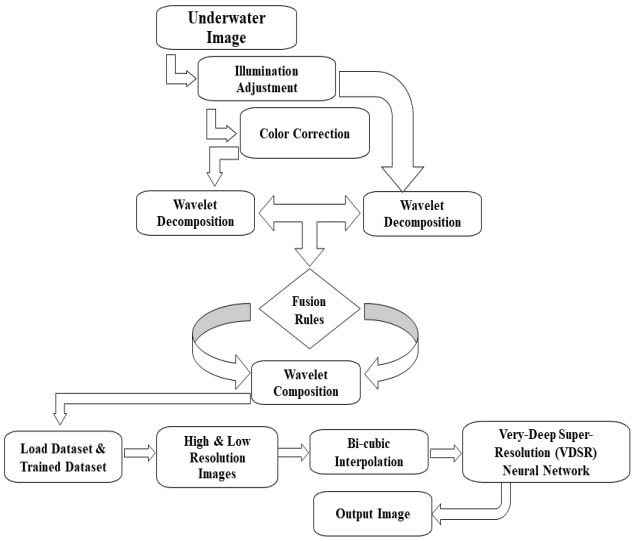


Figure 1: Proposed System Architecture.

3.1. Illumination Adjustment

Illumination adjustment basically is to falling the light on objects and making them visible as they are illuminated, it happens for scaling up the light in the water. Explaining M×N underwater images illumination of a by illuminant is possible, etc. [6]

This is our input image 1 we are working on.



(1)

In the illumination adjustment process for removing illumination, firstly there we have used the grayscale algorithm to convert the original image into a grayscale image.



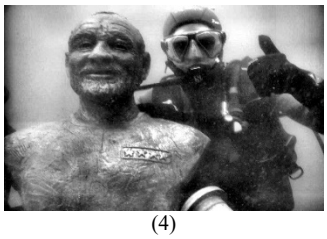
(2)

We then used the Adaptive Histogram Equalization on the grayscale image to control the local contrast of the pixels from the adjoining area, essentially it is used to increase the image contrast.



(3)

In order to fix the image luminance, we subtracted the background from the Adaptive Histogram image and applied the gamma correction algorithm to the image. After that we used median filter on the image, it is generally used for removing noise (unwanted pixel) from the image. The total progression, from background subtracting to image filtering, is repeated twice (three times in total) to change back to an image's RGB color space.



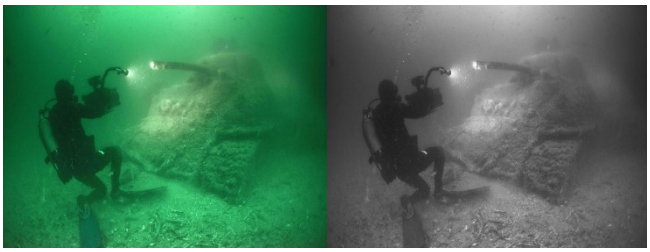
(4)

After that, the three results we got from filtered images are correlated with the processing from the illumination grayscale image of the illumination attuned RGB image.

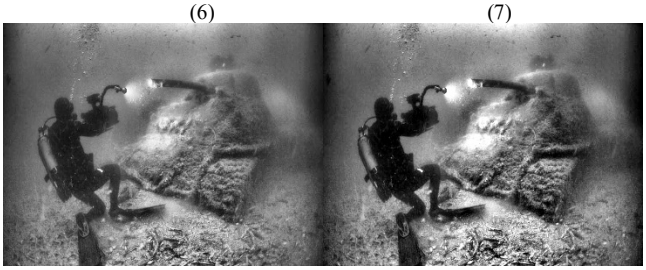


(5)

Here we have shown the same procedure for a another input image for better understanding.



(6) (7)



(8) (9)



(10)

Figure 2: (1) Input image 1, (2) Gray scale conversion of input image 1, (3) Adaptive Histogram Equalized image 1, (4) Illumination adjusted gray scale image 1, (5) Illumination adjusted output image 1, (6) Input image 2, (7) Gray scale conversion of input image 2, (8) Adaptive Histogram Equalized image 2, (9) Illumination adjusted gray scale image 2 (10) Illumination adjusted output image 2

3.2. Color Correction

In the procedure of color correction, the illumination adjusted images' output is used as the input image of color correction.



(1)

Images are influenced by low contrast because of dispersion in the underwater environment. Firstly, the RGB space of the illumination adjusted image is converted into HSV (Hue Saturation Value).



(2)

Then HSV is prepared to point up the maximum visible portion of the output we have with max luminance. It is a useful technique that is used to increase the contrast of an image for balancing image intensity after Histogram Equalization is used on the processed image.



(3)

However, sometimes demolitions disclose parts and present a nonuniform result. For this reason, here we use Adaptive Histogram Equalization on the Histogram Equalized image of the input image of the color enrichment. Adaptive Histogram is the up-gradation process of Histogram Equalization.



(4)

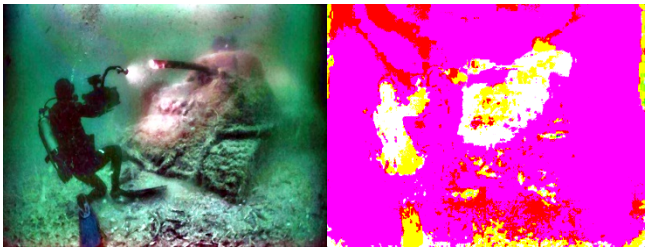
Adaptive Histogram Equalization operates by contemplating only a small component and produces a contrast enhancement of

this part based on the surrounding CDF. Histogram Equalization technique is performed to measure the intensity of the individual structure of the image, this. and especially changes each RGB channels color.



(5)

The total color correction process for input image 2.



(6) (7)



(8) (9)



(10)

Figure 3: (1) Contrast enhancement (Illumination adjusted image) Input image 1, (2) HSV image for image 1, (3) Histogram equalized image for image 1, (4) AHE image for image 1, (5) Contrast enhanced image for image 1, (6) Input image 2 for contrast enhancement, (7) HSV image for image 2, (8) Histogram equalized image for image 2, (9) AHE image for image 2, (10) Contrast enhanced image for image 2.

3.3. Fusion Procedure

Combination procedure fundamentally may be a method of combining apocope data from two or more images into a single image, the yield image will represent more data than the input image. Within the Combination method, Wavelet Changes are fundamentally the developments of the plausibility of Tall pass sifting. The way towards applying the Discrete Wavelet Transform (DWT) can be talked to as a bank of channels. At respectively degree of decay, the flag is separated into tall recurrence and low-frequency elements, and the low-frequency portion can be moreover decayed till we get the perfect determination. The Wavelet change is deliberate to urge awesome recurrence determination for low-frequency elements and tall common determination for high-frequency components.

Here two input images are taken. Firstly, we took illumination adjusted images output and secondly, we took the color corrected image as input. Then we separated the high-frequency and low-frequency elements from the image. Whereas the wavelet transforms both dimension of is used. Firstly, the high pass and low pass filter used on the rows, after that it is used on the columns. At a point it provides horizontal approximation value when a low pass filter is used on the rows and we get horizontal details by using high pass filter. The sub-signal has the highest frequency equivalent to half of the first, created from the low pass filter on the columns of the horizontal approximation and horizontal information, low pass filters and high pass filters are once again used, carrying around four sub-images.[10]: rough picture, horizontal detail, vertical detail, and diagonal detail separately. In figure 4 the one level decomposition of the input picture has shown [1].

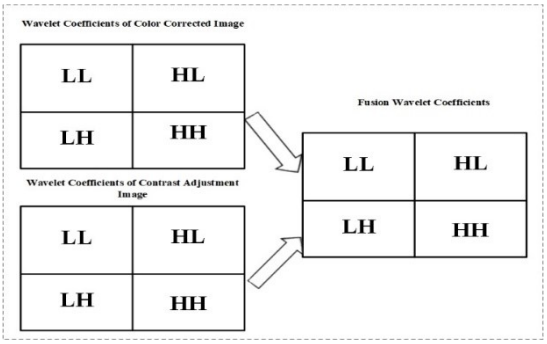


Figure 4: 2D wavelet-based image decomposition of level two.

The pictures are used as an input for their coefficient to be decomposed and composed inversely. Then the wavelet transform is used separately to get the modified image as a final image. Up sampling completes the decomposition of wavelet transformation, and down sampling is used for the converse composition.

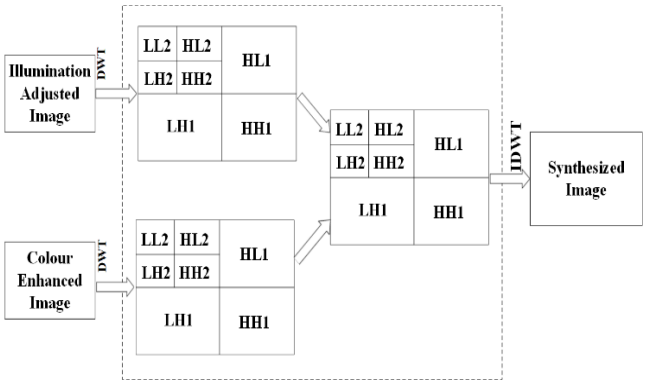
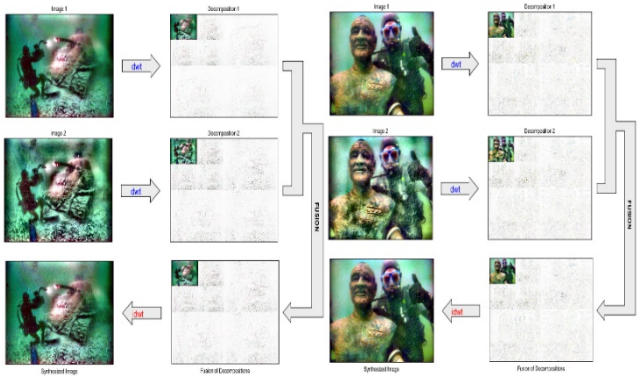


Figure 5: 2D decomposition of level two, fusion of coefficients and Output Image

The output images of illumination adjusted image and the color-enhanced image are used as the input for DWT decomposition in figure 5. From there, no opposite change is applied to the coefficients of its decomposition and fused utilizing the maximum estimations of the coefficients.



(1) (2)

Figure 6: 2D decomposition of level two, fusion of coefficients and synthesized image (1) & (2) of the proposed method



(1) (2)

Figure 7: (1) Fusion Output 1 (2) Fusion Output 2

3.4. Super Resolution

Super-resolution could be a method of imaging that progresses the resolution of an imaging framework, additionally, it could be a prepare of combining a sequence of low-resolution pictures to create the next determination picture or arrangement. In arrange to demonstrate the execution of this strategy, it was being utilized on the stacked dataset(images), compared with diverse strategies there had been prepared the taking after stack dataset(images) utilizing code.

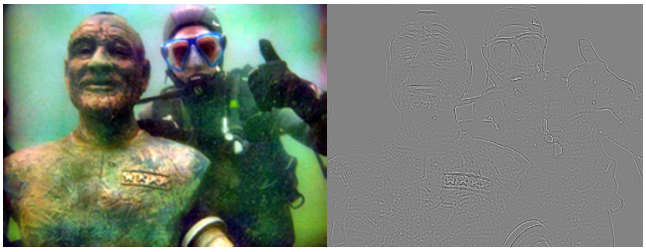
After applying the bi-cubic interpolation procedure, the VDSR (Very-Deep Super Resolution) method has been applied to the image. VDSR refers to a convolutional neural network engineering intended to perform single image super-resolution [11]. The VDSR learns the retailing among low-and high-goals images. This retailing is conceivable on the grounds that low-resolutions and high-resolution images have comparative image content and vary principally in high-recurrence subtleties. VDSR utilizes a leftover learning methodology, implying that the system figures out how to evaluate a lingering image. With regards to super-resolution, a leftover image is a distinction between a high-resolution reference picture and a low-resolution image that has been upscaled utilizing bi-cubic interpolation to coordinate the size of the allusion image. A remaining image holds data of high-frequency subtleties of an image. The residual image from the luminance of a color image is found by the VDSR network. VDSR is prepared to utilize just the luminance channel since human understanding is more sensitive to alternate in brilliance than to alternate in color. For a scale factor, the size of the reference image to the size of the low-resolution image is expected. As the scale factor expands, SISR turns out to be all the more not well presented in light of the fact that the low-goals image loses more data about the high-frequency image content. By utilizing a huge responsive field VDSR takes care of this issue. This model trains a VDSR network with different scale factors utilizing scale augmentation. Scale augmentation improves the outcomes at bigger scope factors on the grounds that the network can exploit the image context from littler scale factors. Also, the VDSR system can sum up to acknowledge images with inter-scale factors.

4. Result and Discussion

The results are described as follows;



(1) (2)



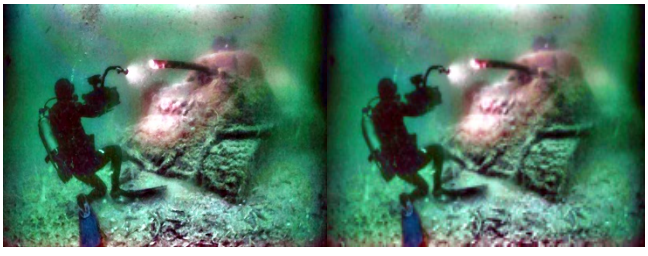
(3) (4)



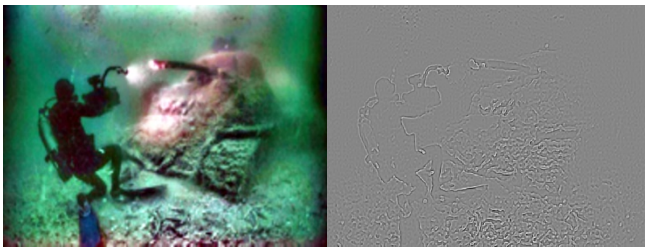
(5) (6)



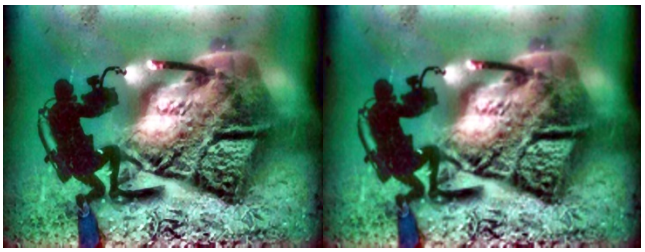
(7)



(8) (9)



(10) (11)



(12) (13)

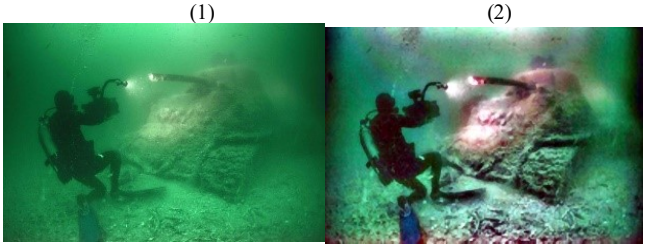


(14)

Figure 8: (1) Fusion output of input image 1, (2) Low Resolution input image 1 (3) Bi-Cubic Interpolation of input image 1, (4) Residual Image from VDSR of input image 1 (5) High-Resolution Image of input image 1 Obtained Using VDSR (6) High-Resolution Results of input image 1 using Bicubic Interpolation (7) Final Output of input image 1, (8) Fusion output of input image 2, (9) Low Resolution of input image 2 (10) Bi-Cubic Interpolation of input image 2, (11) Residual Image from VDSR of input image 2 (12) High-Resolution of input image 2 Obtained Using VDSR (13) High-Resolution Result of input image 2 using Bicubic Interpolation (14) Final Output of input image 2



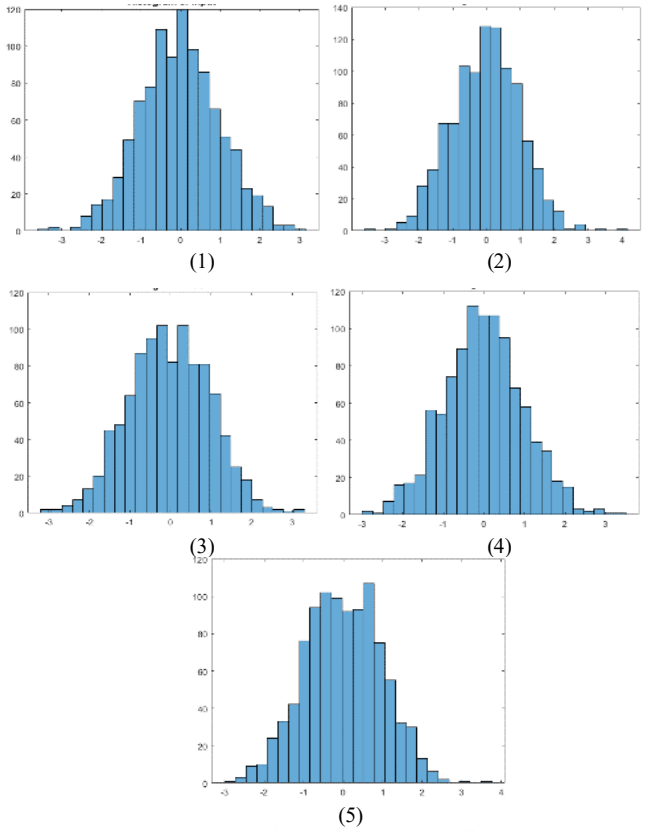
(1) (2)



(3) (4)

Figure 9: (1) Input image 1, (2) Output for input image 1, (3) Input image 2, (4) Output for input image 2 by, applying super-resolution Process.

An Image of Histogram outlines a chart that will be acclimated assess the standard of an image by looking at the values of an image some time recently and after upgrade [11]. The histogram of underwater input image is darker and so the improved output images' histogram has brighter pixel values. It is the mixture of the image modified by illumination and the image corrected by color. As a consequence of adding super resolution, another source image has brighter pixel values.



(1) (2) (3) (4) (5)

Figure 10: Histogram Curve of (1) Input Image, (2) Illumination Output Image, (3) Color Correction Output Image (4) Fusion Output Image (5) Super-Resolution Output Image.

Table 1: Result of comparison table

Image	Proposed				
	Entropy	PSNR	MSE	RMSE	SSIM
Input	7.5844	21.9562	416.8655	98.0544	0.9996
Illumination	7.5812	21.9416	420.6785	106.3258	0.9744
Color Correction	7.9011	21.9430	413.1758	15.7525	0.9535
Fusion	7.5812	21.9343	423.3366	15.8607	0.9507
Super Resolution	7.8738	21.8837	415.2216	15.7598	0.9926

Table 2: Numerical Result for Entropy of output Image and compare with [1] and [12] Methods

Image	Entropy		
	[1]	[12]	Proposed
Output	6.8683	7.5112	7.8738

5. Conclusion

In this article, by illumination improvement and color correction with fusion technique, we defined super resolution-based

underwater image enhancement that will ensure successful clarity enhancement and recovery of blurred underwater image. An image taken underwater faces so many problems like blurring and distraction, attenuation in water etc. For recovering this problem, we have proposed a method that has been used to correct all faced problems. Initially we took a hazy underwater image then we used illumination adjustment and color correction method on the image, after getting two outputs from illumination adjustment and color correction, we have used the fusion technique on those two images and combined them by the fusion technique to get an effective fused image as output. That clarity of the image has been improved after that. Then we have applied the super resolution method on the image obtained from fusion technique. Applying super resolution, the clarification of the image has been more-high. On the other-hand if there is any blur or noise in the image it will be removed and sharpened and we will get a high-resolution image. In Table 1 shows that the performance of the Entropy, Peak Signal-to-Noise Ratio (PSNR), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Structural Similarity Index (SSIM) result for the images which are used for analysis. The comparison between our output image and [1], [12]authors output image in Table 2, and its clearly describes that the output image has a higher entropy value from them, and it's as desired for the quality of a good image. From here we can clearly clarify that we can achieve a more effective clear image from a hazy image.

6. Future Works

Our plan is to focus on addressing the inhomogeneous dispersion and artificial lighting problems in Super Resolution, in future. On the other hand, in addition to hydrophyte robot inspection and IOT, particularly in vision systems, underwater inspection and observation tasks.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We the authors of Department of Electrical and Electronic Engineering, Department of Computer Science and Engineering and Department of Information and Communication Technology, Bangladesh Army University of Engineering and Technology and Bangladesh University of Professionals are very grateful for the tremendous support and guideline during the process of this work.

References

[1] C.O. Ancuti, C. Ancuti, C. De Vleeschouwer, P. Bekaert, "Color Balance and Fusion for Underwater Image Enhancement," IEEE Transactions on Image Processing, **27**(1), 379–393, 2018, doi:10.1109/TIP.2017.2759252.

[2] C. Li, J. Guo, R. Cong, Y. Pang, B. Wang, "Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior," IEEE Transactions on Image Processing, **25**(12), 5664–5677, 2016, doi:10.1109/TIP.2016.2612882.

[3] S.Q. Duntley, "Light in the Sea*," Journal of the Optical Society of America, **53**(2), 214, 1963, doi:10.1364/josa.53.000214.

[4] W. Zhang, G. Li, Z. Ying, "A new underwater image enhancing method via color correction and illumination adjustment," in 2017 IEEE Visual Communications and Image Processing (VCIP), 1–4, 2017, doi:10.1109/VCIP.2017.8305027.

[5] R. Khan, Fusion based underwater image restoration system, 2014.

[6] X. Wang, R. Nian, B. He, B. Zheng, A. Lendasse, "Underwater image super-resolution reconstruction with local self-similarity analysis and wavelet decomposition," in OCEANS 2017 - Aberdeen, 1–6, 2017, doi:10.1109/OCEANSE.2017.8084745.

[7] A.O. and A.Z.T. Kashif Iqbal, Rosalina Abdul Salam, "Underwater Image Enhancement Using an Integrated Colour Model," IAENG International Journal Of Computer Science, **32**(2), 239–244, 2007.

[8] K. Iqbal, M. Odetayo, A. James, R.A. Salam, A.Z.H. Talib, "Enhancing the low quality images using Unsupervised Colour Correction Method," in 2010 IEEE International Conference on Systems, Man and Cybernetics, 1703–1709, 2010, doi:10.1109/ICSMC.2010.5642311.

[9] R.SwarnaLakshmi, B. Loganathan, "An Efficient Underwater Image Enhancement Using Color Constancy Deskewing Algorithm," International Journal of Innovative Research in Computer and Communication Engineering, **3**, 7164–7168, 2015.

[10] R. Singh, M. Biswas, "Contrast and color improvement based haze removal of underwater images using fusion technique," in 2017 4th International Conference on Signal Processing, Computing and Control (ISPCC), 138–143, 2017, doi:10.1109/ISPCC.2017.8269664.

[11] J. Huang, A. Singh, N. Ahuja, "Single image super-resolution from transformed self-exemplars," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5197–5206, 2015, doi:10.1109/CVPR.2015.7299156.

[12] R. Singh, M. Biswas, "Adaptive histogram equalization based fusion technique for hazy underwater image enhancement," in 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 1–5, 2016, doi:10.1109/ICCIC.2016.7919711.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 43-48 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Switching Capability of Air Insulated High Voltage Disconnectors by Active Add-On Features

Mariusz Rohmann*, Dirk Schräder

Siemens Energy Global GmbH & Co. KG, SE GP T SP R&D TC ME, Research and Development, Berlin, 13629, Germany

ARTICLE INFO

Article history:
Received: 08 October, 2020
Accepted: 29 December, 2020
Online: 10 January, 2021

Keywords:
Disconnector
Contact system
Contact finger
Switching capability
Closing time
Opening time
Secondary contact
Passive add-on feature
Active add-on feature
Arcing
Contact erosion
Arc erosion

ABSTRACT

The need of add-on features (secondary contacts) for current paths of air insulated high voltage disconnectors switching capabilities (e.g. bus transfer switching) is introduced. The relevant product components for the switching capability and their functionality is described, which is giving boundary conditions for adding features necessary to achieve the switching capability. Possible features are discussed with necessary properties and performance. Those are separated in passive and active solutions, which are focused. Given solutions successfully applied, are explained and capabilities are shown based on experimental design and testing – where calculations and/or computational analysis are not shown (consequently no resulted data of such approaches), as those have not been used for the given solutions (the disconnector is still a low-cost product within industrial business circumstances, where invests for comprehensive design activities are unfortunately very limited). The testing of the solutions is covered with information for limited switching capability values and measures and/or further, partially theoretical, alternative solutions, for increased values. Also, the testing itself and possible laboratories with certain test execution opportunities and/or challenges is elaborated. The conclusion provides a market view, product users perspectives, in regards of the applicability for passive solutions and the need for active solutions.

1. Introduction

This paper is an extension of work originally reported in Proceedings of the International Conference on High Voltage Engineering and Technology (Central Power Research Institute, India) [1].

The standardized minimum switching capability of air insulated high voltage disconnectors (negligible currents of 0.5 A for rated voltages up to 420 kV as per IEC 62271-102) [2, 3] is frequently exceeded by three main applications:

- no-load conditions of transformers (inductive)
- long high voltage conductors (overhead lines, capacitive)
- bus-transfer cases within substations with more than one bus bar

The bus bar transfer case, being very common, is standardized by the IEC 62271-102, where the latest edition partially considers increased values (bus-transfer voltage and current) [2, 3].

A disconnector in its standard configuration will not be able to switch exceeded values. The damages on the contact system (contact erosion) would be too high to be still operational after such a switching case. The application of add-on features is necessary (secondary contact arrangements, like additional contact fingers), so that the contact system is protected and accordingly switching capabilities are achievable. Those add-on features are design as passive and/or active secondary contact arrangements [4-9].

The Main difference of the passive and active solutions is driven by the operational speed, where passive solutions are fully dependent on the current path design and active solutions are integrating semi or full independency of the current path design.

2. Current Path Operation

2.1. Current Path Types

The disconnector type determines the current path outlining (shape). The known types are groupable in inline disconnectors (the center break, the double side break, the side break, the V-type (based on center break or double side break disconnectors), the vertical break and the knee-type disconnector) and busbar

*Corresponding Author: Mariusz Rohmann, Siemens Energy Global GmbH & Co. KG, SE GP T SP R&D SHA, Nonnendammallee 104, 13629 Berlin, Germany, +49 (173) 7487359, mariusz.rohmann@siemens-energy.com

https://dx.doi.org/10.25046/aj060105

disconnectors (the vertical reach, the semi-pantograph and the pantograph disconnector).

For the inline disconnectors the current path operation occurs horizontally, and a common secondary contact arrangement solution can principally be achieved. For the busbar disconnectors the current path operation occurs vertically – a common secondary contact arrangement is not known to be achievable [4-9].

For the busbar disconnectors, the vertical reach and the semi-pantograph are likely to share a secondary contact solution with their familiars of the inline group (vertical break and knee-type respectively).

The pantograph disconnector is very unlikely to achieve a common secondary contact solution with one of the other mentioned disconnector types – its design principal of the current path is too distinctive.

2.2. Current Path Opening and Closing

Disconnectors are only operational under no-load conditions. The requirements for switching capability are given by residual currents, consequently main contact systems are designed for static current carrying capabilities [1-9].

In order to protect the main contact system (carrying rated continuous currents and short-circuit currents) it is necessary for any secondary contact arrangement to run ahead of the main contact system while the closing operation and to go behind the main contact system while the opening operation of the current path.

This is essential for the protection purpose, as the switching capability in any case results in arcing (lightning) and accordingly arc erosion resistance requirement. An optimized contact system in regards of the main purpose of carrying current will not be able to sustain such arcing erosion, where the main purpose of the secondary contact arrangement is this resistance but will not be suitable for carrying current continuously.

3. Secondary Contact Types

Where the main contact is optimized for low resistance (best current carrying capability while minimized losses, heating or hot spots) and longest lifetime, the secondary contact needs optimization for high resistance and a minimum lifetime (considering the IEC 62271-102 or specific market/customer requirements) [1-9].

Striving for this, the main contact needs:

- To be decoupled in switching operation conditions of the disconnector for being save from damages by arcing (lightning).
- A contact pressure suitable for the current carrying capability, but as small as possible for friction avoidance (no-load operations).
- A material choice suitable for the current carrying capability, but a surface supporting lowest friction (no-load operations).

Usually this results in silver plated copper fingers with applied springs to achieve the necessary contact pressure – being still one of the most economic solutions. Highly controlling the elasticity of the used copper may allow spring elimination but is likely to require silver-graphite-plating (increasing lubricating properties, decreasing friction), which again is more costly due to the manufacturing process efforts.

The secondary contact needs:

- To be decoupled in normal operation conditions of the disconnector for being save from damages by rated continuous currents or short-circuit currents.
- A contact pressure suitable for the switching capability – residual currents need to be carried safely, so that arcing on the main contact cannot occur. Additionally, considering certain numbers of switching operation and occurring erosion, the chosen contact pressure needs to overcome accordingly outline changes (reduction) of the contacts (sacrificial contacts) - preventing jumping or bouncing effects of the secondary contact and any arc-back (reigniting) due to contact separation.
- A material choice suitable for the arc erosion resistance (residual currents)

In dependence of the secondary contact solution, the material choices vary significantly – again under economic aspects, so that any material usage is minimized as much as possible in regards of the used solution.

Used names for known solutions like arcing horn, guiding horn, bus-transfer device or filter-switching device are not used. Those are given mainly by the reasons for occurrence of the residual currents to be switched and are not reflecting the common task [1].

3.1. Passive Secondary Contacts

Passive solutions, depending fully on the operation speed of the disconnector, require materials with high arc erosion resistance. Material suitability is firstly indicated by a high melting point and a high tensile strength, nevertheless the tensile strength of a material is sometimes significantly reduced when heated (even far below the melting point). Tungsten, the element with the highest melting point of the periodic table of the elements, is a proven choice.

Market availability (raw material, alloys, machinability) and implied material costs is another important factor.

Finally, the resistance to corrosion and the conductibility is still of importance, so that mostly copper alloys will be used and respectively proportions needs to be defined. Further optimized material usage will be achieved by inlet applications and possible conductive connection methods need to be considered for the material choices, as well.

Reported application examples have been [1]:

- Passive add-on finger – a copper zirconium alloy finger with a copper tungsten alloy inlet and a copper tungsten alloy stud (high tungsten content – more than 3 quarters of the alloy experienced to be necessary of tungsten)
- Passive add-on needle – copper zirconium alloy crossbars with copper tungsten alloy inlets and ‘tungsten needle’

The examples also successfully applied zirconium, which is another high-melting point element with a high enough tensile strength, also heated. Principally the noble metals are potential candidates: Ruthenium (Ru), Rhodium (Rh), Osmium (Os),

Iridium (Ir), Zirconium (Zr), Hafnium (Hf), Vanadium (V), Niobium (Nb), Tantalum (Ta), Chromium (Cr), Molybdenum (Mo), Tungsten (W), Rhenium (Re) and Technetium (Tc).

3.2. Semi-Active Secondary Contacts

Passive secondary contact may be enhanced by detailed design features for generating air flows advantageous for reducing the arcing time, so to increase the possible switching quantity or increase the switching capability.

For this, the air flow needs to be directed to one of the foot points of the arcing (inception point or arc origin), so that the separation or arc interruption will be facilitated. The enhancing effects are limited, but a significant advantage is the minimum weight effect for consideration of the mechanical behavior of the current path.

One successfully applied example is shown in figure 1, where a simple synthetic material could significantly increase the capability withstanding heating sufficiently.

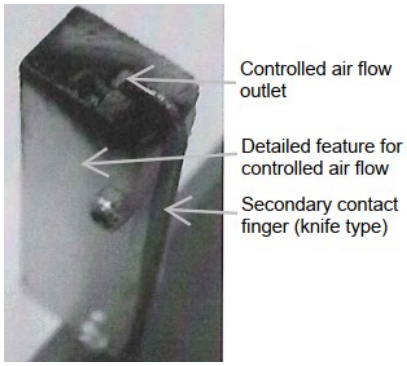


Figure 1: Example of Air Flow Controlling Detailed Feature

The controlled air flow is not only forced into the right direction towards the arcing foot but is also determined for volume and speed by the given passage. As the success was experimental, the air flow details have not been measured, but the switching current was more than doubled (form 6 A to 15 A) at the same switching voltage (21 kV). The achieved switching current was limited by the used laboratory, so that the capability limit of the solution could not be determined.

The finger outline (knife type) allows a relatively long air flow controlling item and therefore passage, which is important for determining the speed of the used volume. The passage is generated as a cuboid-like channel between the detailed design feature and the secondary contact finger by the fixation solution of the feature on the finger. The applicability depends highly on the outline of the used secondary contact, so that the shown example is limited in its applicability accordingly.

Further it represents rare occasions, where the type test execution allowed contractually a limit testing (in this case limited by the laboratories capabilities). The usual situation, as valid for all other shown cases, is a design and testing as per standard and its required values to be fulfilled (mainly IEC 62271-102 [2] or specific customer requirements).

3.3. Active Secondary Contacts

Full active secondary contacts will be designed with elements storing kinematic energy, which is released when required and provide the independent speed – classically those elements will be springs [5-9].

Figure 2 shows an example of such an active secondary contact. Compared to known passive contact systems it could be validated, that the switching capabilities are achievable with using only classical materials – the using of high-melting point materials or accordingly copper alloys was not necessary.

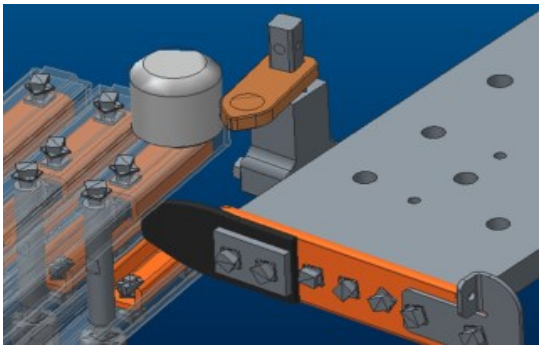


Figure 2: Active secondary contact system (double side break disconnector)

Figure 3 shows the same example of such an active secondary contact at a position of the closing operation, where the secondary contact is ahead of the main contact system. The copper sheet needs to run over the stainless-steel stud, so that the secondary contact goes behind the main contact system while the opening operation.

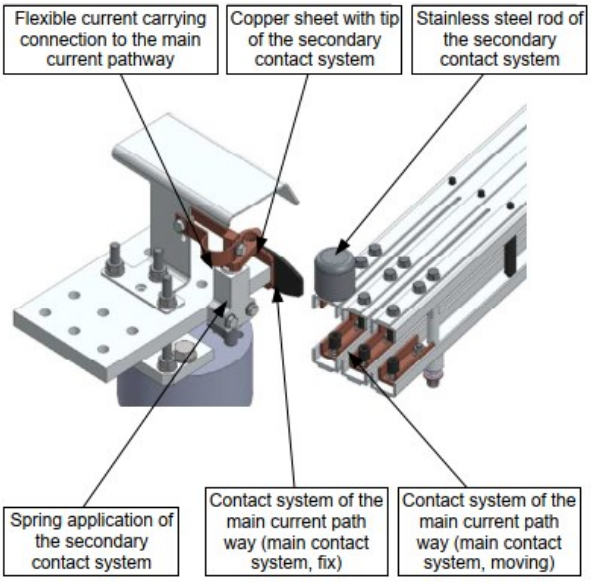


Figure 3: Feature Active secondary contact system in ahead-position

The spring is effectively applied only while the opening operation, so that the kinematic energy is specifically used for the demolition of the arc.

The intensity of the arc erosion while the closing operation is of much lower severity – as the arc arises, while the arcing distance decreases steadily. The arcing time is generally lower and the arc erosion stress on the foot points is less. In contrary, the erosion while the opening operation is of much higher severity – as the arc arises, while the arcing distance increases steadily until a high enough distance for insulation by air is achieved. In dependence on the arc intensity (voltage and current) the arcing time is

maximized in relation to the air gap necessary for demolition – exactly here the spring application comes into effect. Where passive solutions need to withstand this to full extend (only minimizable by dimensioning and eventually material elasticity), active solutions can significantly reduce the time independent speed. Of course, also active solutions require right dimensioning, as the used spring energy needs to fit to the allowed arcing time – achieved air gap at the time of arc ignition.

The spring application of active solutions does not require high-melting point material, which avoids the before mentioned challenges in regards of market availability. This allows again the usage of general core activities of disconnector designer and supplier concerning the current paths – plated copper / contact systems with spring applications.

Additionally, secondary contacts can be equipped with air flow controlling detailed feature. The shown example of figures 1 and 2 gives only little room for this.

3.4. Medium Voltage Switching Units

A further possibility for highest switching capabilities is given by the application of medium voltage switching units for active secondary contact systems. Those can be SF6 switching units or vacuum switching units, where the latter one is by far the preferred one as the first one is getting banned already in certain markets due to its negative environmental effects. Such a unit provides the advantage that in case of closing operations being kept open, no lightning occurs (because of the gap in the chamber of the unit). Important considerations of the application:

- Need for short-circuit resistivity in regards of the disconnector application
- Weight implications for the mechanical operation behavior of the current path – especially if applied to the moving part
- Potential requirement for complex kinematics for the operation of such units (tensioning and releasing unit internal spring applications)
- Protection of such units against environmental effects (e.g. rain or ice) with housings

The so far reported solutions (passive and active) are providing the big advantage against potential switching units application, that the weight implications is only low or eventually moderate and that the environmental protection coverage is mostly already given by the accordingly protection of the main contact system (or by simple and small size increase of the same).

The following earthing switch examples (figures 4 to 6) show known and successfully applied medium voltage switching units, which provide a technical basis being transferable to disconnectors.

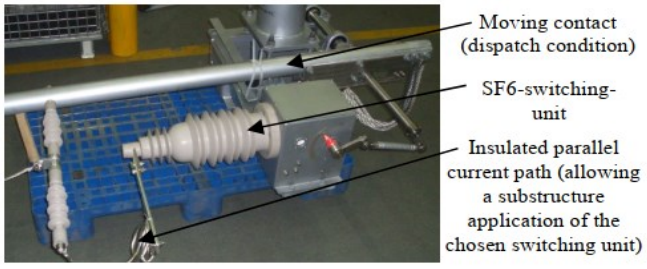


Figure 4: Earthing Switch Application with SF6-Switching-Unit (switching capability beyond class B as per IEC 62271-102)

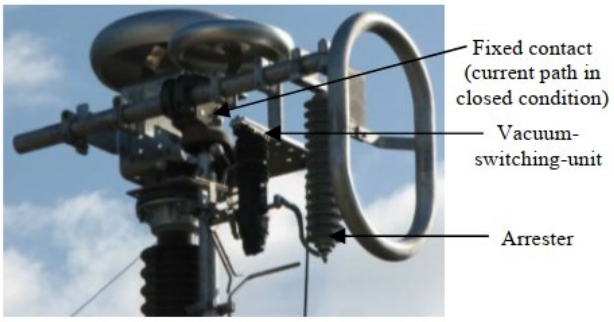


Figure 5: Earthing Switch Application with Vacuum-Switching-Unit and Arrester (switching capability requirement without lightning fulfilled)

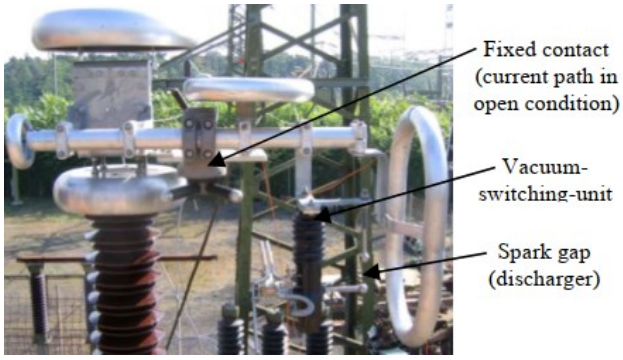


Figure 6: Earthing Switch Application with Vacuum-Switching-Unit and Spark Gap (switching capability requirement without lightning fulfilled)

4. Values and Testing

In practice, as well as per IEC 62271-102, type test confirmations are possible and common [2,3].

Regarding passive secondary contacts a confirmation of the type test for the same type of disconnector is physically evident for increasing disconnector sizes (higher rated voltages), as the arm length of the current path increases and therefore the angular speed increases. This provides beneficial conditions for the secondary contact as the arcing time gets reduced the reference type test is safely applicable. The same is difficult for decreasing disconnectors sizes by reduced current path arm length and consequently reduced angular speed and increased arcing time. Even with the proven margin within a type test, the applicability is not secure. Such kind of confirmation challenges will be increased by changing the disconnector type.

For active secondary contacts the confirmation situation is much more convenient based on the independent speed. The applicability or validity of successful type tested active secondary contact is evident based on the positioning of the secondary contact elements for their function and in relation to the main contact system (protection purpose) in accordance to the reference type test. The disconnector operational speed is not relevant. This is another important advantage of active solutions.

4.1. Switching Capability Values

The reported solutions are classical solutions and followed the edition 1 of IEC 62271-102. Therefore, the switching capability target was usually set to a switching current of 1600A, the limit given with the rule for 80% of the rated continuous current. Further, the switching voltage was mostly set to 100V, as most of

the market requests have been served with accordingly rated voltages (up to 170kV as per edition 1 of IEC 62271-102, but often also for higher rated voltages based on customer acceptance). The reported example of an active solution followed this approach, as it was achieved before the release of edition 2 of the IEC 62271-102 [2,3].

Since edition 2 of IEC 62271-102 [2], the general limit of 1600A is only applicable for rated voltages <245kV. For rated voltages of ≥245kV and ≤550kV, the limit is set to 4000A, which is a significant increase in arcing intensity for the switching capability (even though the percentage approach was reduced to 60%). Further, the switching voltage required is not only based on ranges of the rated voltage, but additionally dependent on the rated continuous current [2, 3].

Following an approach to design a solution for covering a major part of the values with this given current scenario is very difficult, if not impossible with passive secondary contacts. Latest for rated voltages of 245kV and above this will require active secondary contacts, as those provide higher capabilities – beside the advantageous confirmability.

4.2. Test Circuits and Laboratories

Edition 1 of IEC 62271-102 requested a power factor of 0.15 for the test factor and certain conditions for the transient recovery voltages, where the edition to requires a power factor of 0.5 and skipped the part for the transient recovery voltages. On the maintenance for at least 3s after interruption of the power frequency recovery voltage is no change between the editions of IEC 62271-102 [2, 3].

In fact, this is reducing the arc intensity for the type test. However, first experiences indicate that this is far not enough for having passive secondary contacts capable to achieve the before mentioned approach on switching capability values [7-9].

Basically, high power laboratories are providing equipment for testing the switching capabilities of air insulated high voltage disconnectors. The following restrictive points for the test execution of disconnectors must be carefully aligned with such potential laboratories:

- Space constraints – as disconnector test objects easily lead to high installation space necessary, specifically with increasing rated voltage of the test object. Ground fixation can be a challenge regarding the typical outline of a disconnector, which is potentially unusual for the equipment tested in the concerned laboratory. The installation height is another specific aspect of a typical disconnector outline with potential high challenge for the concerned laboratory
- Where the switching current to be tested is generally unproblematic for high power test laboratories, the switching voltage to be tested has carefully to be checked with the laboratory. This is even more stressed regarding the required values for high rated continuous currents following the IEC 62271-102 edition 2 [2] (identifying a high voltage laboratory with capabilities for the switching currents to be tested is very unlikely).
- The time necessary for the test execution is an important factor, as well. As such tests are usually not the core business of accordingly laboratories, the necessary slot size by time needs carefully to be aligned – the laboratory utilization, the throughput of test quantities, will be significantly impacted. The main determination of the overall test execution time is not mainly given by the quantity of operations to be done, but by the necessary time between to operations. This can be called cooling time and has to be determined by the disconnector designer/ supplier. Classically the focus was given by the avoidance of a motor overload (operating mechanism of the disconnector, automated test execution), which is with nowadays' motor performances not an issue. In fact, the focus needs to be given to the heating of the secondary contact – pre-heated secondary contact in view of subsequent operation. A maximum time gap between 2 operations is not regulated by standards and a realistic occurrence in the field will not be within minutes (maybe within days). For the before mentioned passive solutions a minimum time gap of 2 minutes was found to be enough (based on the high-melting point element tungsten). Considering 100 necessary operations (IEC 62271-102) [2], this ends up already to 3 hours, where setting, operation and frequent visual inspection has to be added, so that this part of the test execution easily need already a full day



Figure 7: Switching Capability Type Test Set-Up of a 245kV Center Break Disconnector with a Passive Add-On Finger Solution

Market demands represented in standards or resulting in new versions of standards are especially in the case of IEC globally driven. Based on this the possible design realization may be given already, but not generally. Usually it will be left to suppliers and customers to foster such realization, which will require consideration and alignment with testing facilities for realizations of accordingly type tests. The basis will be for the supplier, as well was for testing facilities, positive economic value adds. Potential volume, utilization and market prices will need to provide a

minimum basis for judging accordingly efforts. Partially even the requesting customers will join in a collaborative approach covering certain costs for realization (invests on disconnector supplier and/or tasting facility sides).

The switching capability testing for air insulated high voltage disconnectors could already be executed successfully with medium voltage high power laboratories (refer to figure 7). On first view this seems to be unusual but was possible since the above restrictive points have been servable.

5. Applicability Conclusion

As in every case of significant changes in standards and regulations there will be a transition period before the market will strictly insist. The established secondary contact solutions for the switching capability realization will therefore persist on the market for a certain period.

The market is already demanding steadily increased rated continuous currents, which is also reflected by the edition 2 of IEC 62271-102 [2], but not comprehensively. This is fostering the increased switching capabilities reflected in the edition 2 of IEC 62271-102 [2], but accordingly market demands run after the rated continuous current demands.

5.1. Passive Solutions

The shown examples are well suitable and proven especially for the rated voltage application <245kV, also in regards of the second edition of IEC 62271-102 [2]. Nevertheless, those solutions are at its limits and may exceed the given capabilities by semi-active detailed features. It requires the management of high-melting point elements for the design and the purchasing.

Consequently, it is reasonable that those passive secondary contact solutions will experience a replacement by simple active secondary contact solutions (e.g. the one reported above), as those allow increased switching capabilities and the avoidance of high-melting point elements.

5.2. Active Solutions

It is expected that active solutions will be the common solution in the feature based on the potentials given above – specifically in case of simple material usage and spring application, as reported above.

Considering the highest switching capability requirements as per edition 2 of IEC 62271-102 [2] – for the highest rated voltages and rated continuous currents – a strong demand for solutions integrating medium voltage switching units (chambers) is expected. Taking into account typical examples of bus-transfer disconnectors, like the pantograph disconnectors, and accordingly expected switching capabilities requirements as per IEC 62271-102 [2] or even above based on customer requirements, it is unlikely to be achievable with kinds of simple active solutions without compromising on functional reliability.

Vacuum chamber applications to the mating contact of pantograph disconnectors of 420kV rated voltage are under investigation and a final design achievement will need to be proven by type testing.

5.3. Earthing Switch Solutions

Earthing switches by its design and purpose must consider distinctive switching capabilities against disconnectors. Already accordingly standards are covering those switching capabilities reasonably distinctive [1-3].

Introduced earthing switch references with passive, active or active solutions considering medium voltage switching units are to be discussed separately. Basically, those known and successfully applied solutions are transferrable to disconnectors, but need specific considerations – e.g. the current path of an earthing switch is not having any duty of carrying rated continuous current, where this is one of the main duties of the current path of the disconnector (also this underlies exceptions for the earthing switch for certain applications) [1-9].

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Mariusz Rohmann thanks his wife Nallely Dzib Soto for her patience and continuous support while his professional efforts and specifically responsibilities and commitments abroad on delegations.

Mariusz Rohmann thanks Guru Moorthy Kurra for his collaboration, teamplay and continuous support while his delegation to India (April 2018 to January 2020) and up to today within his professional responsibilities and tasks.

References

[1] M. Rohmann, D. Schräder, G. Kurra, S. Challa, S. Mogatala, Swicthing Capability of Air-Insulated High Voltage Disconnectors", IEEE, 2019, ISBN 978-1-5386-7577-9/19

[2] IEC 62271-102, High-voltage switchgear and controlgear – Part 102: Alternating current disconnectors and earthing switches, Efition 2.0, 2018.

[3] IEC 62271-1, High-voltage switchgear and controlgear – Part 1: Common specifications for alternating current switchgear and controlgear, Edition 2.0, 2017.

[4] CIGRE Technical Brochure No. 511: Final Report of the 2004 – 2007 International Enquiry on Reliability of High Voltage Equipment, Part 3 – Disconnectors and Earthing Switches, 2012

[5] S.S. Rao, "Switchgear Protection and Power Systems (Theory, Practive and Solved Problems)", Khanna Publishers, 2008.

[6] M. Rohmann, "Analysis of Assembly Times for a Gas-Insulated High Voltage Switchgear on the Level of Components", M. Sc. Thesis with Siemens AG at FernUniversitaet Hagen, 2008.

[7] H. Lindner, Dr. Harry Brauer, Prof. Dr. Constants Lehmann, "Pocket Book of Electro-Technics and Electronics", Fachbuchverlag Leipzig, 1999, ISBN 3-446-21056-3.

[8] G. Balzer, Bernhard Boehle, Kurt Haneke, Hans Georg Kaiser, Rolf Pöhlmann, Wolfgang Tettenborn, Gerhard Voß, „Substations", 9th ed., Cornelsen Verlag Schwann-Giradet Duesseldorf, 1992, ISBN 3-464-48233-2.

[9] Prof. Fleck, P. Kulik, "High and Low Voltage Substations", Verlag W. Girardet Essen, 1975, ISBN 3-7736-0262-6.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 43-50 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Analysis of Standard Uncertainty of Six Degree of Freedom (DOF) Robot

Auttapoom Loungthongkam, Chana Raksiri*

Department of Industrial Engineering, Kasetsart University, Bangkok 10900, Thailand

ARTICLE INFO

Article history:
Received: 08 November, 2020
Accepted: 28 February, 2021
Online: 10 March, 2021

Keywords:
Robotic uncertainties
Robotic errors
Robotic kinematics

ABSTRACT

Robotic arms or industrial robots are a machinery that is widely used in the medical and military industries because it is a flexible, highly accurate and reliable. It is very necessary to work in complex tasks requiring more accuracy than humans can work. This paper presents an estimate of the standard uncertainty of 6 DOF robotic arm, KUKA KR5 ARC robot, and describes the experimental setup of a laser tracker to measure the position of the reflector mirror installed on a robot end-effector. This research describes the method of testing and experimenting to calculate the errors of each joint by using the inverse kinematic model, calculating the actual angle of their joint in comparing it with a nominal joint angle. The Jacobian matrix was applied to calculate the robotic position error. The calculation of uncertainties of each joint was conducted by using the Jacobian matrix to calculate the uncertainty in the robot and the four points testing were designed for estimating the error value and uncertainty value. The results showed that the error and uncertainty of each test point were within the range of the average error and the average uncertainty of the robot specification. The position errors and the position uncertainties of all test points within the robotic moving space were calculated and estimated by the proposed method and model. Therefore, the position error tolerance of each required moving target point must be smaller than the position errors and the position uncertainties that are estimated from this proposed model. These estimated robot linear position end effector uncertainties were used to compare and adjust the robotic path based on the required robotic position target and tolerance control.

1. Introduction

In the modern manufacturing industry, technology is applied, whether it is software or hardware or a combination of the both. The 6 degrees of freedom (DOF) robotic arm is one of the most sought-after combination machinery, as is designed to be flexible for mimicking human arm functions. This is consisting of arm parts and automation control parts that are accurate and precise. However, it is like a common machine. When used for a long time or used in improper conditions, it will result in deterioration. As a result, lower accuracy does not meet the specified features, and also result in a loss of reliability. From such problems, the researchers created the idea of estimating random errors and standard uncertainty of the 6 DOF robot, which researchers have determined that the joints are the most moving parts. As a result, this is the source of the most common errors. This research develops the principles of calculating random error and uncertainty that occurs with every joint and robot.

*Corresponding Author: Chana Raksiri, Bangkok, fengcnr@ku.ac.th

In this research, there are two objectives: 1) Estimate the random error of each joint and estimate the robotic error. and 2) Estimate the uncertainty of each joint and estimate the robotic uncertainty. When the experiment was set up, the laser tracker was applied to measure the position of the reflector installed at the end-effector. This paper is an extension of work originally presented in 2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA) [1].

Over the years, many researches has been conducted in analyzing the kinematic and kinematic errors of robotic arms and other machines, such as: a study of kinematics model of Staubli RX 90 robot 6 DOF robotic arm and position of robot control was calculated and tested by writing the English Alphabet [2]. Focus on machine tool random errors in a 3D workspace and offered new models that help predict product tolerances caused by uncertainties of a machine tool [3]. The new approach with a 2D manifold that reduces the dimensionality of the workspace to improve the efficiency of error compensation for robotic machining [4]. The photogrammetry-based measurement to compensate the

machining errors from parts and validated results [5]. A proposal for a guideline for a 3D-Piezo compensation mechanism unit that could quickly and accurately adjusted the spindle position to optimize robotic machining [6]. The evaluate the deviations for calculating the efficient robotic trajectory from aligned optically scanned point clouds [7] and the analysis of positioning accuracy and the kinematic parameters of robotic end-effector influenced by both internal and external temperature factors [8]. In addition, a study was proposed on improving the performance of the SEIKO D-TRAN RT3200 robot by studying the repeat control [9]. Improvement in the performance of robotic arms to have higher accuracy. Analyzed the positioning error value and then compensate for it [10]. Mathematical modeling to analyze the geometric errors of joint assembly affecting the end-effector of the 6 DOF robot was the one importance factor for robot accuracy [11]. Calibration and determination of the measurement uncertainty of robotic arms is an important issue in the confidence of the robot. There have been several research essays that discuss the principle of the measurement uncertainty such as: The kinematic error model presented by classifying the source of the error value by designing calibration methods compared to conventional calibration [12]. Applied laser interferometer to measure roundness [13]. Proposed the application of laser interferometer for calibration grade 2 gauge blocks and applied it for warhead roundness measurement [14-15]. Calibrated 6 DOF robotic arm using the Circle point analysis (CPA) method to configure a circle to measure and use the Monte Carlo method to find out the measurement uncertainty [16] and used a new method to calibrate the end-effector of 6 DOF robotic arm [17]. As mentioned above, in this research, there are the action plan is as follows: 1) Preparation of equipment and tools, 2) Learnt the involved methods and theories such as principle of the measurement method, kinematics model and the measurement uncertainties 3) experimental design and testing 4) result and conclusion.

2. Materials and Methods

2.1. Robotic arm

Robotic arms are machines that are developed for many characteristics, but at the same time are precise and accurate. This research presents an estimate of the standard uncertainty of 6 DOF robotic arm, focusing on welding robotic arms. The robotic arm used in this research is the KUKA KR5 ARC as shown in Figure 1 and D-H parameters and joints limit as shown in Table 1.

Table 1: KUKA KR5 ARC D-H parameters

Link (n)	Link twist (α_{n-1})	Link length (a_{n-1})	Link offsets (d_n)	Joint angles (θ_n)
1	180˚	0	-400	θ_1
2	90˚	180	0	θ_2
3	0˚	600	0	θ_3-90º
4	90˚	120	-620	θ_4
5	-90˚	0	0	θ_5
6	90˚	0	-115	θ_6
7	180˚	0	0	180º

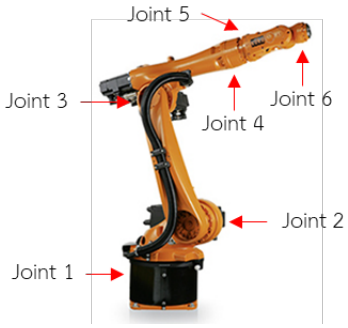


Figure 1: KUKA KR5 ARC robot

2.2. Laser Tracker

Laser tracker is a 3D measurement device, which is a standard metrology in precision, accuracy and reliability. This research uses the FARO laser tracker ION model as shown in Figure 2 to measure the position of the laser reflector installed at the end-effector.



Figure 2: FARO laser tracker ION model

2.3. Kinematic model

Kinematics describes the movement of object points and systems of bodies (groups of objects), regardless of the force that causes movement. Kinematics is a major part of mechanics which is often referred to as the "geometry of movement". The kinetic problem begins by explaining the system's geometry and declaring the initial conditions of the value, position, speed and/or acceleration within the known system. From a geometric point of view, it can locate the speed and acceleration of any unknown part of a system [18].

Robotic kinetics relies on differentials to describe the relationship between joints and links from the base to the end-effectors. The frame attached to the robotic joints is serialized like a chain. The relationship of one frame versus another frame from the bottom up will result in a conversion equation. This is the relationship of the base frame against the tool frame [19].

2.3.1. Forward kinematic

Forward kinematics are the mathematical model to compute the coordinates and directions (homogeneous conversion form) of robot end-effector positions relative to the function of the robot angles of each joints. This paper followed Denavit and Hartenberg (D-H) by choosing the reference frames in a robotic application

that Jacques Denavit and Richard S. Hartenberg proposed. In this convention, the coordinate frame is attached to the joints between the two links so that one conversion involves joint [Z] and the second frame is linked to the link [X], converting coordinates with serial robots that contain *n* links to the robot's equation format. D-H convention determines each conversion A_n with a multiple of fundamental conversions, respectively. The transformation matrix can be written as follows:

$$ {}^{n}_{n-1}T = A_n = \begin{bmatrix} C\theta_n & -S\theta_n & 0 & a_{n-1} \\ S\theta_n C\alpha_{n-1} & C\theta_n C\alpha_{n-1} & -S\alpha_{n-1} & -d_n S\alpha_{n-1} \\ S\theta_n S\alpha_{n-1} & C\theta_n S\alpha_{n-1} & C\alpha_{n-1} & d_n C\alpha_{n-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1) $$

where, A_n is the homogeneous matrix that describes the movement of the frame of each contiguous joint (n-1 and n). θ_n is the joint angle of the *i*th joint, d_n is the link offset of the *i*th joint, a_n is the link length of the *i*th joint, and α_n is link twist of the *i*th joint. $S\theta_n = S_n = \sin(\theta_n)$, $S\alpha_n = \sin(\alpha_n)$, $C\theta_n = C_n = \cos(\theta_n)$ and $C\alpha_n = \cos(\alpha_n)$.

The parameters were replaced in Table 1 in (1), and multiplied all the matrixes in order. It is the last matrix that shows the position and direction of the end-effector compared to the base can be displayed as follows:

$$ {}^{7}_{0}T = A_1A_2A_3A_4A_5A_6A_7 = \begin{bmatrix} n_x & o_x & a_x & P_x \\ n_y & o_y & a_y & P_y \\ n_z & o_z & a_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2) $$

In this research, the linear position of end-effector is determined, which shows the position in the coordinates x, y, and z, as shown in (3).

$$ \begin{aligned} P_x &= C_1[a_1 + a_2C_2 + S_{23}(a_3 + d_6C_4S_5) - C_{23}(d_4 + d_6C_5)] \\ &\quad + d_6S_1S_4S_5 \\ P_y &= -S_1[a_1 + a_2C_2 + S_{23}(a_3 + d_6C_4S_5) - C_{23}(d_4 + d_6C_5)] \\ &\quad + d_6C_1S_4S_5 \\ P_z &= -d_1 - a_2S_2 + C_{23}(a_3 + d_6C_4S_5) + S_{23}(d_4 + d_6C_5) \end{aligned} \quad (3) $$

2.3.2. Inverse kinematic

In the previous section, the robot forward kinematics are the mathematical equations, used to calculate the position and direction of the end-effector frame compared with base frame when the variables joints (q_0, q_1, q_2,…, q_n) are known. On the other hand, to know the variable joints (q_0, q_1, q_2,…, q_n), when determining the homogeneous matrix at the end-effector frame compared to the base frame (${}^{n}_{0}T$), the resulting relationship is called inverse kinematics.

For general cases in robotics, the inverse kinematics calculation of the robot is more complex than the forward kinematics problems because the results can occur in 3 different ways consisting of no solution, unique solution, and many solutions.

2.3.3. Jacobian matrix

The Jacobian matrix is a matrix that shows the relationship between the error of the end-effector and the 6 joints error of the robot. The Jacobian matrix can be found in the analysis of the forward kinematic, as shown in (4).

$$ dX = Jd\theta_i \quad (4) $$

where, $d\theta_i$ are the angle errors of 6 axis (i=1, 2, 3,…,6).
J is Jacobian matrix.

Jacobian's pattern is in the form of a 6x6 matrix, which is based on (5):

$$ J_{6\times6} = [J_1\ J_2\ J_3\ J_4\ J_5\ J_6] $$

$$ J = \begin{bmatrix} J_{11} & J_{12} & J_{13} & J_{14} & J_{15} & J_{16} \\ J_{21} & J_{22} & J_{23} & J_{24} & J_{25} & J_{26} \\ J_{31} & J_{32} & J_{33} & J_{34} & J_{35} & J_{36} \\ J_{41} & J_{42} & J_{43} & J_{44} & J_{45} & J_{46} \\ J_{51} & J_{52} & J_{53} & J_{54} & J_{55} & J_{56} \\ J_{61} & J_{62} & J_{63} & J_{64} & J_{65} & J_{66} \end{bmatrix} \quad (5) $$

The six joints of the robot are all revolute (R). Therefore, the Jacobian matrix can be calculated as follows:

$$ J = \begin{bmatrix} z_{i-1} \times (o_n - o_{i-1}) \\ z_{i-1} \end{bmatrix} \quad (6) $$

where, z_{i-1} is the value of the top 3 elements in column 3 of the matrix T^i_0
o_i is the value of the top 3 elements in column 4 of the matrix T^i_0

When a robot moves, the linear position error for x, y and z direction of the end-effector, can be written in a partial derivative of equation 3 relative to the angle in each of the changing joints. The Jacobian components that are affected by the x, y and z linear position of end-effector in Jacobian matrix as follows:

$$ \begin{aligned} J_{11} &= -S_1[a_1 + a_2C_2 + S_{23}(a_3 + d_6C_4S_5) - C_{23}(d_4 + d_6C_5)] \\ &\quad + d_6C_1S_4S_5 \\ J_{12} &= C_1[C_{23}(a_3 + d_6C_4S_5) + S_2(d_4 + d_6C_5) - a_2S_2] \\ J_{13} &= C_1[C_{23}(a_3 + d_6C_4S_5) + S_{23}(d_4 + d_6C_5)] \\ J_{14} &= d_6S_5(C_1S_{23}S_4 + S_1C_4) \\ J_{15} &= d_6(C_1S_{23}C_4C_5 + C_1C_{23}S_5 + S_1S_4C_5) \\ J_{16} &= 0 \\ J_{21} &= -C_1[a_1 + a_2C_2 + S_{23}(a_3 + d_6C_4S_5) - C_{23}(d_4 + d_6C_5)] \\ &\quad - d_6S_1S_4S_5 \\ J_{22} &= S_1[a_2S_2 - C_{23}(a_3 + d_6C_4S_5) - S_{23}(d_4 + d_6C_5)] \\ J_{23} &= S_1[-C_{23}(a_3 + d_6C_4S_5) - S_{23}(d_4 + d_6C_5)] \\ J_{24} &= d_6S_5(S_1S_{23}S_4 + C_1C_4) \\ J_{25} &= -d_6(S_1S_{23}C_4C_5 + S_1C_{23}S_5 - C_1S_4C_5) \\ J_{26} &= 0 \\ J_{31} &= 0 \\ J_{32} &= -[S_{23}(a_3 + d_6C_4S_5) - C_{23}(d_4 + d_6C_5) + a_2S_2] \\ J_{33} &= -S_{23}(a_3 + d_6C_4S_5) + C_{23}(d_4 + d_6C_5) \\ J_{34} &= -d_6C_{23}S_4S_5 \end{aligned} $$

$J_{35} = -d_6(S_{23}S_5 - C_{23}C_4C_5)$
$J_{36} = 0$

2.4. Principle of measurement the random error of individual robotic joints method

In this research, the principle of the proposed measurement method measures the positions the end-effector in 3 dimensions system (x, y, z) that is affected from the rotation of a joint when the others are locked. Figure 3 shows an example of the measurement principle of joint 2. After the measurement system setup, the program instructs the end-effector to move to the specified position (P_1, P_2, P_3,..., P_n). The laser tracker measures the position according to the specified cycle, the identification of rotation plane and rotation center are shown in section 2.4.1 and 2.4.2.

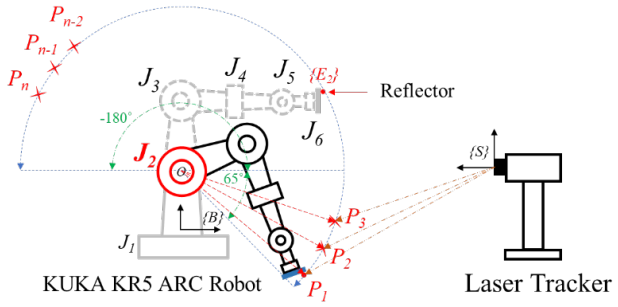


Figure 3: Basic principle of the measurement method for J_2 random joint error setting

Identifying a rotating plane of the robotic arm joints aims to create a circular arc from the rotation of the joints measured in the Cartesian area. This arc has a measurement points m. The rotation plane in the Cartesian space can be calculated as follows:

$$z = Ex + Fy + G \tag{7}$$

The previous equation is where x, y and z are coordinates points in the rotation plane. E, F and G are the rotation plane coefficients.

In the rotation plane, the adjustment of the measurement point (x_i, y_i, z_i) when k=1, ..., m is derived from the minimization problem as follows:

$$J_z = \sum_{k=1}^{m}(z - z_k)^2 \tag{8}$$

2.4.2. Identification of rotation center

In theory, the trajectory of the m point in the arc of the circle of a rigid body, this circle plane is perpendicular to the fixed axis. However, in the practice, there are some factors such as vibration during rotation and the non-standard assembly etc. So, that result in the trajectory of the m points may not be completely circular and assume that the theoretical circular and circle trajectory have very few deviations, thereby, a circle equation from the equation (9) is used to fit these points. So, the rotation plane of robotic arm joints can be calculated. Using the least square method was done in order to identify the circle on the circle plane. A standard form of a circle equation is as follows:

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \tag{9}$$

The previous equation is where (x_c, y_c) is the rotation center. r is the radius of the circle. From (3) can be rewritten as follow:

$$w = x^2 + y^2 = Ax + By + C \tag{10}$$

The previous equation is where A, B and C are the coefficients of the circle center in 3D system. Then, there are the sets of measured (x_i, y_i, z_i) when i=1, 2, 3, …, m is derived from the minimization problem as follows:

$$J_w = \sum_{k=1}^{m}(w - w_k)^2 \tag{11}$$

2.4.3. Experiment setup

This system consists of 3 important parts: 1) robotic arm. 2) laser tracker. and 3) installation of the reflector at the end effector as shown in Figure 4.

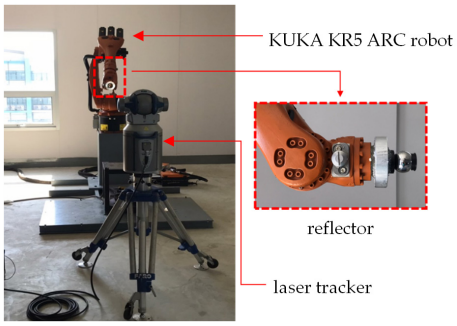


Figure 4: The 6 DOF KUKA KR5 ARC robot experimental setup

- Assign the end-effector move to the first position $P1$.
- Unmeasured joints are locked from moving. After that measured the points P_1, P_2, P_3, …, P_n are fitted with a plane in the Cartesian space.
- On the determined plane, specifies the points on the arc of the circle P_1, P_2, P_3, …, P_n and determines the center of the circle O.
- Measured 30 times repeat for every points m.

2.5. Evaluation for Measurement Accuracy

In a high-accuracy and high-precision measurement system to increase measurement and reliable confidence levels, measurement accuracy is required to increase the confidence and

reliability of the measurement. Therefore, measurement accuracy assessment is required. This evaluation compares the deviation (errors) between frame $\{E_i\}$ and frame $\{E_i'\}$ where, frame $\{E_i\}$ is calculated by the proposed method and frame $\{E_i'\}$ is calculated by forward kinematic of robotic arm.

2.5.1. Joint errors

Joints errors are due to repeatability of measurement and can be calculated as the difference between actual (measurement) and nominal angle. The position of end-effector is converted to the angle of each joint with inverse kinematics. The actual angle as shown in (12-17):

$$\theta_1 = Atan2(-\bar{n}_y, \bar{n}_x) \text{ or } \theta_1 = Atan2(\bar{n}_y, -\bar{n}_x) \quad (12)$$

$$\theta_2 = Atan2(ad - bc, ac + bd)$$
$$\text{or } \theta_2 = Atan2(-ad + bc, -ac - bd) \quad (13)$$

$$\theta_3 = Atan2(d, c) + Atan2\left(\pm\sqrt{c^2 + d^2 - e^2}, e\right) \quad (14)$$

$$\theta_4 = Atan2(a, b) \text{ or } \theta_4 = Atan2(a, b) + a_1 \quad (15)$$

$$\theta_5 = Atan2\left(\pm\sqrt{a^2 + b^2}, c\right) \quad (16)$$

$$\theta_6 = Atan2(a, b) + Atan2\left(\pm\sqrt{a^2 + b^2 - c^2}, e\right) \quad (17)$$

for θ_2. $a = -d_1 + d_6\bar{n}_z$, $b = -\left(a_1 + d_6\bar{n}_x C_1 - d_6\bar{n}_y S_1\right)$, $c = -d_4 C_3 + a_3 S_3$ and $d = a_3 C_3 - d_4 S_3 + a_2$

for θ_3. $c = 2a_2a_3$, $d = -2a_2d_4$ and $e = a^2 + b^2 - (a_3^2 + d_4^2 + a_2^2)$

for θ_4 and θ_5.

$$a = \left(p_y C_1 + p_x S_1\right)/d_6, b = \Big((p_z + d_1)C_{23} - (a_2 + a_3)S_{23} - a_3 + (p_y(C_{123} - C_{1-2-3}))/2 + \left(p_x(S_{123} - S_{1-2-3})\right)/2\Big)/d_6$$

and $\Big(c = (p_z + d_1)S_{23} - a_1C_{23} - a_2C_3 + \Big(p_y(S_{123} - S_{1-2-3})\Big)/2 + \left(p_x(C_{123} - C_{1-2-3})\right)/2 - d_4\Big)/d_6$

for θ_6. $a = C_4$, $b = C_5S_4$ and $c = o_yC1_+o_{yx}S1$

Therefore, the joint errors can be calculated as follow:

$$d\theta_i = \frac{1}{g}\sum_{k=1}^{g}(\theta_{in} - \theta_{ia}) \quad (18)$$

where, θ_{in} are the nominal angle.
θ_{ia} are the actual (measurement) angle.
g is number of measurement points.
i are the robotic joints (i=1, 2, 3, …, 6).

2.5.2. Position errors

In this research, the linear position error for x, y and z direction of the end-effector (*dx*, *dy*, *dz*) are determined. Therefore, the position errors in the *x*, *y* and *z* axes at the end-effector as shown in (19-21) respectively.

$$dx = J_{11}d\theta_1 + J_{12}d\theta_2 + J_{13}d\theta_3 + J_{14}d\theta_4 + J_{15}d\theta_5 + J_{16}d\theta_6 \quad (19)$$

$$dy = J_{21}d\theta_1 + J_{22}d\theta_2 + J_{23}d\theta_3 + J_{24}d\theta_4 + J_{25}d\theta_5 + J_{26}d\theta_6 \quad (20)$$

$$dz = J_{31}d\theta_1 + J_{32}d\theta_2 + J_{33}d\theta_3 + J_{34}d\theta_4 + J_{35}d\theta_5 + J_{36}d\theta_6 \quad (21)$$

From that (19-21) can calculate the total error at the end-effector as follows:

$$e_t = \sqrt{d_x^2 + d_y^2 + d_z^2} \quad (22)$$

2.6. Measurement uncertainty

In metrology, measurement uncertainty is an expression of statistical distribution of values caused by measured quantities. Therefore, measurement uncertainty is used in combination with the measurement results to reflect the actual value of the measurement quantity. Especially, when the measurement results are applied to various quality criteria, the results must be applied to the quality criteria. According to international agreements, this uncertainty is based on probability and shows the incomplete of the quantity and it is a non-negative parameter.

Multiple repetitive measurements and calculating standard deviations from re-measuring are among the most common practices in estimating measurement uncertainty. Whether it is a full or partial measurement, it can be repeated. The estimation of the uncertainty received is a standard deviation of the repetitive measurement results, which is a statistical process. It is called type A uncertainty. The type B uncertainty is estimated from other deviation source such as material certificates, specifications, and long-term experience-based assessments [20].

The Expression of Uncertainty and Confidence in Measurement, M3003 [21] and Fundamental Parameter and Application describes how to calculate standard uncertainty.

2.6.1. Type A uncertainty

Type A uncertainty is estimated using statistical principles. By performing measurements, conditions are subject to repeated measurement conditions to see the original iteration or view the distribution of the measured average. The arithmetic mean, or average, of the results should be calculated. If there are *n* independent repeated values for a quantity *Q* then the mean value $\bar{q}$ is given by:

$$\bar{q} = \frac{1}{n}\sum_{i=1}^{n} q_i = \frac{q_1 + q_2 + \ldots + q_n}{n} \quad (23)$$

Then the estimated fragmentation of data is calculated from the standard deviation of the *n* values can compute as follows:

$$\sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(q_j - \bar{q}\right)^2} \quad (24)$$

Equation (24) gives the standard deviation for sampling from all populations. However, based on the results of a single measurement sample, approximate, *s(qj)*, can be made from the standard deviation σ of the entire population of possible values from the relation:

$$s(q_j) = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(q_j - \bar{q}\right)^2} \quad (25)$$

The average *q* will be derived from the exact sample number of n and, therefore, its value will not be the exact average if the infinite

sample number is taken. This uncertainty is called the standard deviation of the mean that can compute as follow:

$$s(\bar{q}) = \frac{s(q_j)}{\sqrt{n}} \quad (26)$$

2.6.2. Type B uncertainty

Estimating type B uncertainty is the other way around. In other words, it uses information from various sources that are academically reliable to consider the assessment. The error element from a source called the systematic error is valued as a type B uncertainty. Therefore, estimating the type B uncertainty requires a lot of knowledge of measurement techniques. In order to be able to identify a source of uncertainty, much of the systematic error is complete. The estimation of uncertainty needs to consider the source of uncertainty values as follows:

- Deviation Instability compared to the reference standard stipulated in the reported calibration uncertainty.
- Calibration of measuring instruments includes accessories and any drift or instability in measuring instruments or readings.
- Resolution and uncertainty during measurement.
- The operational procedure.
- Mistakes caused by the operator.
- The effects from environment.

2.6.3. Combined uncertainty

The type A and B uncertainty values are part of the overall measurement uncertainty value, which can be calculated for the combined standard uncertainty as shown in (27):

$$u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^2\, u^2(x_i)} \equiv \sqrt{\sum_{i=1}^{N} u_i^2(y)} \quad (27)$$

where $u_c(y)$ is the combined uncertainty, c_i is the sensitivity coefficient and $u(x_i)$ is the standard uncertainty.

2.6.4. Expanded uncertainty

The combined uncertainty that is approximated is considered to have a certain level of confidence which is not suitable enough for laboratory use. Therefore, it is necessary to multiply combined uncertainty with coverage factor k. The expanded uncertainty as follows:

$$U = k.u_c(y) \quad (28)$$

2.6.5. Robot uncertainty

In this research, the uncertainties position errors for x, y and z direction of the end-effector (U_x, U_y and U_z) are determined. Therefore, the uncertainties position errors in the x, y and z axes at the end-effector (U_x, U_y and U_z) as the function of joint uncertainty (U_1, U_2, U_3, U_4, U_5 and U_6) are shown in (29-31) respectively.

$$U_x = J_{11}U_1 + J_{12}U_2 + J_{13}U_3 + J_{14}U_4 + J_{15}U_5 + J_{16}U_6 \quad (29)$$
$$U_y = J_{21}U_1 + J_{22}U_2 + J_{23}U_3 + J_{24}U_4 + J_{25}U_5 + J_{26}U_6 \quad (30)$$
$$U_z = J_{31}U_1 + J_{32}U_2 + J_{33}U_3 + J_{34}U_4 + J_{35}U_5 + J_{36}U_6 \quad (31)$$

From that (29-31), it can calculate the total uncertainty at the end-effector as follows:

$$U_t = \sqrt{U_x^2 + U_y^2 + U_z^2} \quad (32)$$

3. Results

The research focuses on the study estimating the standard uncertainty of 6 DOF KUKA KR5 ARC robot, considering the 3 key parts of which parameter values used in calculations derived from the experiment as shown in section 3.1-3.2 consist of robotic errors and robotic uncertainties.

3.1. Robotic error

The errors of the robot in this research are divided into two parts: joint errors and position errors.

3.1.1. Average Joint errors

Based on the experiments measuring data from section 2.4, the average joint errors are calculated according to (18).

Joint 1 average error ($d\theta_1$) equals 0.101°.
Joint 2 average error ($d\theta_2$) equals –0.287°.
Joint 3 average error ($d\theta_3$) equals 0.070°.
Joint 4 average error ($d\theta_4$) equals –0.097°.
Joint 5 average error ($d\theta_5$) equals 0.026°.
Joint 6 average error ($d\theta_6$) equals 0.140°.

3.1.2. Average Position errors

When average joint errors are known, the average position errors at the end-effector in x, y and z axis can be calculated according to (19-21).

Average Position error in x axis (d_x) equals –0.357 mm.
Average Position error in y axis (d_y) equals –0.303 mm.
Average Position error in z axis (d_z) equals 0.383 mm.

So, instead of position errors at the end-effector in x, y and z axis in (21), the average robot error (e_t) is equal to 0.605 mm.

3.2. Robotic uncertainty

The uncertainties of the robot in this research are divided into two parts: joint uncertainties and position uncertainties.

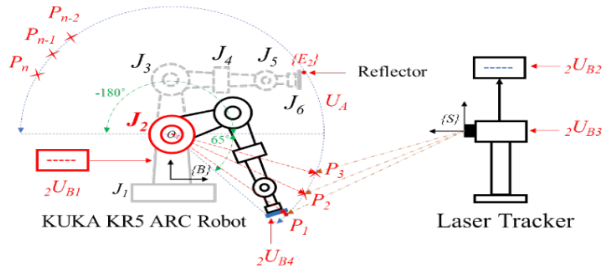


Figure 5: Basic principle of the measurement method and sources of uncertainty

3.2.1. Joint uncertainties

The research is based on the principle of the measurement method and then calculates the average position errors and the standard deviations of position errors of the robotic joints. These

errors show the accuracy, and the standard deviations show the precision of the joints. After that, these values are used to determine the type A standard of joints 1 to 6 of the robot.

Figure 5 shows the basic principle of the measurement method and the sources of uncertainty. In this research, the sources of uncertainty consist of five sources, that uncertainties are type A uncertainty ($_iU_A$) and four sources of type B uncertainty ($_iU_{Bj}$):

where, $_iU_A$ are the uncertainty from repeatability.
$_iU_{B1}$ are the uncertainty from resolution of the robot.
$_iU_{B2}$ are the uncertainty from resolution of laser tracker.
$_iU_{B3}$ are the uncertainty from accuracy of laser tracker.
$_iU_{B4}$ are the uncertainty from thermal effect at the reflector fixture.
i are the robotic arm joints.
j are the number of the sources of uncertainty.

Table 2 provides a preview of the sources of the uncertainty and the calculation of the total uncertainty of the 1st joint. The uncertainty of each joint are as follows:

The uncertainty of joint 1 (U_1) equal ±0.204º.
The uncertainty of joint 2 (U_2) equal ±0.149º.
The uncertainty of joint 3 (U_3) equal ±0.076º.
The uncertainty of joint 4 (U_4) equal ±0.126º.
The uncertainty of joint 5 (U_5) equal ±0.140º.
The uncertainty of joint 6 (U_6) equal ±0.115º.

3.2.2. Position uncertainties (Robot uncertainty)

When the uncertainties of each joint are known. The uncertainties in x, y and z axis at the end-effector can be calculated according to (29-31), the results as the follow:

The uncertainty in x axis (U_x) equal ±0.225 mm.
The uncertainty in y axis (U_y) equal ±0.613 mm.
The uncertainty in z axis (U_z) equal ±0.366 mm.

So, instead of position uncertainties at the end-effector in x, y and z axis in (32), the average robot uncertainty (U_t) is equal to ±0.748 mm.

The research results showed that the KUKA KR5 ARC robot used in the experiment had the average robot error which (e_t) is equal to 0.605 mm. and with the average robot uncertainty (U_t) is equal to ±0.748 mm. Then, examples of four test points (P_1, P_2, P_3 and P_4) are used to determine the estimated error and uncertainty in each joint from propose model. The position errors and position uncertainties at the end-effector in x, y and z axis can be calculated according to (19–21) and (29–31). The results of the estimated robot position errors and the position uncertainty are shown in the Table 3.

4. Discussion

Objective 1: Measure random errors of each joint and the robot end-effector linear random error in x, y and z direction and average error can be calculated. When comparing the robot error with the specified accuracy, the robot error was found to be slightly higher than the accuracy value. The results of the research process can be compensated in the control system, as it can reduce the error value.

Objective 2: Estimate the uncertainty of each joint and estimate the robot uncertainty. The results show that the uncertainty of each joint at 95% confidence level was quite high and when calculating the average uncertainty of the robot it was quite high too. The high average uncertainty will result in a difference of the robotic error. The biggest uncertainty is the uncertainty from repeatability. It was between 86.21% and 94.48%. This suggests that the uncertainty from repeatability has a significant effect on the total uncertainty, so it is appropriate to consider this factor.

Table 2: The standard uncertainty of joint 1

Type	Value (*degree*)	Distribution	Divisor	Uncertainty contribution u_i (*degree*)	Sensitivity coefficient c_i	Effective DOF V_{eff}
$_1U_A$	0.1019	Normal	1	0.1019	1	20
$_1U_{B1}$	0.0019	Rectangular	$\sqrt{3}$	0.0011	1	∞
$_1U_{B2}$	0.0095	Rectangular	$\sqrt{3}$	0.0055	1	∞
$_1U_{B3}$	0.0035	Rectangular	$\sqrt{3}$	0.0020	1	∞
$_1U_{B4}$	0.0002	Rectangular	$\sqrt{3}$	0.0001	1	∞
$_1U_C$		-		0.0083		∞
$_1U_r$		Normal k=2		**0.2042**		∞

Table 3: The estimated position errors and uncertainties for the example of four test points

Points	Target position (mm)			Position errors (*mm*)				Position Uncertainties (*mm*)			
	x	y	z	d_x	d_y	d_z	e_t	U_x	U_y	U_z	U_t
P1	125	125	125	– 0.033	– 0.012	0.445	**0.447**	±0.164	±0.163	±0.177	**±0.291**
P2	225	225	225	– 0.118	0.182	0.414	**0.467**	±0.164	±0.163	±0.177	**±0.291**
P3	225	25	225	– 0.125	– 0.077	0.465	**0.488**	±0.159	±0.161	±0.185	**±0.292**
P4	25	25	25	– 0.050	– 0.093	0.199	**0.225**	±0.168	±0.159	±0.178	**±0.292**

Therefore, the results from Table 3, the position errors and the uncertainties of all test points within the robot moving space are calculated and estimated by the proposed method and model. Therefore, the position error tolerance of each required moving target point must be smaller than the position errors and the uncertainties that are estimated from this proposed model.

5. Conclusions

In this research, it describes the method of testing and experimenting to calculate the errors of each joint, and calculating the actual angle of their joint to compare with a nominal joint angle. The Jacobian matrix was applied to estimate the robot position error. The estimation of the uncertainties of each joint by using Jacobian matrix was done to estimate the robot position uncertainty. The research results showed that the KUKA KR5 ARC robot used in the experiment average robot error (e_t) was equal to 0.605 mm. and with the average robot uncertainty (U_t) being equal to ±0.748 mm.

The example of the four test points (P_1, P_2, P_3 and P_4) determined the estimated error and uncertainty in each joint from proposed model. The position errors (e_t) was 0.447mm, 0.467 mm, 0.488 mm and 0.225 mm respectively. The position uncertainty (U_t) was ±0.291 mm, ±0.291 mm, ±0.292 and ±0.292 mm. The position errors and the uncertainties of all test points within robot space moving were calculated and estimated by the proposed method and model. Therefore, the position error tolerance of each required test point must be smaller than the position errors and the uncertainties that are estimated from this proposed model.

Acknowledgement

This research was supported by the National Science and Technology Development Agency, Thailand, Project no. FDA-CO-2560-4832-TH.

References

[1] A. Loungthongkam, C. Raksiri "A Development of Mathematical Model for Predictive of The Standard Uncertainty of Robot Arm" International Conference on Industrial Engineering and Applications (ICIEA)., 2020. doi: 10.1109/ICIEA49774.2020.9101979

[2] B. Panomruttanarung, W. Pornsukvitoon, J. Pakkawanit "A Study of Forward and Inverse Kinematics for 6-Link Robot Arm (Staubli RX90)" The Journal of KMUTNB., **27**(2), 241-252, 2017. DOI: 10.14416/j.kmutnb.2017.03.013

[3] K. G. Ahn, D. W. Cho "An analysis of the volumetric error uncertainty of a three-axis machine tool by beta distribution" International Journal of Machine Tools and Manufacture., **40**(15), 2235-2248, 2000. doi: 10.1016/S0890-6955(00)00048-1ff

[4] Z. Weidong, L. Guanhua, D. Huiyue, K. Yinglin "Positioning error compensation on two-dimensional manifold for robotic machining" Robotics and Computer Integrated Manufacturing., **59**, 394-405, 2019. doi: 10.1016/j.rcim.2019.05.013

[5] J.D. Barnfather, M.J. Goodfellow, T. Abram "Development and testing of an error compensation algorithm for photogrammetry assisted robotic machining" Measurement., **94**, 561-577, 2016. doi: 10.1016/ j.measurement.2016.08.032

[6] U. Schneider, M. Drust, A. Puzik, A. Verl "Compensation of Errors in Robot Machining with a Parallel 3D-Piezo Compensation Mechanis" Procedia CIRP., **7**, 305-310, 2013. doi: 10.1016/j.procir.2013 .05.052

[7] J.D. Barnfather, T. Abram "Efficient compensation of dimensional errors in robotic machining using imperfect point cloud part inspection data" Measurement., **117**, 176-185, 2018. doi: 10.1016/j.measurement. 2017 .12.021

[8] R. Li, Y. Zhao "Dynamic error compensation for industrial robot based on thermal effect model" Measurement., **88**, 113-120, 2016. doi: 10.1016/j.measurement.2016.02.038

[9] P. Roopyai, B. Panomruttanarung, "Positioning Error Reduction in Robotic Manipulator SEIKO D-TRAN RT3200 Using Repetitive Control" The Journal of KMUTNB., **28**(2), 299-312, 2018. DOI: 10.14416/j.kmutnb.2018.03.006

[10] Y. Zeng, W. Tian, W. Liao, "Positional error similarity analysis for error compensation of industrial robots" Robotics and Computer-Integrated Manufacturing., **42**, 113-120, 2016. doi: 10.1016/j.rcim.2016.05.011

[11] C. Raksiri, K. Pa-im, S. Rodkwan, "An Analysis of Joint Assembly Geometric Errors Affecting End-Effector for Six-Axis Robots" Robotics., **9**(2), 1-13, 2020. doi: 10.3390/robotics9020027

[12] L. Ma, P. Bazzoli, P. Sammons, R.G. Landers, "Modeling and calibration of high-order joint-dependent kinematic errors for industrial robots" Robotics and Computer-Integrated Manufacturing., **50**, 153-167, 2018. doi: 10.1016 /j.rcim.2017.09.006

[13] T. Dokyor, "An application of Michelson laser interferometer for roundness measurement" M.S. Thesis, Kasetsart University, 2008.

[14] A. Loungthongkam, "A study an application of Michelson laser interferometer for calibration block gauges" M.S. Thesis, Kasetsart University, 2011.

[15] A. Loungthongkam, "High precision technique of laser interferometer for warhead roundness measurement" Journal of Industrial and Intelligent Information., **3**(2), 158-162, 2015. doi: 10.12720/jiii.3.2.158-162

[16] J. Santolaria, "Uncertainty estimation in robot kinematic calibration" Robotics and Computer-Integrated Manufacturing., **29**(2), 370-384, 2013. doi: 10.1016/j.rcim.2012.09.007

[17] H.-N. Nguyen, J. Zhou, H.-J. Kang, "A New Full Pose Measurement Method for Robot Calibration" Sensors., **13**(7), 9132-9147, 2013. doi: 10.3390/s130709132

[18] A.A. Shabana, Dynamics of Multibody Systems, Cambridge University Press, 2013.

[19] W. Toojinda, Industrial Robot Analysis and Control, Chula Press, 2017.

[20] JCGM, Evaluation of measurement data – Guide to expression of uncertainty in measurement, Joint Committee for Guides in Metrology, 2008.

[21] United Kingdom Accreditation Service, The Expression of Uncertainty and Confidence in Measurement, M3003, 2019

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 51-57 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Green Blocks Made of Recycled Construction Waste using Recycled Wastewater

Elgaali Elgaali*, Adel Al Wazeer

Civil Engineering Program, Higher Colleges of Technology, Dubai Men's College, Dubai, 10000, UAE

ARTICLE INFO

Article history:
Received: 07 December, 2020
Accepted: 24 February, 2021
Online: 10 March, 2021

Keywords:
Sustainability
Waste
Recycling

ABSTRACT

This study tests the feasibility of manufacturing concrete blocks made of recycled materials. The paper is an extension of work originally presented in ASET conference in Dubai. The paper, depicts and analyzes how the characteristics of the blocks (strength/durability) are affected by the presence of recycled concrete ingredients (recycled aggregate (RA)) and recycled water (RW). The recycled materials (RA and RW) were mixed in 16 different configurations; from each one 10 samples were prepared for testing. In each concrete configuration the RA and RW gradually replaced the fresh materials at 25%, 50%, 75%, and 100%. The RA moderately impacted the bearing capacity but significantly impacted the durability. The results show that using recycled aggregate decreases the bearing capacity by 22% (at the 100% replacement), and the recycled water slightly affected the bearing capacity (at the 100% replacement). To boost the durability, the ground granulated blast furnace slag (GGBS) was used, in the concrete mix, instead of the ordinary Portland cement (OPC). The GGBS was used at 3 magnitudes: 25%, 50%, and 75% of OPC. As a result the carbon foot-print footprint (1000 kg/m^3) was significantly lowered. Besides, the strength and durability of the blocks are reasonably enhanced. Generally, producing blocks from recycled materials is economical and feasible. The use of GGBS helps to lower the carbon footprint and enhance the strength and durability.

1. Introduction

The volume of construction waste (CW) produced in the countries of the Gulf Cooperation Council (GCC), especially in the UAE, is increasing due to on-going developments activities and construction projects. Dubai alone produces huge amount of waste from construction activities (5,000 tons/day). This amount accounts for around 70% of the total solid waste generated [1]. Almost the whole amount of the CW is dumped into landfills posing hazards to the local and regional environment. Therefore, there is a necessity to devise methods for utilization of this huge amount of the construction waste produced [2].

Worldwide many studies have been conducted looking for alternatives to reduce construction waste and hence preserve the natural resources [3], [4]. Completely recycling construction waste, is one of best alternatives have been tested. Recycling of construction waste preserves the virgin natural resources and reduces its negative impacts on the environment. Construction waste recycling gains importance because it provides the construction industry with huge amount of ingredients for construction applications (aggregate and sand). The recycled aggregate and sand are widely used now in the construction industry applications; road subbase, back-filling, and pipelines bedding material [5], [6]. Significant amount of research efforts have been directed to study the properties of these materials (sand and aggregate) and their effects on construction quality [7]-[14]. In general, compared to the fresh aggregate, the RA was found to have higher water absorption, lower bulk density, and lower specific gravity. The water absorption is higher due to the fact that the RA absorbs water available for hydration of cement. It is reported that water absorption greatly affects the workability and the strength of the concrete [8]. Opposite to that, a recent research reported that with thorough cleaning the RA can have low water absorption [9]. However, the outcome of the most of the research done show that the properties of concrete made of recycled materials (recycled concrete, RC) vary in wide limits, sometimes are even opposite, but generally, compared to the normal fresh concrete (NC), the RC was found to have reduction in compressive strength and increase in water absorption and porosity (decrease in durability) [11]-[13]. Nonetheless, the mechanical properties of concrete structures (cubes, cylinders, and beams) made from recycled materials (RM) is well documented,

*Corresponding Author: Elgaali, P.O. Box: 15825, Dubai, UAE, 971 552493085, eelgaali@hct.ac.ae

however blocks made from RM is not yet widely known. Therefore, the main objective of this paper is to examine the properties (bearing capacity and durability) of concrete blocks made of recycled materials.

Globally and in the UAE, hollow blocks are widely used as a building material in the construction of non-load-bearing walls. Blocks are produced with hollow centers to reduce weight and improve insulation. They are characterized to be light weight and require low maintenance. The block industry is expected to absorb huge amount of the recycled materials.

This paper also aims to test the feasibility of using recycled wastewater in blocks' manufacturing. Dubai generates huge quantity of wastewater every day (around 4 MGD). Presently, this large quantity of wastewater is treated, recycled, and widely reused in different activities such as: irrigation of the public greeneries, stabilization of loose soil, and more recent casting of concrete [15].

For long time, fresh water has been widely used in concrete industry. Unfortunately, nowadays fresh water, around the globe, neither comfortably obtainable nor affordable. Therefore, recycling and reusing of wastewater in concrete industry have been under investigation for a while. According to EN 1008 (2002) and ASTM C 94 the recycled wastewater is well qualified to be used in concrete mixing (Table 1) [16], [17]. Besides, the cost of the recycled water is less than the one of the potable water, especially in this region where the potable water is produced by desalination of the sea water. The properties of concrete - setting, hardening, strength, and durability - were proven to be drastically affected by the characteristics of the water used in concrete casting. However, few studies so far have focused on studying the effect of the recycled water on the concrete quality [1]. As such, this research aims to contribute to the body of literature that examines how the RA and RW impact the quality (strength/durability) of blocks made of recycled materials.

Reducing the carbon foot print (CFP) of the concrete industry was one of the goals of this study. The relatively high carbon foot print (100 kg/m^3) that cement (OPC) put on environment is already documented [9]. One approach to alleviate such high CFP is by replacing the OPC with a material has lower carbon foot print. The ground granulated blast furnace slag (GGBS), is one of the materials that is widely used to replace the cement to lower the carbon footprint and enhance the durability [9]. The GGBS is a secondary product derived from manufacturing process of steel. Therefore, it is used here, as a green material, to make the concrete more durable and reduce the CFP. Hence, excluding the tiny amount of the OPC that was not replaced, the concrete mixtures configured in this study practically could be considered completely recycled concrete or environmentally called "green concrete'.

The recycled aggregates and water (RA and RW) were used to prepare concrete samples according to the mix design given in Table 2. Sixteen mixes of concrete were designed and 160 specimens were prepared. In the concrete mixes 25%, 50%, 75%, and 100% of the fresh aggregate and water were replaced by recycled aggregate and recycled water. In order to enhance the durability and carbon foot print 25%, 50%, and 75% of the OPC in the concrete mixes were replaced by GGBS. The results are presented and discussed in section 3 of this paper.

2. Materials and Methods

2.1. Recycled Aggregate Materials

The recycled aggregate used in this research, was obtained from a recycling plant for demolition and construction waste in Dubai, UAE. Nowadays, almost the whole amount of the construction and demolition waste is recycled in aggregate of high-quality for customers across the country Figure 1.



(a) Input Materials



(b) Output Material

Figure 1: Recycling of Construction Waste in Dubai, UAE

In order to determine the quality of the RA, several samples were collected (grabbed), processed, cleaned, and mixed in representative samples for laboratory testing. Then the following tests were conducted:

- Sieve Analysis.
- Water absorption and Specific gravity.
- LA abrasion.

Figure 2 shows the sieve analysis result. The figure shows that the RA does not meet the specifications set by Dubai Municipality. In order to use this materials, it has to be blended to meet the specifications.

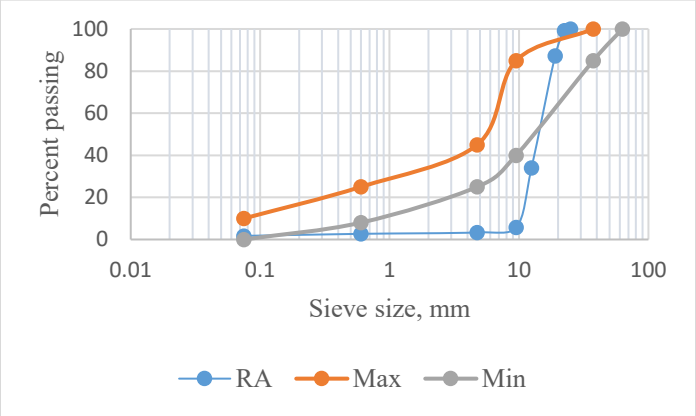


Figure 2: Sieve analysis results

Figure 3 shows the result of the water absorption and specific gravity test. Five representative samples were tested. The figure shows that the RA meets the specifications set by ASTM's construction standards. The values of water absorption are less than 2%; the values of the specific gravity are greater than 2.6% (the permissible limits set by ASTM).

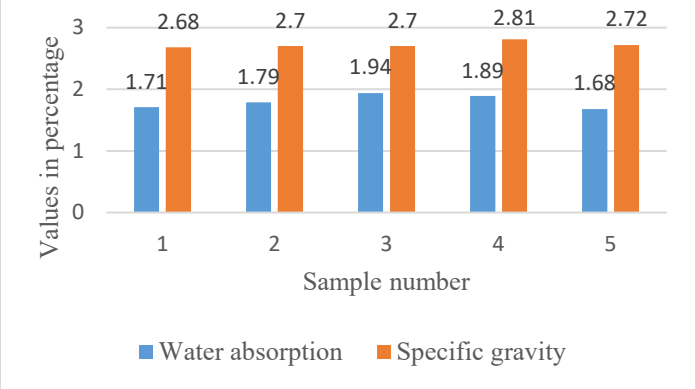


Figure 3: Water absorption and specific gravity test results

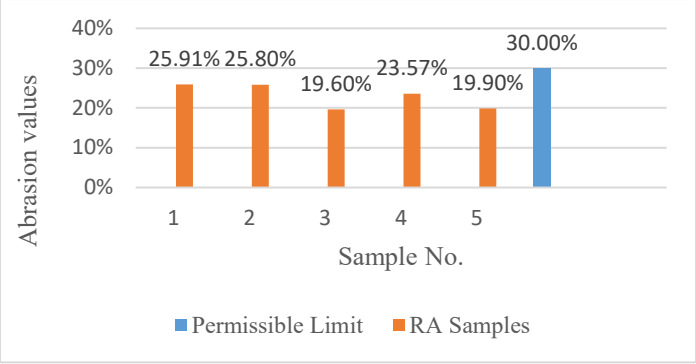


Figure 4: LA abrasion test results

Figure 4 shows the result of the LA abrasion test. Five representative samples were tested. The figure shows that the RA meets the specifications set by ASTM. The values of abrasion are less than 30% (the permissible limits set by ASTM).

However, the average measured density and water absorption of the recycled aggregate used in this study were found to be 2.51 t/m^3 and 1.8%, respectively.

2.2. Recycled Wastewater

The water used in the concrete mixing in this research, was obtained from sewage treatment plant in Dubai. The wastewater was conventionally treated. In the conventional way, the wastewater is treated through different stages starting from primary treatment to advanced (tertiary) treatment. The advanced treatment is designed mainly to remove the nitrogen and the phosphorus to control eutrophication [18]. Compared to the standards (Chloride 1000, Sulfate 2000, Alkali carbonates and bicarbonates 1000, and Total dissolved solids 2000 mg/L), the characteristics of the recycled water used in this study were found to be acceptable by the masonry blocks standards authority in Dubai [1]. Laboratory analysis results reported the following characteristics of the recycled water: pH 7.704 (standard usually ranges from 7 to 9), Chloride 480, Sulfate 197, Alkali carbonates and bicarbonates 132, and Total dissolved solids 1126.5 mg/L. Table 1, shows the physical and chemical properties of the treated water used in this research.

Table 1: Physical and Chemical Properties of Treated Wastewater

	Units	Min	Max	Average
Temp	(°C)	18	25	20
pH		7.59	7.84	7.704
Cond.	(uS/cm)	1,915	2,400	2203.82
TDS	(mg/l)	1,007	1,218	1126.5
Turbidity	(NTU)	0.6	2.9	1.7
TSS	(mg/l)	2	7	4
VSS	(mg/l)	1	6	2
COD	(mg/l)	23	33	27.0886
sCOD	(mg/l)	24.2	31.4	26.975
cBOD5	(mg/l)	1.64	3.92	2.31
NH4-N	(mg/l)	0.28	5.85	2.56
NO2-N	(mg/l)	0.59	2.16	1.37
NO3-N	(mg/l)	2.74	3.35	3.08
Total P	(mg/l)	3.61	4.61	4.21
Chlorides	(mg/l)	441	532	480
Sulfates	(mg/l)	185	204	197
T-ALK	(mg/l)	119	149	132

Total Hardness	(mg/l)	248	300	278

(Courtesy: Dubai Municipality, Dubai, 2017)

2.3. Mixes Design

The study was done in two phases: phase I where the concrete mixes were made of non-processed recycled materials (not sieved or cleaned). Phase II where the concrete mixes were made of processed materials (sieved and cleaned). However, the recycled materials were used to prepare four sets of concrete mixes according to the mix design given in Table 2. Each set consists of four mixes. In total sixteen mixes of concrete were prepared following the same proportions that used in the commercial manufacturing of the blocks. Usually, the blocks mix consist of the following quantities: coarse aggregate: 146 kg of 3/8" and 113 kg of 3/16"; Water: 20 kg; sand: 56 kg; and cementitious materials: 32 kg. As required by the central laboratories in Dubai, out of each mix, ten specimens (blocks) were prepared for testing (160 specimens in total). The blocks are 6-inch hollow blocks (400 X 150 X 200 mm) (Figure 5). This size is selected, because it is the most widely used one in the UAE.

Table 2: Test Matrix

Material Substituted	Percentage			
Recycled Aggregates	25%	50%	75%	100%
Recycled Water	25%	50%	75%	100%
Recycled Water/Aggregate	25%	50%	75%	100%
GGBS	25%	50%	75%	100%

(The GGBS was replaced in a mix of 100% recycled water and 25% of recycled aggregate)

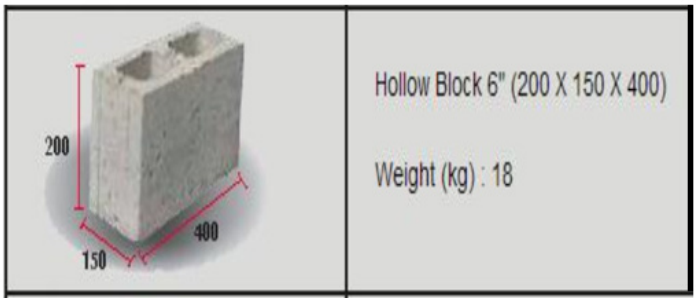


Figure 5: 6-inch hollow block dimensions and weight

In the concrete mixes the RA and RW gradually replace the fresh ones at the following percentages: 25%, 50%, 75%, and 100%. Table 2 shows these replacements more clearly in details. In order to determine the extent of the impact of the recycled materials on the characteristics of the blocks, a mix was prepared from fresh materials and used as a control mix (baseline).

The bearing capacity of each block was tested on a 3000 kN capacity Wizard basic hydraulic machine (Figure 6). Load was applied gradually and continuously until failure.



Figure 6: Testing the block's bearing capacity

3. Results and Discussion

3.1. Effects of Recycled Aggregate

In Phase I, the results show a gradual reduction the axial strength of the block corresponding to the gradual increase in the percentage of the recycled aggregate. It is depicted in Figure 7 that the failure in the bearing capacity (strength) of the blocks ranges from 23% to 46%. The highest reduction was recorded to be associated with the blocks made of 75% replacement. It worth mentioning that the high contents of sand in the RA used, led to the failure of the mix with 100% replacement at early stage (casting stage). However, the results presented are highly supported by some other studies conducted previously [8]-[10]. They documented that the gradual reduction in the axial strength of the mixes directly related to the increase of the percentage of RA. They also depicted that any replacement of RA below 20% has minor effect on the axial strength.

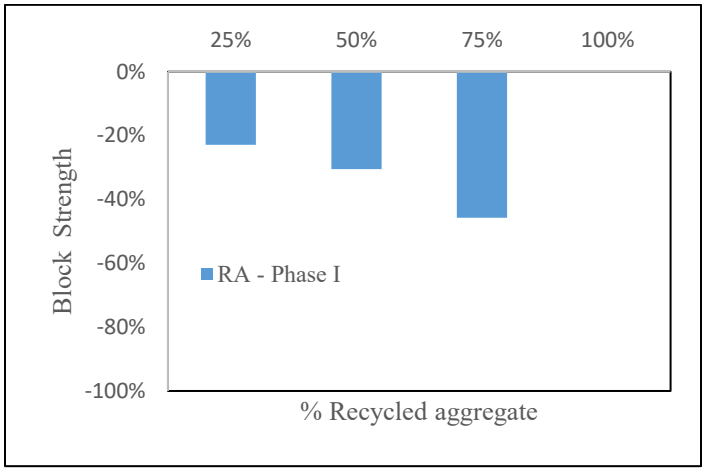


Figure 7: Effect of % recycled aggregate on bearing capacity

In phase II, the results support that the gradual reduction the axial strength of the block corresponding to the gradual increase in the percentage of the recycled aggregate. For better illustration, the results of both phase I and Phase II are compared in Figure 8. The figure shows how the RA impacts the compressive strength of the blocks. The results of phase I show how the presence of sand and dust in the samples of the RA used, led to a significant reduction in the compressive strength extending from 23% to 46%. The results of Phase II show considerable improvement in the blocks quality (Figure 8). At 25% replacement, the block's quality improved from 23% drop to 9% drop in in the strength; at 100% replacement quality of the blocks improved from 100% drop to 22% drop in the strength. In general the magnitude of enhancement in strength extends from 14% to 88%. The enhancement in the blocks' quality is mainly caused by the deep cleaning the RA undergone, in phase II, which results in high removal of sand and dust.

In general, the results presented confirm the inverse relation that governs the magnitude of the block's strength and the percentage of the recycled aggregate (RA).

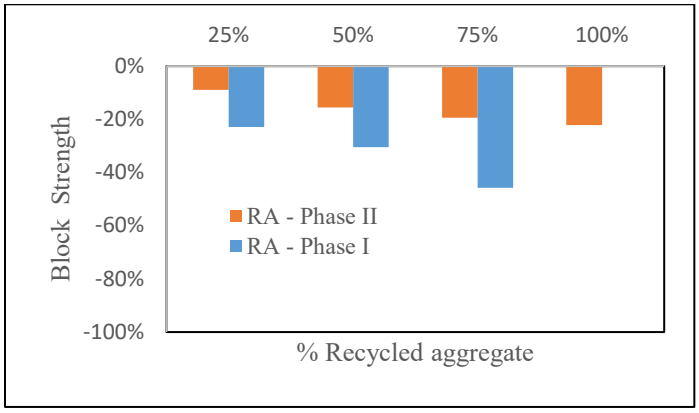


Figure 8: Effect of % recycled aggregate on bearing capacity

3.2. Effect of Recycled Sand

In Phase I, the results depict how the block's strength was drastically dropped when the percentage of the recycled sand (RS) increased. It is seen in Figure 9 that the failure in the bearing capacity (strength) of the blocks ranges from 90% to 95% when sand is replaced by 25% and 50% respectively.

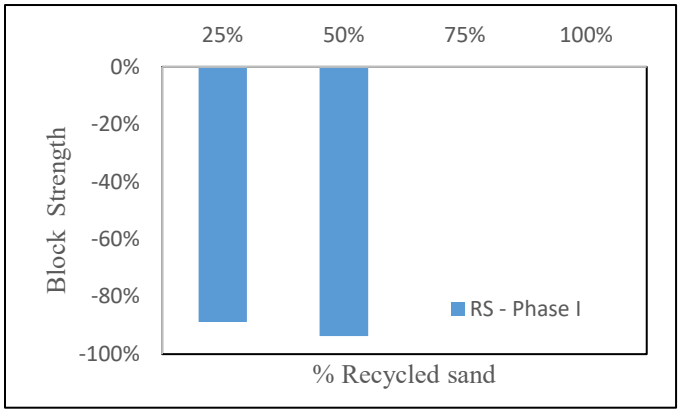


Figure 9: Effect of % recycled sand on bearing capacity

The presence of clay and silt at high levels in the recycled sand used in this study, caused the mixes with 75% and 100% recycled sand fail at early stage (casting stage). This result is supported by previous research [9].

The results of phase II show how the quality of the RS hindering the process of making sustainable blocks from completely recycled materials. For better illustration, the results of both phase I and Phase II are compared in Figure 10. The figure shows how the RS impacts the compressive strength of the blocks. The results of phase I show how the presence of clay and silt in the samples of the RS used, led to a significant reduction in the compressive strength. The results of Phase II show slight improvement in the blocks quality (Figure 10). At 25% replacement, the block's quality improved from 90% drop to 43% drop in in the strength; at 50% replacement quality of the blocks improved from 95% drop to 69% drop in the strength. In general, the enhancement in the blocks' quality is of small magnitude signifying that using RS at any percentage leads to lower the blocks' quality.

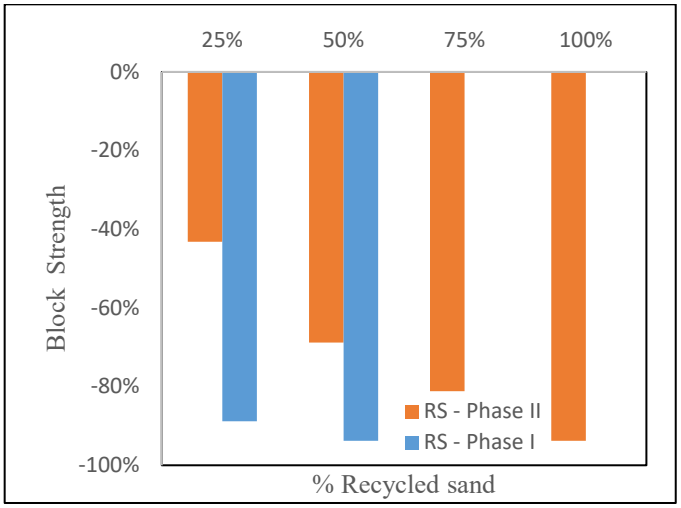


Figure 10: Effect of % recycled sand on bearing capacity

3.3. Effect of Recycled Wastewater

Figure 11 shows that replacing the fresh water in concrete mixes with 100% of RW has very minor impact on the block's quality (bearing capacity). This result signifies the high quality of the RW that characterized by low concentrations of sulfates and chlorides. The deleterious effects of the sulfates and chlorides on the weathering and the durability of concrete are well documented. The high concentrations of the total dissolved solids (TDS) in any water, used for concrete mixing, tend to make the concrete mix less durable [19]. However, the main properties of the RW water used in this study, are presented here for better understanding of the direction of the results in this section. The properties include pH of 7.704, Chloride of 480 mg/l, sulfate of 197 mg/l, Alkali carbonates and bicarbonates of 132 mg/l, and Total dissolved solids of 1126.5 mg/l; compared to standard values of 7 – 9, 1000, 2000, 1000, and 2000 respectively. The results presented, clearly show that the RW is characterized by high quality and can be safely used in concrete industry.

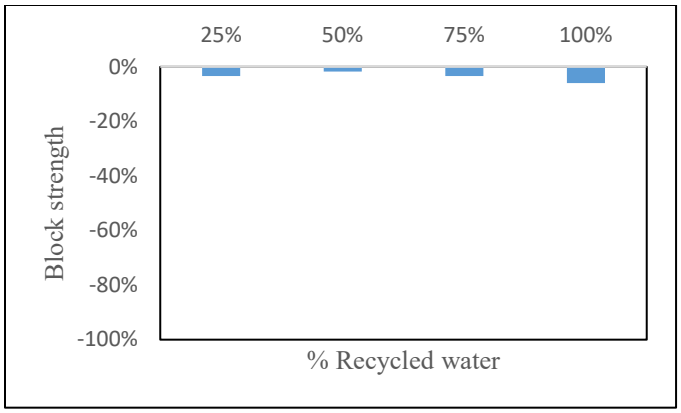


Figure 11: Effect of percentage of recycled water on block's bearing capacity

3.4. Effects of Sulfates and Chlorides

To study the effects of salts on the blocks' quality, the concentrations of the sulfates (SO_3) and chlorides (Cl) were determined and compared to the standards developed by the British standards (BS 1881: Part 124). In 10 specimens the average concentrations of the chlorides and sulfates are determined to be 0.04 mg/l and 0.25 mg/l respectively. The BS set maximum concentration level (MCL) of 0.05%, by mas of concrete, for the chlorides, and 0.5%, by mass of concrete, for the sulfates. Thus, the result obtained here clearly show that the levels of both SO_3 and Cl do not exceed the standards set by the BS. However, both phases I and II show similar results regarding the levels of the chlorides and sulfates.

3.5. Effects of GGBS

Tests conducted to examine the effects of replacing the ordinary cement (OPC) by GGBS. In the concrete mix with 25% RA, 25%, 50%, and 75% of the OPC were replaced by GGBS. As shown in Figure 12, replacing the OPC with 75% GGBS had no effect on the durability, besides no improvement had been shown on the block strength.

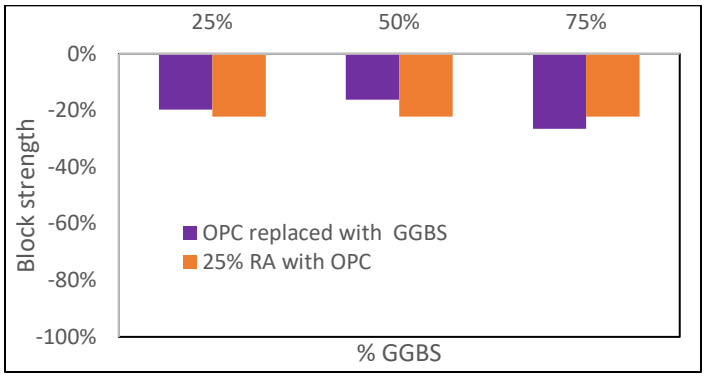


Figure 12: Effect of percentage of GGBS on block's bearing capacity

The other two mixes (25% and 50%) both showed improvement in the block strength. As shown (Figure 12), the quality of the blocks improved and the reduction in the axial strength decreased from 22% to 19% (at 25% GGBS) and from 22% to 16% (at 50% GGBS). As reported, the magnitude of improvement falls in the range of 3% - 6%. However, the result show that when the GGBS is more than 50% its influence is negligible.

In general, GGBS was found to boost the durability in any concrete mix. Therefore, to manufacture blocks with high quality, from recycled materials, the blocks industry is encouraged to use GGBS. Besides, the economical (lower cost) and environmental (lower carbon foot print) benefits obtained when GGBS is used instead of OPC.

4. Summary and Conclusions

This study tests how concrete blocks of high quality could be manufactured from ingredients recycled from construction waste (concrete). The paper, depicts and analyzes how the characteristics of the blocks (strength/durability) are affected by the presence of recycled concrete ingredients (recycled aggregate (RA)) and recycled water (RW). The RA moderately impacts the bearing capacity but significantly impacts the durability, while the RW negligibly impacts both. In general it was found that producing blocks from recycled materials is economical and feasible. However, the following specific conclusions are drawn:

4.1. Recycled Aggregate

How the RA affects the qualities of the blocks (strength/durability) was investigated and depicted. The presence of dust on the surface of the RA led to significant drop in the blocks qualities (strength/durability). The magnitude of the drop was estimated to fall in the range of 23% to 46%. When the RA is furtherly sieved, processed, and cleaned the blocks' quality improved a lot. The magnitude of the improvement was estimated to fall in the range of 14% - 63%. It is documented that using RA at magnitude less than 20% has intangible impact on the quality of the blocks.

4.2. Recycled Sand

How the recycled sand (RS) affects the qualities of the blocks (strength/durability) was investigated and depicted. The impacts of RS on blocks; quality was enormous. The impact of RS is very huge ranges between 43% - 95% reduction in the compressive strength. The RS was found to have stronger impact on the blocks' quality when compared to RA and RW. It is documented that using RS at any quantity heavily affects the blocks' quality.

4.3. Recycled Wastewater

How the recycled water (RW) affects the qualities of the blocks (strength/durability) was investigated and depicted. The RW used, is of high quality that matches the standards set by the international and local (Dubai) organizations that manage the masonry blocks industry. The very low levels of the sulfates and chlorides caused the quality of the blocks to be insensitive to the recycled water. However, it can be stated that the high quality of the RW almost has no impact on the blocks' quality (strength and durability). It is documented that using RW at magnitude less or equal to 100% has intangible impact on the quality of the blocks.

4.4. Sulfates and Chlorides

The concentrations of the sulfates (SO_3) and chlorides (Cl) in the blocks, were determined and compared to the standards developed by the international (BS) and local (Dubai)

organizations that manage the masonry blocks industry. The levels of the SO_3 and CL was found to meet the standards set internationally and locally. However, the current levels (0.04 mg/l and 0.25 mg/) have no impacts on the blocks quality.

4.5. GGBS

Using the ordinary cement (OPC) in concrete mixes, adversely affects the local and regional environment (increase the carbon footprint). Also, using recycled materials in concrete mixes has huge impact on the ddurability. Therefore, GGBS was tested to replace the OPC to preserve the blocks' quality (durability) and save the environment (reduce CFP). Replacing the OPC with GGBS by percentage exceeds 50% has no positive effect on the quality of the blocks. The outcome of this study documents that using GGBS, in block manufacturing process, helps to lower the carbon footprint and enhance the strength and durability.

Conflict of Interest

No conflict of interest in publishing this paper.

Acknowledgment

The authors would like acknowledge the contribution of Dubai Men's College (DMC) – Higher Colleges of technology, and the following students: Abdulla Taher, Hussain Habib, Mohammad Fardan, Saif Nasser, and Zayed Almadhanni.

References

[1] E. Elgaali, A. Al Wazeer, "Green blocks made of recycled waste phase II," Advances in Science and Engineering Technology International Conference, ASET, Dubai, UAE, 1-5, 2020, DOI: 10.1109/ASET48392.2020.9118227.

[2] E. Elgaali, A. Alsharid, A. Julfar, "Green blocks made of recycled waste," Advances in Science and Engineering Technology International Conference, ASET, Dubai, UAE, 1-5, 2018, DOI: 10.1109/ICASET.2018.8376773

[3] A. Al-Hajj, K. Hamani, "Material waste in the UAE construction industry: main causes and minimization practices," Architectural Engineering and Design Management, **7**(4), 2011, doi: 10.1080/17452007.2011.594576.

[4] V. Tamab, M. Soomro, A. Evangelistadc, "A review of recycled aggregate in concrete applications (2000–2017)," Construction and Building Materials, 172, 272-292, 2018, doi: 10.1016/j.conbuildmat.2018.03.240

[5] R. OBE, J. Brito, R. Silva, C. Lye, " Use of recycled aggregates in geotechnical applications," Sustainable Construction Materials, 419-450, 2019, doi: 10.1016/B978-0-08-100985-7.00011-X.

[6] V. Panchal, V. Kulkarni, A. Kulkarni, A. Kumar, "Use of construction demolition waste in pavement," International Journal of Advanced Research in Science, Engineering and Technology, **4**(12), 4956-4964, 2017, ISSN: 2350-0328.

[7] R. V. Silva, J. R. Jiménez, F. Agrela, J. de Brito, "Real-scale applications of recycled aggregate concrete," New Trends in Eco-efficient and Recycled Concrete, 573-589, 2019, doi: 10.1016/B978-0-08-102480-5.00021-X.

[8] Said Kenai, Waste and supplementary cementitious materials in concrete, Woodhead, 2018.

[9] E. Mohamed, E. Elgaali, "Sustainable concrete made of construction and demolition wastes using recycled wastewater in the UAE" Journal of Advanced Concrete Technology-Japan, 10, 110-125, 2012. DOI:10.3151/jac .10.110

[10] A. Hassan, A. Nabil, H. Abdulrazaq, F. Juma, M. Salem, Developing Design Criteria for Utilization of Construction and Demolition Wastes in Construction Applications, Thesis, Higher Colleges of Technology, 2019.

[11] C. Pellegrino, F. Faleschini, C. Meyer, "Recycled materials in concrete," Developments in the Formulation and Reinforcement of Concrete, 19-54, 2019, doi: 10.1016/B978-0-08-102616-8.00002-2.

[12] E. Ledesma, J. Jiménez, J. Ayuso, J. Rodriguez, "Maximum feasible use of recycled sand from construction and demolition waste for eco-mortar production - Part-I: Ceramic masonry waste," Journal of Cleaner Production, **87**(1), 692-706, 2015, DOI: 10.1016/j.jclepro.2014.10.084.

[13] P. Chindaprasirt, T. Cao, Reuse of recycled aggregate in the production of alkali-activated concrete. Handbook of Alkali-Activated Cements, Mortars and Concretes, Woodhead, 2015.

[14] M. Malešev,V. Radonjanin, S. Marinković, " Recycled concrete as aggregate for structural concrete production," Sustainability, **2**(5), 1204-1225, 2010 doi: 10.3390/su2051204

[15] C. Tortajada, "Contributions of recycled wastewater to clean water and sanitation sustainable development goals," npj Clean Water, 3, 22, 2019, doi: 10.1038/s41545-020-0069-3

[16] EN 1008, British Standards, Mixing water for concrete specification for sampling, testing and assessing the suitability of water, including water recovered from processes in concrete industry, as mixing water for concrete. 2002.

[17] ASTM C 94, Standard test specification for ready-mixed concrete, American Society for Testing and Materials, Philadelphia, 1994

[18] M. J. Hammer Sr., M. J. Hammer Jr., Water and Wastewater Technology, Prentice Hall, 2003.

[19] A. Sales, F. R. De Souza, "Concrete and mortars recycled with water treatment sludge construction and demolition rubble," Construction and building materials, 23, 2326-2370, 2009, doi:10.1016/j.conbuildmat.2008. .11.001

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 54-83 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Power Converters and EMS for Fuel Cells CCHP Applications: A Structural and Extended Review

Nganyang Paul Bayendang*, Mohamed Tariq Kahn, Vipin Balyan

Department of Electrical, Electronic and Computer Engineering (DEECE), Cape Peninsula University of Technology (CPUT), Bellville, Cape Town, 7535, South Africa

A R T I C L E I N F O

Article history:
Received: 25 December, 2020
Accepted: 19 April, 2021
Online: 12 May, 2021

Keywords:
CCHP System
Energy Management Systems
Fuel Cells
Hybrid Power
Power Converters
Renewable Energy

A B S T R A C T

Fuel Cells (FCs) and Combined Cooling, Heating and Power (CCHP) systems are becoming very popular due to their environmental friendliness and immense applications. This extended review paper commenced by introducing the rampant South Africa's electricity crisis as the basis for the study, followed by some structural analyses of up to forty-four miscellaneous power electronics converters case studies applicable to fuel cells including at least sixteen FCs energy management systems (EMS) applicable case studies. The review rationale is to determine innovative best practices that can be applied to devise an efficient power converter and EMS for an energy efficient FC CCHP system. From these analyses, it is realized that each power converter and EMS scheme has its merits and demerits depending on the targeted applications and most importantly the research project objectives − that is, whether the goals are to reduce costs, enhance efficiency, reduce size, boost performance, simplicity, durability, reliability, safety etc. Therefore, the conclusion drawn is, there is no "one size fits all" approach, as all the various reviewed case studies reported relatively good results based on their chosen schemes for their targeted applications. Notwithstanding, this review highlights are, i) the interleaved boost converter and variants as well as ii) the maximum power point tracking (MPPT) technique; are the most widely used schemes, as they are reasonably effective and simple to implement. The contributions brought forward are i) an apt single reference study that presents a quick topological insight and synopsis of assorted FCs power converters as well as EMS and ii) our proffered FC CCHP system undergoing research to offer an innovative energy efficient solution for basic household energy needs such as electricity, heating, cooling and lighting.

1. Introduction

Compounded with re-occurring electrical energy and power problems in South Africa and at large Africa; this extended article to [1], extensively investigates with emphasis on fuel cells (FCs), some assorted research on power converters and energy management systems / storage (EMSs) techniques. Currently, there's up to 2 hours of daily electricity rolling blackouts in South Africa (RSA), due to RSA national energy utility company (ESKOM) inability to generate enough energy to meet its local electricity demands. This is due to the legacy apartheid energy system being over-stretched with more underprivileged areas / users now having access to electricity, old energy infrastructures being upgraded, poorly designed and in-efficient new energy infrastructures, poor technical maintenance, inadequate technical abilities, corruption as well as illicit business and political reasons.

Thus, various optional renewable energy sources, especially solar and wind powers are being commissioned to augment and or stabilize RSA national grid supply and also for personal use. Alternative energy sources such as i) FCs − which produce power and heat as well as water when fueled with H_2 and O_2 and ii) thermo-electricity − which simply generates a) electricity based-on Seebeck effect and b) heat / cold based-on Peltier effect, are of interests. In this regard, our research focuses on FCs and thermo-electricity; however, these clean energy sources need supporting technologies and techniques to operate well. In light of this, we review applicable best practices that can be developed to execute an energy efficient fuel cell alternative power / energy system for domestic and commercial combined cooling, heating and power (CCHP) applications − since electricity, heat / cold and light are the most commonly used forms of energy in most homes in RSA. FCs CCHP systems are versatile, clean and becoming very trendy;

*Corresponding Author: NP Bayendang, CPUT DEECE, +27765404896, bayendangn@cput.ac.za

https://dx.doi.org/10.25046/aj060308

thus, discussed next are miscellaneous power converters and EMS research applicable to FCs and by extension FC CCHP systems.

A fuel cell stack closed-circuit voltage is a function of the FC activation, concentration and Ohmic losses governed by the Nernst equation. This FC loaded condition draws more current resulting to a voltage dip/drop as a result of delay of Hydrogen / Oxygen flow and this phenomenon is called the fuel cell fuel starvation. This voltage drop and the load current together with other FC parameters, establish the degree and duration of the problem. This issue is tackled diversely as analyzed briefly in [1].

In [1]-[36], DC-DC power converters are paramount to either boost and or reduce DC power sources (e.g FCs) and then sustain a consistent power thereafter. FCs typically produce low DC voltage but high current, which combined with the possibility to fluctuate when connected to a load, demands power regulations to the DC bus and ultimately to the different DC and or AC loads.

In DC-DC power converters, with the exception of linear regulators, the fundamental three switching kinds are the step-down, step-up and step-up / step-down − from which various and or improved versions are derived to give isolated derivatives with one or multiple switches including the soft-switching versions.

The voltage output from step-down DC-DC power converters, is always lesser than its voltage input. Contrarily, the voltage output from a step-up DC-DC converter is always more than its voltage input. The step-down / step-up power converter based-on its duty cycle value; can respectively either lower (duty cycle < 0.5) or increment (duty cycle > 0.5) the converter voltage output.

As per [37]-[70], EMSs are essential to manage the power converters to ensure maximum power is transferred to the load, the bus is stable and the energy / power supply system is efficient.

2. Power Converters

Power converters are required to buck, boost and provide regulated voltage to the DC bus. In what follows and summarized in Table 1, are miscellaneous case studies that structurally analyzed in brief details, some power converters for use with FCs.

2.1. DC-DC Power Converters Architectures for Fuel Cells Applications

As presented in [23], power sources based on fuel cells are now trendy devices. They offer reliability, flexibility as well as efficiency through multi-stack topologies. To access the market requires simplifying further the FCs design and its supporting components, which among others include the power converters which ensure the output voltage is regulated. Their research thus focused on DC-DC power converters by giving an inclusive outline on the interfaces of power converters for use in aircraft, railways, automotive and small static areas such as households.

The significance of selecting the correct power converters topology and the related technology is crucial, as its facets allow thermal compatibility with various methods for integrating the DC-DC power converters to the fuel cells. These topological and technological features that have been examined and displayed in Figures 1a -1f with highlights in Figures 1e -1f, are some popular power converters topologies. In their study, they indicated how connecting a fuel cell stack / DC-DC power converters in parallel and in series increase the current and voltage outputs respectively.

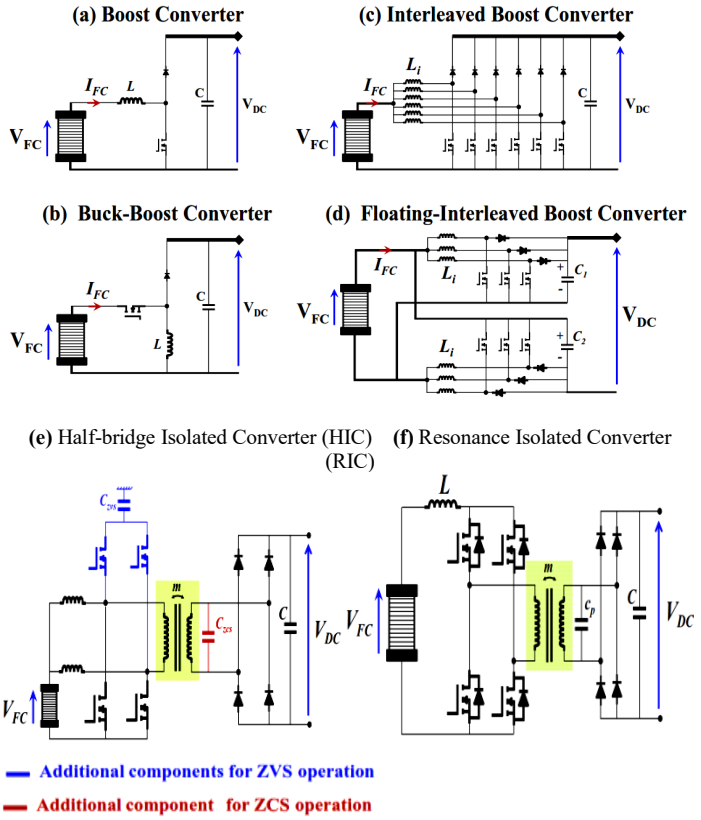


Figure 1: DC-DC Power Converters Architectures for Fuel Cells Applications (adapted from [23])

Explained in their research and portrayed in Figures 1c and 1d, are non-isolated multi-phase boost converters, which are mainly appropriate for applications that require low DC bus voltage. The interleaved topologies shown, meet the prerequisite for curbing low FC ripple currents. The depicted standard interleaved boost converter (IBC) shown in Figure 1c and the floating interleaved converter (FIC) in Figure 1d, show similar merits. Z-sources inverters (ZSI) were also articulated, in which their features and merits make them suitable choices for 3-phase electric drives − for instance automotive / railway applications. Furthermore, the study indicated that the isolated converters based-on high frequency planar transformer (which according to [23], only one quantity was left in the market in 2014), is beneficial in high DC link voltage applications such as railway. Contrarily, the isolated converters give a low efficiency for medium power applications. However, the soft-switching function enables the enhancement of the converter efficiency but at the cost of using supplementary components in the converter configurations. These improvements are shown in Figures 1e and 1f − whereby the half-bridge isolated converter (HIC) and the current-fed full-bridge resonant isolated converter (RIC), respectively illustrate the zero-voltage switching (ZVS) and zero current switching (ZCS) operations, in which both increase the efficiency by reducing the devices switching losses. The technological review on the other hand focused on the new wideband-gap semiconductor materials and the utilization of Silicon Carbide (SiC) and Gallium Nitride (GaN) devices with low

on-resistance, high power densities and high speed switching with less losses, which could transcend to major improvements in the power converters performance. Nowadays, GaN devices are suited for low/mid power applications, whereas SiC technology is more desirable for designing high power FCs DC-DC converters.

2.2. State-of-the-Art Fuel Cells DC-DC Converters

According to [13], fuel cells are current intensive sources that have become popular. The study presented various suitable topologies of DC-DC power converters for FCs output voltage conditioning. The goal of the main DC-DC power converter between the FC and DC bus was emphasized, which demands the power converter be designed/operated using high efficiency, high voltage ratio and high density with low-cost. As a result, their paper highlighted some tips in this regard as well as the positives and negatives. Some of the schemes are shown in Figures 2a - 2c.

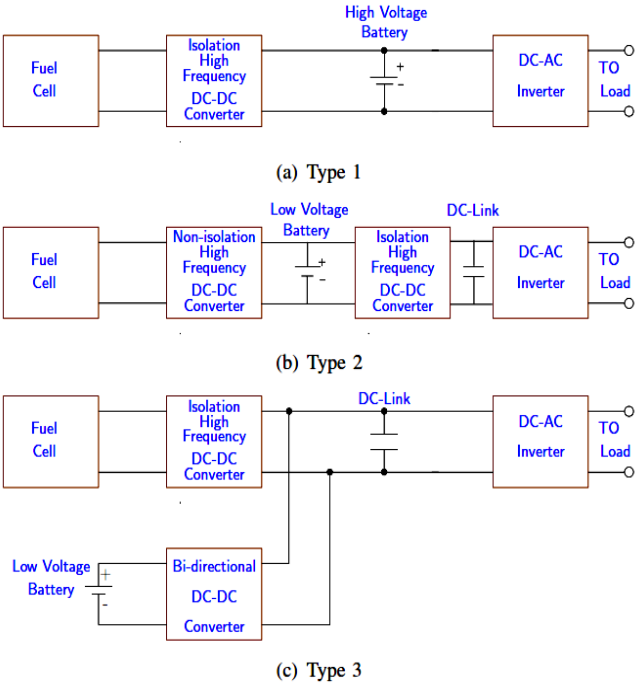


Figure 2: Fuel Cells DC-DC Power Converters (adapted from [13])

2.3. A Soft Switched Push-pull Current-fed Converter for FC Applications

Examined in [14], a zero-voltage switching (ZVS) current-fed push pull DC–DC converter is suggested for fuel cells power generation system as pictured in Figure 3. In the study, the auxiliary circuit in this converter supplies ZVS condition for all converter switches which reduces switching losses and further absorbs at turnoff instances, the voltage surge across the switches.

This merit, enhances the converter efficiency and reduces the size and converter weight – which henceforth enables the implementation of a very simple control circuit based on pulse width modulation (PWM). This setup was then used to analyse and validate the operation of the converter using a lab prototype. The projected DC-DC power converter experimental results, the various operating modes as well as their corresponding timing waveforms, are presented in detailed in their published paper.

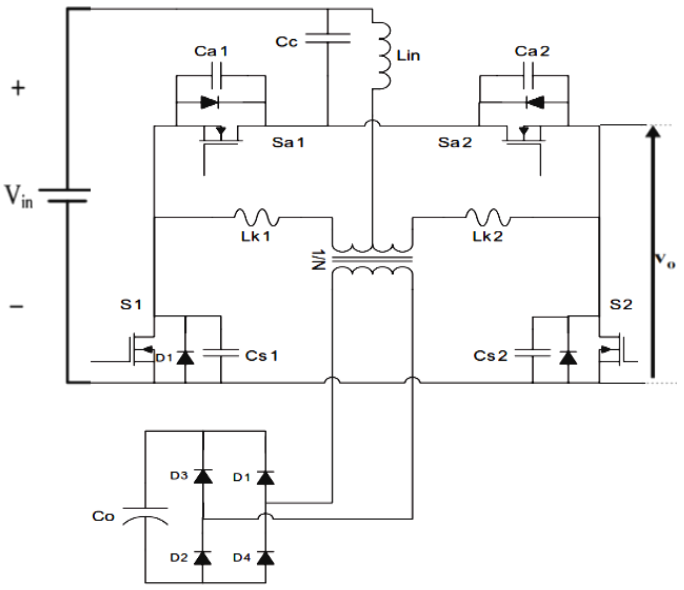


Figure 3: The Proposed Soft Switched Push-pull Current-fed Converter (adapted from [14])

2.4. Topology of FC Hybrid Power Source for Efficient Operation and High Reliability

Proposed in [15], is a fuel cell hybrid power source (HPS) topology with the attribute to curb the ripple current of the FC inverter system. The ripple current usually occurs at the DC port of the FC HPS when operating the inverter system − which is connected to the grid or which supplies AC motors in vehicular applications. As a result, if the alleviation measures are not implemented, this ripple current is propagated back to the FC stack. The suggested FC HPS has other good performance features; such as the maximum power point tracking (MPPT), high steadfastness in operation during transient power pulses and finally improved energy efficiency in peak power applications.

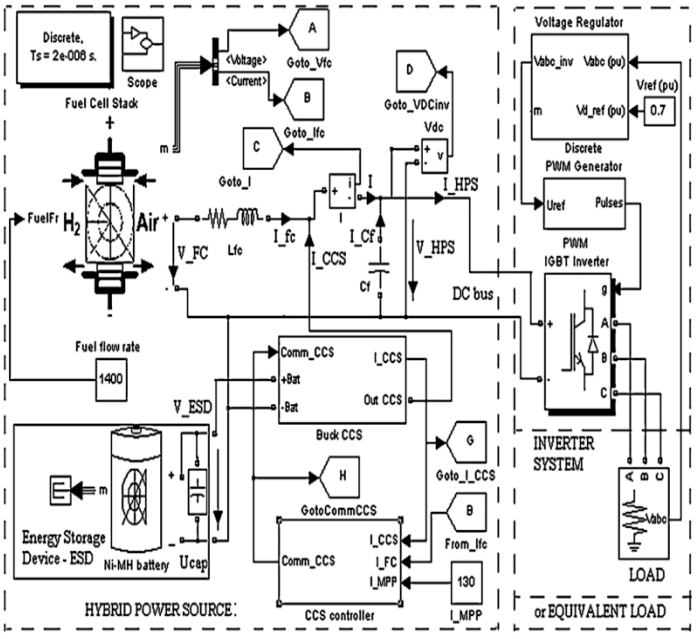


Figure 4: The Proposed Hybrid Power Source Topology (adapted from [15])

To mitigate the ripple, this approach made use of an inverter system powered directly from the FC stack with a controlled buck current source that was used as the low power source. The low frequency (LF) ripple reduction is rooted in active control, whereby the anti-ripple current is injected in HPS output node and this has the LF power spectrum similar to the inverter ripple. In light of this, the fuel cell ripple current was curbed by the designed active control. Indicators defined to evaluate the mitigation ratio of the LF harmonics were used to assess the ripple current alleviation performances. The relatively good performances shown were attained with the use of a hysteresis current control, but better if a devoted nonlinear controller is used, which can be designed in two ways as follows i) simulation trials that assist to draw the attributes of ripple mitigation ratio verse fuel cell ripple current and ii) fuzzy logic controller (FLC). The ripple factor was ~ 1% in both cases. Figure 4 depicts the study.

2.5. Power Flow Control Methods for Ultra-capacitor Bidirectional Converter in DC Micro-grid

Postulated in [16], distributed generation (DG) in the form of DC micro-grids depicted in Figure 5a, has recently attracted increasing research interests. A bidirectional DC-DC converter (BDC) shown in Figure 5b, is required to incorporate renewable energy resources and energy storage devices such as an ultra-capacitor (UC) to the DC bus of a DC micro-grid to sustain the charging and discharging states of the ultra-capacitor. For the quick dynamic response of the ultra-capacitor, a bidirectional voltage-fed setup is suitable, though for a broad input voltage fluctuation of the ultra-capacitor, this setup manifests a greater circulating power flow and greater conduction losses in the end. Presented in this study are a comprehensive overview on the numerous modulation schemes that are employed to manage the power flow of the bidirectional voltage-fed DC-DC converter for the ultra-capacitor applications. An in-depth analysis of the bidirectional converter investigating the impact of the circulating power flow interval was developed and analytical methods such as the conventional phase-shift control (CPC) modulation were applied to develop alternative modulation schemes to advance the efficiency and performance of the bidirectional power converter.

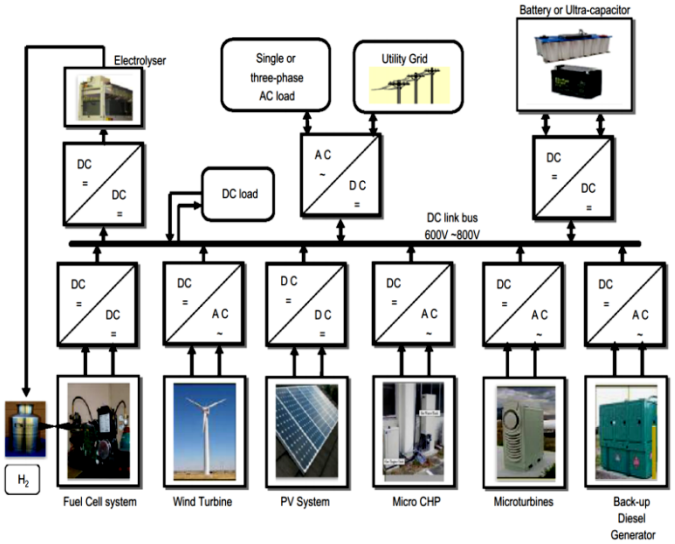


Figure 5 (a): DC Micro-grid Distributed Generation (adapted from [16])

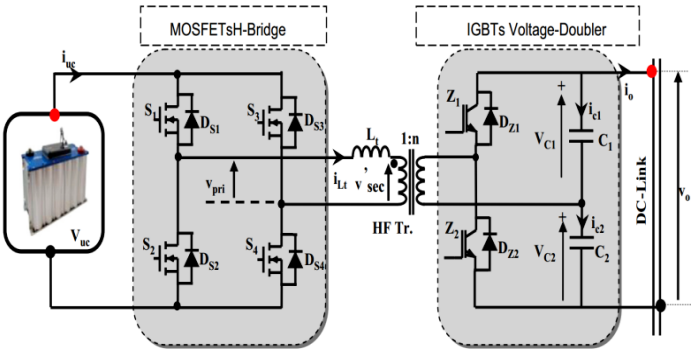


Figure 5 (b): BDC Power Flow Control Methods for Ultra-capacitor in DC Micro-grid (adapted from [16])

2.6. Fuel Cell and Power Converter Models in MatLab / Simulink

Studied in [19], MatLab and Simulink were employed to model a power converter and PEM FC. Depicted in Figure 6, the first section of the research discussed the methodology for an accurate model for the fuel cell stack, as well as its static and dynamic behaviors − which form a crucial aspect in the design of electrical power generation founded on fuel cells. The technique applied was simulated annealing (SA) optimization algorithm, which justifies its customization to meet the goal of a speedy convergence to institute the correct values for the fuel cell parameters. The correlation between the simulation and the experimental results proved that the suggested model provided an accurate depiction of the static and dynamic behaviors for the PEM FC. The second section of the study engaged on feasible architectures that can be tailored for the DC-DC power converter.

The preferred topology must be suitable to take control and optimize the operation point of the fuel cell; as a result, the soft switching attests to be particularly fitting, especially the series resonant topology converters − because it reduces the switching losses and consequently increasing the efficiency. This converter execution can be explained as follows: the supplied voltage by the stack, which is normally low (29V - 42V), is changed to a constant and high amplitude, in this case, a 400VDC bus is used to generate power to the grid via an inverter. The high frequency (HF) transformer is a boost voltage transformer, which is as well used to offer galvanic isolation between the low and high voltage levels of the circuits. In the primary side of the transformer, the resonant converter with its inductor-capacitor (LC) series resonant circuit, provides the sinusoidal waveforms of voltage and current. The circuit resonant frequency is determined by choosing suitable values for the L and C elements, from which the FC DC voltage is initially inverted to AC in the primary side of the HF transformer and then rectified to DC on the secondary side. The PEM FC is protected from the ripple voltage and current the converter produces by the LC filter in the primary side, which as well stores the DC bus energy. The secondary side LC filter reduces the ripple voltage and current to the load. In conclusion, the simulation results were correlated using actual data acquired from a commercial system. As a result, it was justified that, the hybridization of a suitable power converter using a well-defined controller in conjunction with a well-optimized FC stack model, makes FC good for power generation.

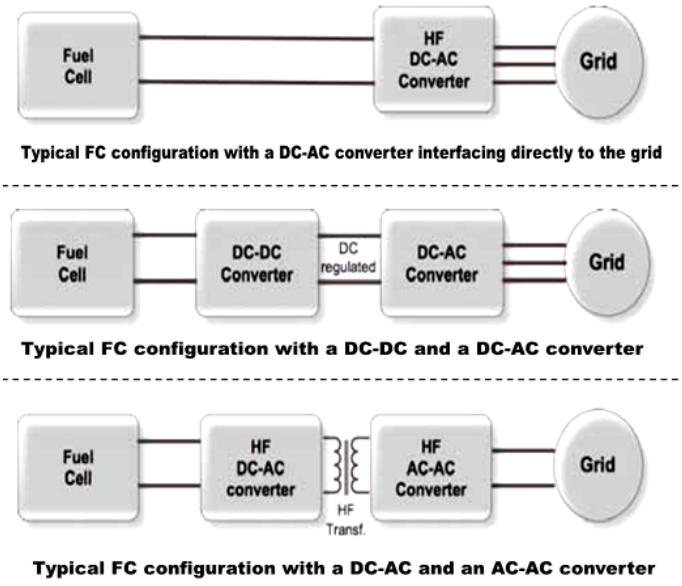


Figure 6: FC Power Converter Models in MatLab / Simulink (adapted from [19])

2.7. High Voltage DC-DC Boost Converter Suitable for Varying DC Voltage Sources

Researched in [11], is a high voltage step-up converter appropriate for varying voltage sources such as photovoltaic (PV) and by extension fuel cells as well as thermo-electric generators (TEGs). Different varying voltage boost sources were assessed to institutes their limits, from which a multi-phase tapped-coupled inductor DC-DC boost converter that can attain high voltage boost ratios from a variable power supply (PV in this case) and without adversely compromising the performance, was then postulated as pictured in Figure 7. The suggested converter achieved minimal voltage and current ripples at both the input and output as well as exhibited reasonable performance at high power levels making it preferred for high power applications. The simulated and practical results correlated to confirm the research.

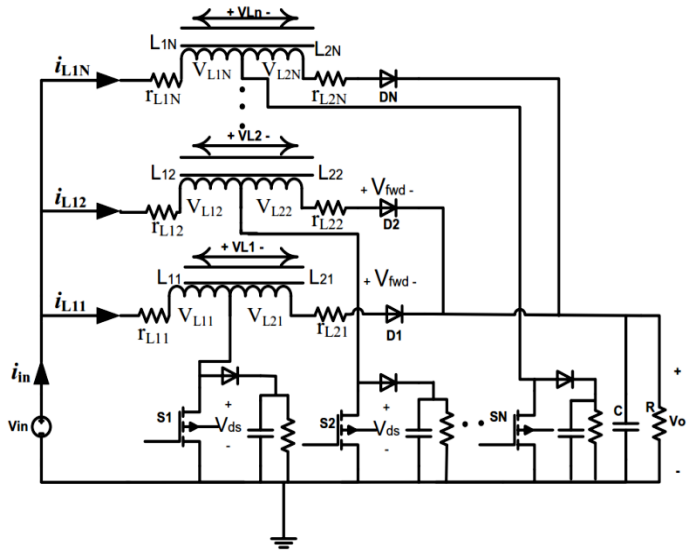


Figure 7: High Voltage DC-DC Boost Converter Suitable for Varying Voltage Sources (adapted from [11])

2.8. High Power Efficient DC-DC Buck Converter Suitable for Varying Voltage Sources

According to [10], in a varying power generation source such as wind (likewise solar-cells and TEGs as well as fuel cells), the power converter efficiency is one of the crucial aspects for the performance of the system. In such systems, the DC-DC step-down converter is usually used for high power systems. Taking into account the cost and efficiency of a converter, their research focus was mainly on the devise of enhanced buck converter topologies with interest on the (inductor, capacitor and diode) LCD converter depicted in Figure 8a – which is to be used for a peak power standalone wind power generation system. A (resistor, capacitor and diode) RCD and an improved RCD buck converter could also remove the voltage spikes; however, it unfortunately further depletes the stored voltage amplitude when the power is switched-off − this is because C_1 discharges the voltage stored through R_1. Therefore, the need for a LCD version.

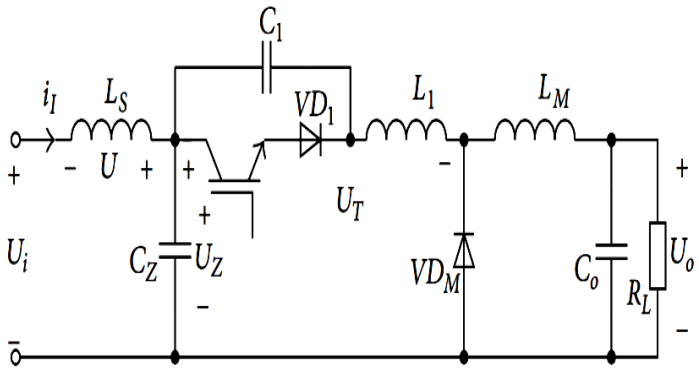


Figure 8 (a): High Power Efficiency DC-DC Buck Converter Suitable for Varying Voltage Sources (adapted from [10])

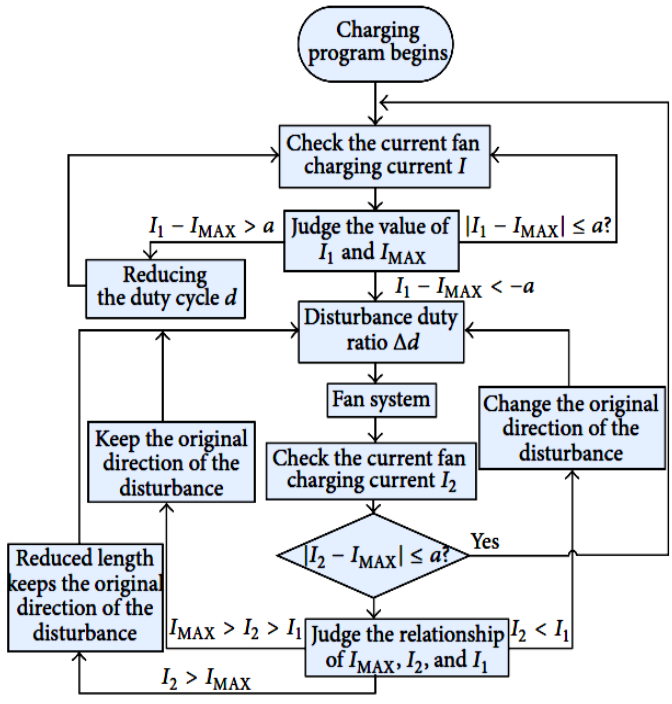


Figure 8 (b): DC-DC Buck Converter Climbing Mountain MTTP Algorithm Flowchart for Varying Voltage Sources (adapted from [10])

This issue was addressed by using a better improved LCD buck converter as shown in Figure 8a. This version does not have any resistor but an inductor with negligible energy loss as displayed in the test result. This LCD architecture is also known as zero-voltage switching (ZVS). At resonance, L_1 and C_1 cancel out, reducing the voltage spike and increasing the switching speed. The practical designed LCD converter is based-on Figure 8a and instead of IGBTs, it uses multi-MOSFETs in parallel to boost the current and the switching speed of the converter. By employing MPPT, the practical utmost efficiency of the designed 2kW DC - DC step-down converter was approximately 96%.

MPPT algorithm ensures the maximum power generated stays constant by monitoring the desired reference output with the generated output and adjusting the duty cycle or PWM signal to the active switch(es) of the power converter. The common MPPT techniques according to [10], includes i) optimum tip speed ratio, ii) power curve control and iii) climbing mountain – the latter was used in the study and its flowchart is illustrated in Figure 8 b.

2.9. High Gain IBC for Fuel Cells Applications

As researched by [9], distributed generation most capable technologies is fuel cell and to design a high efficiency power system using fuel cell, a fitting DC-DC converter is necessary. Among the different DC-DC converters, interleaved converters with switched capacitor are considered a preferred topology for FC systems because of reduced ripple currents in the input and output circuits, quicker transient reaction, small electromagnetic emissions, enhanced efficiency and reliability. This improved conversion efficiency is attained by dividing the output current into 'n' parts, to significantly eliminate I^2R losses and inductor losses. The research aim was to devise and implement a high gain interleaved boost converter based-on switched capacitors (to improve converter voltage gain) for fuel cell systems. In the interleaved converter proposed, the front-end inductors are magnetically cross-coupled to enhance the electrical performance and reduce the weight and size. By using switched capacitors interfaced with FCs, MatLab and Simulink were used to simulate an interleaved converter, from which a prototype was developed to validate the simulation results. Figure 9 depicts the IBC.

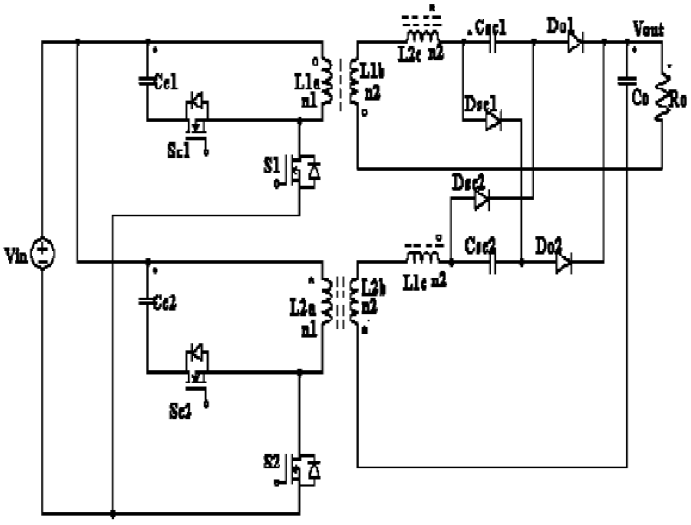


Figure 9: High Gain IBC for Fuel Cells Applications (adapted from [9])

2.10. High Efficiency Isolated Boost Converters for High-Power Low-Voltage FC Uses

As investigated in [8], fuel cells power systems, as portrayed in Figure 10a, show significant output impedance which reduces the output voltage with increased in the output power; as a result, system peak power is attained at converter smallest input voltage. In light of this, a new low-leakage inductance low-resistance to low-voltage high-power isolated boost converter design technique, was presented as shown in Figure 10b. By optimizing the transformer design and circuit lay-out, very low levels of parasitic circuit inductance were attained. Power MOSFETs fully rated for recurring avalanche, were used to eliminate primary side voltage clamp circuits and switch on-state losses. Furthermore, extensive interleaving of the primary and secondary transformer windings, reduced the transformer proximity effect losses.

Silicon Carbide rectifying diodes are not prone to reverse recovery and therefore allow fast diode turn-off, hence were used to further reduce losses. As illustrated in the study, test results from a 1.5kW full-bridge step-up converter confirmed theoretical analysis and demonstrated a very high efficiency. The maximum efficiency was up to 98% whereas the worst-case efficiency with maximum power and at minimum input voltage was ~ 96.8%.

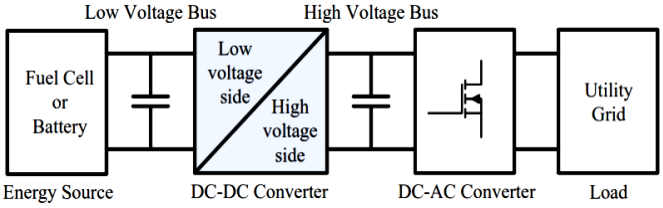


Figure 10 (a): Fuel Cell Power Supply System (adapted from [8])

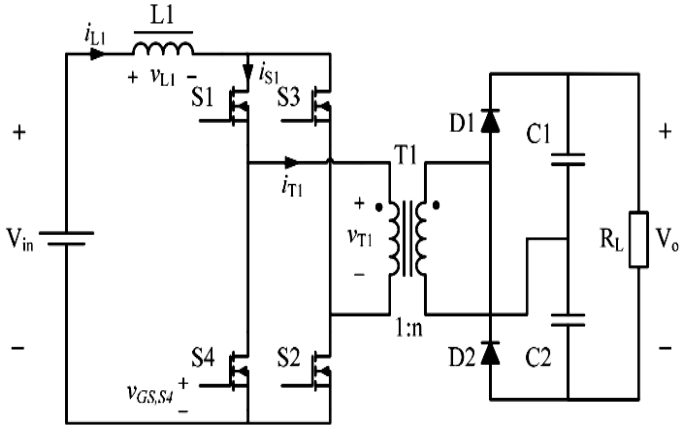


Figure 10 (b): FC High-Power Low-Voltage High Efficiency Isolated Boost Converters (adapted from [8])

2.11. High Power Buck-Boost DC-DC Converters: Automotive Power-train Applications

Investigated in [12], is a high-power buck-boost DC-DC converter for use in the power-trains of hybrid cars as shown in Figures 11a and 11b. To enable smooth transitions between both energy transfer directions, a special digital control strategy was implemented. Equipped with this feature, the converter can obtain the energy management in the electric power-train.

The digital control provides full protection against over-voltage, over-current and over-temperature. Two efficient prototypes of 24kW and 70kW bidirectional (buck-boost) DC-DC converters were developed and evaluated. The presented measurements show that higher voltages for the power-train and storage battery assure higher efficiency due to lower current losses by the use of IGBTs. Using integrated liquid cooling of up to 85°C with very low losses, a high-power density of up to 5W/cm³ was achieved. Characterization data of the converter and measurements in the target application (a hybrid fuel cell car) with test parameters and values of passive components used, are shown in the study as well as what happens when the converter transits from boost to buck mode. Finally, presented in the full manuscript are voltage and current dynamics as well as the efficiencies and output power in the various operation modes.

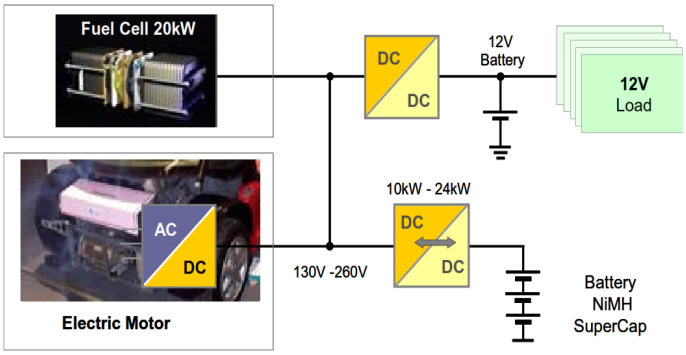


Figure 11 (a): Concept Overview (adapted from [12])

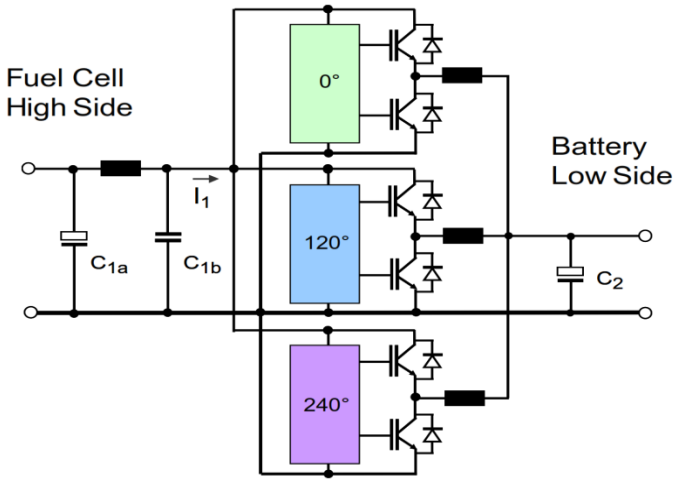


Figure 11 (b): High Power Buck-Boost DC-DC Converters: Automotive Power-train Applications (adapted from [12])

2.12. *PEM FC System with DC-DC Boost Converter: Design, Modelling and Simulation*

Indicated in [3], fuel cells as exemplified in Figure 12a, are regarded as one of the most proficient devices for standalone and grid connected distributed generations (DGs), due to their environmental friendliness, modularity and high energy potential capability. The drawbacks in the extensive use of FCs are their i) sluggish dynamic response to abrupt load changes and ii) costly installation. Their research focused on the simulation of dynamic behaviour of a Nexa 1.2kW PEM FC using DC-DC step-up converter, which was correlated with cascaded 2-stack FC model.

The performance of the basic DC-DC boost converter as power converter for the Nexa TM 1.2kW PEM FC model was analyzed for changing loads to manage the power flow for improved performance. As the FC pressure or temperature rises, the power density of the FC stack also increases for rising loads; therefore, to analyse this dynamic behaviour for changes against temperature, an advanced parametric model based on circuit simulator PSpice for a class of PEM FC was also developed in addition to the fuel cell models based on thermodynamics and electrochemical equations. The fuel cell performance is governed by its electrical and thermal efficiencies − the electrical efficiency of the fuel cell relies on the fuel cell activation and concentration losses besides the natural Joule heating (Ohmic loss), whereas the thermodynamic efficiency relies on the fuel cell fuel processing, water management and the fuel cell system's temperature control.

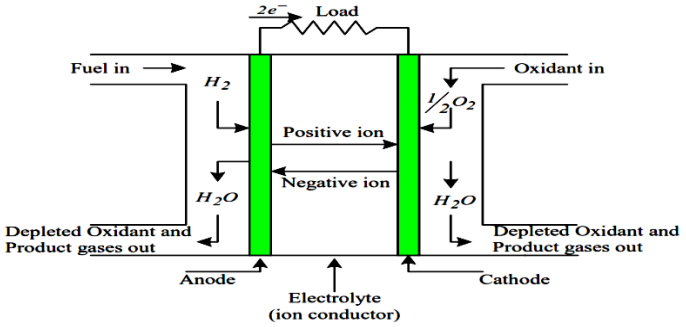


Figure 12 (a): Fuel Cell Overview (adapted from [3])

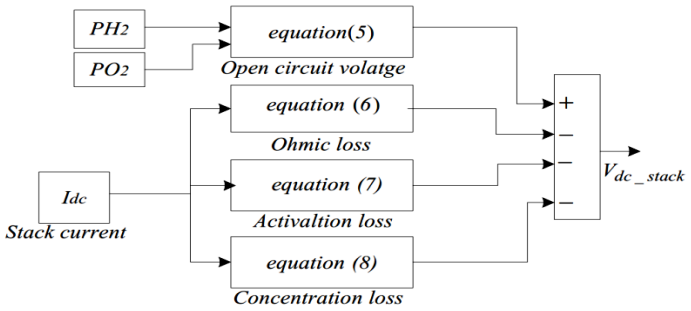


Figure 12 (b): PEM FC System with DC-DC Boost Converter: Design, Modeling and Simulation (adapted from [3])

All these factors were taken into consideration when the design was done using Matlab / Simulink as in Figure 12b. It was noticed that for instantaneous loads variation from 0.6 - 1.1kW, the fuel cell current and voltage took about 50 - 70ms to attain a new steady state. This delay is known as the fuel cell fuel starvation phenomenon – this makes the fuel cells non-linear and should not be operated, because the electrolyte membrane of the FC can be destroyed. The FCs must be operated only in its linear region. The DC-DC converter used was a basic boost converter with PI controller, which gave better performance for load variations without using any storage devices. A constant bus voltage of 80V was maintained in the converter output, regardless of changes in the load and fuel cell terminal voltages. Steady state error was reduced to zero by the PI controller. In their conclusion, operating fuel cells with a basic step-up converter using PI controller, can give better performance for standalone / grid connected low power applications. This claim is evident in the measured and simulated results shown in their full research paper.

2.13. Methodology to Design FC Based Systems Power Converters: A Resonant Approach

Presented in [6], is the evaluation, devise and implementation of a fuel cell-based power generation scheme, which necessitates suitable selection of i) the FC model and ii) the power electronics converters shown in Figure 13a. The fuel cell model used is semi-empirical based on PEM FC Mark 1020 with static and dynamic properties as well as the FC limited current and voltage supply ratings − irrespective of the converter used. The power converter employed a resonant technique that provides high frequency operation, minimal component stresses, soft switching etc.

The power converter controller was split into two functions, namely: i) the voltage controller − which stabilizes the converter output voltage during loading fluctuations and ii) the PEM controller − which enhances the performance by maintaining the PEM FC in its optimal point of operation. The outcome confirmed that the researched converter is a good choice to enhance the efficiency of a PEM FC, because it permits a sufficient control of the power delivered by the fuel cell while sustaining the requirements dictated by the load to maximize the gains with soft-switching control. The FC DC is converted either to DC then to AC (DC-DC-AC) or FC DC to direct AC or FC DC to DC and AC-AC isolated by a transformer. The operation mode of the DC power conversion can be divided into i) linear, ii) switching and iii) soft-switching or resonant. The four non-isolation buck, boost, buck-boost and Cuk converters can be respectively converted to forward, boost, flyback and Cuk; by adding isolation transformer and when the isolation converters use more than a switch; it could be a push-pull, half bridge or full bridge as portrayed in Figure 13b. The results showed that the chosen power converter is suitable to improve the PEM FC efficiency, as it allows proper control of the power delivered by the FC, by satisfying the criteria enforced by the load regulation with minimal losses as a result of adopting soft switching.

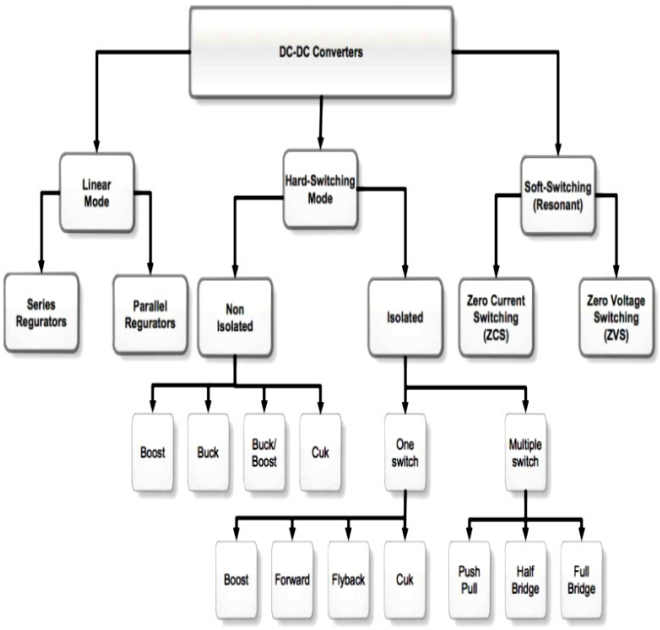


Figure 13 (a): DC-DC Converters Overview (adapted from [6])

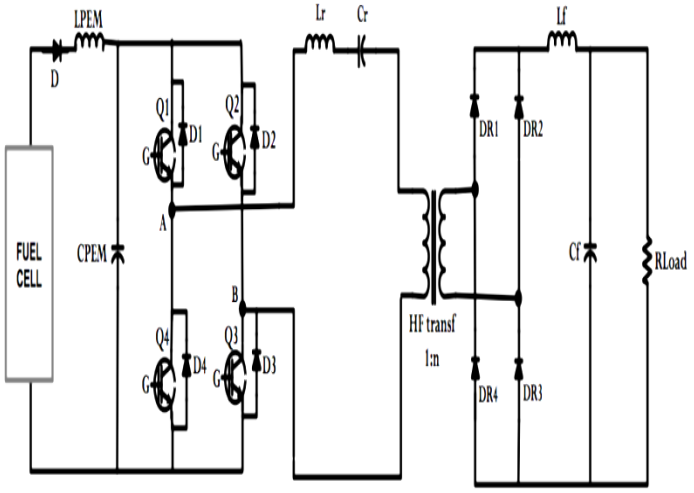
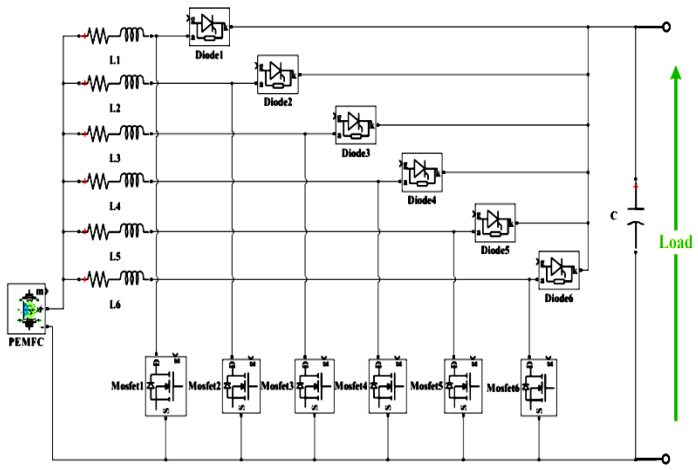


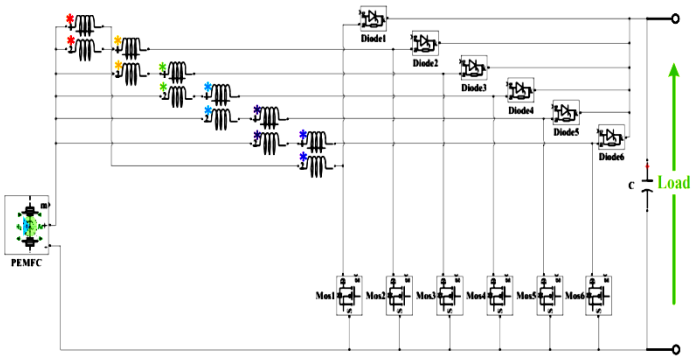
Figure 13 (b): Methodology to Design FC Based Power Converters Systems: A Resonant Approach (adapted from [6])

2.14. Design and Control of a 6-phase IBC Based-on SiC Semiconductors with EIS Functionality for Fuel Cell Electric Vehicle

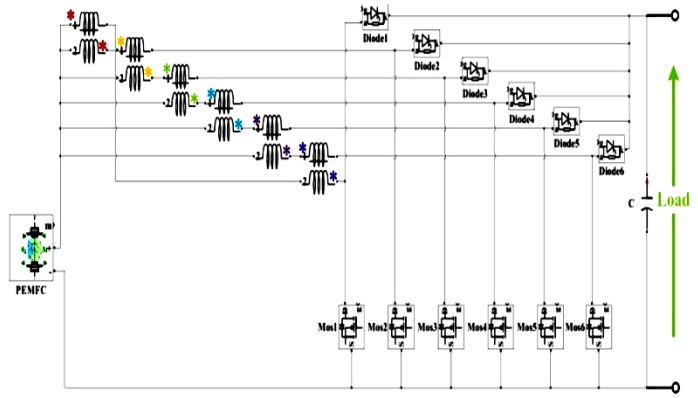
Researched by [20] indicates that, in today's FC Electric Vehicle (FCEV), DC-DC converter is paramount to step-up the PEM FC output voltage to a high level (400 - 700V). As a result, the research aim was to design a 6-phase IBC based on SiC semiconductors and inverse coupled inductors of cyclic cascade structure having high compactness, high efficiency and high voltage gain ratio. The reliability and durability have to be enhanced to advance the consumption and commercialization of FC technologies. Electrochemical Impedance Spectroscopy (EIS) is typically used for PEM FC's diagnosis. To eliminate additional equipment and sensor, the on-line EIS detection functionality incorporated with the control technique of the suggested PEM FC linked to the DC-DC step-up converter was also investigated. The interleaved topology helped decreased the FC current ripple to ensure an extended FC lifespan. Furthermore, the multi-phase topology shared the high input current, hence reducing Joule heating, which allays the electrical stress of the power switches; thus, this redundancy ensures the reliability and robustness of the converter. The magnetic core design is also critical, as it controls the amount of ripple; as a result, the three types Uncoupled (UC)-IBC, Direct Coupled (DC)-IBC, and Inverse Coupled (IC)-IBC were experimented. The SiC-based semiconductors increased the switching frequency and decreased power losses. The on-line EIS detection functionality was integrated with sliding mode control (SMC) of the postulated DC-DC step-up converter. Fuel cells most common problems of membrane drying and flooding were estimated based on PEM FC's equivalent electric circuit model. The real-time hardware-in-the-loop (HIL) validation of the proposed converter was achieved. MicroLabBox (embedded real-time processor with (Field Programmable Gate Array (FPGA)), was used as the real-time platform for prototyping. In all, a 21kW PEM FC's voltage model was developed as the power source and the HIL framework provided in real-time − a benefit to monitor the converter's dynamic working process that was not viable with the offline simulation. Figures 14a-14c summarize the study.



(a). 6-phase IBC (UC-IBC)



(b). 6-hase IBC based on direct coupled inductor of cyclic cascade structure (DC-IBC)

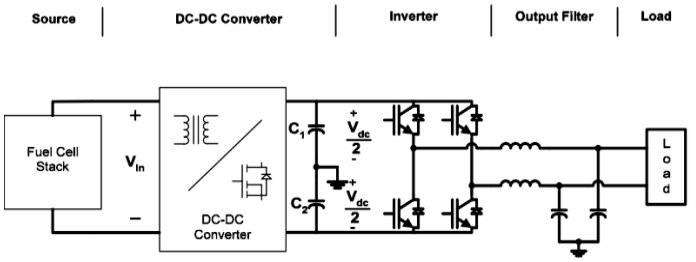


(c). 6-hase IBC based on inverse coupled inductor of cyclic cascade structure (IC-IBC)

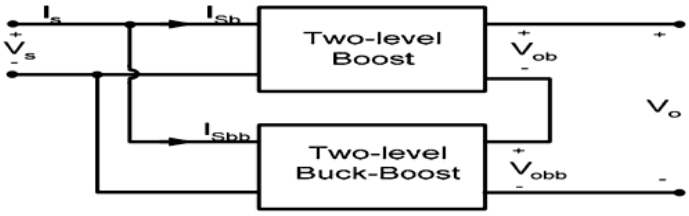
Figure 14: Six-phase IBC Based on SiC Semiconductors with EIS (adapted from [20])

2.15. Design Considerations for DC-DC Converters in Fuel Cells Systems

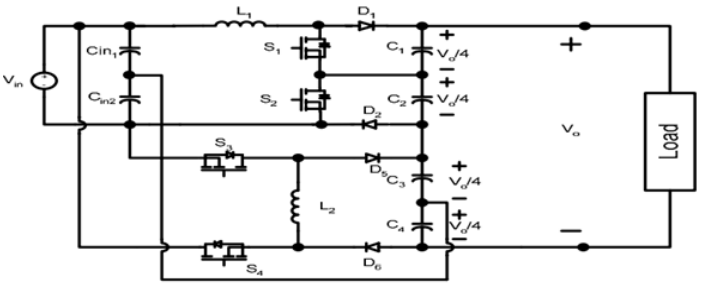
As examined in [4], the development of alternative energy sources, has been improved by the fast increase of fossil fuel costs along with a rise in environmental education – which include but unlimited to FCs, wind, solar and ocean tidal-wave power. Among them, FCs due to their high modularity, efficiency and basic design have received increased interests in recent years. However, their low voltage output and wide variation from unloaded to fully loaded, demands the need of a power converter to interface the FC to its loads. In light of this, their research was undertaken, in which design considerations were attained analytically and experimentally verified to enable an efficient and stable fuel cell as well as power converter system. Further to the design guidelines, investigated also were new power converter layouts that do not need the use of transformers to accomplish a large voltage gain. The general outcomes are means of i) mathematical analysis and ii) experimental prototypes, that contributed to the lessening of the cost and size of the power converter as well as to raise the efficiency of the system. It was discovered that when the FC load current is not purely DC, the Hydrogen usage of the stack increased and the power output of the FC decreased. This effect importance is a function of the ripple current frequency as demonstrated in their full research.



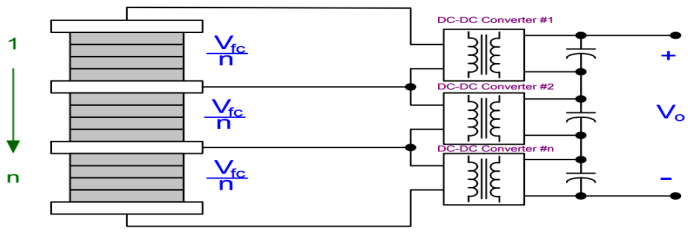
(a) Typical fuel cell power conversion system for residential applications



(b) Block diagram of the proposed DC-DC converter



(c) Circuit schematic of the proposed two level boost and buck-boost converter



(d) Proposed sectioned / taped FC stacked with modular DC-DC converter

Figure 15: Design Considerations for DC-DC Converters in FCs Systems (adapted from [4])

Furthermore, by using analytical and experimental methods, it was demonstrated that for load currents with low frequency ripple (<1kHz), the Hydrogen usage increased up to 7% whereas the power output of the FC decreased up to 30%. In addition, if the frequency of the ripple current is high, >20kHz, the Hydrogen utilized by the FC also increased in the range from 1 - 3%, whereas its power output dropped by 5%. It was further realized that the FC thermal performance was not rigorously affected by high frequency ripple currents presence – due to discontinuous operation mode. It was also found that the FC internal impedance can considerably affect the dynamics of the DC-DC converter. Also, the diminished power left during purging of the FC stack has been shown to be another possible cause for instability. To allay these problems, super-capacitors were connected in parallel to the FCs and a method to compute the value of the super-capacitor to attain stability was derived. A 30W boost converter system experimental results confirmed the validity of the suggested solution. Finally, good dynamic behaviour and stability were proven to be feasible with the use of super-capacitors connected to the output of the FCs. To lessen the cost and volume of the system, a high gain transformer-less DC-DC converter was researched – it employed a two level boost and a two level buck-boost converter in cascade to achieve a high voltage gain and low input ripple current, which contributed to lower electromagnetic interference (EMI). Experimental results demonstrated the viability of the DC-DC converter and showed a possible voltage gain of 5. Normally, FCs are constructed by stacking many cells which limit the generated power to the weakest cell in the stack. In addition, if one or more FCs fail, the entire system must be overhauled. To address these shortcomings, a new modular FC stack and DC-DC converter were pioneered – the FC stack was partitioned into different sections with autonomous operations. This has increased system reliability at a reduced output power should a section failed. Additionally, the generated power from the system was optimized by adjusting the drawn current from each section based on the voltage they produced, which resulted in a 10-14% extra power generation. Common mode noise due to transients was also noticed and was resolved by using shielded transformers. Figures 15a - 15d exposit the highlights of the study.

2.16. *An Overview of Various Fuel Cell DC-DC Converters*

According to [24], fuel cells are now becoming the preferred alternative renewable energy source, as their power production process is not affected by fluctuating environmental factors, contrary to solar cells and wind power plants. However, fuel cells produce low DC output voltage which requires stepping-up and interfacing them to the DC bus. Thus, the need for DC-DC boost power converters, which could be interleaving to help minimize the power ripples as well as bidirectional to charge storage devices such as a battery. In this regard, their paper discussed various interleaved (2, 3, 4 and 6 phases) and bidirectional (non-isolated and isolated) DC-DC boost converters architectures. The non-isolated BDC DC-DC boost converters covered include i) buck-boost converter, ii) cascaded buck-boost converter, iii) CUK converter, iv) SEPIC-ZETA converter and v) switched capacitor. The isolated BDC DC-DC boost converters covered include i) dual half bridge (DHB) and ii) dual active full bridge (DAFB). Figure 16 illustrates the research overview and the conclusion drawn is interleaved boost converters improve power ripples and the more the interleaving, the better the ripple reduction; however, the more costly and bulky it becomes due to the many components used. BDC can additionally charge storage devices and furthermore, the isolated types can offer galvanic protection in high power applications; however, their large size due to the extra isolating transformer, makes them unsuitable for portable and or compact applications.

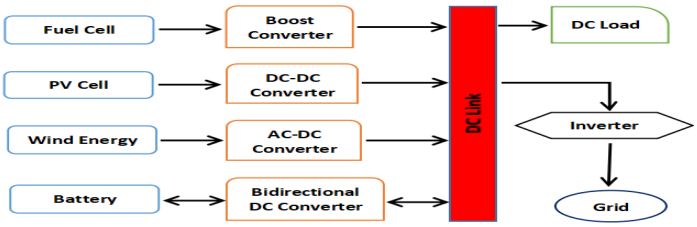
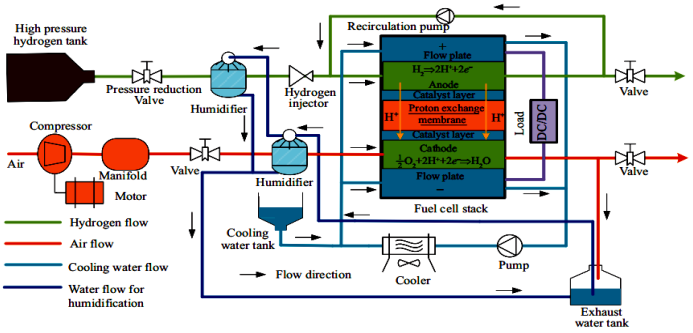


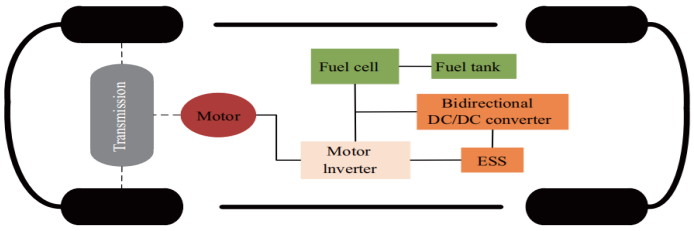
Figure 16: An Overview of Various Fuel Cell DC-DC Converters (redrawn from [24])

2.17. *Challenges and Developments of Automotive Fuel Cell Hybrid Power System and Control*

As assessed in [25], fuel cells are the future replacement for internal combustion engine in vehicles, though the current costs and Hydrogen supply infrastructure are the limiting factors. In their analysis, they noted that FCs in hybrid power systems have energy control, inertia, power, model and optimization problems which were summarized briefly with emphasis on the electro-chemical reactions, dynamics and the core parameters affecting FCs efficiency and durability. Their review concludes by highlighting that fuel cells have various challenges and the best solution is one that is inclusive by incorporating various hardware and software solutions to optimize fuel cells costs, performance and longevity. Figure 17a exemplifies a typical FC system and Figure 17b illustrates a simplified FCEV architecture.



(a) Typical fuel cell system



(b) Fuel cell electric vehicle power-train

Figure 17: Fuel Cell System and FCEV Power-train (adapted from [25])

2.18. *Experimental Study and Performance Analysis on High Power Fuel Cell System*

In [26], it's affirmed that PEM fuel cell for use in vehicles requires high power density — normally during starting-up and accelerating. As a result, their study presented an experimental research of a 100kW fuel cell power supply system with focus on measuring the system parameters such as voltage, current, temperature, pressure and hydrogen consumption. Two test set-ups were used; a i) rated and ii) cycle working condition tests. In the former, the system operates for an hour at a rated point with constant working conditions and the outcomes revealed stable operations when working constantly at the rated output power. In the latter, the test is conducted based-on their specified national standard, in which the fuel cell voltage is regulated to be a fixed value and the output current is varied with the working conditions. The conclusion arrived at is the researched fuel cell power engine reached 80kW at rated power with peak power exceeding 100kW. Figures 18a and 18b exemplify the researched FC power schemes.

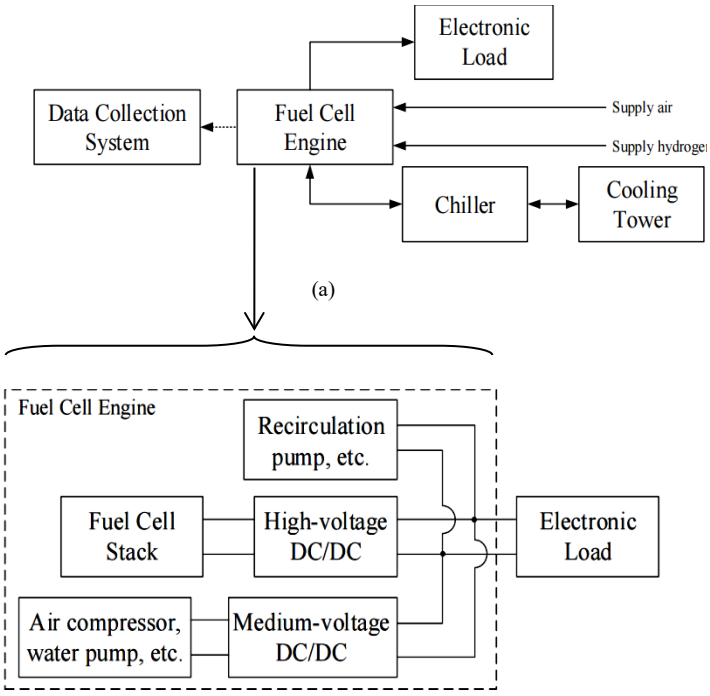
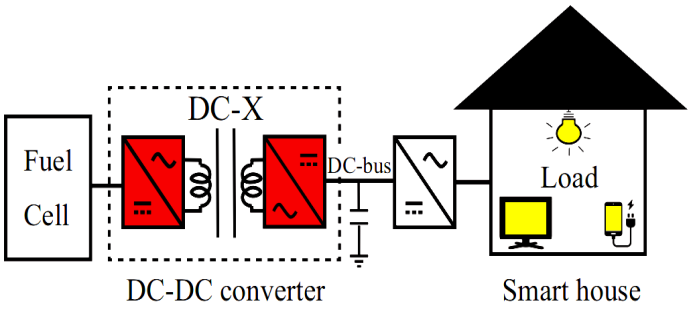


Figure 18: (a) The Power System Overview and (b) Fuel Cell Power Engine (adapted from [26])

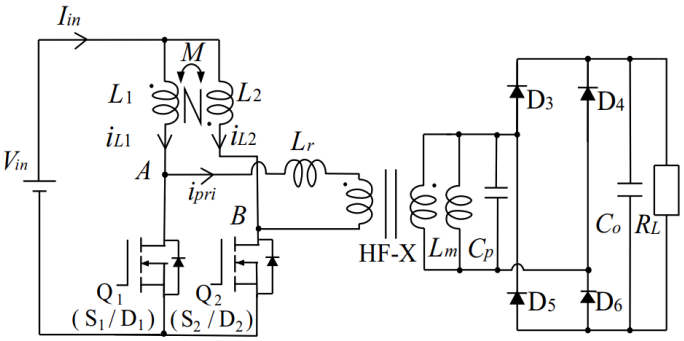
2.19. *Coupled Inductor-assisted Current-Fed Snubber-less Zero-Current-Soft Switching High Step-Up DC-DC Converter for Fuel Cell Power Interface*

Presented in [27], is a ZCS current-fed isolated DC-DC boost converter for a fuel cell smart home power system. To avoid ripple current from damaging the fuel cell electrodes and to ensure good boost voltage ratio, their design incorporated a magnetically coupled interleaved inductors with a 180◦ phase shift and a small number of passive components. A 50/60W 1MHz prototype based on 600V GaN-HFETs, was performance tested focusing on the design efficiency, ripple and voltage ratio.

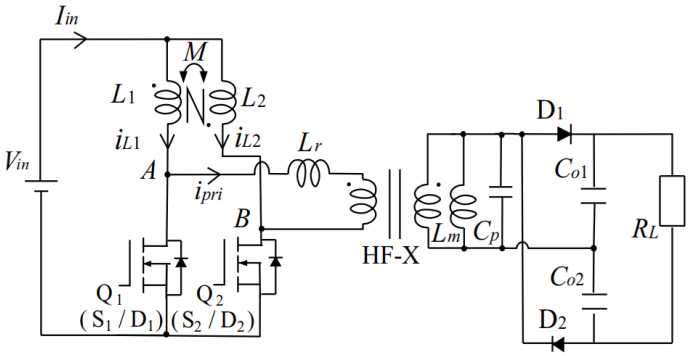
Figure 19a illustrates the concept design and Figures 19b and 19c, portray the DC-DC converter with a full bridge and voltage doubler outputs respectively. Normally, the primary side of the high-frequency transformer (HF-X) windings constitutes a two-phase current-fed high-frequency resonant inverter with complementary Q1 / Q2 as the active switches and L1 / L2 as the coupled inductors. HF-X leakage (Lr) and magnetizing inductances (Lm) produce the multi-resonant transitions with Cp, to generate the high voltage boost ratio and quasi-resonant sub-interval for soft commutations. The leakage inductance operates as a snubber inductor to reduce the high di/dt rate at Q1 and Q2 switched-on transitions. For a higher boost ratio, the full bridge rectifier output can be substituted with a voltage doubler rectifier. Their findings showed that the postulated power converter can attain a snubber-less ZCS commutations, a greater voltage boost ratio and low power ripple; making it suited for smart homes use.



(a) Concept overview



(b) High step-up dc-dc converter with full-bridge rectifier

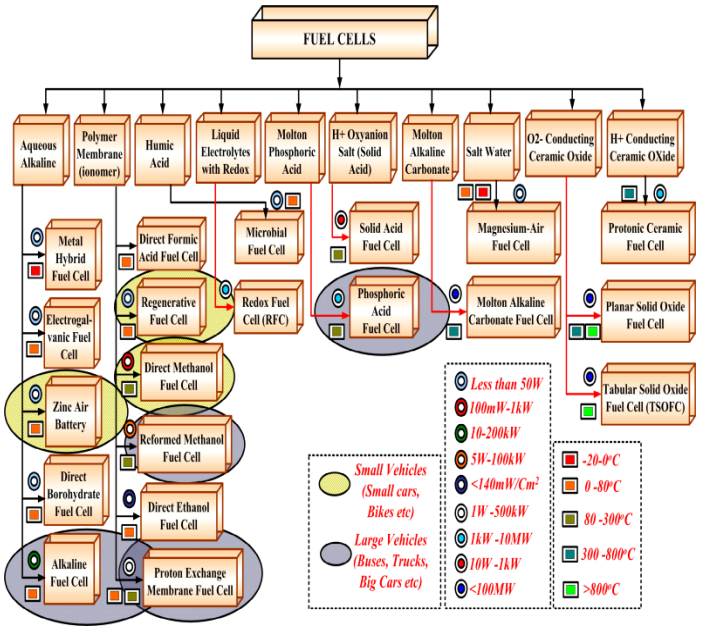


(c) High step-up dc-dc converter with voltage doubler rectifier

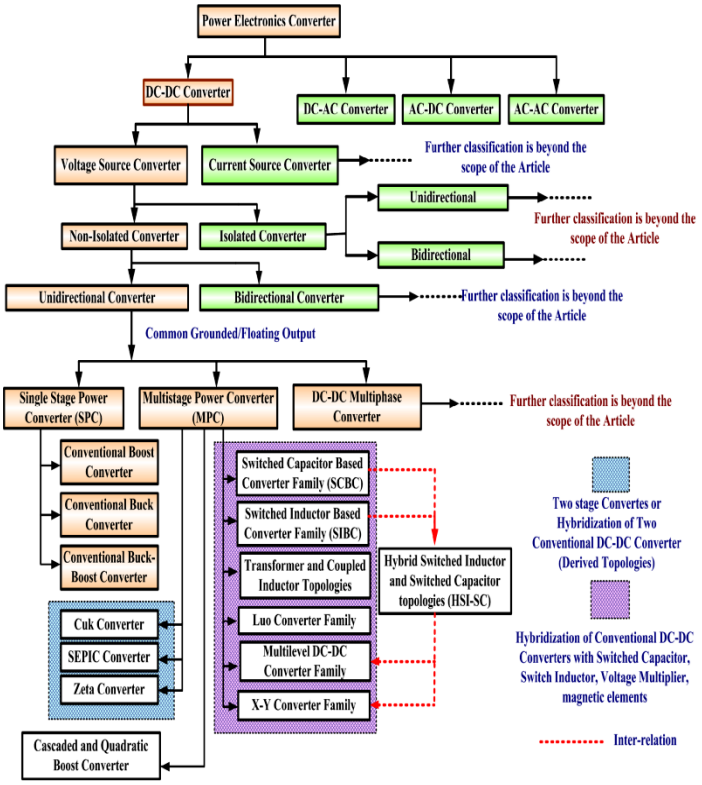
Figure 19: System Concept Design and High Step-up DC-DC Converter Types (adapted from [27])

2.20. Survey of DC-DC Non-Isolated Topologies for Unidirectional Power Flow in Fuel Cell Vehicles

Investigated in [28], is an outstanding research on fuel cells power-trains and power converters. Its extensively discussed in details the theoretical and architectural frameworks of fuel cells. Figures 20a and 20b respectively portray in totality a summary of fuel cell types and power electronics converters classifications.



(a) Fuel cells classification

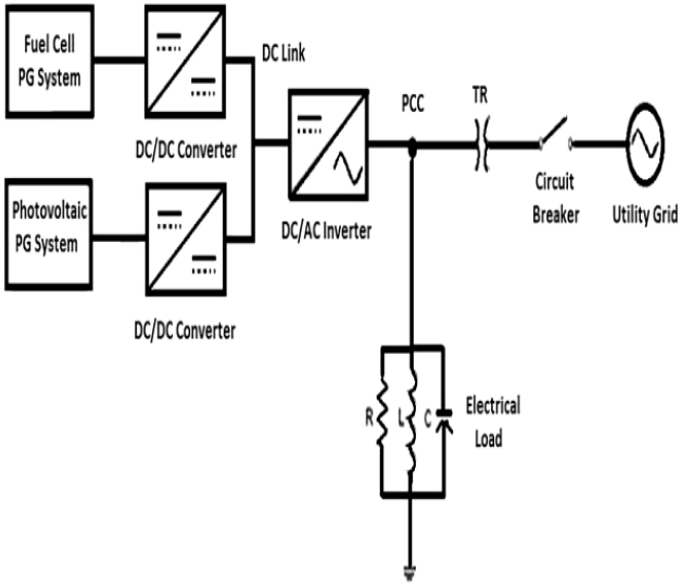


(b) Power converters classification

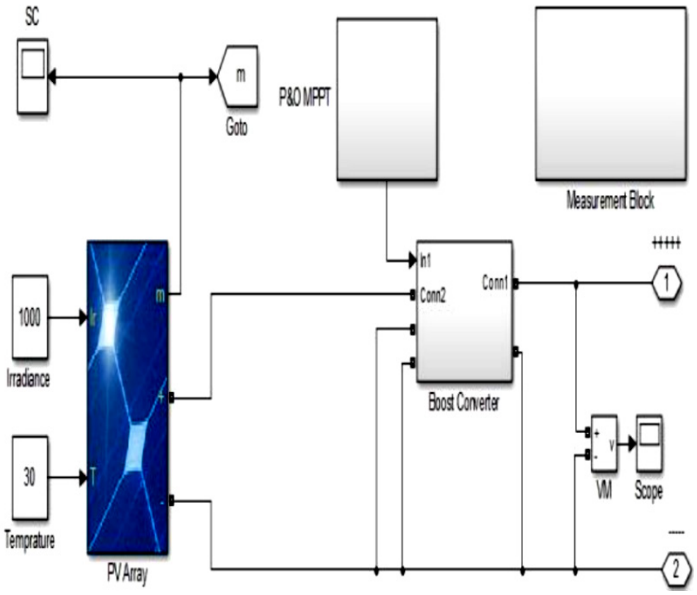
Figure 20: Fuel Cells and Power Converters Classifications (adapted from [28])

2.21. Performance Analysis of PV and Fuel Cell-based Grid Integrated Power System

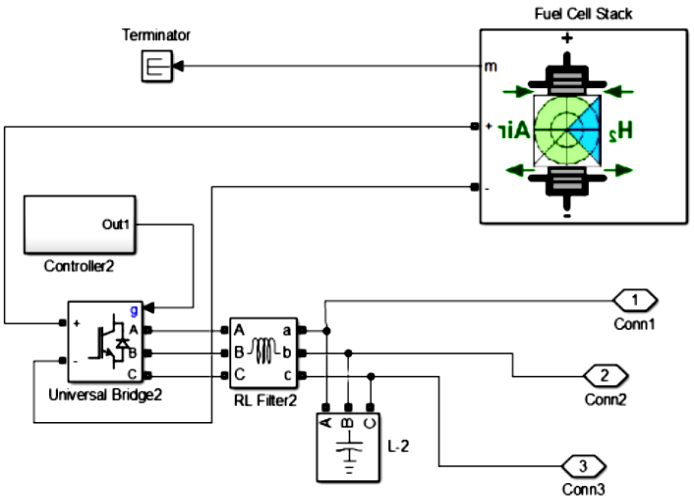
Studied in [29], is a smart grid power generation system constituting solar cells and solid oxide fuel cells (SOFC) hybrid system as shown in Figure 21a. The SOFC augments the PV system during power outages due to fault and non-sunny periods. Alternative to SOFC; biomass and wind power systems can be used by integrating them with phase lock loop (PLL) to maintain a constant output supply. The power electronics made used of a DC-DC converter, a three phase DC-AC inverter for interfacing to the electrical grid and AC loads. In addition, is a LC filter to eliminate unwanted signals in the power system. The energy management techniques used include P&O MPPT as well as reference frame theory and PLL to enable a reliable power supply. Figures 21b - 21c respectively show the PV and SOFC schemes.



(a) Concept overview



(b) Solar cell power system modeled with Matlab

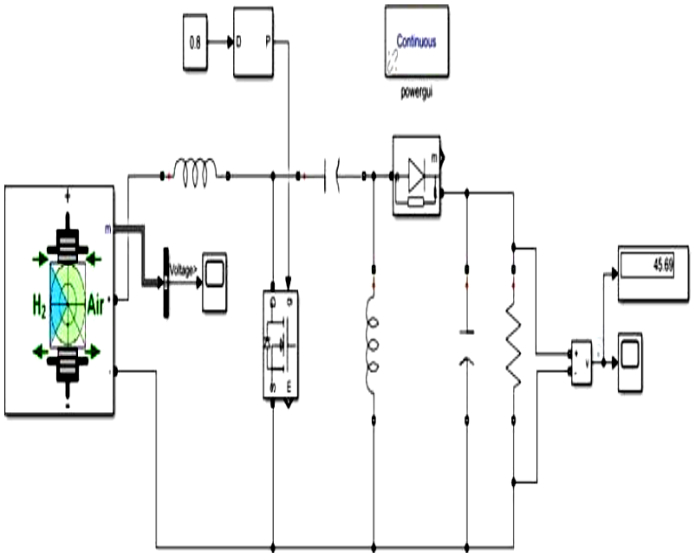


(c) Fuel cell power system modeled with Matlab

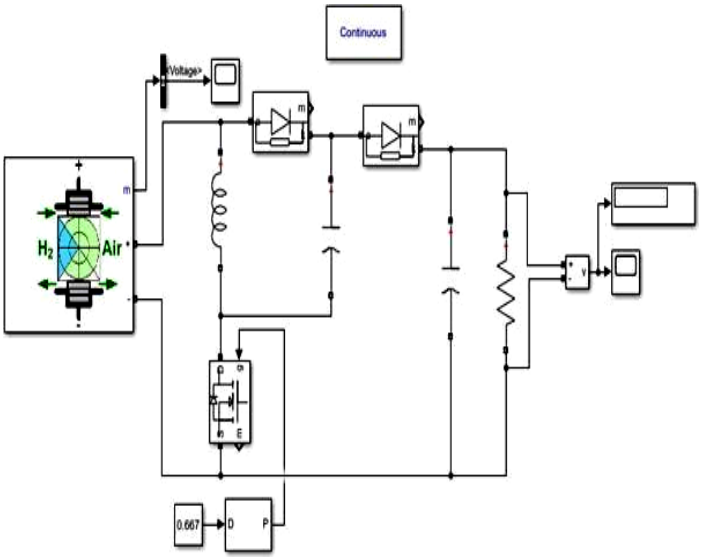
Figure 21: System Concept Design, PV and Fuel Cell Power Systems Depictions (adapted from [29])

2.22. *Modeling and Simulation of DC-DC Converters for Fuel Cell Systems*

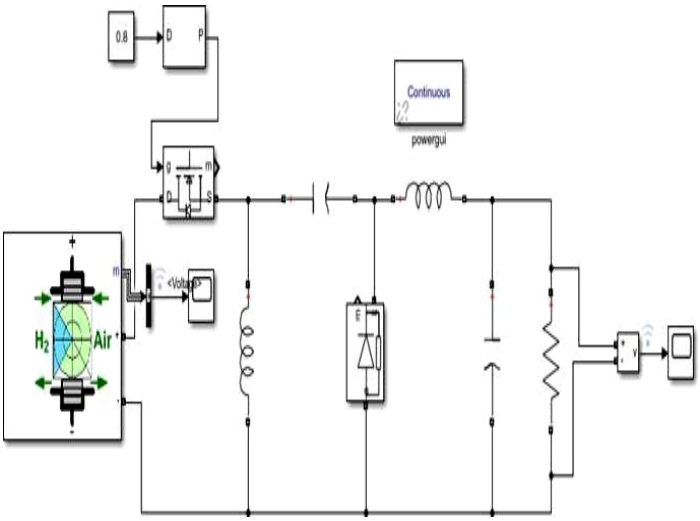
Affirmed in [30], FC is the future renewable energy source, especially for portable applications. Fuel cells as a result of their low output voltage, require highly efficient power converters; thus, their research using MatLab, focused on the modeling and simulations of four types of DC-DC converters, namely i) boost, ii) SEPIC, iii) LUO and iv) ZETA. Their study was tested using the same fuel cell output voltage of 12V, connected to each converter input with each converter output voltage set at 48V. It was found that the ZETA topology offers the best total harmonic distortion (THD), followed by LUO, SEPIC and Boost with respective THD of 31.22%, 53.83 %, 65.38 % and 80.22 %. It was furthermore concluded that the ZETA topology THD performance can be improved with the addition of more filtering components. Figures 22a - 22c exemplified the SEPIC, LUO and ZETA DC-DC power converters modeled using Matlab Simulink.



(a) Fuel cell with SEPIC DC-DC converter simulation model



(b) Fuel cell with LUO DC-DC converter simulation model



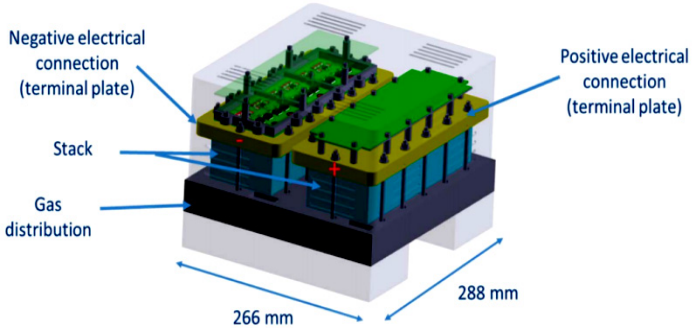
(c) Fuel cell with ZETA DC-DC converter simulation model

Figure 22: Simulink Model of FC with SEPIC, LUO and ZETA DC-DC Power Converters Topology (adapted from [30])

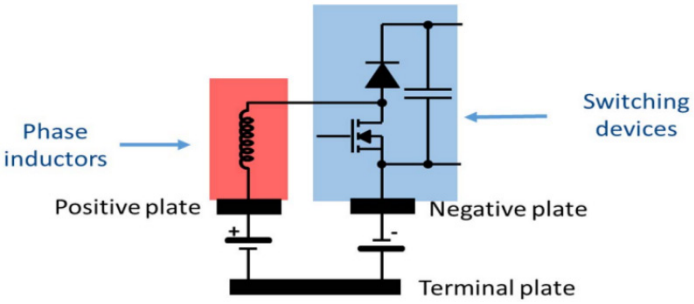
2.23. *Smart Fuel Cell Module (6.5 kW) for a Range Extender Application*

Researched in [31] using SolidWorks, is a 6.5kW fuel cell model with a mechanically integrated 6-phase interleaved DC-DC boost converter for electric vehicles applications. The design constraints were such that the power converter was mounted on the fuel cell terminal plates and cooled using the same FC cooling system as shown in Figures 23a and 23b. The choice of the power converter topology was driven by the simplicity of its design, since the converter must fit on the FC terminals as well as the fuel cell configuration. As a result, the classic boost converter was chosen, as it employed the minimal components count and furthermore the phase inductors and switching devices can be respectively connected directly on the FC positive and negative plates as pictured in Figure 23b. Continuous conduction mode was chosen for the converter and the ripple was minimal.

To conclude their study, the measured converter efficiency was >95% for a minimum output power of 1.5kW and output voltage of 240V. Future work for an aircraft use was considered.



(a) Fuel cell concept design modeled with SolidWorks



(b) Integration of the power converter on the fuel cell plates

Figure 23: SolidWorks Model of the FC with Integrated Power Converter (adapted from [31])

2.24. *Power Converter Topology for Conditioning a Fuel Cell Battery Voltage*

Stated in [32], their research conditioned the output voltage of a FC battery using DC-DC boost converter. The main novelty was to substitute the classic boost converter inductor (L) with a inductor-capacitor-inductor (LCL) filter topology as shown in Figure 24. The output voltage was then controlled using a sliding mode strategy including a load impedance observer. The simulated results showed good performance with varying loads.

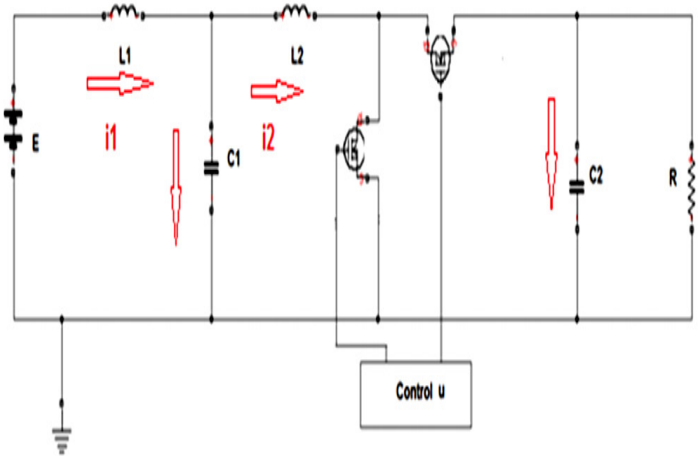
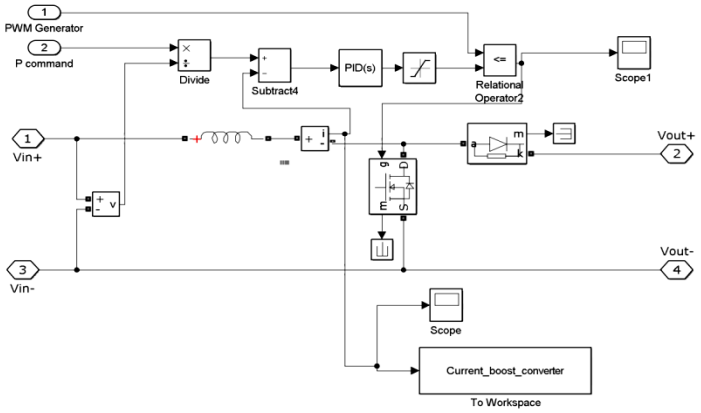


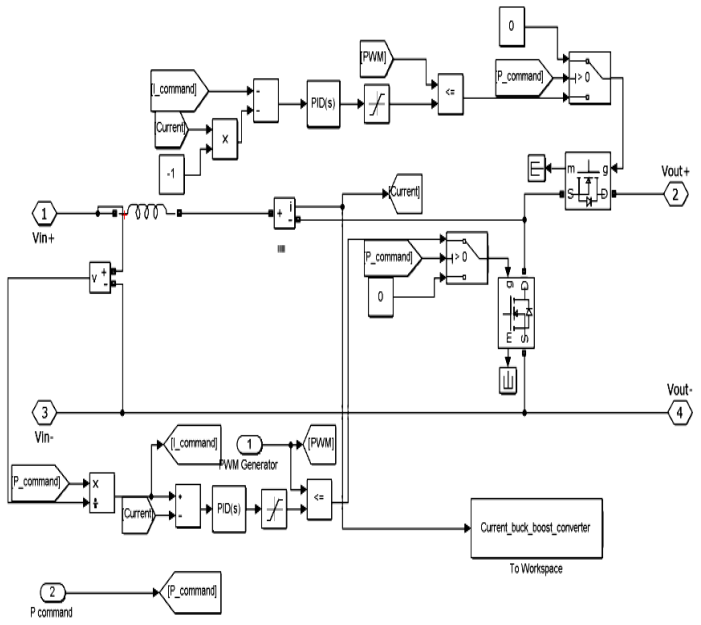
Figure 24: LCL Boost DC-DC Power Converter (adapted from [32])

2.25. *Modeling and Simulation of an Aerodrome Electrical Power Source Based-on Fuel Cells*

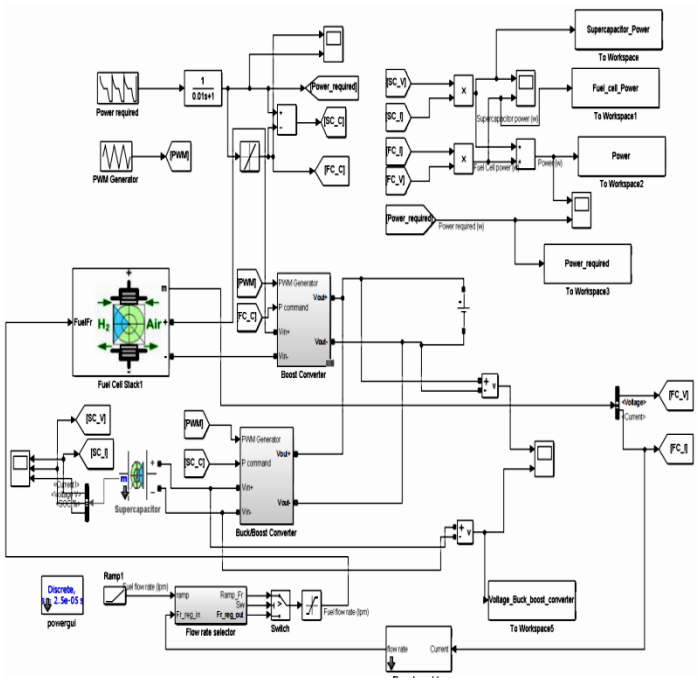
Examined in [33], reducing fuel usage and emissions such as NO_x, is aviation present challenge. As a result, there is the need to develop latest power sources using non-polluting sources such as Hydrogen fuel cells. Their work presented the modeling and simulation of a potential configuration for a hybrid aerodrome fuel cell power source. Their postulated architecture consists of a fuel cell stack, DC to DC set-up converter, super-capacitor and a buck-boost converter. The fuel cell and super-capacitor are respectively connected to a boost and buck-boost converter as correspondingly shown in Figures 25a and 25b. With this set-up, the fuel cell with slower power dynamics, supplies the bulk of the power during steady state operation, whereas the super-capacitor with a faster power dynamics, assist the fuel cell during peak power transient demand as well as stores power from the DC bus. The suggested configurations were simulated using MatLab, Simulink and Simscape Power Systems and it can be summed that the hybrid power aerodrome source shown in Figure 25c, can work efficiently, enabling its use for such long-term applications.



(a) Fuel cell with boost DC-DC converter



(b) Super-capacitor with buck-boost DC-DC converter

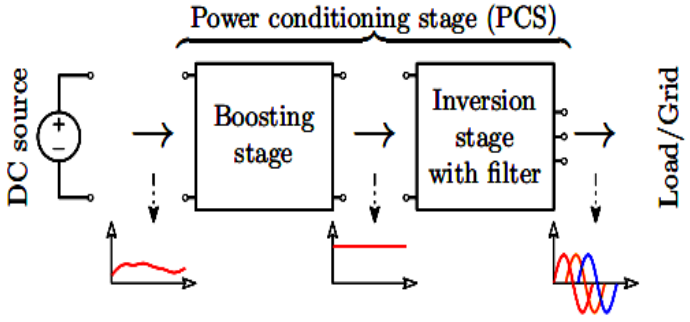


(c) Fuel cell and super-capacitor with boost and buck-boost DC-DC converters

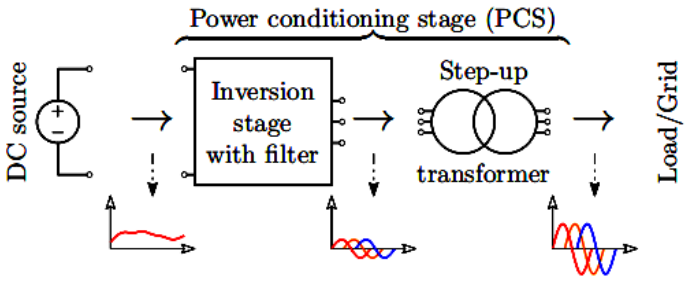
Figure 25: Hybrid Fuel Cell and Super-capcitor Power System with Respective Boost and Buck-boost DC-DC Converters (adapted from [33])

2.26. *Current-fed Modular Multilevel Converter (CMMC) for Fuel Cell and Photovoltaic Integration*

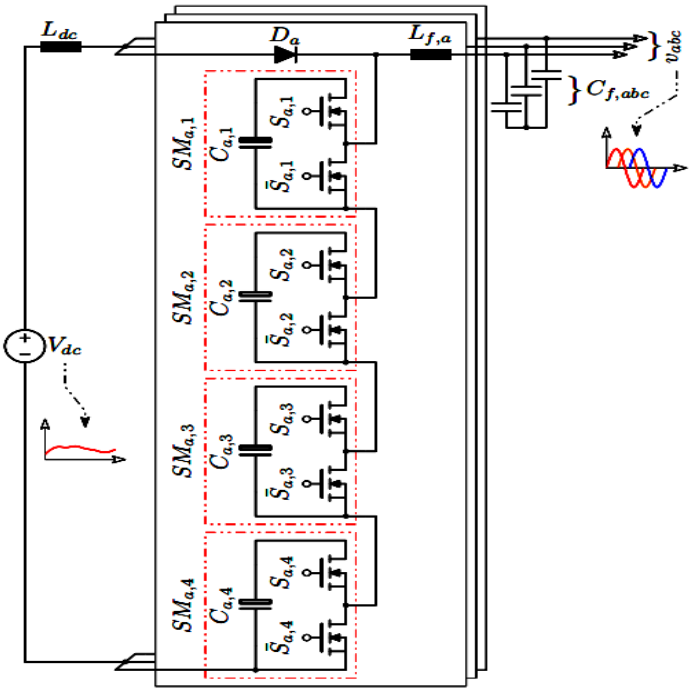
Indicated in [34], is a CMMC single-stage solution to interface a low voltage photovoltaic and fuel cells DC power supplies to a higher voltage AC load and or grid. Usually, power conditioning stage (PCS) in the form of modular multilevel converters, have been used in various low to high voltage applications with good results; however, their two-stage configuration makes them bulky, hence the need for CMMC — whereby the boosting capability is integrated within the inversion, making it a single-stage DC-AC converter / inverter with additional redundancy and modularity. This enables it use in low voltage applications, where low voltage MOSFETs with low ON-state resistance can be used to increase the power conversion efficiency. A 10kW three-phase CMMC using PLECS, was simulated to verify its functionality. Figures 26a and 26b depict the traditional two-stage DC-AC using (a) a boost converter before the inversion and (b) a step-up transformer after the inversion. Figure 26c exemplifies the single-stage CMMC.



(a) Two-stage DC-AC with boost converter before the inversion



(b) Two-stage DC-AC with step-up transformer after the inversion



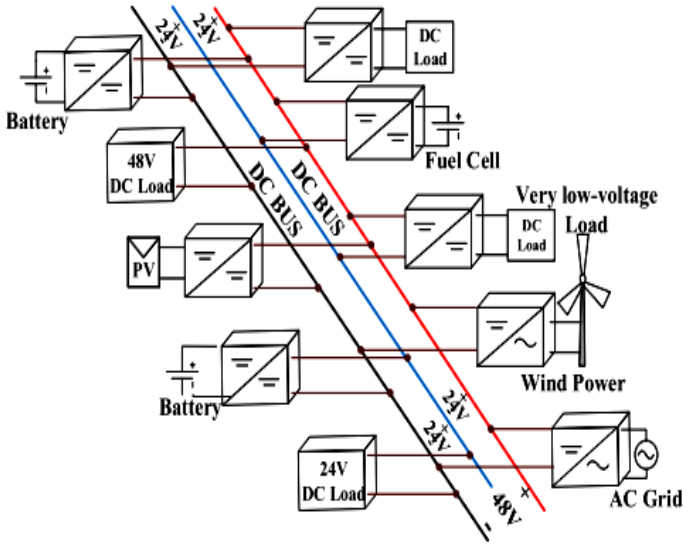
(c) Single-stage DC-AC three-phase CMMC with four sub-modules per phase

Figure 26: Two-stage Traditional and Single-stage CMMC DC-AC Inverters (adapted from [34])

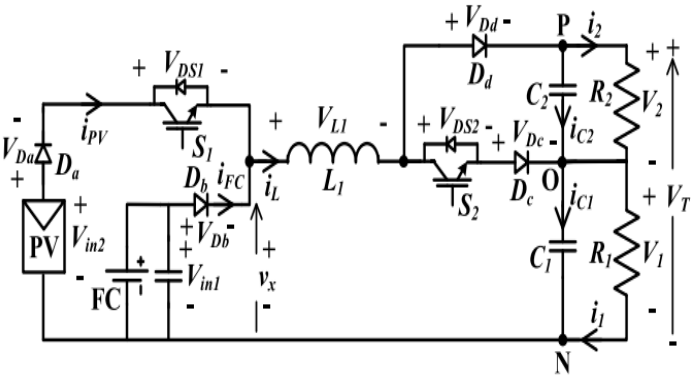
2.27. *Novel Four-Port DC–DC Converter for Interfacing Solar PV–Fuel Cell Hybrid Sources with Low-Voltage Bipolar DC Micro-grids*

Presented in [35] is a bipolar DC micro-grid (BDCMG) power supply system based on a novel 4-port dual-input dual-output DC-DC converter to interface fuel cells, PV and wind power sources to a low voltage BDCMG. Usually, a BDCMG requires several traditional DC-DC power converters to supply power to the BDCMG poles; however, their researched model in addition to being reliable and efficient, is also compact and unidirectional. It can also function as a single-input dual-output converter as well as with two degrees of freedom using its two switches. Furthermore, the duty cycle changes has no effects on the converter dynamic model; thus, the converter can be controlled with just one controller in different modes, making it less complex. By deriving a small signal model for each operating mode, the converter control system was designed. MPPT was used to track the PV voltage and inductor current without needing an extra PV sensor. Its steady and dynamic states operations were validated using close and open loops results. In-lined with both simulations and

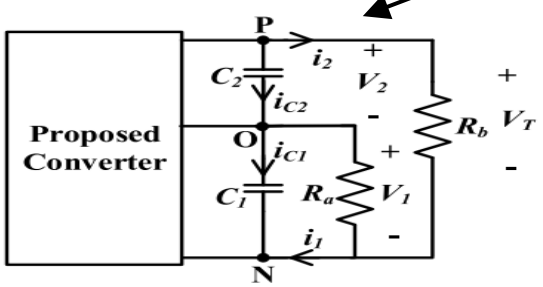
theoretical analyses, they observed that the 24V pole voltage and the photovoltaic power are maintained under different conditions (such as during solar irradiation fluctuations and transient load power demands); thus, validating the converter design performance and reliability. The converter was found to have a 93% peak efficiency and ~87% rated power efficiency. Figure 27a exemplifies the proposed BDCMG power scheme and Figure 27b depicts the converter topology whereas Figure 27c shows a different load configuration.



(a) Low-voltage (48-V) BDCMG system.



(b) Converter topology

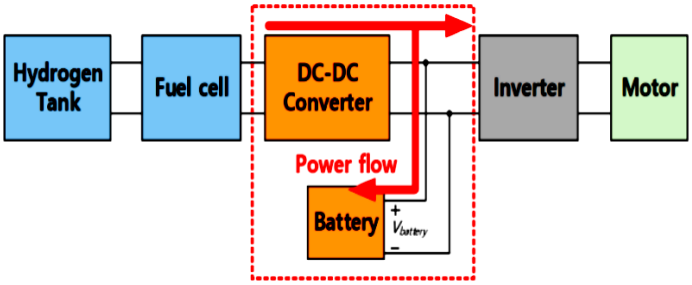


(c) Converter topology with alternative load representation

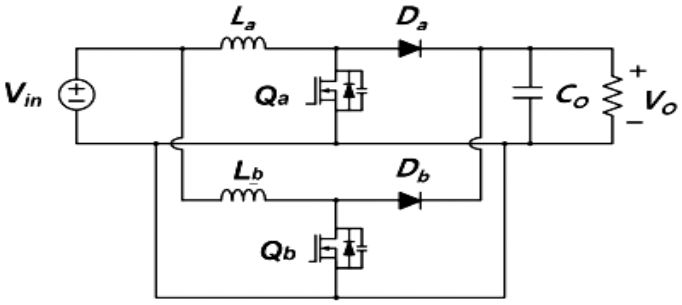
Figure 27: BDCMG System and Converter Topologies (adapted from [35])

2.28. *Study on Boost Converters with High Power-Density for Hydrogen-Fuel-Cell Hybrid Railway System*

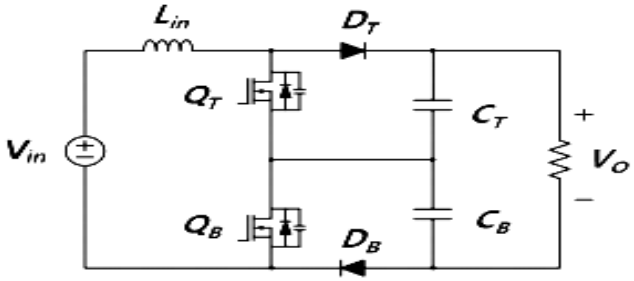
Investigated in [36] is a high power hybrid hydrogen fuel cell railway system portrayed in Figure 28a − with focus on designing an efficient and high power density DC-DC converter, since fuel cells are normally low DC power sources and can not supply the needed 1500V to drive the inverter input needed for the railway traction AC motors. Therefore, two DC-DC power converters, namely the interleaved boost converter shown in Figure 28b and the three-level boost converter depicted in Figure 28c, were researched to determine the most suitable DC-DC boost power converter architecture. Taking into considerations and also using optimal design methods were the boost inductor, output capacitor and power semiconductor devices performances with respect to the hybrid railway specifications. Both power converters designs were verified with a 600V input and 1200V / 20kW output setup and the results concluded that the three-level boost converter performed better in-terms of the efficiency, power density and dynamic current response. As a result, it was chosen as the most suitable topology for the hybrid hydrogen fuel cell rail system.



(a) Propulsion system for hybrid hydrogen-fuel-cell railway system



(b) High step-up DC-DC interleaved boost converter (IBC)



(c) High step-up DC-DC three-level boost converter

Figure 28: Concept Design with High Step-up DC-DC IBC and Three-level Boost Converter (adapted from [36])

Table 1 summarizes the fuel cells power converters studies reviewed − in which the major highlights, advantages and disadvantages of each where applicable, are briefly recapitulated.

Table 1: Power Converters Studies Examined Summary (adapted from [1])

Power Converters	Highlights, Merits and Demerits
Study 2.1 (A. Kolli *et al*, 2015) [23]	Various FCs DC-DC power converters setups. Emphasis on different types of interleaved converters for high, medium and low power uses. FCs in parallel /series raise output power.
Study 2.2 (M. Kabalo *et al*, 2010) [13]	FC vehicles cutting edge DC-DC converters. High voltage ratio, compactness and efficiency with affordability, should be used to implement power converters. Presented different schemes.
Study 2.3 (M. Delshad & H. Farzanehfard, 2011) [14]	ZVS current fed push-pull DC-DC converter. When power is off, voltage surge across the switch is absorbed. This improve its efficiency and compactness to enable basic PWM control.
Study 2.4 (N. Bizon, 2011) [15]	A new architecture of FC HPS for efficient functioning and better steadfastness. HPS with active MPPT and hysteretic current controls were used to minimize ripple current from FC.
Study 2.5 (O.A. Ahmed & J.A.M. Bleijs, 2013) [16]	For an UC in DC micro-grids, a bidirectional voltage-fed setup is preferred for quick dynamic response, though for a broad input voltage instability at the UC, there is greater circulating power flow and conduction losses.
Study 2.6 (A. Carvalho *et al*, 2011) [19]	Modeled a PEM FC using MatLab. Noted the preferred model must take control and optimise the FC operation points. Soft switching based on series resonant and SA was used, as it reduces switching losses and boost efficiency.
Study 2.7 (F.M. Mwaniki, 2014) [11]	Multi-phase tapped-coupled inductor suited for varying high power DC-DC converter uses. Showed minimal input & output power ripples.
Study 2.8 (Y. Huangfu *et al*, 2015) [10]	High power efficiency step-down converter for discrete wind power supply scheme, akin PV. Achieved a 2kW supply with 96% efficiency with step-down ZVS/LCD scheme with MPPT.
Study 2.9 (R. Seyezhai *et al*, 2013) [9]	Interleaved converters with switched capacitor are considered the suitable topology for FC systems, because of reduced ripple power in the input and output, quicker transient reaction, small EMI, enhanced efficiency and reliability.
Study 2.10 (M. Nymand & M.A.E. Andersen, 2008) [8]	A new low-leakage inductance low-resistance design approach to low-voltage high-power isolated boost converters. Poorest efficiency at minimum input voltage with maximum power was ~97%. The maximal efficiency was ~98%.
Study 2.11 (B. Eckardt *et al*, 2005) [12]	FC automotive power-train application using high current buck-boost DC-DC converter with digital control to render apt protection against over-current, over-voltage & over-temperature.
Study 2.12 (A. Kirubakaran *et al,* 2009) [3]	PEM FC setup with DC-DC step-up converter: Design, modeling and simulation. For instant load fluctuation from 0.6 – 1.1kW, the FC current and voltage took ~50 - 70ms (fuel starvation) to attain a new steady state. The altering voltage was tracked with PI controller.
Study 2.13 (M.T. Outeiro & A. Carvalho, 2013) [6]	A method to devise power converters for fuel cell rooted schemes using resonant technique. Independent voltage and PEMFC controllers. Enhanced FC efficiency by managing FC P_{out}.
Study 2.14 (H. Wang, 2019) [20]	Devise and management of a 6-phase IBC rooted in SiC with EIS functionality for FC HEV. IBC dynamic model with HIL real-time.
Study 2.15 (L.M.P. Fanjul, 2006) [4]	Design deliberations for DC-DC converters in FC schemes. Used analytical and experimental schemes to achieve a steady and efficient FC & power converter system. A modular FC stack and DC-DC converter were pioneered by dividing it into autonomous optimal sections.
Study 2.16 (D. Ravi *et al*, 2018) [24]	IBC and BDC were researched. IBC improves power ripples. The more the interleaving, the better the ripple reduction; though, the more costly. BDC can charge storage devices and furthermore, the isolated types offer galvanic protection in high power uses; however, their large size makes them unfit for portable uses.
Study 2.17 (J. Gao *et al*, 2019) [25]	FCs have various challenges and the best solution is one that is inclusive with various hardware and software solutions to optimize better FCs costs, performance and longevity.
Study 2.18 (H. Liu *et al*, 2020) [26]	Investigated a high power fuel cell system. Two test setups were used; a i) rated and ii) cycle working condition tests and found that fuel cell power engine reached 80kW at rated power with the peak power exceeding 100kW.
Study 2.19 (R. Miyazaki *et al*, 2020) [27]	A current-fed snubber-less ZCS FC high step-up DC-DC converter was studied. It achieved a greater voltage boost ratio and low power ripple, making it suitable for smart homes use.
Study 2.20 (M.S. Bhaskar *et al*, 2020) [28]	Reviewed extensively and comprehensively in theory and topologically, the different types of fuel cells with focus on the use of fuel cells in FCEV power-trains. Miscellaneous types of power converters were also assessed in details.
Study 2.21 (K.S. Rathode *et al*, 2019) [29]	Researched a hybrid PV and FC system. The power electronics used a DC-DC converter, a three phase DC-AC inverter for interfacing to the electrical grid and AC loads with P&O MPPT as well as reference frame theory and PLL to enable a reliable power supply system.
Study 2.22 (S. Kavyapriya & R.K. Kumar, 2020) [30]	Modeled and simulated four step-up power converters schemes. Found that the ZETA topology offers the best THD, followed by LUO, SEPIC and Boost with THD of 31.22%, 53.83 %, 65.38 % and 80.22 % respectively.
Study 2.23 (P. Bazin *et al*, 2020) [31]	Implemented a smart FC with built-in DC-DC power converter. The classic boost converter with 6-phase interleaving was chosen, as it fitted well, efficient & offered least parts used. The efficiency was >95% for a nominal output power of ~1.5kW and output voltage of 240V.
Study 2.24 (A. Gonnet *et al*, 2019) [32]	Studied power converter topology for FC battery voltage conditioning. The classic boost converter inductor was replaced with a LCL filter. Gave good performance at varying loads.
Study 2.25 (J. Corcau *et al*, 2019) [33]	Modeled & simulated a hybrid aerodrome FC power source consisting of a FC stack, a boost and buck-boost DC to DC converters as well as super-capacitor to provide clean, stable, peak power and energy dynamics during transients.
Study 2.26 (A. Abdelhakim & F. Blaabjerg, 2020) [34]	Proposed a CMMC single-stage solution to interface a low voltage PV and fuel cells DC power supplies to a higher voltage AC load or grid. This offers better performance and is less bulky, contrary to a two-stage boost converter.
Study 2.27 (P. Prabhakaran & V. Agarwal, 2020) [35]	Presented a BDCMG power supply system based on a novel 4-port dual-input dual-output DC-DC converter to interface fuel cells, PV and wind power sources to a low voltage BDCMG. The converter was reliable, compact, versatile and unidirectional with a 93% peak efficiency and a ~87% rated power efficiency.
Study 2.28 (H.S. Youn *et al*, 2020) [36]	Investigated a high power hybrid hydrogen FC railway system with focus on designing an efficient and high power density DC-DC converter. Two DC-DC power converters, namely the IBC and three-level boost converter were researched to determine the most suitable DC-DC boost power converter architecture. The three-level boost converter out performed the IBC in terms of efficiency, power density and dynamic current response and was chosen.

3. Energy Management Systems / Storage (EMSs)

In [37]-[70], EMS simply deals with the partial or overall management / control of a device, a section or the entire system – that is, from when, where and how the energy / power is generated, used, processed, converted and or stored. Furthermore, some housekeeping such as thermal management is carried-out as well. The management performed could be i) on-demand (upon users requests or executions as per system dynamics dictates), ii) on-schedule (pre-programmed to do certain routine tasks at a particular time) and iii) artificial intelligence (based-on machine learning). Usually, a dedicated microcontroller and or power management chip or an adequate computing platform is used to optimally process and execute advanced control algorithms that i) manages power generation devices (fuel cells, solar-cells, wind-farms, TEGs, etc) and supporting systems (water pumps or fans), ii) manages power conversion switching devices (switch ON and Off or pulsing the power ICs or MOSFETs or IGBTs etc as required), iii) monitoring energy storage devices (batteries, super-capacitors / ultra-capacitors etc), iv) controlling the end user applications (e.g. HEV) and finally v) housekeeping (temperature monitoring, timestamp etc) and interacting with the system processes optimally to ensure the closed loop power generation / energy conversion and storage processes are efficient, affordable, quicker, safer and reliable. Examined in what follows are some case studies on power and energy conversions management schemes applicable to FCs and suitable for FC CCHP systems.

3.1. MIL, SIL and PIL Tests for MPPT Algorithm

Investigated by [37], a boost converter is necessary to convert DC voltage to another DC voltage (DC-DC). In their research, solar energy was harvested by PV array and tracked for continuous power generation using model based MPPT technique. The converter contained a MOSFET as the converter switch, which is managed by PWM signal. Once the MOSFET switch is ON, the energy from the PV module is stored in the inductor and the reverse biased diode disengages the output from the PV generator while the output capacitor supplies current to the load. Conversely, when the MOSFET switch is OFF, the inductor is in a discharge state and forward biases the diode to engage the output to the PV generator. The PV panel voltage and inductor voltage (discharging state) combine to give the output voltage, which is always more than input voltage, hence boost conversion. The study was systematically performed in three stages as follows i) model-in-the-loop (MIL), ii) software-in-the-loop (SIL) and iii) processor-in-the-loop (PIL) as depicted in Figure 29 − in which an algorithm with customized variable step was modeled and connected to a simulated PV panel and a boost converter. The MPPT model was simulated first using Simulink and the process is called MIL as shown in Figure 29. The result acquired using MIL test under STC was asserted in the study and as presented in the steady state, the PV power is equal to 60.54W which is the highest power of the Solarex MSX-60 panel under STC (1,000W/m² and 25°C). The study first demonstrated the MIL test result when the irradiance was raised from 500 to 1000 W/m², then reduced to 800W/m² and finally to 600W/m² − the tailored algorithm gave quicker response during irradiance changes and the steady-state oscillations were almost negligible.

The algorithm was changed from MIL format to SIL and the same irradiance test pattern repeated again to test the MPPT tracking and the same result was achieved similar to that of MIL.

Finally, the code was changed to PIL format and the same irradiance test pattern repeated again to test the MPPT tracking using hardware-in-the-loop as illustrated in Figure 29, which also gave the same result. This concludes that the implemented MPPT algorithm is accurate, as all the three algorithm formats gave the same outcomes. This approach can be applicable also to fuel cells.

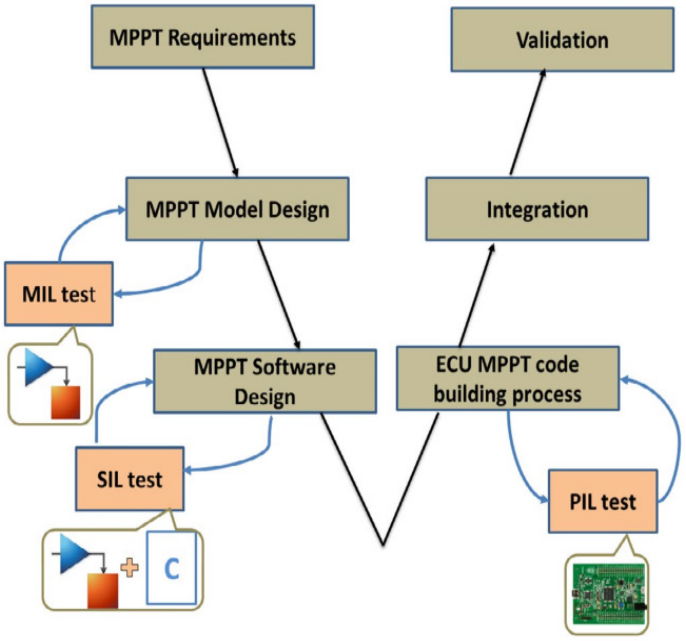


Figure 29: MIL, SIL and PIL Tests for MPPT Algorithm (adapted from [37])

3.2. Review on EMS for FCs Hybrid Electric Vehicle: Issues and Challenges

According to [38], different ways of using a battery to supplement a fuel cell to reliably supply power without experiencing the fuel cell fuel starvation phenomenon were investigated. The basic rationale in the literature was to formulate various types of power converters and energy management systems / storage (EMS) governed by different control strategies − which include but not limited to the followings techniques a) fuzzy logic [39], b) power frequency splitting [40], c) space dynamic equation, d) deterministic dynamic programming [41], e) neural network optimization algorithm [42], etc.

Furthermore in [38], super-capacitor (SC) instead of a battery, was used to supplement a fuel cell. Super-capacitors are known to have very high power density (relative to a battery or FC), enabling it to react faster in transient conditions of brief high current demand. This method requires as well various topologies of power converters and EMS governed by different control techniques, which include but unlimited to the following methods a) differential flatness controls [43], b) polynomial control technique [44], c) wavelet-based load-sharing algorithm [45], d) fuzzy logic [46], e) wavelet adaptive linear neuron (WADALINE) [47], f) adaptive optimal control algorithm (AOCA) [48], etc.

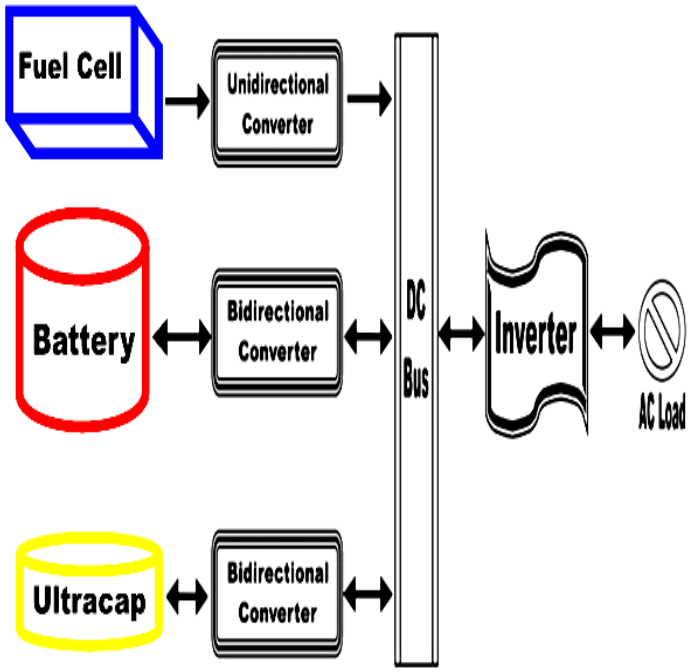


Figure 30: Fuel Cell, Battery and Ultra-capacitor Hybrid Power System (redrawn from [38])

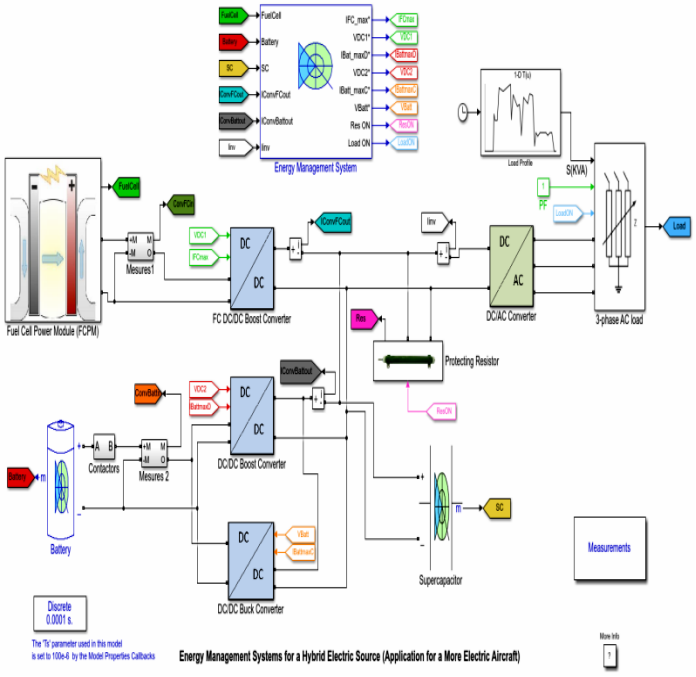


Figure 31: Hybrid Fuel Cell, Battery and Super-cap with EMS (adapted from [57])

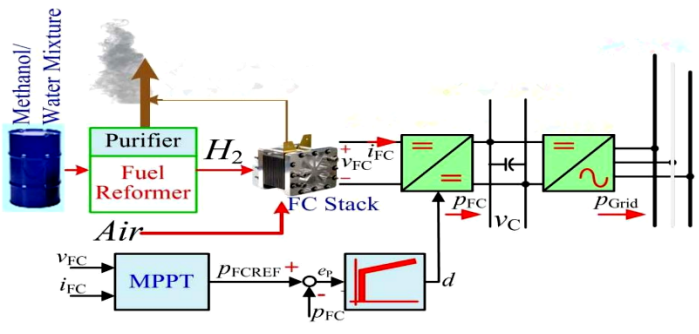
Finally, the third setup as stated in [38], involves all three – the FC, battery and finally the super-capacitor all connected in parallel. This setup requires as well various topologies of power conversion and EMS governed by different control methods, which include but unlimited to the following approaches a) proportional integral (PI) regulator [49], b) fuzzy logic [50], c) various FC, battery and SC configurations [51], d) traction control method [52], e) flatness control technique [53], f) PWM control [54], state machine strategy [55] PI and nonlinear sliding mode controllers [56] etc. The fuel cell, battery and super-capacitor technique is the most effective and widely used, as it provides both high energy and high power densities, as well as storage when needed. Figure 30 summarily depicts this technique.

3.3. A Comparative Study of EMS Schemes for a FC Hybrid Emergency Power System of More-Electric Aircraft (MEA)

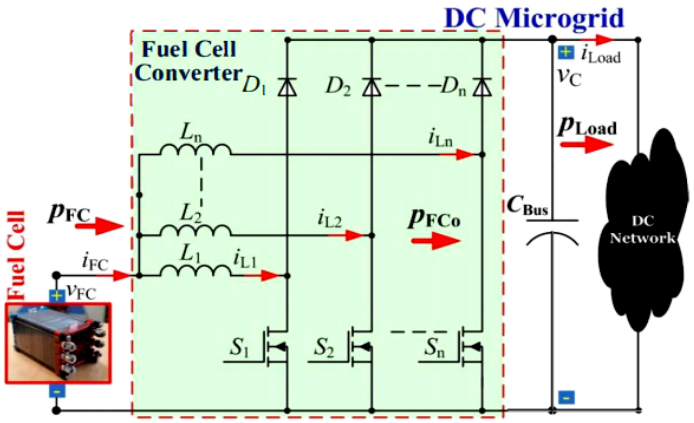
Researched in [57], an articulation of assorted EMS for a fuel cell-based emergency power system of a More-electric aircraft was presented. Akin to Figure 30, the fuel cell hybrid system comprises of a FC, Li-ion battery and super-capacitor, together with DC-DC converters and DC-AC inverter as shown in Figure 31. The EMS techniques comparatively studied include those used in FC vehicle applications such as the proportional integral (PI), the state machine, the fuzzy logic /frequency decoupling, the equivalent consumption minimization and the rule-based fuzzy logic strategies. The main metrics used to compare the various EMS strategies performance are the i) the H_2 consumption, ii) state of charge of the batteries / super-capacitors and iii) general system efficiency. Lastly, a novel technique using the wavelet transform of their instantaneous power, was used to measure the tensions on each energy source to determine the impact on their life cycle. Simulation models as well as an experimental test setup were developed to simulate and practically verify the study.

3.4. Model-Free Control of Multi-phase IBC for FC / Reformer Power Generation

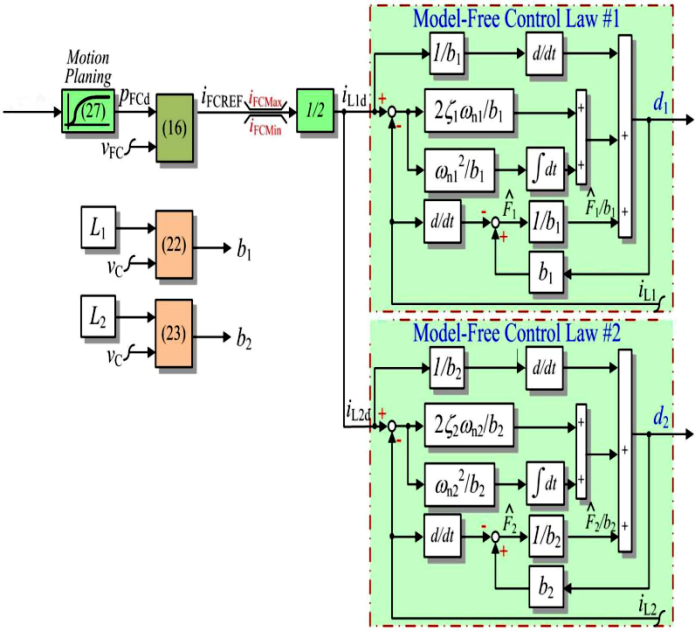
Fuel cells require power converters to boost their low DC output voltage, as well as a control mechanism to optimize its operation. According to [58], the regulation parameters are set using a linear method to assess the convergence problem; as a result, they developed further a model free control (MFC) to manage the fuel cell power for DC micro-grid applications. In their approach, a 2-phase interleaved boost converter was implemented to address the non-linear control problem. Relative to PI and flatness control techniques, a MFC is simple and don't need precise info of the DC micro-grid parameters, though MFC still needs to know the power converter inductances value. The simulated design was done using dSPACE MicroLabBox and practically tested using a 50V 2.5kW PEMFC with two 2.5kW converters connected in parallel to the FC output and both tests correlated with excellent performance. Figure 32a depicts the FC power plant overview and Figures 32b and 32c respectively represent the IBC architecture and a two-phase MFC technique.



(a) FC / reformer power plant for grid connected applications



(b) Multi-phase parallel IBC for FC applications

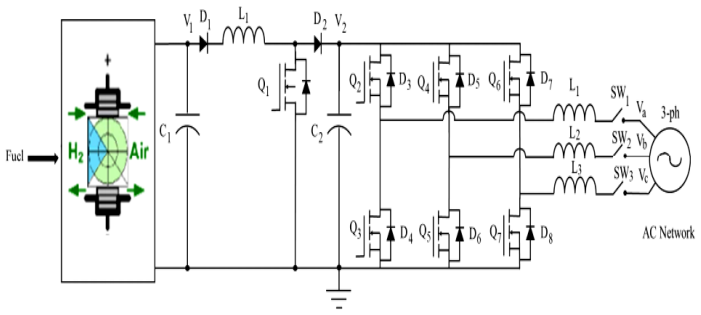


(c) Suggested MFC of the FC power for the multi-phase FC power converter

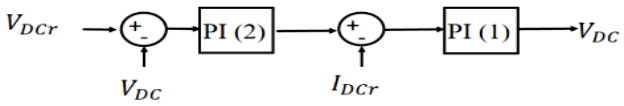
Figure 32: FC Power Plant, Power Converter and MFC (adapted from [58])

3.5. Control and Grid Connection of a FC Power System

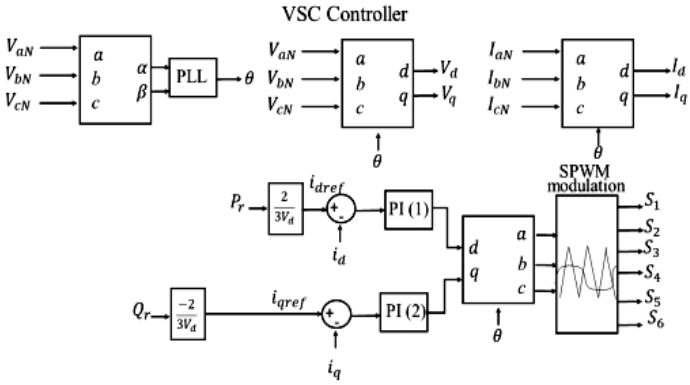
As now known, FC is gaining traction in micro-grids and other applications due to their environmental friendliness. Studied in [59] is 900 cells (0.7V per cell) 625V PEMFC stack connected to a 3-phase electrical network using a 700V DC-DC conventional boost converter and a 420V voltage source converter (VSC) DC-AC inverter. PI linear controllers are used in the power converter to monitor the voltage/current and to regulate the electrical dynamics needed to reliably supply power to the grid. The VSC regulates autonomously the active and reactive powers injected to the grid, using two linear control loops PI(1) and PI(2) and a sinusoidal pulsed width modulation (SPWM) scheme. MatLab / Simulink was used to model and simulate the design and the PI controller could reach steady state within 50ms. The VSC inverter controller was able to reach steady state within 30ms when the active and reactive powers were doubled. Figure 33a illustrates the FC electrical network and Figures 33b and 33c, respectively denote the PI and VSC controllers strategy.



(a) The fuel cell stack, power converter and 3-phase grid



(b) The boost converter PI controller sketch

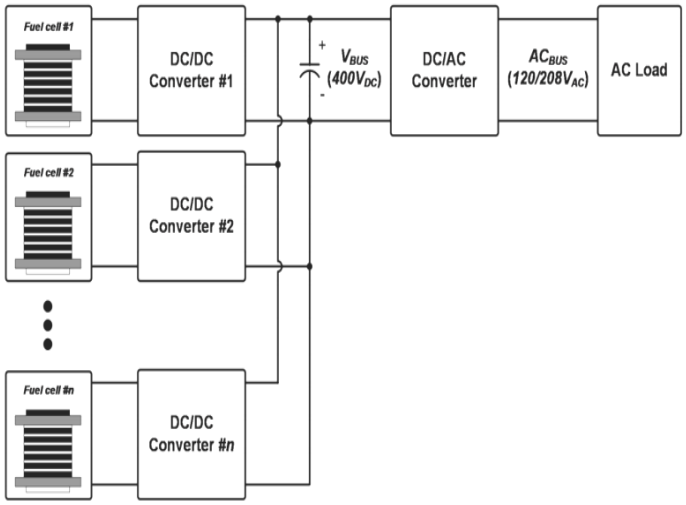


(c) The VSC control and modulation strategy

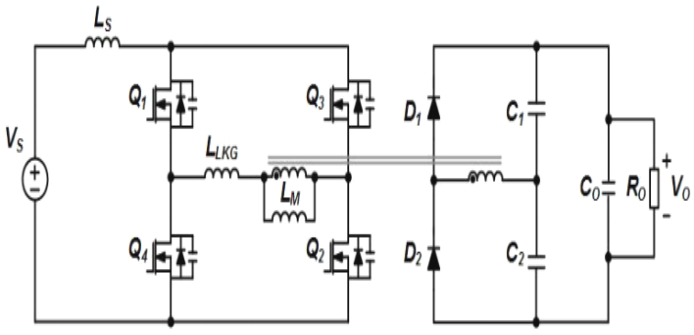
Figure 33: The FC, Power Converter and Electrical Grid with PI and VSC Controllers Schemes (adapted from [59])

3.6. A Novel Control Scheme for High Efficiency FC Power Systems in Parallel Structure

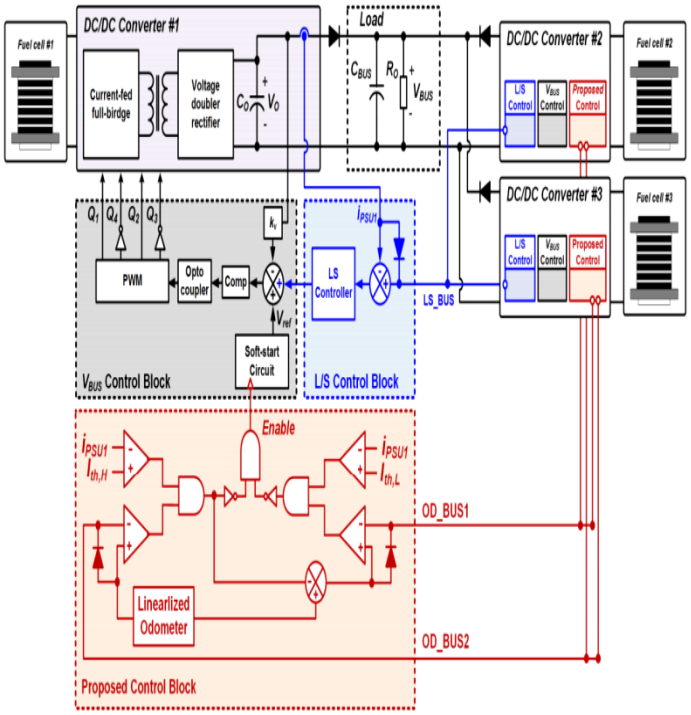
Discussed in [60] is a basic control technique for greater efficiency power converters of a FC distributed generation (DG) system shown in Figures 34a-34c. Usually, multiple FCs and power converters are connected in parallel to meet the power rating required for a FC DG systems. However, power systems have three main losses; namely core, switching and conduction losses − the switching and core losses are insensitive to load fluctuations, whereas the conduction loss is proportional to power output. Therefore, when power systems work under light-load conditions, the switching and the core losses can significantly contribute to the total losses, as the conduction loss will be small. As a result, the traditional paralleling approach entails the power system operates the same irrespective of the load size, making the power system in-efficient at light-load (small current) conditions, due to the predominantly switching and core losses. Therefore, the parallel system efficiency under light-load is enhanced by changing accordingly the quantity of parallel power units to meet just the light-load demand − doing so substantially reduce the switching and core losses, as less power units will be operating and more can be added under heavy-load. Three 300W units were paralleled to achieve a 900W FC DG efficient system.



(a) Parallel FC DG system



(b) Current-fed isolated full-bridge power converter

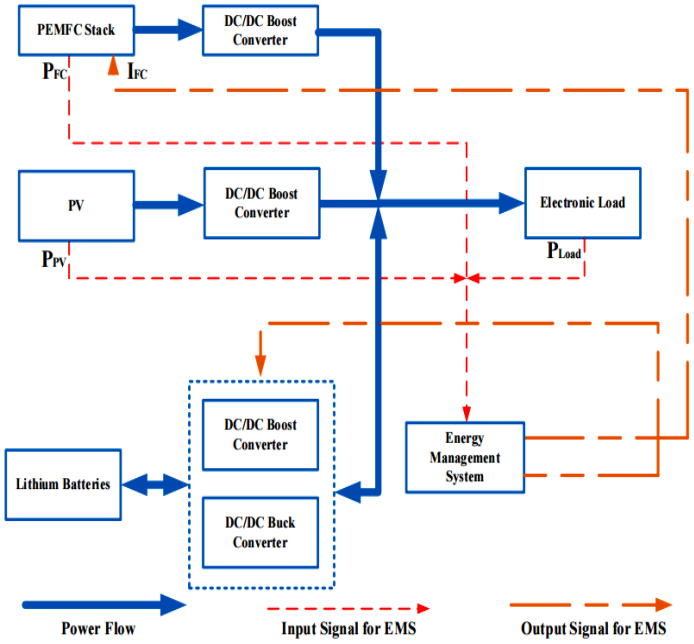


(c) Proposed control scheme block diagram

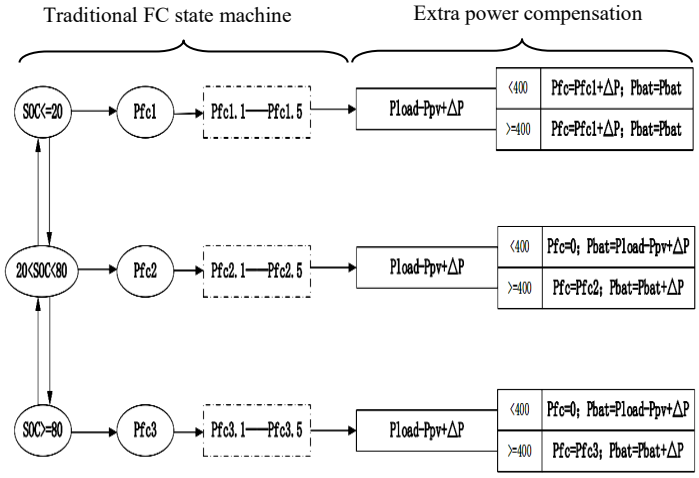
Figure 34: The Proposed FC System, Power Converter and Controller Scheme (adapted from [60])

3.7. An EMS Strategy Based-on State Machine with Power Compensation for PV-PEMFC-Li-ion Battery Power System

Investigated in [61] is a hybrid power supply system constituting a PEMFC, PV and auxiliary Li-ion battery for electric vehicles as exemplified in Figure 35a. A conventional FC and solar cell DC-DC boost converters are used and for a Li-ion battery, a boost-buck power converter is employed. To efficiently coordinate the different power / energy sources and stabilize the DC bus voltage, a state machine EMS control technique with power compensation was employed. The rationale is to minimize the frequency of the PEM FC power output variations and ensuring the Li-ion battery charges and discharges within the ideal intervals. Figure 35b illustrates the state machine EMS used to adjust the FC voltage and Li-ion SoC to attain optimal results.



(a) FC, TV and Li-ion Power System Overview



(b) Optimized FC state machine EMS with power compensation

Figure 35: Proposed Power System and State Machine EMS (adapted from [61])

3.8. Development of a Fuzzy-Logic-Based EMS for a Multi-port Multi-operation Mode Residential Smart Micro-grid

Demonstrated in [62] is an advanced grid-tied residential smart micro-grid composing of a fuel cell, solar cell and battery bank to supply the local loads using both electric and magnetic buses. Typically, an electric bus comprising of multiple converter based micro-grids is used; however, this setup is costly and bulky with numerous and large conversion stages; thus, the addition of a common magnetic bus with multi-port converters circumvent these shortcomings and furthermore isolate the conversion ports. Their hybrid architecture with EMS translates to a centralized quicker and versatile system. The suggested micro-grid was capable of working in multiple grid-tied and off-grid modes using a fuzzy logic energy management unit (EMU) controller to choose the proper mode of operation − taking into cognizance short and long-term energy generation and usage. The micro-grid operation performance was enhanced using synchronized bus-voltage balance control technique. The executions of the micro-grid and EMU were experimentally tested for three different cases of the residential load in grid-connected and off-grid modes. The energy distribution and cost analyses for each case show the merits of the EMU for both the grid and user. The various control schemes for each of the power converters/inverters is detailed in the full text. Figure 36 exemplifies and summarizes the research.

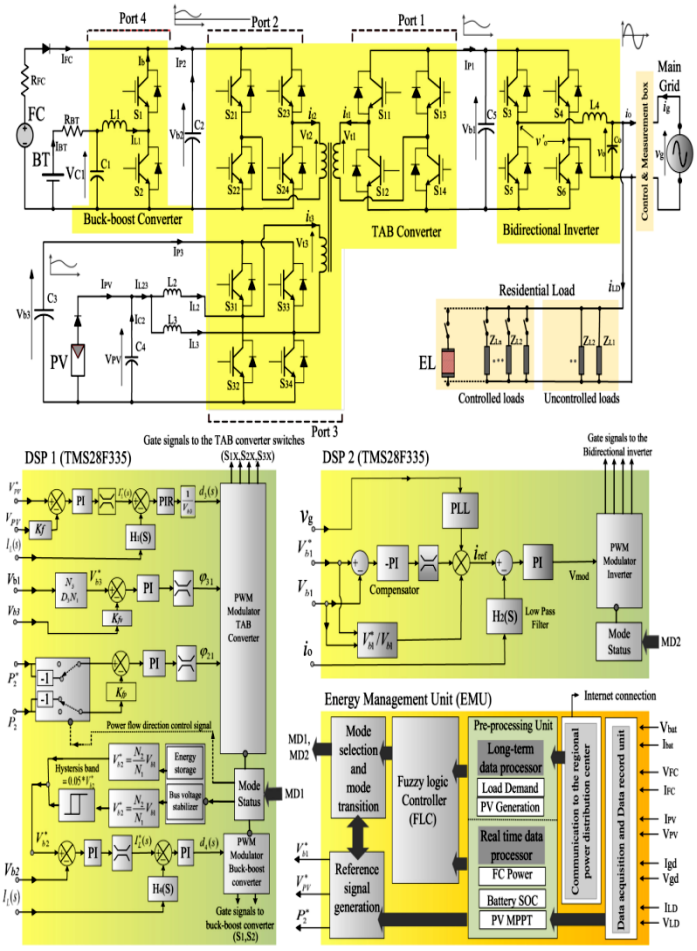
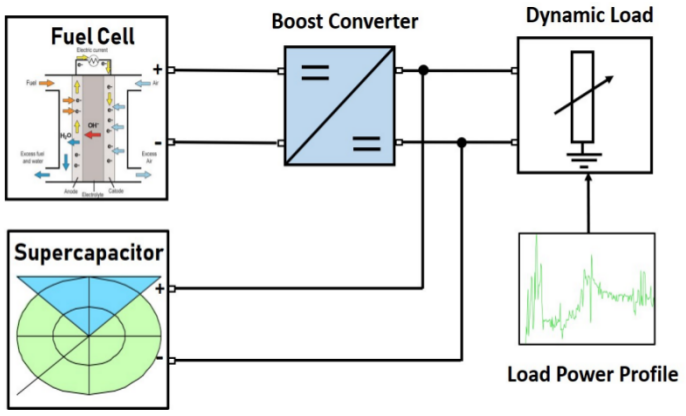


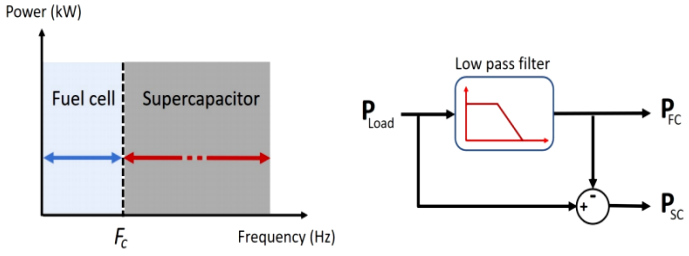
Figure 36: Proposed Smart Micro-grid including Converters, Controllers and EMU (adapted from [62])

3.9. Frequency Separation-based Power Management Strategy for a FC-Powered Drone

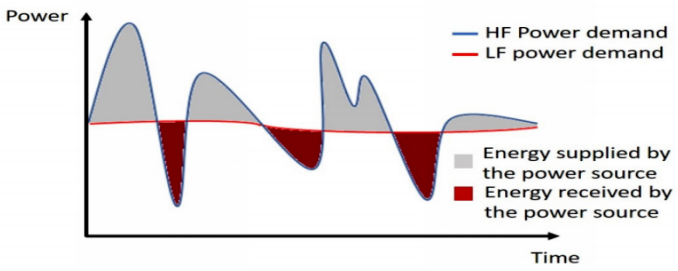
Studied in [63] is a hybrid FC and super-capacitor with a DC-DC boost converter power system for drones depicted in Figure 37a. The EMS control technique exploited is routed-in frequency separation-based technique whereby the required power is shared between the energy sources − in this case, the FC and super-capacitor. Depicted in Figures 37b and 37c, the drone flight load profile is divided into low and high frequency components, in which the FC connected to the DC-DC boost converter is controlled to handle the low frequency dynamics whereas the supercapacitor handles the high frequency dynamics during peak power demands as expatiated in Figure 37c. The system was simulated using a real power profile from a small hexacopter experimental flight test and the results justify the EMS was capable of minimizing the fuel cell power variations with the supercapacitor handling all of the transient / peak power demands, consequently prolonging the FC lifetime and drone flight periods.



(a) Drone FC and super-capacitor power system overview



(b) Frequency power sharing concept

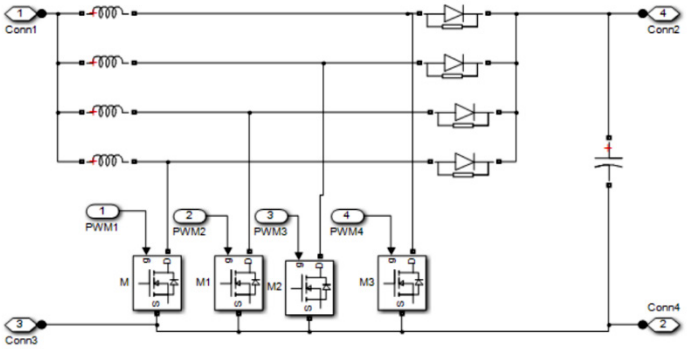


(c) Power frequency sharing

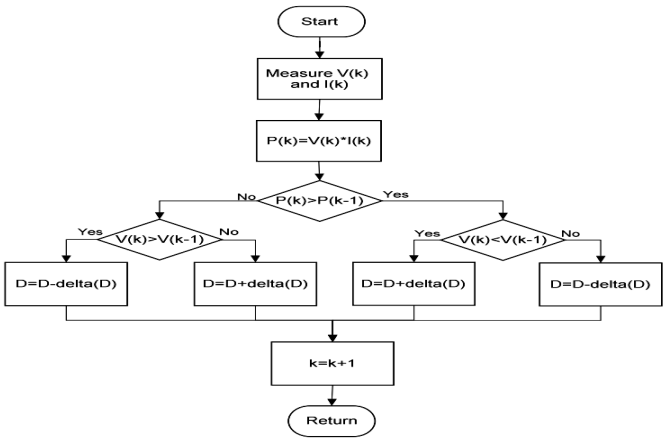
Figure 37: Proposed Drone Hybrid Power System and Frequency Separation EMS (adapted from [63])

3.10. MPPT Control of an IBC for a PEM FC Applications

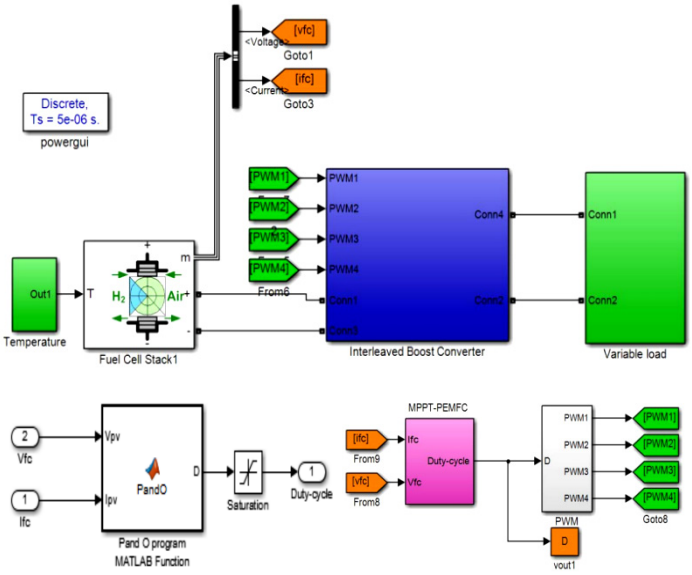
Stated in [64] is simply how FC power can be stepped-up using a four-phase IBC and controlled efficiently using a MPPT P&O EMS. The IBC further reduced the FC voltage and current ripples whereas the MPPT ensured max power is extracted from the fuel cell. Figures 38a-38c summarize the implementation.



(a) Four-phase IBC Simulink model



(b) Perturb and Observe MPPT

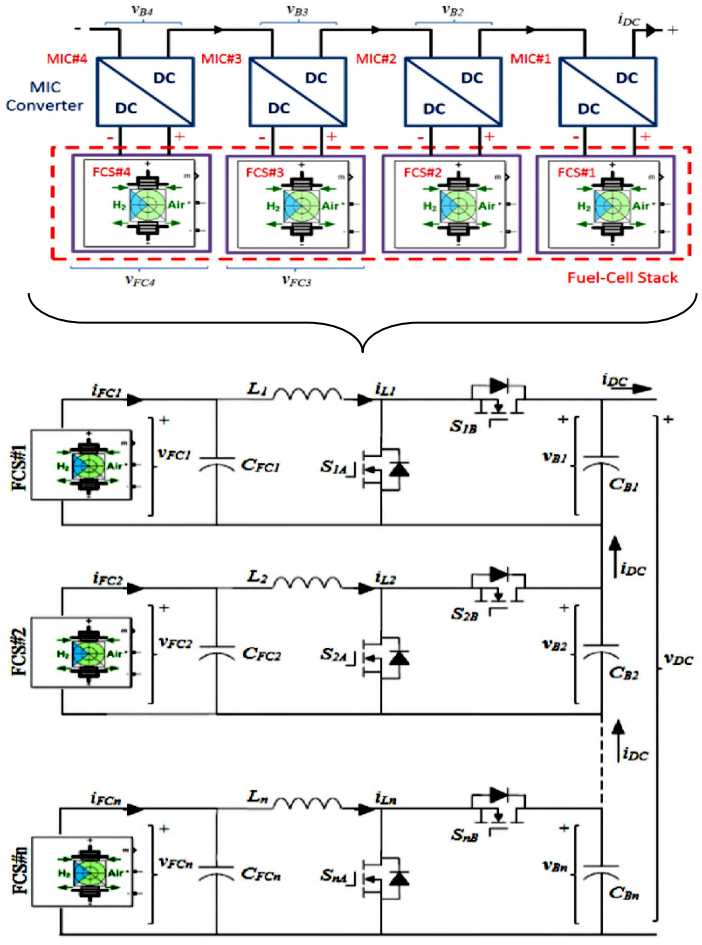


(c) IBC with MPPT Simulink model

Figure 38: FC and Four-phase IBC with MPPT P&O EMS (adapted from [64])

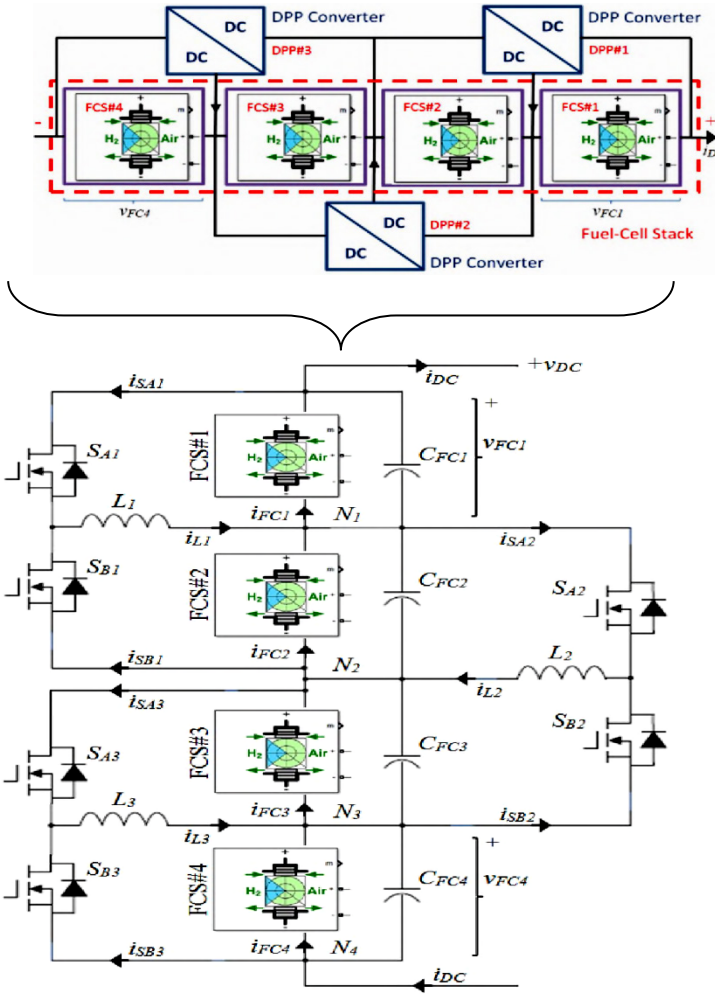
3.11. Power Flow Control via Differential Power Processing to Enhance Reliability in Hybrid Systems based on PEM FC

Presented in [65] is an interesting study on interconnecting fuel cells to obtain maximum output power. Two interconnection techniques; namely, i) modular integrated converter (MIC) and ii) differential power processing (DPP) including their power converters (synchronous switching bidirectional buck-boost) and EMS technique (MPPT P&O / Hill Climbing (HC)) were modeled using Matlab and discussed in details. As depicted in Figure 39a, the MIC topology has each FC connected in parallel to its own separate converter and each converter is in turn connected in series. This allows several converter topologies and control schemes to be implemented independently. However, the main disadvantages are the number of converters used, the cost involved and poor conversion efficiency − as 100% of the power produced by each FC sub-module is processed. As displayed in Figure 39b, the DPP architecture simply has two FCs connected in parallel with a single buck-boost converter, thereby reducing the total number of power converters by one. Further advantages include simplicity, speed, affordability and improved efficiency − since only a fraction of the FC sub-module generated power is processed. However, the main disadvantage is the complexity in the control technique implementation, since the active balance needs to communicate with each FC sub-module to apply MPPT.

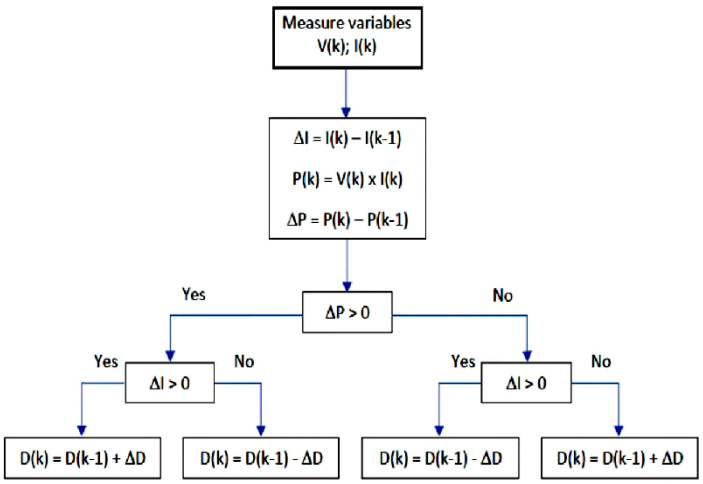


(a) MIC topology showing FCs and converters interconnections

Using Matlab / Simulink, MPPT P&O algorithm illustrated in Figure 39c was implemented on the DPP converter to achieve a fast control loop. The simulation verified the merits of the DPP topology and the MPPT P&O algorithm convergence technique.



(b) DPP topology showing FCs and converters interconnections

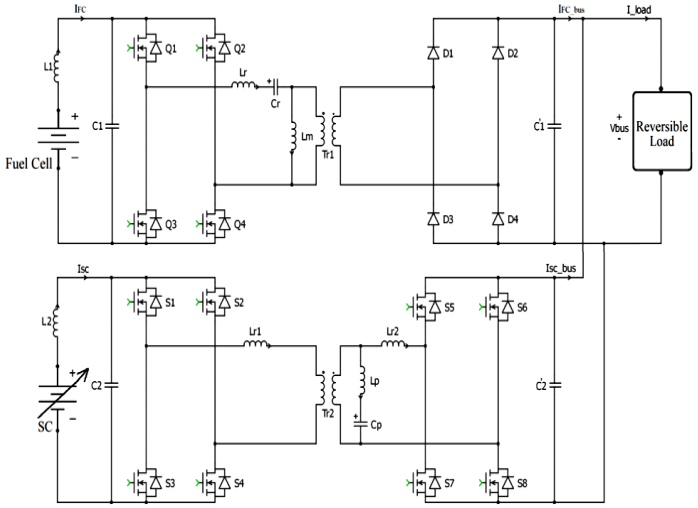


(c) HC / P&O MPPT control algorithm

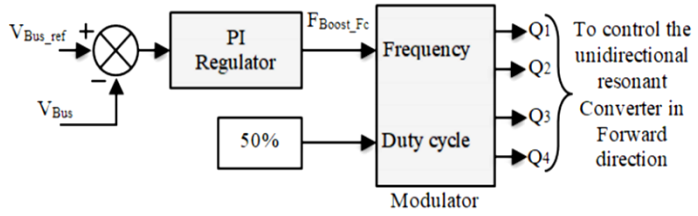
Figure 39: MIC, DPP and MPPT P&O Techniques (adapted from [65])

3.12. EMS in a Multi-source System using Isolated DC-DC Resonant Converters

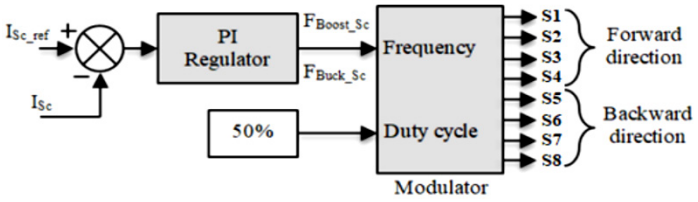
Described in [66] is a hybrid FC and super-capacitor with PI controller power system. The research purpose was to develop a control mechanism for the fuel cell and super-capacitor DC-DC resonant power converters to share the system power according to their dynamic responses. Usually, the FC has a slower power dynamic relative to the super-capacitor; therefore, initially the super-capacitor will handle the instant peak power demands whereas the FC provides the bulk of the power during steady state. The isolated LLC resonant converter connects the FC to the DC bus, whereas the super-capacitor connects to the DC bus via the bidirectional resonant converter which charges and discharges the super-capacitor. The modeling and simulations were done using Matlab / Simulink with PLECS and the findings affirm the merits of using resonant converters − which also offer isolation and reduced switching losses. Portrayed in Figure 40a is the FC and super-capacitor hybrid power system and Figures 40b and 40c respectively lucubrate FC and super-capacitor PI controllers.



(a) FC and super-capacitor hybrid system



(b) Fuel cell voltage control loop

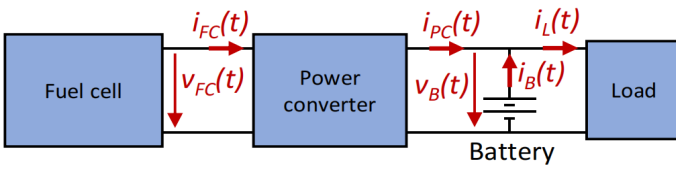


(c) Super-capacitor's bidirectional control loop

Figure 40: FC, Super-capacitor Hybrid Power System with PI Controller (adapted from [66])

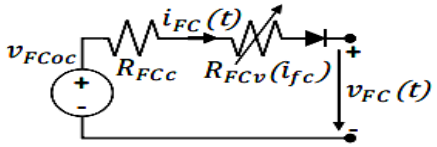
3.13. EMS Optimization for a FC Hybrid Vehicle based on Power Losses Minimization

Indicated in [67], fuel cells hybrid vehicle is a suitable alternative to internal combustion engine vehicles, as they are environmentally friendly. Their research thus proposed an energy management optimization technique for the power distribution system, to increase the driving range of fuel cell hybrid vehicles. The hybrid energy system constitutes a FC connected to a power converter which in turns connects concurrently to a battery and a DC load as represented in Figure 41a. The suggested control optimization algorithm is based on minimizing the energy losses in the system. Using Simulink / PSIM; the losses, costs, size and mass were evaluated, in which it was found that lower FC power and higher battery capacity offer low energy losses and low consumption; whereas maximum fuel cell power and lowest battery capacity provide minimum costs, mass and size. The results correlated other studies in the literature. Figures 41b and 41c depict the basic FC and battery models respectively and Figure 41d shows the optimization algorithm high level overview.



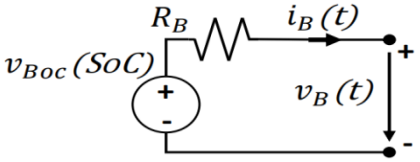
(a) Fuel cell, power converter and battery hybrid power overview

$$v_{FC}(t) = v_{FCoc} - R_{FCc} \cdot i_{FC}(t) - R_{FCv}(i_{FC}) \cdot i_{FC}(t) \quad (1)$$

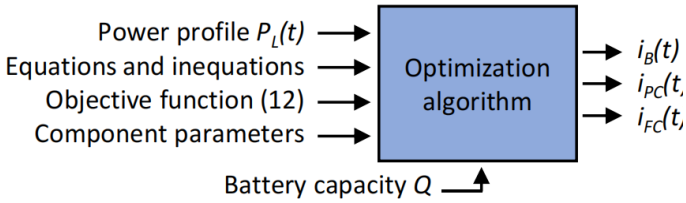


(b) Super-capacitor simplified model

$$v_B(t) = v_{Boc}(SoC) - R_B \cdot i_B(t) \quad (2)$$



(c) Battery basic model

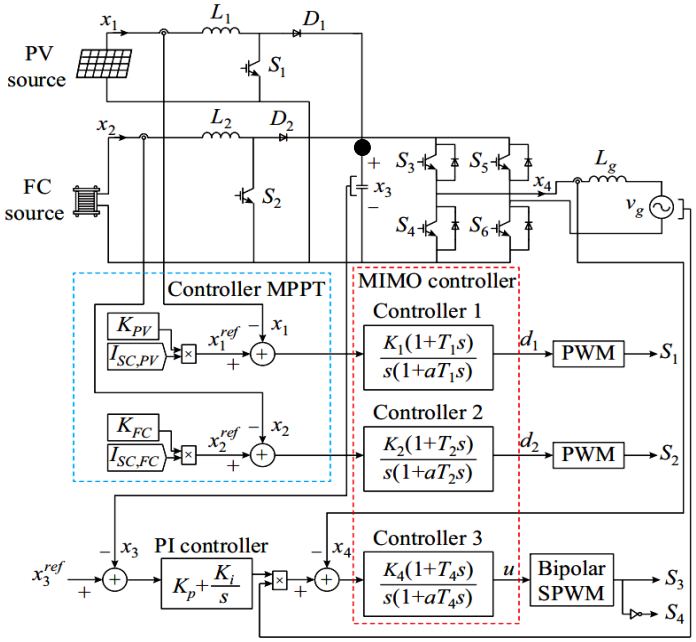


(d) Optimization algorithm input and output data

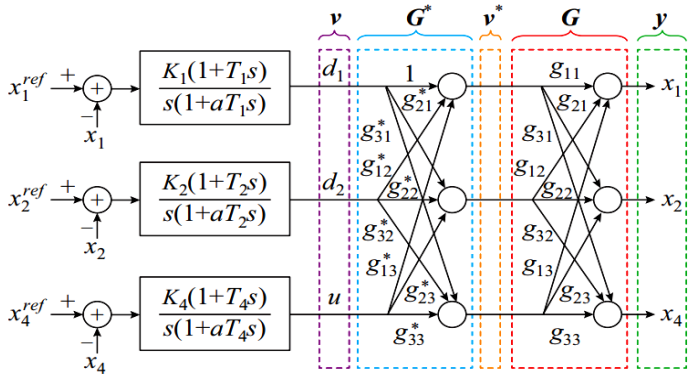
Figure 41: System Overview, Fuel Cell and Battery Models and Optimization Algorithm Overview (adapted from [67])

3.14. Dynamic Modeling and Closed-loop Control of Hybrid Grid-connected Renewable Energy System with Multi-input Multi-output Controller

Proposed in [68] and summarized in Figure 42a, is the use of multi-input multi-output (MIMO) technique to dynamically model and closed-loop control a hybrid grid-tied renewable energy system. The system constitutes a solar cell and FC each respectively connected to their boost converters which are in turn connected in parallel to a single-phase H-bridge inverter to supply an AC load. The system employed the traditional MPPT and PI control techniques compensated by the MIMO network detailed in Figure 42b. Using the system transfer functions frequency response, the MIMO controller gains are tuned. MatLab and Simulink were used to simulate and analyse the designed MIMO controller accuracy and effectiveness and from the results, MIMO is quick and stable at various functional points, having a negligible steady-state error as well as with a grid THD of ~1.48% in accordance with standards of distribution networks.



(a) Suggested control structure for grid-connected PV/FC hybrid energy system

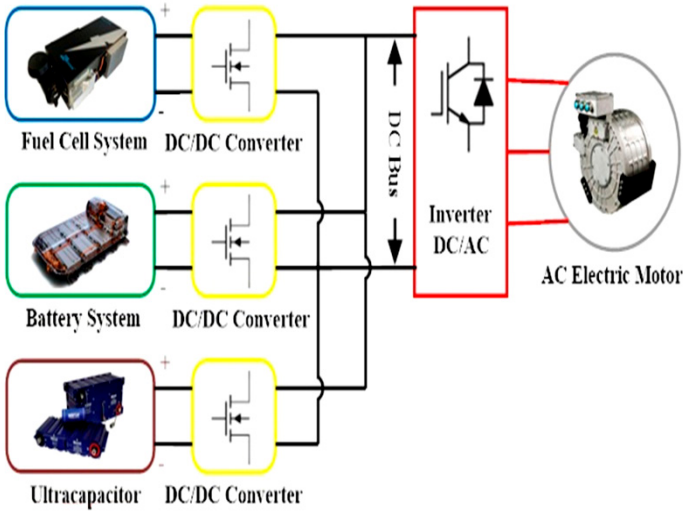


(b) Compensation network

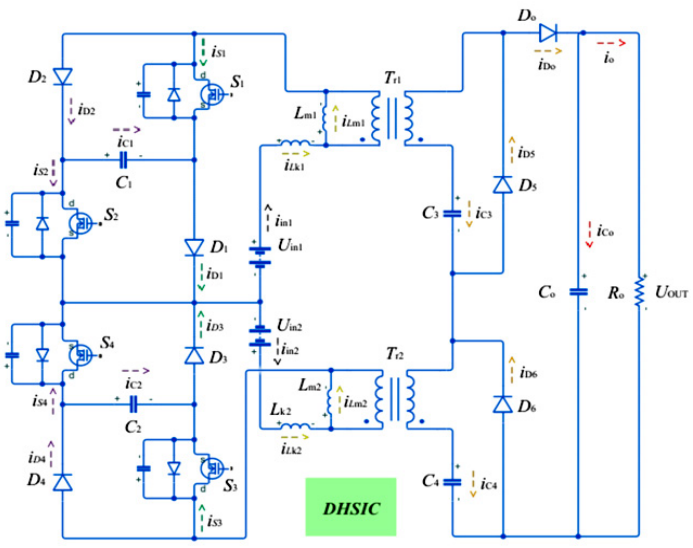
Figure 42: Proposed FC and PV Hybrid Energy System and Compensation Network (adapted from [68])

3.15. FCEVs — A Brief Review of Current Topologies and EMS Strategies

Articulated in [69], advancement in technology and new international policies on electric / hybrid electric vehicles are becoming trendy. In light of this, their research focused on fuel cells and energy storage devices as well as power converters and EMS techniques to sustain hybrid electric vehicle dynamic power demands. Different fuel cell energy / power configurations and power converters topologies were assessed and the highlights are presented in Figures 43a and 43b − respectively a fuel cell with storage devices and dual-input high step-up isolated converter. Furthermore, miscellaneous EMS strategies were examined with focus on energy efficiency, usage of hydrogen and sub-systems decay involved. The pros and cons of rule-based, learning-based and optimisation-based EMS strategies were discussed and the conclusion is to hybridize modern and existing strategies to eliminate the uncertainties regarding EMS techniques robustness.



(a) FC, battery and ultra-capacitor hybrid electric vehicle configuration

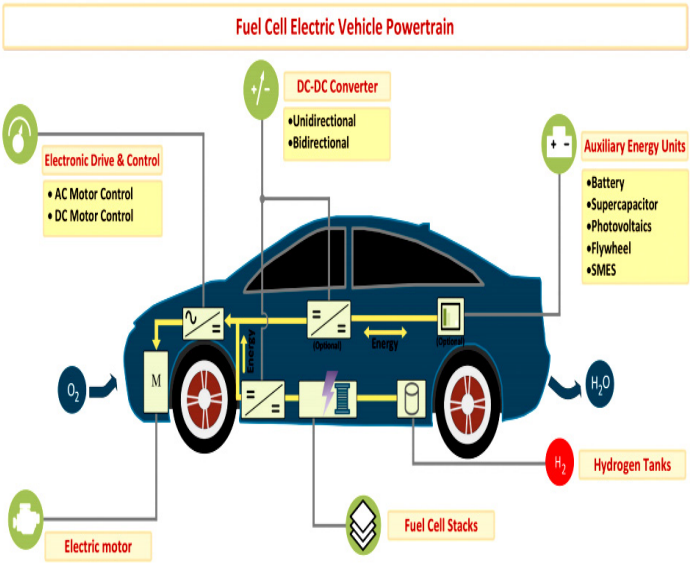


(b) Dual-Input High Step-Up Isolated Converter (DHSIC)

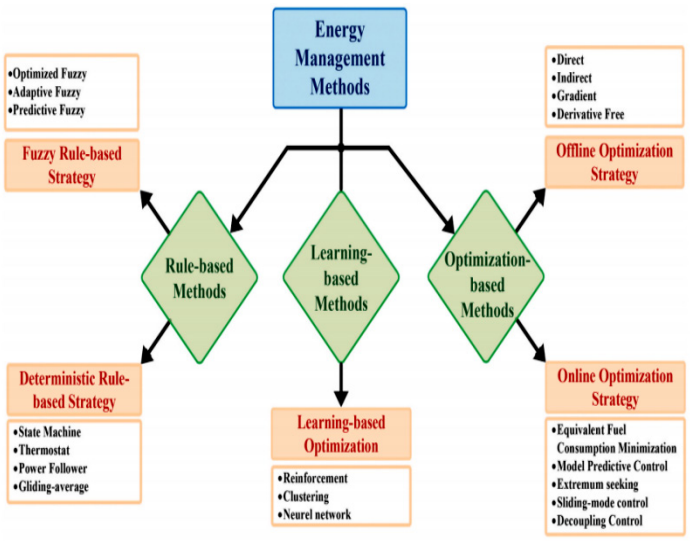
Figure 43: FC Hybrid Electric Vehicle and DHSIC Scheme (adapted from [69])

3.16. A Review and Research on FCEVs: Topologies, Power Electronic Converters, EMS Methods, Technical Challenges, Marketing and Future Aspects

Analyzed extensively in [70], fuel cells are the future especially for FCEV as shown in Figure 44a. In this respect, a comprehensive study of types of FCs with electric motors are explained with focus on their areas of applications, diagnostic properties and working environments. Furthermore, power converters which boost the FC voltages to drive different motor topologies used in FCEVs, are elaborated based on their structural frequency of use, their architecture and difficulty. Summarized in Figure 44b, assorted FCEVs power converters EMS schemes and technical challenges were examined and the final closing remarks highlighted the present status and future prospects using significant number of marketing and target data.



(a) FCEV power transmission system with auxiliary power supplies



(b) FCEVs EMS classification of technical challenges and system problems

Figure 44: FCEV Power Transmission System and EMS Schemes and Challenges (adapted from [70])

Table 2: EMS Case Studies Examined Summary

EMS Research	Highlights, Merits and Demerits
Study 3.1 (S. Motahhir *et al*, 2017) [37]	MIL, SIL and PIL tests for MPPT algorithm. Implemented MPPT algorithm on each and all three formats reasonably gave similar results.
Study 3.2 (N. Sulaiman *et al*, 2015) [38]	Extensive analysis on FC fuel starvation and EMS schemes for FC HEV: In-depth FC issues, challenges & solutions were presented.
Study 3.3 (S.N. Motapon *et al*, 2014) [57]	Implemented simulated and experimental test frameworks for relative analyses of various EMS methods for a FC hybrid power system.
Study 3.4 (P. Mungporn *et al*, 2019) [58]	Developed further a model free control (MFC) to manage the fuel cell power for DC micro-grid applications. In their approach, a 2-phase interleaved boost converter was implemented to address the non-linear control problem and the simulated and practical results correlated.
Study 3.5 (G.G. Suárez-Velázquez *et al*, 2020) [59]	Used VSC to autonomously regulate the active and reactive powers injected to the grid, via a sinusoidal SPWM strategy and two linear control loops PI(1) and PI(2). The VSC was able to reach steady state within 30ms when the active and reactive powers were doubled.
Study 3.6 (Y. Jeong *et al*, 2019) [60]	The efficiency of parallel FC systems under light-load is enhanced by changing accordingly the quantity of parallel power units to meet the light-load demand and substantially reduce the switching / core losses. Three 300W units were paralleled to achieve a 900W efficient system.
Study 3.7 (Y. Zhang *et al*, 2019) [61]	A state machine EMS control technique with power compensation was employed to curb the frequency of the PEM fuel cell power output variations and ensuring the Li-ion battery charges & discharges within the ideal periods.
Study 3.8 (M. Jafari *et al*, 2019) [62]	Demonstrated an advanced grid-tied household smart micro-grid consisting of a fuel cell, solar cell and battery bank to provide the local loads using both electric and magnetic buses. The magnetic bus with the multi-port converters augments the electric bus and further isolates the conversion ports. The EMU controller was tested for three distinct scenarios of the home load and the distribution of energy and cost analyses for each case, show the EMU merits.
Study 3.9 (M.N. Boukoberine *et al*, 2020) [63]	The EMS method used is routed-in frequency separation-based scheme, whereby the required power is shared between the energy sources.
Study 3.10 (E.M. Barhoumi *et al*, 2020) [64]	FC power can be stepped-up using a four-phase IBC and controlled efficiently using a MPPT EMS scheme to reduce power ripples.
Study 3.11 (K.S. Artal-Sevil *et al*, 2020) [65]	Presented MIC and DPP connections of FCs with power converters. MPPT was used on the DPP converter to achieve a fast control loop.
Study 3.12 (M. Arazi *et al*, 2020) [66]	Developed a mechanism for FCs and SCs DC-DC resonant power converters to share the system power using their dynamic responses.
Study 3.13 (A. Martin-Lozano *et al*, 2020) [67]	Proposed an EMS optimization technique for the power distribution system to increase the driving range of fuel cell hybrid vehicles. The results correlated other studies in the literature.
Study 3.14 (M. Salimi *et al*, 2021) [68]	Suggested the used of MIMO technique to dynamically model and closed-loop control a FC hybrid grid-tied renewable energy system.
Study 3.15 (I.S. Sorlei *et al*, 2021) [69]	Used FCs and energy storage devices as well as power converters and EMS techniques to sustain hybrid electric vehicle dynamic power.
Study 3.16 (M. İnci *et al*, 2021) [70]	Reviewed extensively FC types with electric motors with focus on their uses, diagnostic properties and finally working environments.

Table 2 summarizes the reviewed EMS FC applicable studies.

4. Conclusion

Energy insecurity and electrical energy in particular, is a progressive pressing societal crisis in South Africa and Africa. In this regard, this article articulated a structural review of forty four different power converters and EMS research case studies to reasonably choose and develop a suitable FC power converter and EMS scheme for a hybrid FC CCHP system for households / commercial applications. From the review, it was observed that the power converters based on IBC / variants and as well isolated boost converters were of interests. IBC are simple, more robust, good for ripples reduction and peak power applications. However, the fundamental IBC topology is not isolated and adding isolation transformers offers protection but increases the costs and size. Likewise, EMS techniques can be grouped under rule-based, learning-based and optimization-based but the most popular EMS strategy used with power converters are the MPPT and PI controller. Furthermore, FCs can also be modularized with each FC sub-module having its own power converters and EMS scheme to increase the system efficiency. In sum, there is no method that is flawless − choosing a particular approach and trading-off different features depend on the targeted applications and the research objectives; whether to maximize efficiency, robustness, safety, performance etc and minimize costs, size, noise, complexity etc. For our research project, power converters based-on IBC variants and BDC with EMS based on MPPT and or PID controller for use with FC, Li-ion battery, ultra-capacitor and thermo-electric devices are considered to investigate further the CCHP system postulated in Figure 45 to conclude the study.

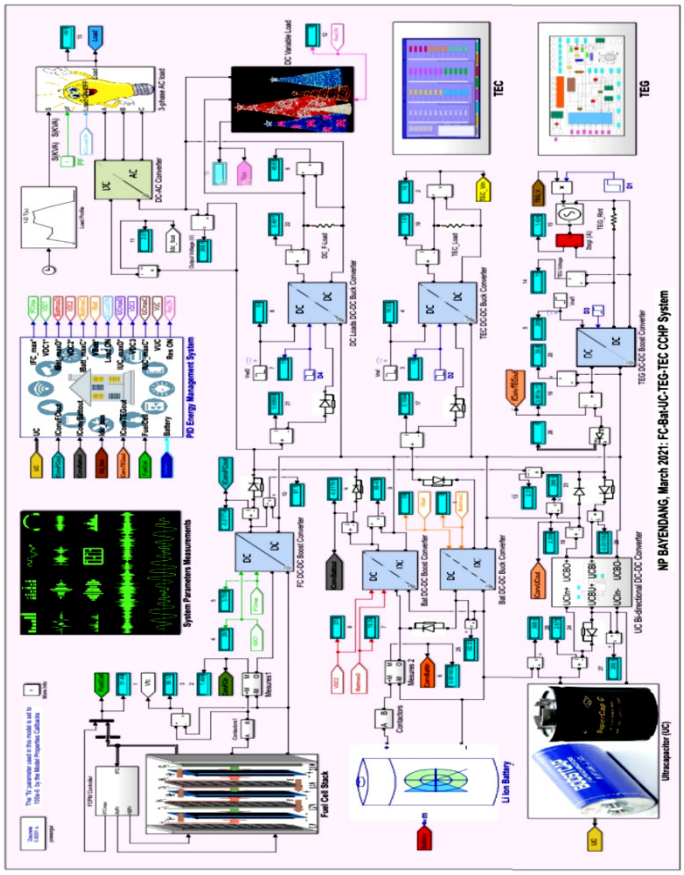


Figure 45: Postulated Fuel Cell CCHP System Undergoing Research

Glossary

AC	Alternating Current
AOCA	Adaptive Optimal Control Algorithm
BDC	Bidirectional DC-DC Converter
BDCMG	Bipolar DC Micro-grid
CCHP	Combined Cooling, Heating and Power
CMMC	Current-fed Modular Multilevel Converter
CPC	Conventional Phase-shift Control
CPUT	Cape Peninsula University of Technology
DAFB	Dual Active Full Bridge
DC	Direct Current
DC-IBC	Direct Coupled-Interleaved Boost Converters
DG	Distributed Generation
DHB	Dual Half Bridge
DHSIC	Dual-Input High Step-Up Isolated Converter
DPP	Differential Power Processing
EGS	Energy Generation System
EIS	Electrochemical Impedance Spectroscopy
EMI	Electro Magnetic Interference
EMS	Energy Management Systems / Energy Management & Storage
EMU	Energy Management Unit
FC	Fuel Cell
FCEV	Fuel Cell Electric Vehicle
FCS	Fuel Cell Stack
FIC	Full-bridge Isolated Converter
FLC	Fuzzy Logic Controller
FPGA	Field Programmable Gate Array
GaN	Gallium Nitride
HC	Hill Climbing
HF	High Frequency
HIC	Half-bridge Isolated Converter
HIL	Hardware In the Loop
HPS	Hybrid Power Source
IBC	Interleaved Boost Converters
IC-IBC	Inverse Coupled-Interleaved Boost Converters
LF	Low Frequency
MFC	Model Free Control
MIC	Modular-Integrated Converter
MIL	Model In the Loop
MIMO	Multi-Input Multi-Output
MPPT	Maximum Power Point Tracking
PCS	Power Conditioning Stage
PEM	Proton Exchange Membrane or Polymer Electrolyte Membrane
PI	Proportional Integral
PIL	Processor In the Loop
PLL	Phase Lock Loop
PV	Photovoltaic
PWM	Pulse Width Modulation
RIC	Resonance Isolated Converter
SA	Simulated Annealing
SC	Super-capacitor
SiC	Silicon Carbide
SIL	Software In the Loop
SOC	State of Charge
SMC	Sliding Mode Control
SPWM	Sinusoidal Pulsed Width Modulation
TEG	Thermoelectric Generator
THD	Total Harmonic Distortion
UC	Ultra-capacitor
UC-IBC	UnCoupled-Interleaved Boost Converters
VSC	Voltage Source Converter
WADALINE	Wavelet Adaptive Linear Neuron
ZCS	Zero Current Switching
ZSI	Z-Source Inverter
ZVS	Zero Voltage Switching

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors thank CPUT and HySA UWC for the funding.

References

[1] N.P. Bayendang, M.T. Kahn, V. Balyan, I. Draganov, "CCHP Systems Analysis with Emphasis on Fuel Cells, Thermoelectricity and Power

Converters," 2020 5th International Conference on Smart and Sustainable Technologies, SpliTech 2020, 2020, doi:10.23919/SpliTech49282.2020.9243720.

[2] S. Vigneshwaran, R. Vijayalakshmi, "High efficiency DC/DC buck-boost converters for high power DC system using adaptive control," American-Eurasian Journal of Scientific Research, **11**(5), 381-389, 2016. doi: 10.5829/idosi.aejsr.2016.11.5.22957.

[3] A. Kirubakaran, J. Shailendra, R.K. Nema, "The PEM fuel cell system with DC/DC boost converter: Design, modeling and simulation," International Journal of Recent Trends in Engineering, **1**(3), 2009, doi: 01.ijepe.01.01.05.

[4] L.M.P. Fanjul, Design Considerations for DC-DC Converters in Fuel-cell Systems, PhD Thesis, Texas A&M University, 2006.

[5] M. Grotsch, M. Mangold, A. Kienle, "Analysis of the coupling behaviour of PEM fuel cells and DC-DC converters", Energies, **2**, 71-96, 2009, doi: 10.3390/en2010071.

[6] M.T. Outeiro, A. Carvalho, Methodology of Designing Power Converters for Fuel-cell Based Systems: A Resonant Approach, New Developments in Renewable Energy, 2013, http://dx.doi.org/10.5772/54674.

[7] K. Suh and A.G. Stefanopoulou, "Coordination of Converter and Fuel Cell Controllers," in Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, Limassol, Cyprus, 2005, doi: 10.1109/.2005.1467076.

[8] M. Nymand, M.A.E. Andersen, "A New Approach to High Efficiency in Isolated Boost Converters for High-Power Low-Voltage Fuel Cell Applications," in 2008 13th Power Electronics and Motion Control Conference, Poznan, Poland, 2008, 127-131, https://doi.org/10.1109/EPEPEMC.2008.4635255.

[9] R. Seyezhai, R. Anitha, S. Mahalakshmi, M. Bhavani, "High gain interleaved boost converter for fuel cell applications," Bulletin of Electrical Engineering and Informatics, **2**(4), 2013, 265-271, doi: https://doi.org/10.11591/eei.v2i4.192.

[10] Y. Huangfu, R. Ma, B. Liang, Y. Li. "High power efficiency buck converter design for standalone wind generation system," International Journal of Antennas and Propagation, Volume 2015, Special Issue: Recent Advances in Theory and Applications of Electromagnetic Metamaterials, http://dx.doi.org/10.1155/2015/751830.

[11] F.M. Mwaniki, High voltage boost DC-DC converter suitable for variable voltage sources and high power photovoltaic application, MEng Dissertation, University of Pretoria, 2014, http://hdl.handle.net/2263/37320.

[12] B. Eckardt, M. März, A. Hofmann, M. Gräf and J. Ungethüm, "High Power Buck-Boost DC/DC Converter for Automotive Power-train Application," in 2005 Official Proceedings of the International Conference of Power Electronics, Intelligent Motion, Power Quality: PCIM 2005 Europe, Nürnberg, Germany, 2005, 685, http://publica.fraunhofer.de/documents/N-46644.html.

[13] M. Kabalo, B. Blunier, D. Bouquain, A. Miraoui, "State-of-the-Art of DC-DC Converters for Fuel Cell Vehicles," in 2010 IEEE Vehicle Power and Propulsion Conference, Lille, France, 2010, doi: 10.1109/VPPC.2010.5729051.

[14] M. Delshad, H. Farzanehfard, "A new soft switched push pull current fed converter for fuel cell applications," Energy Conversion and Management, **52**(2), 917–923, 2011, https://doi.org/10.1016/j.enconman.2010.08.019.

[15] N. Bizon, "A new topology of fuel cell hybrid power source for efficient operation and high reliability," Journal of Power Sources, **196**(6), 3260–3270, 2011, doi: 10.1016/j.jpowsour.2010.11.049.

[16] O.A. Ahmed, J.A.M. Bleijs, "Power flow control methods for an ultra-capacitor bidirectional converter in DC micro-grids – a comparative Study," Renewable and Sustainable Energy Reviews, **26**, 727–738, 2013, https://doi.org/10.1016/j.rser.2013.06.021.

[17] A. Kolli, A. Gaillard, A. De Bernardinis, O. Bethoux, D. Hissel, Z. Khatir, "Design and Control of a High Frequency Six-phase Interleaved DC/DC Converter based on SiC Power MOSFETs for a PEMFC," in 2015 6th International Conference on Fundamentals and Development of Fuel Cells (FDFC 2015), Toulouse, France, 2015.

[18] A. Kolli, A. De-Bernardinis, Z. Khatir, A. Gaillard, O. Béthoux and D. Hissel, "Part-load Control Strategy of a 20kW SiC Power Converter for Embedded PEMFC Multi-stack Architectures," IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society, Yokohama, Japan, 2015, 4627-4632, doi: 10.1109/IECON.2015.7392821.

[19] A. Carvalho, M.T. Outeiro, MatLab in Model-Based Design for Power Electronics Systems, Engineering Education and Research Using MatLab, 2011, doi: 10.5772/21144.

[20] H. Wang, Design and Control of a 6-phase Interleaved Boost Converter based on SiC semiconductors with EIS Functionality for Fuel Cell Electric Vehicle. PhD Thesis, UBFC and UTBM, Belfort, France, 2019.

[21] J.W. Pratt, L.E. Klebanoff, K. Munoz-Ramos, A.A. Akhil, D.B. Curgus, B.L. Schenkman, Proton Exchange Membrane Fuel Cells for Electrical Power Generation On-Board Commercial Airplanes, SAND2011-3119 Unlimited Release Printed, 2011.

[22] J.L. Casteleiro-Roca, A.J. Barragán, F. Segura, J.L. Calvo-Rolle, J.M. Andújar, "Fuel cell output current Prediction with a hybrid intelligent system," Complexity, Volume 2019, Special Issue: Advanced Controls in Complex Clean Energy Devices, Subsystems, and Processes, https://doi.org/10.1155/2019/6317270.

[23] A. Kolli, A. Gaillard, A. De Bernardinis, O. Bethoux, D. Hissel, Z. Khatir, "A review on DC/DC converter architectures for power fuel cell applications," Energy Conversion and Management, **105**, 716-730, 2015, https://doi.org/10.1016/j.enconman.2015.07.060.

[24] D. Ravi, S.S. Letha, P. Samuel, B.M. Reddy, "An Overview of Various DC-DC Converter Techniques used for Fuel Cell based Applications," 2018 International Conference on Power Energy, Environment and Intelligent Control (PEEIC), Greater Noida, India, 2018, 16-21, doi: 10.1109/PEEIC.2018.8665465.

[25] J. Gao, M. Li, Y. Hu, H. Chen, Y. Ma, "Challenges and developments of automotive fuel cell hybrid power system and control," Science China Information Sciences, **62**(5), 2019, doi: https://doi.org/10.1007/s11432-018-9690-y.

[26] H. Liu, L. Xu, Z. Hu, J. Li, H. Jiang, M. Ouyang, "Experimental Study and Performance Analysis on High Power Fuel Cell System," IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 2020, 2031-2035, doi: 10.1109/IECON43393.2020.9254802.

[27] R. Miyazaki, T. Mishima, C.M. Lai, "Coupled Inductor-assisted Current-Fed Snubber-less Zero-Current-Soft Switching High Step-Up DC-DC Converter for Fuel Cell Power Interface," 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), Kobe, Japan, 2020, 797-801, doi: 10.1109/GCCE50665.2020.9291776.

[28] M.S. Bhaskar, V.K. Ramachandaramurthy, S. Padmanaban, F. Blaabjerg, D.M. Ionel, M. Mitolo, D. Almakhles, "Survey of DC-DC Non-Isolated Topologies for Unidirectional Power Flow in Fuel Cell Vehicles," in IEEE Access, 8, 178130-178166, 2020, doi: 10.1109/ACCESS.2020.3027041.

[29] K.S. Rathode, S.K. Sharma, S. Shringi, "Performance Analysis of PV & Fuel cell based Grid Integrated Power System," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, 970-975, doi: 10.1109/ICCES45898.2019.9002090.

[30] S. Kavyapriya, R.K. Kumar, "Modeling and Simulation of DC-DC Converters for Fuel Cell System," International Journal of Engineering and Advanced Technology (IJEAT), **9**(3), 2020, doi: 10.35940/ijeat.C5754.029320.

[31] P. Bazin, B. Beranger, J. Ecrabey, L. Garnier, S. Mercier, "Smart fuel cell module (6.5 kW) for a range extender application," 2020 22nd European Conference on Power Electronics and Applications (EPE'20 ECCE Europe), Lyon, France, 2020, 1-8, doi: 10.23919/EPE20ECCEEurope43536.2020.9215856.

[32] A. Gonnet, S.G. Jorge, C. Busada, J. Solsona, "Power converter topology for conditioning a fuel cells battery voltage," 2019 Argentine Conference on Electronics (CAE), Mar del Plata, Argentina, 2019, 9-14, doi: 10.1109/CAE.2019.8709157.

[33] J. Corcau, L. Dinca, I. Adochiei, T. L. Grigorie, "Modeling and Simulation of an Aerodrome Electrical Power Source Based on Fuel Cells," 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 2019, 1-4, doi: 10.1109/EHB47216.2019.8969870.

[34] A. Abdelhakim, F. Blaabjerg, "Current-fed Modular Multilevel Converter (CMMC) for Fuel Cell and Photovoltaic Integration," 2020 IEEE 21st Workshop on Control and Modeling for Power Electronics (COMPEL), Aalborg, Denmark, 2020, 1-6, doi: 10.1109/COMPEL49091.2020.9265695.

[35] P. Prabhakaran, V. Agarwal, "Novel Four-Port DC–DC Converter for Interfacing Solar PV–Fuel Cell Hybrid Sources With Low-Voltage Bipolar DC Microgrids," in IEEE Journal of Emerging and Selected Topics in Power Electronics, **8**(2), 1330-1340, 2020, doi: 10.1109/JESTPE.2018.2885613.

[36] H.S. Youn, D.H. Yun, W.S. Lee, I.O. Lee, "Study on Boost Converters with High Power-Density for Hydrogen-Fuel-Cell Hybrid Railway System", Electronics, **9**(5), 771, 2020, doi: https://doi.org/10.3390/electronics9050771.

[37] S. Motahhir, E.G. Abdelaziz, A. Derouich, "MIL and SIL and PIL tests for MPPT algorithm," Cogent Engineering, **4**(1), 2017, https://doi.org/10.1080/23311916.2017.1378475.

[38] N. Sulaiman, M.A. Hannan, A. Mohamed, E.H. Majlan, W.R. Wan Daud, "Review on energy management system for fuel cell hybrid electric vehicle: Issues and challenges," Renewable and Sustainable Energy Reviews, **52**, 802–814, 2015, https://doi.org/10.1016/j.rser.2015.07.132.

[39] D. Xiao, Q. Wang, "The Research of Energy Management Strategy for Fuel Cell Hybrid Vehicle," in 2012 International Conference on Industrial Control and Electronics Engineering, Xi'an, China, 2012, 931–934, doi: 10.1109/ICICEE.2012.247.

[40] N. Ouddah, M. Boukhnifer, A. Raisemche, "Two Control Energy Management Schemes for Electrical Hybrid Vehicle," in 2013 Proceedings of the 10th International Multi-conference on Systems, Signals & Devices (SSD13), Hammamet, Tunisia, doi: 10.1109/SSD.2013.6564135.

[41] O. Sundstrom, A. Stefanopoulou, "Optimal Power Split in Fuel Cell Hybrid Electric Vehicle with Different Battery Sizes, Drive Cycles and Objectives," in 2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, Munich, Germany, 1681–1688, doi: 10.1109/CACSD-CCA-ISIC.2006.4776894.

[42] C.J. Xie, S.H. Quan, Q.H. Chen, "Multiple Model Control for Hybrid Power System of Fuel Cell Electric Vehicle," in 2008 IEEE Vehicle Power and Propulsion Conference, Harbin, China, 2008, doi: 10.1109/VPPC.2008.4677527.

[43] P. Thounthong, S. Pierfedereci, J.P. Martin, M. Hinaje, B. Davat, "Modelling and control of fuel-cell super-capacitor hybrid source based on differential flatness control," IEEE Transaction on Vehicle Technology, **59** (6), 2700–2710, 2010, doi: 10.1109/TVT.2010.2046759.

[44] A. Tani, M.B. Camara, B. Dakyo, Y. Azzouz, "DC/DC and DC/ACconverters control for hybrid electric vehicles energy management ultra-capacitors and fuel cell," IEEE Transactions on Industrial Informatics, 9(2), 686–696, 2013, doi: 10.1109/TII.2012.2225632.

[45] M. Uzunoglu, M.S. Alam, "Modelling and analysis of an FC/UC Hybrid vehicular power system using a novel-wavelet-based load sharing algorithm," IEEE Transaction on Energy Conversion, **23**(1), 263–272, 2008, doi: 10.1109/TEC.2007.908366.

[45] D.S. Efstathiou, A.K. Petrou, P. Spanoudakis, N.C. Tsourveloudis, K.P. Valavanis, "Recent Advances on the Energy Management of a Hybrid Electric Vehicle," in 2012 20th Mediterranean Conference on Control & Automation (MED), Barcelona, Spain, 2012, 896–901, doi: 10.1109/MED.2012.6265752.

[47] Y. Ates, M. Uzunoglu, O. Erdinc, B. Vural, "A Wavelet-ADALINE Network based Load Sharing and Control Algorithm for a FC/UC Hybrid Vehicular Power System," in Proceedings of the 2009 International Conference on Clean Electrical Power, Capri, Italy, 2009, 591–94, doi: 10.1109/ICCEP.2009.5211976.

[48] C.H. Zheng, C.M. Lee, Y.C. Huang, W.S. Lin, "Adaptive Optimal Control Algorithm for Maturing Energy Management Strategy in Fuel–cell/Li-ion/Super-Capacitor Hybrid Electric Vehicles," in 2013 9th Asian control conference, Istanbul, Italy, 2013, 1–7, doi: 10.1109/ASCC.2013.6606091.

[49] M.A. Hannan, F.A. Azidin, A. Mohamed, "Multi-sources model and control algorithm of an energy management system for light electric vehicles," Energy Converters Management, 62, 123–30, 2012, https://doi.org/10.1016/j.enconman.2012.04.001.

[50] J.S. Martinez, D. Hissel, M.C. Pera, M. Amiet, "Practical control structure and energy management of a test bed hybrid electric vehicle." IEEE Transactions on Vehicular Technology, **60**(9), 4139–4152, 2011, doi: 10.1109/TVT.2011.2169821.

[51] E. Schaltz, P.O. Rasmussen, "Design and Comparison of Power Systems for a Fuel Cell Hybrid Electric Vehicle," in IEEE Industry Applications Society Annual Meeting," Edmonton, AB, Canada, 2008, 1–8, doi: 10.1109/08IAS.2008.184.

[52] V. Paladini, T. Donateo, A. De Risi, D. Laforgia, "Super-capacitors fuel-cell hybrid electric vehicle optimization & control strategy development," Energy Conversion and Management, **48**(11), 3001-3008, 2007, https://doi.org/10.1016/j.enconman.2007.07.014.

[53] M. Zandi, A. Payman, J.P. Martin, S. Pierfederici, B. Davat, F. Meibody-Tabar, "Energy management of a fuel cell/super-capacitor/battery power source for electric vehicular applications," IEEE Transactions on Vehicular Technology, **60**(2), 433–443, 2011, doi: 10.1109/TVT.2010.2091433.

[54] H. Fathabadi, "Novel fuel cell/battery/super-capacitor hybrid power source for fuel-cell hybrid electric vehicles," Energy, 143, 467 – 477, 2018, https://doi.org/10.1016/j.energy.2017.10.107.

[55] Q. Li, H. Yang, Y. Han, M. Li, W. Chen, "A state machine strategy based on droop control for an energy management system of PEMFC-battery, super-capacitor hybrid tramway," International Journal of Hydrogen Energy, **41**(36), 16148 – 16159, 2016, https://doi.org/10.1016/j.ijhydene.2016.04.254.

[56] O. Kraa, H. Ghodbane, R. Saadi, M.Y. Ayad, M. Becherif, A. Aboubou, M. Bahri, "Energy Management of Fuel Cell/ Super-capacitor Hybrid Source Based on Linear and Sliding Mode Control," in the International Conference on Technologies and Materials for Renewable Energy, Environment and Sustainability, TMREES15, Energy Procedia, 2015, 74, 1258 – 1264, https://doi.org/10.1016/j.egypro.2015.07.770.

[57] S.N. Motapon, L. Dessaint, K. Al-Haddad, "A comparative study of energy management schemes for a fuel-cell hybrid emergency power system of More-Electric Aircraft," IEEE Transactions on Industrial Electronics, **61**(3),1320-1334, 2014, doi: 10.1109/TIE.2013.2257152.

[58] P. Mungporn, B. Yodwong, P. Thounthong, B. Nahid-Mobarakeh, N. Takorabet, D. Guilbert, N. Bizon, P. Kumam, C. Kaewprapha, "Model-Free Control of Multiphase Interleaved Boost Converter for Fuel Cell/Reformer Power Generation," 2019 Research, Invention, and Innovation Congress (RI2C), Bangkok, Thailand, 2019, 1-6, doi: 10.1109/RI2C48728.2019.8999919.

[59] G.G. Suárez-Velázquez, G.E. Mejia-Ruiz, P.M. García-Vite, "Control and Grid Connection of Fuel Cell Power System," 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, Mexico, 2020, 1-5, doi: 10.1109/ROPEC50909.2020.9258729.

[60] Y. Jeong, R.A.L. Rorrer, B. Lee, J. Park, "A Novel Control Scheme for High Efficiency Fuel Cell Power Systems in Parallel Structure," 2019 IEEE Energy Conversion Congress and Exposition (ECCE), Baltimore, MD, USA, 2019, 940-946, doi: 10.1109/ECCE.2019.8912856.

[61] Y. Zhang, Y. Huangfu, W. Liu, L. Guo, "An energy management strategy based on state machine with power compensation for photovoltaic-PEMFC-lithium battery power system," 2019 IEEE International Conference on Industrial Technology (ICIT), Melbourne, VIC, Australia, 2019, 1675-1680, doi: 10.1109/ICIT.2019.8843696.

[62] M. Jafari, Z. Malekjamshidi, D.D. Lu, J. Zhu, "Development of a Fuzzy-Logic-Based Energy Management System for a Multiport Multioperation Mode Residential Smart Microgrid," in IEEE Transactions on Power Electronics, 34(4), 3283-3301, 2019, doi: 10.1109/TPEL.2018.2850852.

[63] M.N. Boukoberine, Z. Zhou, M. Benbouzid, T. Donateo, "Frequency Separation-based Power Management Strategy for a Fuel Cell-Powered Drone," 2020 2nd International Conference on Smart Power & Internet Energy Systems (SPIES), Bangkok, Thailand, 2020, 209-214, doi: 10.1109/SPIES48661.2020.9242975.

[64] E.M. Barhoumi, P. Chukwuleke Okonkwo, I. B. Belgacem, M. Zghaibeh, "MPPT Control of an Interleaved Boost Converter for a Polymer Electrolyte Membrane Fuel Cell Applications," 2020 International Conference on Electrical and Information Technologies (ICEIT), Rabat, Morocco, 2020, 1-5, doi: 10.1109/ICEIT48248.2020.9113228.

[65] J.S. Artal-Sevil, J.A. Domínguez-Navarro, C. Bernal-Ruiz, A. Coronado-Mendoza, "Power Flow Control through Differential Power Processing to improve reliability in hybrid systems based on PEM-Fuel Cell," 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), Monte-Carlo, Monaco, 2020, 1-12, doi: 10.1109/EVER48776.2020.9243125.

[66] M. Arazi, A. Payman, M. B. Camara, B. Dakyo, "Energy management in a Multi-source System using isolated DC-DC resonant converters," 2020 22nd European Conference on Power Electronics and Applications (EPE'20 ECCE Europe), Lyon, France, 2020, 1-7, doi: 10.23919/EPE20ECCEEurope43536.2020.9215662.

[67] A. Martin-Lozano, A. Barrado, A. Rodriguez-Lorente, A. Lázaro, C. Fernández, "Energy Management System Optimization for a Fuel Cell Hybrid Vehicle based on Power Losses Minimization," 2020 IEEE 14th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), Setubal, Portugal, 2020, 402-408, doi: 10.1109/CPE-POWERENG48600.2020.9161482.

[68] M. Salimi, F. Radmand, M. H. Firouz, "Dynamic Modeling and Closed-loop Control of Hybrid Grid-connected Renewable Energy System with Multi-input Multi-output Controller," in Journal of Modern Power Systems and Clean Energy, **9**(1), 94-103, 2021, doi: 10.35833/MPCE.2018.000353.

[69] I.S. Sorlei, N. Bizon, P. Thounthong, M. Varlam, E. Carcadea, M. Culcer, M. Iliescu, M. Raceanu, "Fuel Cell Electric Vehicles — A Brief Review of Current Topologies and Energy Management Strategies", Energies, **14**(1), 252, 2021, doi: https://doi.org/10.3390/en14010252.

[70] M. İnci, M. Büyük, M.H. Demir, G. İlbey, "A review and research on fuel cell electric vehicles: Topologies, power electronic converters, energy management methods, technical challenges, marketing and future aspects," Renewable and Sustainable Energy Reviews, **137**, 2021, doi: https://doi.org/10.1016/j.rser.2020.110648.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 58-64 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Methodology for Calculating Shock Loads on the Human Foot

Valentyn Tsapenko[*,1], Mykola Tereschenko[1], Vadim Shevchenko[1], Ruslan Ivanenko[2]

[1]*Faculty of Instrumentation Engineering, National Technical University of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, 03056, Ukraine*

[2]*The Ukrainian Scientific and Forensic Expertise of the Security Service of Ukraine, Kyiv, 03113, Ukraine*

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 14 February, 2021
Online: 10 March, 2021

Keywords:
Foot
Musculoskeletal System
Step Cycle
Locomotion
Biomechanical Parameters
Support
Elastic Characteristics
Coefficients of Capacity
Shock Load

ABSTRACT

The leading place among diseases of the musculoskeletal system is occupied by various feet deformations. Clinical movement analysis and posturological examination are required to objectively assess the distribution for load caused by the weight of human body on the feet and its locomotion effect. In normal conditions, the foot is exposed to elastic deformations. When analyzing the foot loads, it`s necessary to consider shock loads as one of dynamic load types. The foot is the first to perceive the shock impulse by support reaction, and the further nature for interaction with the environment directly depends on its functional capabilities. However, the foot supporting properties haven`t been fully researched. The purpose for this research is to increase the accuracy of estimating the human foot biomechanical parameters, by assessing the dynamic impact, namely short-term shock loads by step cycle relevant phases. This goal is solved by developing a method of static-dynamic load analysis, which allows to estimate dynamic and shock loads on foot and is reduced to determining the capacity coefficients, dynamic and shock loads. In the course of studies, conducted in this research, it was found that the maximum contact per unit time has front section (repulsion phase), then - the rear section (landing phase) and the smallest - the foot middle section (rolling phase), the greater speed and length step – so the greater shock loads coefficient, and their peak falls on the front and rear sections. The practical significance of the obtained results is to improve the existing methods of researching biomechanical parameters by comprehensively assessing by standing and gait features, foot step cycle and support properties.

1. Introduction

The current stage of development for instrument making actualizes the applied aspect for human locomotions analysis. Diagnosis of posture wouldn`t be complete without measuring and assessing the state of support-spring foot properties. Clinical analysis for movement and posturological examination (body position examination) are required to objectively assess the effect by foot load distribution on locomotion. In modern biomedical engineering, a fairly promising area is clinical analysis of motor activity - the research of various pathologies of gait and main rack, using biomechanics methods. There is no doubt about the importance of assessing the functional state by musculoskeletal system in patients with orthopedic and neurological profile [1].

In recent years, the number of diseases, injuries and pathologies of the musculoskeletal system (MSS) is growing steadily, which significantly affects the quality of life. Analysis of human gait is subject by many scientific studies. Various diseases and injuries of supporting organs are often accompanied by serious functional disorders of musculoskeletal system, decreased muscle strength and tone, loss of ability to normal movements, which ultimately leads to disability and subsequent disability [2, 3]. The foot is structural segment in musculoskeletal system, which provides it`s stomato-motor function, and is integral morphofunctional object that human motor function depends [2]. By exteroreceptors located on plantar surface, and it collects information about body mass fluctuations and directs it to the central nervous system (CNS), which will coordinate postural stability [4].

[*]Corresponding Author: Valentyn Tsapenko, 37, Prosp.Peremohy, Kyiv, Ukraine, 03056, capenko.valik@ukr.net

The foot - is the first, most loaded part of musculoskeletal system, which makes contact with the support, redistributes the reaction force to the higher segments for musculoskeletal system and performs an important spring function, provides stability of the lower limb and adhesion to support surface [5]. There are three main functions of foot:

- spring - the ability to elastic stratification under the action of load and restore its shape after removal of the latter;
- balancing - participation in the regulation of posture and positional activity when standing and walking;
- push (repulsive) - the transfer of acceleration of the total center of mass (TCM) for all body during locomotion [3].

During life, the functional parameters of foot change to one degree or another. First of all, they relate to its spring and repulsive functions. In the case by violations of symmetry for load distribution on foot, as well as by reducing its elastic characteristics - reduced damping properties, which leads to increased shock loads and increased vibration effects on the entire MSS [6]. Currently, longitudinal and transverse flattening of the feet, which can be both independent and in combination with other deformities, is one of the most common orthopedic diseases. According to the WHO, 75% of people have some pathological changes in the feet, which the most common is flat feet (changes in the shape of foot, that is characterized by a decrease in the height of its transverse and longitudinal arches). This deformity is the cause of many serious MSS diseases, which often lead to disability [2, 3].

Analysis of relevant literature sources [6-12] shows the current state of injuries biomechanics and foot deformities. In particular, a significant number of scientific publications aimed at researching the condition by lower extremities of different groups for population and identifying the relationship between foot deformity and other structural disorders of MSS were considered. Researchers note that the problem of early diagnosis of injuries and diseases for feet is relevant in the choice of prevention methods, treatment, orthosis and their effectiveness evaluation. Given that foot bears the main load, the violation of its functions is reflected in entire musculoskeletal system functioning and can lead, in the future, to a number of chronic diseases. The research of human foot in the dynamics (with gradual or sharp variable load on it) is a key factor in the diagnosis for its functional state, which allows you to identify abnormalities and determine the necessary set of treatment measures [6-12]. There are methods for determining the biomechanical properties for foot soft tissues, based on the assessment changes in the pressure values applied from the outside, sufficient to block the arterial vascular bed of tissues, determined by discoloration for skin surface or disappearance of arterioles in blood vessels. However, these methods have a number of limitations, in particular, can`t be used in the assessment of gait. During walking, there are additional short-term shock loads of the anterior and posterior shock, in the damping of which may involve various biological media, in particular, in addition to musculoskeletal, fluid blood-lymphatic media play a significant role [6]. Also in scientific publications described the types and variants of gait at various deformations, scoliosis and number of skeleton congenital deformations. The authors of some publications point to the imbalance and lack of synchrony in the muscles for lower leg and foot, changes in TCM, instability when walking and rapidly progressing pain in various pathologies [11, 12]. However, among the variety of scientific approaches to addressing this issue, the research for distribution the foot load during the full step cycle in dynamics, not fully studied. Known methods don`t allow to diagnose functional changes in the foot that occur when the load changes, as well as to determine the individual physiological characteristics of lower extremities, which limits the use of such methods. In addition, in domestic and foreign literature not enough attention is paid to quantitative assessment for parameters of the foot elasticity, while the latter analysis will improve the MSS accuracy diagnosis.

At overloads of arch supporting systems - functions of foot are broken, the motor stereotype as a whole is distorted, there are undesirable redistributions of forces and overloads which are transferred to other parts of MSS therefore there are pathologies. In such cases, the foot works not as an elastic-elastic system, but as an elastic-plastic system, with its inherent residual deformation. The most common cause of the latter is overload associated with functional impairment of foot arches, i.e. recompensation of arches is expressed in a particular injury [6]. Violation for arches formation to the feet occupies 26.4% of all orthopedic pathology and up to 81.5% for all deformities of lower extremities in children [12]. Researches show that in children`s these disorders can be partially or completely corrected, while in adulthood - these disorders are extremely difficult to correct, sometimes impossible. It has also been proven that the absence of physiological arches and disturbances for ankle axis joint leads to development of pathological processes in the large joints of lower extremities and spine, and is often the cause of pain [5, 6]. With foot flat-valgus deformity, in addition to muscle weakness and ligaments, the shape and foot bones ratio are disturbed, the reference vector is displaced laterally (to foot outer edge). This changes the nature of gait, is the cause of increasing the dynamic load on the entire musculoskeletal system [2].

Based on this, various foot deformities, in particular flat-valgus, should be considered as primary link in unstable gait formation, which affects the violation of MSS, and leads to various lesions of the latter. The vast majority of modern tool systems are physical rather than linear, which complicates the search for solutions on an elementary basis and sometimes makes it impossible. It`s in this statement that modern instrumental analysis develops. For linear systems, the minimum change in the initial state - causes the corresponding changes by its final state. Otherwise, for nonlinear systems, small differences in the initial conditions lead to unpredictable changes in final state. For example, a person undergoing static analysis doesn`t know how much and how to change their motor behavior during dynamic analysis [4]. In this regard, morphofunctional diagnosis of the foot is an essential element in the prevention of MSS number disorders, and the applied aspect for analysis of human locomotion, in modern biomedical engineering, is very relevant and promising.

2. Materials and methods

2.1. Formulation of the problem

Clinical analysis of movement and posturological examination are required to objectively assess the distribution of load caused by

the weight of human body on feet and its effect on locomotion. There is baropodometry method, which allows you to objectify the research of biomechanical parameters of foot, taking into account the static and dynamic components. This allows for more in-depth research of locomotions and to detect a certain frequency of identical or pressure similar images, which don`t fully correspond to the known metatarsal formula [1-5]. Thus, the research for biomechanics of the lower extremities in norm and at various deformations is necessary and very promising in terms of assessing the state by MSS and forecasting its dynamics.

In order to increase the accuracy for research of biomechanical parameters of foot, it`s necessary to take into account the influence of dynamic, namely short-term shock loads caused by body weight on the corresponding phases by the step cycle.

2.2. The foot impact load calculating method

The article task - to identify the initial stages of injuries and foot deformations, is solved by developing a method of calculating the dynamic load taking into account the short-term foot impact.

The foot ability to withstand various loads is due not only to biomechanical perfection, but also the properties of its constituent tissues [2]. Normally, due to vaulted structure of the foot and its spring function, up to 70% of the acceleration is damped and amortized. In normal development of the musculoskeletal system, the load is distributed as follows: through the body of the heel talus, navicular and cuneiform bone, then on the heads of I-III metatarsals, forming an external longitudinal arc (As shown in Figure 1) [3]. The foot biomechanics and its functions in different phases of the step cycle are different. Mitigation of the inertial load during walking and running is carried out by articular ligament difficult complex, which connects the 26 main foot bones, where there are 5 longitudinal bones and the transverse arch. The heel, talus and metatarsal and metatarsal bones form a kind of arch - a spring that can shrink and straighten under the loads action. Body weight is normally evenly distributed on the front and back of the foot. These divisions are connected in a single kinematic chain by inter-articular ligaments, as well as a strong elastic tendon - plantar aponeurosis, which, like a spring, returns to normal position spread out under the load of the foot arch [1-5].

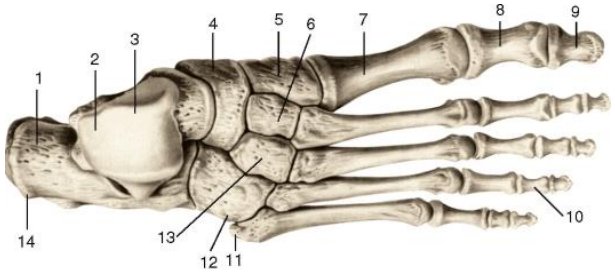


Figure 1: The human foot structure [3]

1 - heel bone; 2 - talus block; 3 - talus; 4 - navicular bone; 5 - medial wedge-shaped bone; 6 - intermediate wedge-shaped bone; 7 - I metatarsal bone; 8 - proximal phalanx; 9 - distal (nail) phalanx; 10 - middle phalanx; 11 - hump V of the metatarsal bone; 12 - cuboid bone; 13 - lateral wedge-shaped bone; 14 – humerus

The gait main functional unit - cycle of the step, that time from the beginning of limb contact with the support to the next same contact with the same limb. The normal locomotor cycle consists of two bipolar and two portable phases. The average step cycle time, at normal speed, is close to 1*s*. The step cycle, for each limb, consists of three main periods: the support period, the transfer period and the double support period. Research in the field of biomechanics has shown that the forces applied during movement have vertical and horizontal components. Gait is a complex cyclic movement associated with the repulsion of the body from the support surface and its movement in space. The efforts made are dynamic. Characteristic of normal gait is the constant storage of support on one or two limbs [2].

The basis for any locomotor act is the support interactions, i.e. short-term contact of a MSS certain link with the support, as a result of which there are forces capable of changing the motion center of the body's TCM. Support interactions have all the physical features for shock loads (short duration, significant increase in the modulus of force, etc.), so such interactions can be considered as shock [6].

The algorithm for calculating shock loads is:

- Step cycle registration;
- Determination of biomechanical parameters for step cycle - load values, speed, contact areas of support surfaces, half-step length, etc.;
- Determination for dynamic capacity coefficient of the lower extremities;
- Determination of the dynamic load factor;
- Determination for impact load factor.

2.3. Determination of step cycle biomechanical parameters

During the movement in each step cycle, there are two phases of double and single support. The moment of time t_{fi} of the maximum vertical pressure $P_{v.max}$ s called the forward shock. The time t_{bi} of the rear shock always coincides with the maximum $P_{l.max}$ of longitudinal forces that move the body forward, at this time t_{bi}, the maxima $(P_{v.max} + P_{l.max})$ of the vertical and horizontal forces are summed. Between the moments of time $\Delta t=(t_{bi} - t_{fi})$ of the front and rear shock of the lower extremity there is a damping failure $\Delta P_f=P_{v.min}$, which corresponds for minimum vertical pressure. When the overall center of gravity rises higher, the pressure on the support is directed forward, replaced by the pressure directed backwards, i.e. the braking is replaced by a push. In healthy people, the step reference period is $t_{rp} = 0{,}64\pm0{,}016$ *s*, the transfer period $t_{tp} = 0{,}36\pm0{,}014$ *s*, the two-support period $t_{trp} = 0{,}135\pm0{,}010$ *s*. The duration of support time on the right and left foot usually differs by ±5% and is [2]:

$$\Delta t_{tp} = (24—29)\% \cdot t_k \quad (1)$$

where t_k - step time.

Given the different pace, these rates are estimated as a percentage for entire step cycle: the period of rolling through the foot is equal to [2]:

$$\Delta t_r = (45—51)\% \cdot t_k \quad (2)$$

The period for the lower limb over the support transfer [2]:

$$\Delta t_{tp} = (31—41)\% \cdot t_k \quad (3)$$

Bipolar step period [2]:

$$\Delta t_{trp} = (8—24)\% \cdot t_k \quad (4)$$

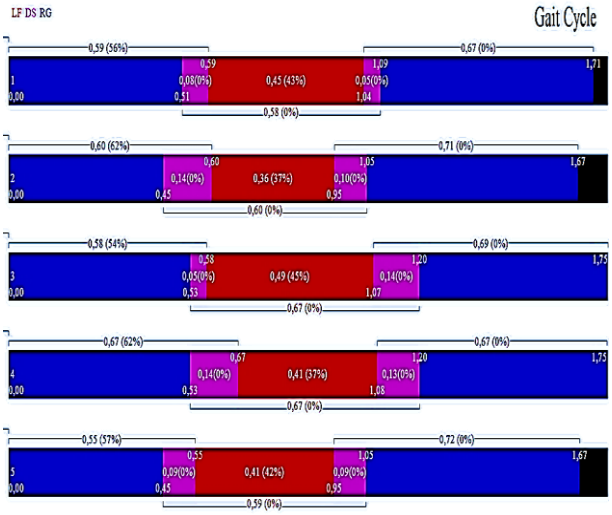


Figure 2: The human foot structure [2]

In fig. 2 [2] presents step cycle graphical representation. Red indicates the period of support on left lower limb, blue - on the right, pink indicates the period of double support (two limbs), black characterizes the contact lack. After analyzing the step cycle, we can conclude that the period of transfer for one limb is equal to single support period on the opposite limb [2].

In order to identify the leading limb, we introduce the dynamic capacity coefficient [2]:

$$K_{dS} = P \cdot t \quad (5)$$

where P – foot pressure on the support surface with a single support (kg/m^2),

t – the duration of the support period (s).

In previous researches [2] was determined the dynamic coefficient q, which is calculated as ratio of force transmitted to the resistance during locomotion to body weight m:

$$q = \frac{P}{m} \quad (6)$$

The following values for dynamic coefficient are calculated and determined q:

- for slow walking $q \leq 1$
- for fast walking $q \leq 1,5$
- for running $q \leq 1,8$

In the first iteration, the dynamic coefficient q will be equal to one (the beginning of the movement), so the pressure for foot P will be proportional to the body weight m. Given the latter - the maximum value of the coefficient for dynamic capacity will be the longest period of single support. In order to identify the leading limb, we write the ratio of the coefficients K_{ds} for the left and right lower extremities, and obtain [2]:

$$\frac{K_{dSL}}{K_{dSR}} = \frac{t_L}{t_R} \quad (7)$$

where $K_{dSl/r}$ – dynamic capacity coefficients for the left and right lower extremities, respectively.

Twenty-five children aged 3 to 17, with an average body weight of 45 kg, were selected as the subjects for this research. The research tool base was a baropodometric platform, measuring 0,4×1,8×0,02 m, (As shown in Figure 3) with appropriate software.

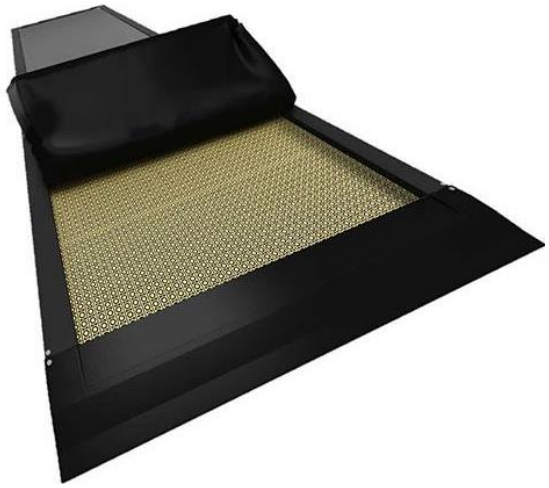


Figure 3: Digital baropodometric platform [3]

The main technical characteristics of the system are given in the table 1 [3].

Table 1: Biometric system technical parameters

Parameters	
Measurement error	± 5%
Scan type	Matrix
Resolution, dpi	x, y=9,6 z=16
Pressure on a point, kg/m^2	max $150 \cdot 10^4$
Measurement frequency, staff/s	60
Sensor type	High-resistance sensor with active matrix
Number of sensors on 1 m^2	400
Sensor dimensions, m	0,025
Power supply	AC input: 100-240V - 1A, 50-60Hz, DC output: 12V – 2,5A
Compatibility with operating systems	Windows 10-64 bit
Interface	USB 3,0

The method of baropodometry was used in the research. The platform consists of 4 active modules, measuring 0.4x0.4 *m* and 4 passive. 6400 sensors are installed in each active module (the total number of sensors on the platform is 25600) [1]. The platform works in two modes: static and dynamic.

Baropodometry of static position determines the distribution of load zones, outlines the perimeter of the support polygon, fixes the centers for foot position and the projection of TCM and its displacement, calculates the percentage of support surface and pressure force, including limb overload or pelvic rotation. Static analysis is considered as a geometric model that relates basic biomechanical parameters, correlating with the information obtained from the morphology of the sole (podocontourmetry), and with the reflection for foot pressure obtained in the dynamic analysis. Quantitative values of pressure imprints, detected after the research and presented in the section of static analysis, allow you to notice any possible asymmetry or deviation from the physiological state. The most important of these are divided into "load values" (correlated mainly with pressure information) and "offset values" (correlated mainly with spatial information) - they are interrelated [6].

Baropodometry in the dynamics shows how the pressure is distributed during the rolling to each foot. The point of landing, contact and shock normally have a clear sequence, speed and strength. The graphical representation of the movement can clearly track the stability of the joints, lateral or medial deviations of movement. In the process of analysis, cycles of steps with time characteristics of mono-support and double support are recorded. Elongation of the foot with dynamic support, its expansion in the anterior part during movement are determined. The full step cycle is analyzed for three prints, and a half-step for two. Having a complete step cycle is always a more informative parameter than single prints. When walking, there can also be a supporting mid-foot polygon (As shown in Figure 4), i.e. the moment at which the body transfers the load from one limb to another during the double support. This parameter is crucial in the kinematic reconstruction of stability disorders, which can be caused by incorrect contact, because it allows to detect the interaction of possible stability disorders [4].

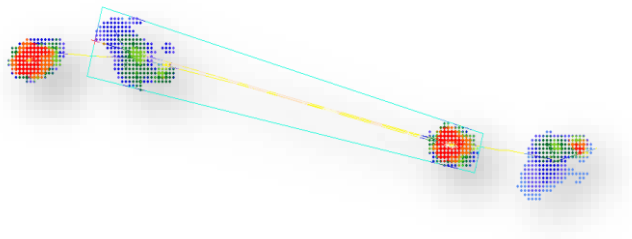


Figure 4: Supporting mid-foot polygon [4]

The research of the load distribution on foot plantar surface. According to formula above, determine the dynamic capacity coefficient *q* (6). After the corresponding calculations it was obtained: for 40% of the group the left lower extremity has the maximum period of support, in the rest of the group - the right. Further calculations were performed for a larger group, i.e. on the left support.

3. Results and discussions

In order to analyze the distribution of dynamic load on the support surface, let`s introduce the coefficient of dynamic load K_{dl} [2]:

$$K_{dl} = P_i \cdot S_i \cdot t_i \tag{8}$$

where P_i – the value of the load on the i-th characteristic of the cycle for foot step (the period of push, roll and landing), kg/m^2;

S_i - characteristics area, m^2;

t_i - duration of the characteristic, *s*.

Impact is a load that is applied in a very short period of time. For example, the shock load occurs when one body falls on another or when the pressure between the bodies in question changes rapidly. It should be borne in mind that if considered a body with a certain mass m, such as a weight of 20 *kg*, the weight of the weight will not change before, after or at the time of fall. This statement is true only when it comes to gravitational mass, but a thorough world research of the phenomena shows that any body also has an inert mass. Researches show that the inertial mass is equal to the gravitational mass, and when it comes to shock loads, they are created not by gravitational but by inert mass. From the general course of physics, the term "load" is not used, but the concept of "force" is used. In this case, all forces can be divided into external and internal. In the course of theoretical mechanics, the theory of resistance of materials, the theory of elasticity, the theory of strength, etc. external forces acting on a particular structure are considered as loads, and internal forces - as stress. It`s assumed that the sum of external forces is equal to the sum of internal forces, this ultimately allows you to make an equilibrium equation for the selected system [2, 6]. In this research considered the human body with mass *m*, as a concentrated load Q_{st}, which according to Newton's second law is calculated:

$$Q_{st} = m \cdot g \tag{9}$$

where *m* – body weight, *kg*;

g – acceleration of gravity.

Due to the support interaction of body weight with foot, in the latter there are certain ratios of elastic forces - the internal force field, which counteracts the occurrence of plastic deformations by support caused for body weight. If the support conditions change (for example, when changing posture), then the configuration for the force field of elastic forces in the considered system changes accordingly. In this case, if a person doesn`t move, i.e. the speed of its movement relative to the considered frame of reference is “0”, then his body still creates a load - static. Here, a human foot is considered as a support that acts with force equal in value to the reference force of the human body *F* (load caused by the mass of the body applied to the support) and opposite in direction, such a force is called the support reaction R_d. When $R_d=F$ – the system is in equilibrium. After the start of motion with a certain acceleration *a* there is a vertical force of inertia directed opposite to the acceleration. Accordingly, if the inertia force is directed downwards, the load caused by the body mass on the resistance

increases, because under the influence of acceleration, the static load Q_{st} is replaced by dynamic Q_d:

$$Q_d = Q_{st} \cdot \frac{a}{g} \tag{10}$$

Relation *a/g* is a measure of mechanical overload, which determines the change in physical condition of the body. If the load is applied instantly - the reference force will be the maximum value, because the body is mass *m*, at a certain point in time *t* will reach a fairly high speed *v*, therefore, the calculations must take into account the gravitational, inert mass of the body and the inert mass of the system (body + support). Therefore, when we consider the types of loads on the human foot, it is necessary to consider shock loads as one of dynamic loads types, which differ from static ones in that the maximum stresses should take into account the forces of inertia.

If we consider the body of person with mass *m* as a physical body that creates a load, and his foot - as a body that receives this load with the occurrence of certain stresses, then the interaction of the body with the foot - the speed of gravity of both bodies, according to accepted frame of reference, do not change. This allows us to consider the loads and stresses caused by them as those that are due only to the gravitational interaction. The foot as a supporting structure is the first to perceive the shock pulse p of the reference reaction, and the further nature of the interaction with the environment directly depends on its functional capabilities.

$$p = m \cdot v \tag{11}$$

where *v* – speed of movement n-th phase of the step cycle, *m/s*.

In this research, we consider a rectilinear (translational) motion, i.e. one for the description of which it`s sufficient to consider the movement for only one material point, which in this case coincides with the body gravity center. The method of calculating the coefficient of dynamic load on the foot (8) is described above. We introduce the coefficient K_{st}, which takes into account the total area of foot support *S* and the time of its contact *t* with the surface:

$$K_{St} = S \cdot t \tag{12}$$

The coefficient value K_{st} will be maximum at the maximum area of foot support, taking into account the period of support. The total amount of foot support reactions R_d, which create shock loads will be equal to the value of effective impact force Q_{shl} body weight. Based on this, it`s fair to say:

$$P = Q_{shl} = R_d \tag{13}$$

where $P = \sum P_i$ – total value of the foot load.

And from equation (12) we obtain the coefficient K_{shl} shock load for a certain phase for step cycle, which can be written as:

$$K_{shl} = K_{St} \cdot R_d \tag{14}$$

Under normal conditions, the foot is exposed to elastic deformations, i.e. it can be considered as an elastic system with one degree of freedom. Then, theoretically, if known the time *t* during which the pulse will be transmitted from one body to another, the impact force can be calculated by the formula:

$$Q_{shl} = \frac{p}{t} \tag{15}$$

However, in practice it`s very difficult to determine the pulse transmission time, because it depends on many different factors and can vary from microseconds to seconds. Therefore, to accurately determine the value of time, and hence the impact force is quite difficult.

To solve this problem, accept the following assumptions:

- in this research, it will consider the foot as an elastic system with one freedom degree. That is, we believe that all the deformities of the foot will be in the area of elastic, restored later.
- deformations of the considered element for design from loadings extend on all length of an element, obey Hooke's law and are proportional to deformations arising at static application of loading from the same body and in the same place.
- the proportionality of dynamic and static deformations is determined by dynamic coefficient *q* (6).

Given this, formula (14) will look like:

$$Q_{shl} = q \cdot \frac{m \cdot v}{t} \tag{16}$$

Taking into account expression (6) and carrying out the transformation, formula (16) is written as follows:

$$Q_{shl} = \frac{P \cdot v}{t} \tag{17}$$

where *P* - the total value of the load on the foot.

Given expression (13), it can be write the formula for calculating the coefficient of shock loads as:

$$K_{shl} = S \cdot P \cdot v \tag{18}$$

Researched for the coefficient of shock loads for above group of subjects. The calculation was performed by left support. Together with the calculation for this coefficient determined the length of the half-step *l*. Figure 5 shows a graph of the coefficient of foot impact loads K_{shl} on the length of the half-step *l*.

The half-step *l* calculated as the distance from the center of heel in the first phase of the step cycle to the center of heel to the same foot in the next phase, *m*. The graph shows that the maximum value of shock loads coefficient (0,584) corresponds to the maximum value of the half-step length (0,7 *m*), and therefore we can conclude

that the greater for half-step length - than the greater the coefficient of foot shock loads.

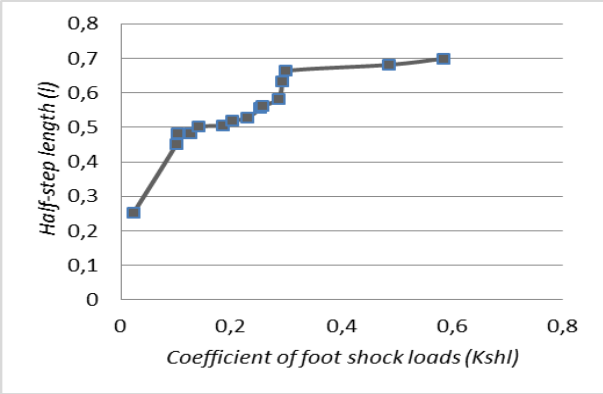


Figure 5: Graph for dependence of the coefficient of foot shock loads from the half-step length

4. Conclusions

The problem of early diagnosis for injuries and deformities the feet is relevant when choosing prevention methods, treatment, orthosis and evaluation for their effectiveness. An important design feature of the foot is its vaulted structure, which provides spring, balancing and repulsive functions. The foot, under normal conditions, is exposed to elastic deformations, i.e. it can be considered as an elastic system with one degree of freedom. It should be emphasized that the human musculoskeletal system is quite complex, but the foot as a supporting structure and part of this system is first to perceive the shock impulse of the support reaction, and its functionality directly affects the further nature for interaction with the environment. The supporting foot properties haven`t been fully researched - to increase the accuracy of their evaluation methods, it`s necessary to take into account the dynamic influence, namely short-term shock loads, on the step cycle relevant phases.

Proposed new method of analysis for stato-dynamic load, which allows to estimate dynamic and shock foot loads and is reduced to the determination for corresponding coefficients. It`s established that the maximum contact per unit time has front part (repulsion phase), then the rear part (landing phase) and the smallest - the middle foot part (rolling phase), the greater the speed and stride length, the greater the coefficient of shock loads, and their peak on the anterior and posterior departments, therefore, these areas need special attention when choosing methods of prevention, treatment and orthosis.

The practical significance for obtained results is to improve the existing methods of researching foot biomechanical parameters, by comprehensively assessing the features of standing and walking, step cycle and support properties. Considered method of researching the functional properties for lower extremities, and in the future, will be consistently improved and expand the range of its application. In particular, in order to assess the dynamic stresses arising in the feet caused by shock loads, additional researchies will be conducted using the measurement by force plate, which will increase the effectiveness of prevention, treatment and orthosis of MSS segments.

References

[1] V. Tsapenko, M. Tereshchenko, “Analysis of the influence of difference between lower findings on biomechanical parameters of walking,” Bulletin of Kyiv Polytechnic Institute. Series Instrument Making, 2019, doi:10.20535/1970.57(1).2019.172034.

[2] V. Tsapenko, M. Tereshchenko, G. Tymchik, S. Matvienko, V. Shevchenko, “Analysis of Dynamic Load on Human Foot,” in 2020 IEEE 40th International Conference on Electronics and Nanotechnology, ELNANO 2020 - Proceedings, 2020, doi:10.1109/ELNANO50318.2020.9088788.

[3] V. Tsapenko, M.F. Tereshchenko, G.S. Tymchik, “Models of evaluation of biomechanical parameters of lower extremities in children,” KPI Science News, 2019, doi:10.20535/kpi-sn.2019.1.158812.

[4] V. Tsapenko, M. Tereshchenko, “Baropodometric method of foot biomechanics research”, in XVI All-Ukrainian scientific-practical conference of students, graduate students and young scientists Efficiency and automation of engineering solutions in instrument making, 341–344, 2020.

[5] V. Tsapenko, M. Tereshchenko, “A comprehensive method for the study of biomechanical parameters of human foot”, in XII Opening of the All-Ukrainian Scientific and Practical Conference Looking into the future of instrument making, 333-336, 2019.

[6] A. Vitenson, K. Petrushanskaya, B. Spivak, A. Matveeva, G. Gritsenko and I. Sutchenkov, “Peculiarities of biomechanical structure of walking of healthy children of different age groups”, Russian Journal of Biomechanics, **17**(1),78-93, 2013, doi: 10.15593/RZhBiomeh/2016.2.05.

[7] Krumka, Kasapova, “Dynamic loads induced by human motion”, in Proc. Bauhaus Summer School in Forecast Engineering: Global Climate change and the challenge for built environment, 1-22, 2014.

[8] V.V. Lashkovsky, M.I. Ignatovsky, “Quantitative assessment of pedobarographic data in plano-valgus deformity of the foot in children”, Medical news, **7**, 69-71, 2012.

[9] N. N. Rukina, M. Yu. Ezhov, Yu. I. Ezhov, “Characteristic features of load distribution in foot zones under various biomechanical conditions”, Bulletin of the Ivanovo Medical Academy, **17**(3), 32-36, 2012.

[10] B. Sh. Minasov, S.P. Gutov, A.R. Bilyalov, “Evaluation of static and dynamic biomechanical parameters of the lower extremities in normal and degenerative-destructive disorders of the feet”, Medical Bulletin of Bashkortostan, 62-66, 2011.

[11] P. Levinger, G.S. Murley, C.J. Barton, M.P. Cotchett, S.R. McSweeney, H.B. Menz, “A comparison of foot kinematics in people with normal- and flat-arched feet using the Oxford Foot Model,” Gait and Posture, 2010, doi:10.1016/j.gaitpost.2010.07.013.

[12] Z. Taha, M.S. Norman, S.F.S. Omar, E. Suwarganda, “A Finite Element Analysis of a Human Foot Model to Simulate Neutral Standing on Ground,” in Procedia Engineering, 2016, doi:10.1016/j.proeng.2016.06.240.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 65-74 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Evaluation the Effects of Climate Change on the Flow of the Arkansas River – United States

Elgaali Elgaali[*,1], Zeyad Tarawneh[2]

[1]*Civil Engineering Program, Higher Colleges of Technology, Dubai Men's College, Dubai, 10000, UAE*

[2]*Department of Civil Engineering, The Hashemite University, Zarqa, 13115, Jordan*

ARTICLE INFO

Article history:
Received: 11 December, 2020
Accepted: 16 February, 2021
Online: 10 March, 2021

Keywords:
Climate change
River flow
Impacts

ABSTRACT

The behavior of rivers' hydrology and flow under changing climate has been an objective of interest for long time. In this study the impacts of climate change on streamflow of the Arkansas River will be investigated. The paper is an extension of work originally presented in ASET conference in Dubai. The Arkansas River is a crucial element in the economy of the Colorado state in the USA. It is a vital transportation channel and main source of water for irrigated agriculture. In order to understand the direction and magnitude of climate change, the changes in the monthly flow regimes of the Arkansas River were projected using two future climate scenarios. The projections extend over 100 years (2000 – 2100). The projections were carried out in the period from April to September because this is the period of the river's significant runoff. For better presentation the monthly flows were aggregated and presented on decadal time scale. Project stream flow is simulated using a neural network that was developed to autonomously model the relationship between different flow levels and the resultant changes in temperature and precipitation. In general, the projections depict a rise in the magnitude of the flow in the river. In general the increases concurred with the patterns of temperature and precipitation projected for the region. Noticeably, the high temperatures cause the precipitation to melt earlier shifting the peak flow to April instead of June. Statistical analysis show that in the future the current levels of flow would be surpassed more frequently. The probability of exceedance fluctuates between from month to month – reaching its peak in April-July; before retreating to a very low level in August and becoming almost negligible in September. Overall, the results reveal profound implications for regional water resource planning and management.

1. Introduction

The Arkansas River is a crucial element in the economy of the state of Colorado, in the western United States. It is a vital transportation channel and main source of water for irrigated agriculture. Water sources for the river mainly from snowmelt.

Snow melt is a critical component of the water cycle in the American West, providing the region with at least 50% of its annual runoff, and up to 80% in some years. Concerns are growing regarding the impacts the altered climate might have on snowpack and the region's water cycle overall. A few studies projected the region to have profound changes in minimum winter temperatures, summer average temperatures, snowfall, snow-melt, and growing season rainfall quantities [1]-[5]. Several other studies have reported that spring and early summer temperatures are expected to increase while the quantity of snowpack in spring is expected to decrease [6]-[8]. There are other several earlier studies of snowmelt dominated systems show similar seasonal shifts in snowmelt runoff as a result of warmer temperatures and a shorter snow accumulation period [8]-[11].

If the changes in climate take place as projected, it is expected to have profound impacts on the hydrology of the Arkansas River and hence the economy of the region. Many studies, using historical data analyses, have explored the impacts of climate change on water in this region where water is already under stress [12]–[19]. The outcome of these studies reported that, even though, the direction and magnitude of change in climate is well documented (increase in temperature) but there is no consensus on the direction and magnitude of the impacts on the region. And this attributed to uncertainty regarding the changes in precipitation's

[*]Corresponding Author: Elgaali, P.O. Box: 15825, Dubai, UAE, 971 552493085, eelgaali@hct.ac.ae

patterns under warmer climate. But, precipitation is the main driving factor of the changes in streamflow. Almost all of the studies conducted in this region, have reported that snow is expected to melt early shifting the peak runoff and the growing season to take place earlier. But none of them reported increases in streamflow. However, the changes in precipitation regimes are expected to cause the snowpack to decrease and/or to melt (shift) earlier. Therefore, streamflow is expected to decrease and the peak emanates earlier shifting the growing season towards winter season [6], [13], [16]. Some studies on the region, projected a reduction in streamflow up to 30% below the historical recorded flow [18]. More recent studies using data sets extracted from GCMs support this result projecting a reduction in streamflow ranges between 10 – 30% by the end of the 21st century [6], [13]-[14]. The only increase in streamflow resulted from increase in precipitation was reported by Groisman and others [20].

It is noticeable, from the outcome of all these studies that there is no consensus on the magnitude and direction of the change in the precipitation and hence the river flow under the changing climatic conditions. This may be attributed to the nature of the climate and hydrologic models used (coarse spatial and temporal resolution). The coarse resolution bounds the general circulation models (GCMs) to reproduce a similarly complex spatial environment that simulates the actual precipitation. Besides, the discrepancies in the precipitation projections are larger than the ones in the temperature projections [21]-[23].

Scale is crucial for climate and hydrologic modeling. It is reported that the signal of climate change on monthly patterns of runoff is stronger than the annual ones [23], [24]. Projections at finer scales is necessary to better evaluate the impacts the changing climate might have on the hydrology in the region [25]. Therefore, in this research, streamflow is linked to climate change scenarios on a monthly scale and aggregated to annual and decadal scale.

Series of climate parameters (scenarios) have been projected to simulate the changes in climate in the future. Most of these scenarios were developed assuming equilibrium (change in climate caused by jumps in CO2 levels in the atmosphere sometime in the future). In fact, climate is expected to have linear trend of change (transient) following the linear trend of CO2 on the earth [26], [27]. Even though, transient scenarios have not been widely used in studies of climate change. As such, it is necessary to explore their impact, as:

• Transient studies provide deeper insight on trends in climatic change and annual variability;

• Transient projections can indicate the potential rate of change, which is crucial in determining how to respond and adapt to that change, and;

• Transient simulations may give a more accurate picture of the likelihood of when certain critical thresholds will be crossed [27].

This research aims to demonstrate how transient scenarios are better able to evaluate the implications the changing climate might have on river flow at scale. These findings can be particularly useful in supporting or facilitating more informed and higher quality decision-making for water planning and management.

2. Description of the Study Area

The subject of this study is the Arkansas River Basin in Colorado which is enclosed by the Rocky Mountains to the west and Kansas to the east and expanding towards New Mexico and Oklahoma to the south. (Figure 1), the surface area of the basin is approximately 72,742 km2 (28,415 square miles) and accounts for about 27% of the land area of the state of Colorado. The river's headwaters can be found at over 3,050m (10,000 feet) above sea level, in the vicinity of Leadville, CO. From there, its elevation drops rapidly as the river flows out of the mountains near Pueblo, CO and continues eastward towards the border with Kansas border, near the town of Holly, where the elevation drops to roughly 1,036 m (3,400 feet).

The basin sees a wide range of temperatures and precipitation due to variations in topography. The temperature can see average annual lows of about 2ºC at the basin's highest point in the mountains and can reach average highs of about 12°C in the lower valley. Seasonal variations are also significant, illustrated by the. the average frost free season (0ºC) ranging from a low of 85 days at Leadville to a high of 167 days at Canon City.

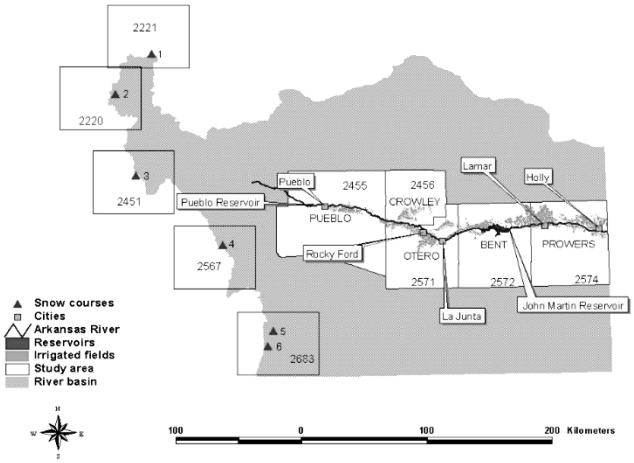


Figure 1: Features of the study area in the Arkansas River Basin, Colorado

Precipitation also demonstrates significant variability throughout the year, ranging from 229-305 mm annually in the middle and eastern parts of the basin, to 406-508 mm in the center and east, to a range of 406-508 mm in the west, and as much as 1143 mm at the highest elevations in the mountains. At these heights, much of the precipitation manifests as snow, the runoff of which accounts for the bulk of the region's annual water supply. Consequently, the size of the water supply is dependent on the volume of winter snowpack and can shift from year to year. However, in general, on average, over 60% of the annual runoff takes place in late spring and mid-summer between April and July, with a further 20% occurring in later summer and early autumn between August and October [28].

3. Data

The underlying data for this study is in two parts: (1) historical climate scenarios and (2) future climate scenarios. It was sourced

from the Vegetation-Ecosystem Modeling and Analysis Project (VEMAP) [29] which collected extensive climate data for the contiguous United States and includes historical data dating back to 1895 and up to 1993, supplemented by projections from two GCM-based scenario models covering the period from 1994-2099.

The historical scenarios were based on data records of varying length from 1,200 stations for the 105-year period between 1895 to 2000 as well as shorter records from approx. 6,000 to 8,000 stations for the 50-year period between 1951 to 2000. Both GCM models, the HAD model and the CCC model, further output climate scenarios with the underlying assumption that carbon dioxide concentrations are progressively increasing at a rate of 1% every year (transient). The VEMAP divided the continental US into a grid with 0.5° x 0.5° cells and used these to generate scenarios that would demonstrate the impact of factors such as topography and local ecosystems on climate [29]. The project sponsor, the National Center for Atmospheric Research (NCAR), used a downscaling technique, spatial interpolation to topographically adjust both the historical and projected climate data to fit the small grid cells. The downscaling process took into account the impact of local topography on climate parameters. Figure 2 shows adjusted projections (downscaled) from both models as well as the mean +/- 1 standard deviation for each parameter.

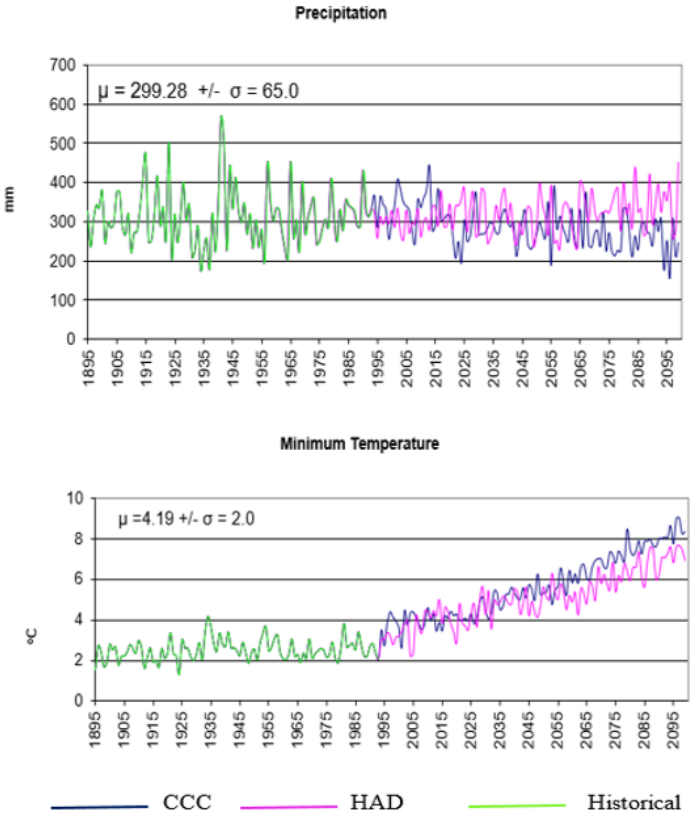


Figure 2: Historical and the future scenarios of Precipitation and Temperature

The historical data, sourced from the U.S. Geological Survey (USGS) [30], captures flow levels for the Arkansas River near Pueblo, on a monthly basis between 1900 and 2000. Figure 3 shows the average monthly distribution of the river flow with data collected over 30 years.

As shown in Figure 1, climate data (precipitation & temperature) was collected from five snow courses (grid cells) in the Rocky Mountains a sample which represent overall runoff in the region. Each individual cell contains a minimum of one course upon which runoff forecasts were based. Since the accuracy of the model's simulations depends heavily on the accuracy of the input data as well as the temporal and spatial scale, the selected snow courses were vetted to ensure accurate and reliable records that reflect the natural runoff and correlated highly snowmelt levels across the entirety of the region [28]. The selected courses, as well as their grid cell number are listed in the adjacent table below (Table 1).

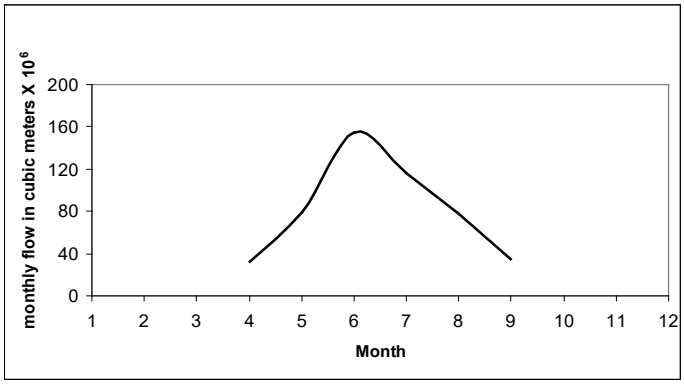


Figure 3: Seasonal Cycle of the Arkansas River at Pueblo

Table 1: Characteristics of Snow Courses

Station	Latitude (deg)	Longitude (deg)	Elevation (m)	Cell #
Apishapa	37.33	105.07	3,040	2683
Brumley	39.08	106.53	3,222	2220
Fremont Pass	39.38	106.20	3,465	2221
Prophyry	38.48	106.33	3,271	2451
South Colony	37.97	105.53	3,294	2567
Whiskey Creek	37.22	105.12	3,117	2683

4. Methods

In this paper, in order to study the changes in the river flow that would accompany any climatic changes in the future, we used climate scenarios (precipitation and temperature) extracted from two General Circulation Models (GCM). Namely, the Hadley Centre for Climate Prediction and Research (HAD) and the Canadian Climate Centre (CCC). The data from these two GCMs is of transient nature (gradual change) and high spatial and temporal resolution. To account for spatial variability, the climate data was scaled-down and then used to quantify the impacts the climate change might have on the monthly flows of the river (Arkansas River). The paper focused to explore the changes in streamflow during the period of April-September, because this is

the growing season in the region, and the time during which the effective runoff is usually generated. The monthly flow in the river was simulated using artificial neural network (ANN). The neural network models have been proved to be capable to transform minimum number of inputs into an output [31] - [33]. The neural network model was used for the minimum number of the input data sets it requires. The model was calibrated and validated using 100-years of historical data records. To account for the extrapolation, the historical data was reinforced by extreme events that supposed to resemble the climate conditions under global warming. The ANN model was then used to simulate the river flow under global warming conditions. The output (simulations) from both models (GCM and ANN) are compared to historical records (baseline) to estimate the quantity and extent of the change.

4.1. Modeling Streamflow

The model developed can be described as feedforward artificial neural network (ANN). It is three-layer network; in the middle layer there is a function that map the relationship between the inputs in the first layer (precipitation (PPTa) and temperature (T)) and the outputs in the last layer (streamflow (Qr). The ANN is well formulated in the following form:

$$Q_r(t) = f(PPTa, T) \quad (1)$$

Here, Qr represents the average monthly streamflow whereas PPTa represents precipitation (accumulated from October of the preceding year, up to each individual month of the year after). To illustrate, the value of PPTa in April would be equal to the cumulative levels of snowpack between October and April. The variable T, meanwhile, represents the average temperature (calculated between April and each individual month in the model). Again, to illustrate, the value of T for May is equal to the average temperature between and April and May. For the month of April itself, we substituted the average March temperature. Accumulating precipitation and temperature was found to contain stronger signals of climate variability. The input data to the ANN was normalized to fall in the range (-1, 1). Normalization is required to remove geometrical biases and equally distribute importance of each input.

4.2. Model Testing and Validation

Data records of 100 years (1900 -2000) were used in training, validating, and testing the neural network. Sixty years of data records 1(900-1960) were used for model's training, 14 years (1961-1975) were used for validation, and 24 years (1976-2000) were used for testing.

The developed neural network was found favorably capable to simulate the river flow (output) when the precipitation and temperature were used as inputs. The parameters shown in Table 2 summarize the validity of the model. The correlation coefficient (R) is usually used to show the degree of connection between two variables. The values of R range from zero as minimal to one as optimal (0, 1). The Root Mean Square Error (RMSE) is used to determine the magnitude of departure (residual variance) of simulated values from measured ones; the zero value indicates the least magnitude of departure (optimal value).

However, the values of the presented parameters (R and RMSE) document the high correlation between the measured and predicted values and hence the validity of the developed model.

Table 2: Summary of the Model Validation and Testing

Month	Training		Testing	
	R	RMSE	R	RMSE
April	0.634	0.17	0.562	0.24
May	0.790	0.10	0.770	0.18
June	0.863	0.11	0.847	0.15
July	0.899	0.06	0.740	0.13
August	0.852	0.06	0.781	0.12
September	0.904	0.07	0.769	0.17

5. Results and Discussion

5.1. Climatic changes

To determine the directionality of the climate change in the region, the main features of the climate scenarios generated from the two GCMs are presented. For better presentation the climate scenarios were aggregated into decadal monthly mean.

Fig. 4 shows the changes in the average monthly temperatures under the two scenarios (HAD and the CCC). The GCMs predicted slight increase in temperature (the region is projected to be warmer). There is a very clear gradual increase in temperature. The change in temperatures during the winter (Dec – Feb) is predicted to be comparatively higher than in the summer (June – August). The changes in temperature look high when compared to the very low historical series of temperatures in winter.

During the focus period (April–September) the temperature average is predicted to increase by 5°C in the 2090s, with a growth rate of 0.45°C every decade. Generally, the 2090s is predicted to be the warmest decade. The figure shows that the CCC scenario predicted higher temperatures in the region compared to the HAD scenario. However, the temperature projections, generated by both scenarios, are similar in trend of change.

Fig. 5 shows the changes in monthly precipitation in decades. It is shown that the HAD scenario predicted large and sudden changes in the average monthly precipitation. Despite the variation, the figure shows a clear pattern in the decadal mean. The figure shows an increase in the decadal mean from the 2010s to 2030s. The increase was gradual at a rate of 1.4% per decade. The 2040s and 2050s noticeably experienced a drop in the decadal mean. The drop was gradual and rapid at a rate of 1.4% per decade. Then a gradual increase took place again. There was rapid and gradual increase in the decadal mean from 2060s to the 2090s with a growth rate of 2.2% every decade. The projected drop in precipitation could be attributed to changes in natural, large-scale features of the climate: the El Niño Southern Oscillation (ENSO), El Niño, and La Niña [20]-[24].

In general, under the HAD scenario, the increase of the precipitation above the historical base line is expected to be of magnitude of 36% in the spring and of 25% in the summer. On the other hand, it is shown that the CCC scenario predicts a drier

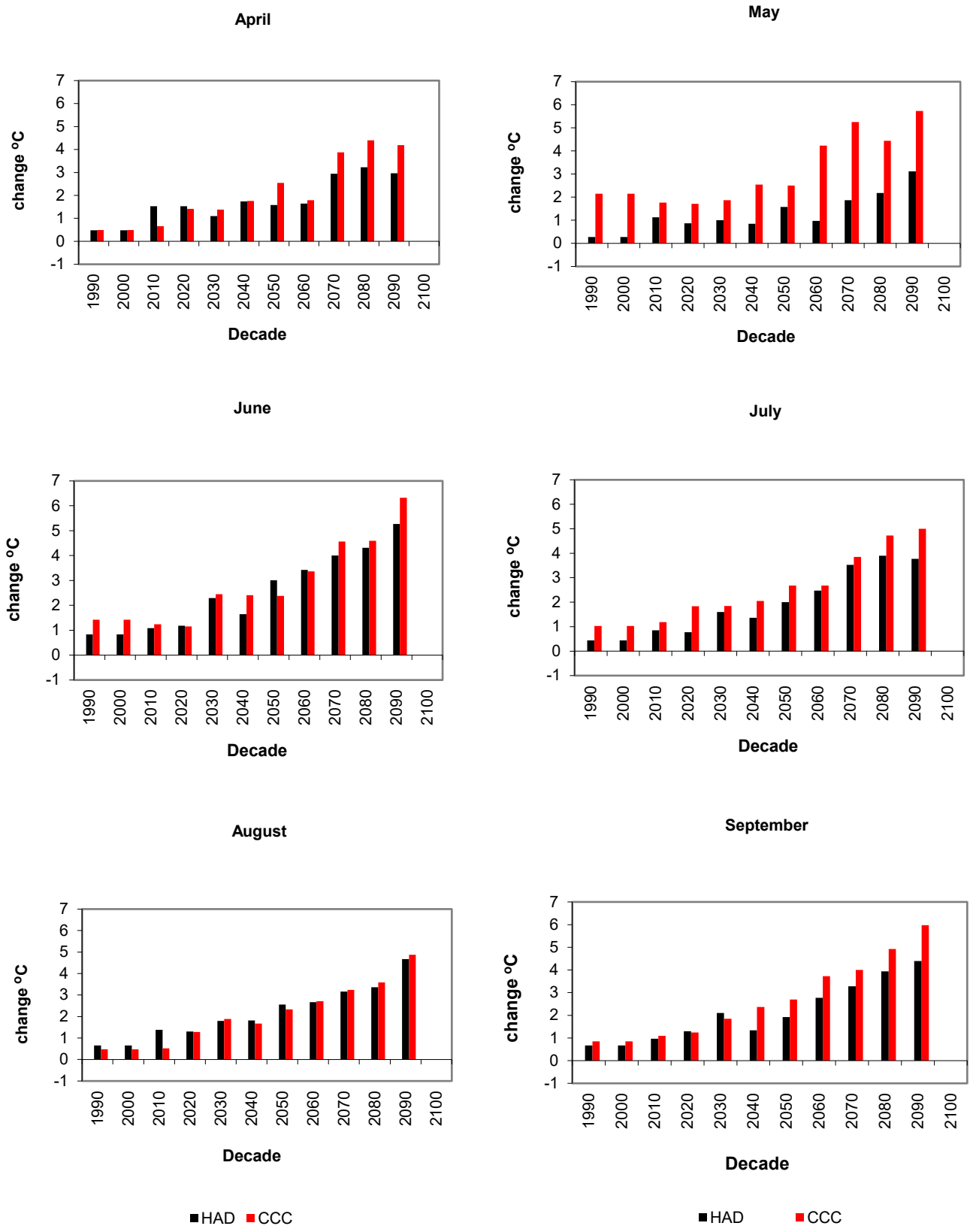


Figure 4: Departures, form historical distribution, of temperature projections from the GCMs: HAD/CCC

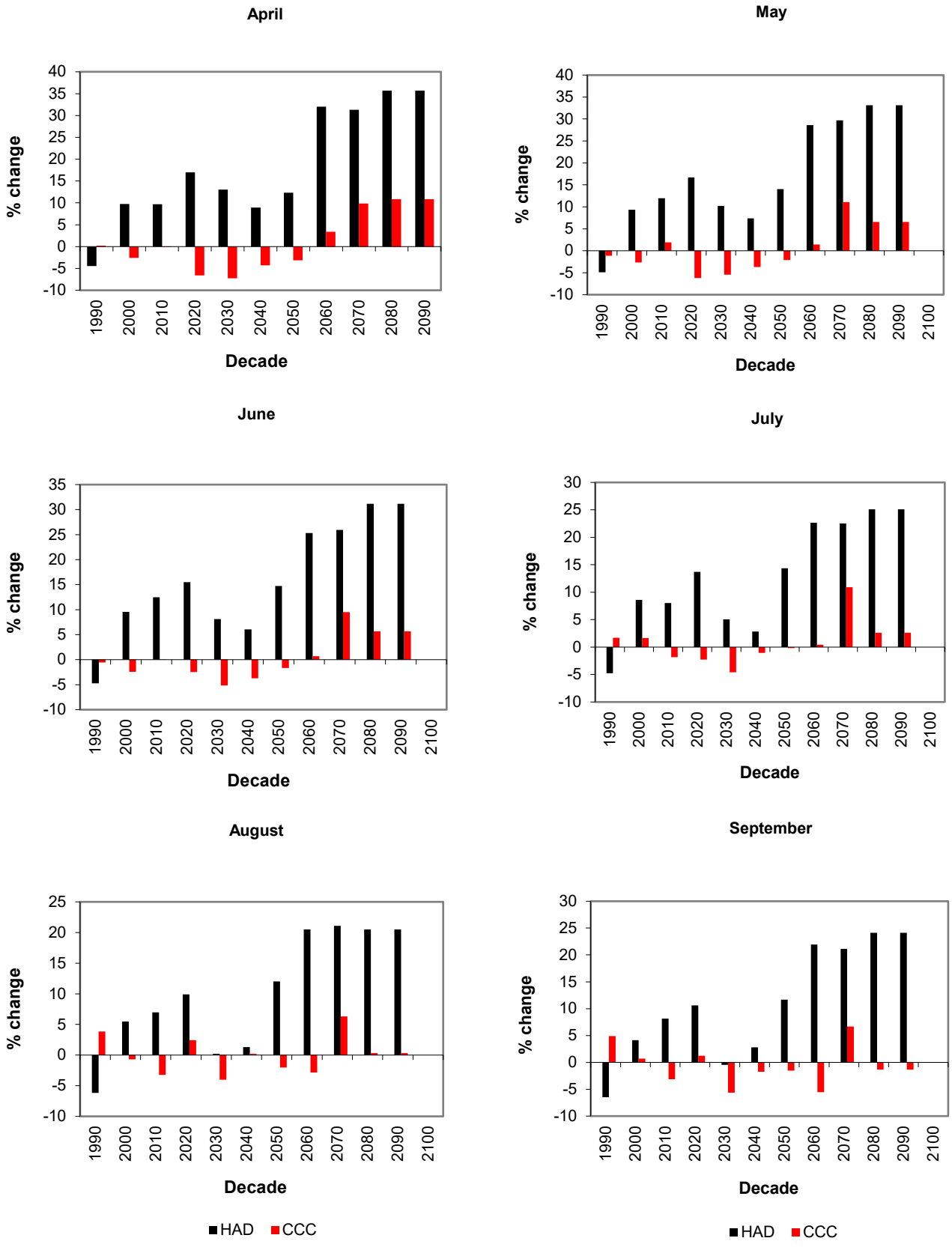


Figure 5: Departures form historical distribution of precipitation projections from the GCMs: HAD/CCC

climate for the region. The projections of the CCC show no clear departure from the historical levels except for the 2080s where a the region was projected to have significant increase (9% above the baseline). Worth to mention, the driest climate is expected during the 2040s with a departure of 7%, below the historical levels, in April, 4% in August, and 47% in October. However, the CCC tends to project precipitation patterns with no clear trend and less variability compared to the HAD. Generally, the two scenarios, show poor similarity in the trend and the extent of change in the precipitation projections.

5.2. Impacts on Streamflow

Figure 6 shows the projections of streamflow. The projections are compared to the historical flow distribution. In spring season, under the HAD scenario, driven by the projected increases in precipitation shown in Figure 5, the results show a large increases in future monthly streamflow. Compared to historical records, the average monthly streamflow increased by 87% for the 2030s and 2050s. The increase jumped to an average of 200% for the 2070s, 2080s, and 2090s. In summer, the increases are projected to be less ranging from 39% for the 2030s to 74% for 2090s. In the fall, dampened by the increase in temperature, the increase in streamflow dropped to 5% in 2030s and to 15 % in 2090s. In general the results show that the largest departures (increases) in flow occur in late spring (April) through early summer (June). This is attributed to the early melting of snow in the mountains. The combined effect of high temperatures and increase in winter snow leads to early snowmelt. However, as a result of change in climate, the peak flow is expected to occur in April rather than June.

The CCC scenario projects a drier climate than the HAD one. Driven by the projected decreases in precipitation shown in Figure 5, the results show that, for 2030s and 2050s, the future monthly streamflow experiences a very slight decrease (5%) in the summer and negligible change in the spring and the fall. For the 2070s and 2090s, the results show a considerable increase (65%) in the spring and negligible change in the fall. The fluctuations (decreases and increases) in projections of the streamflow are attributed to the annual and decadal variability in the projected changes in climate. The results presented here is widely supported by other several earlier studies conducted on the region [7], [9], [13]-[15].

Figure 7 shows the frequency analysis results of the monthly flow during the focus period (April – September). The results are presented in terms of flow duration curves (FDC). In general, FDC are used to depict the percentage of time the flow in the future exceeds the current levels of flow (historical). Under the CCC scenario the percent of exceedance is almost nil for all months. Under the HAD scenario the percent of exceedance is high for all months (April – September). The flow duration curve falls above the historical (baseline) curve, indicating that the river flow in the future exceeds the current flow value (historical). The probability of exceedance differs from month to month; it is relatively high in the months of April to July and low in August. In September the probability of exceedance almost nil because the two flow curves (historical and future) look similar and almost coincide. The variability in the projected flow is greater in April-July, indicated by the steep flow curves, and stable in August and September inferred by the gentle flow curves.

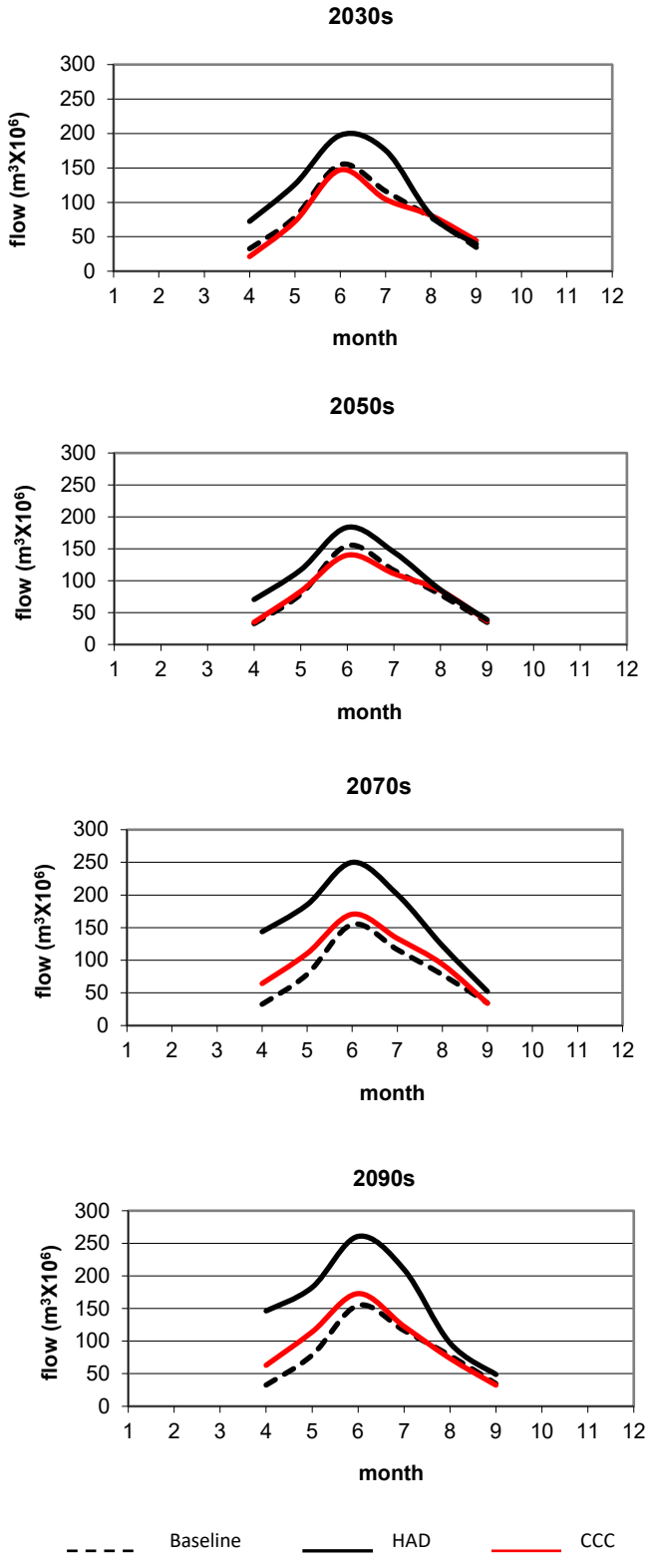


Figure 6: Historical and projected streamflow of the Arkansas River

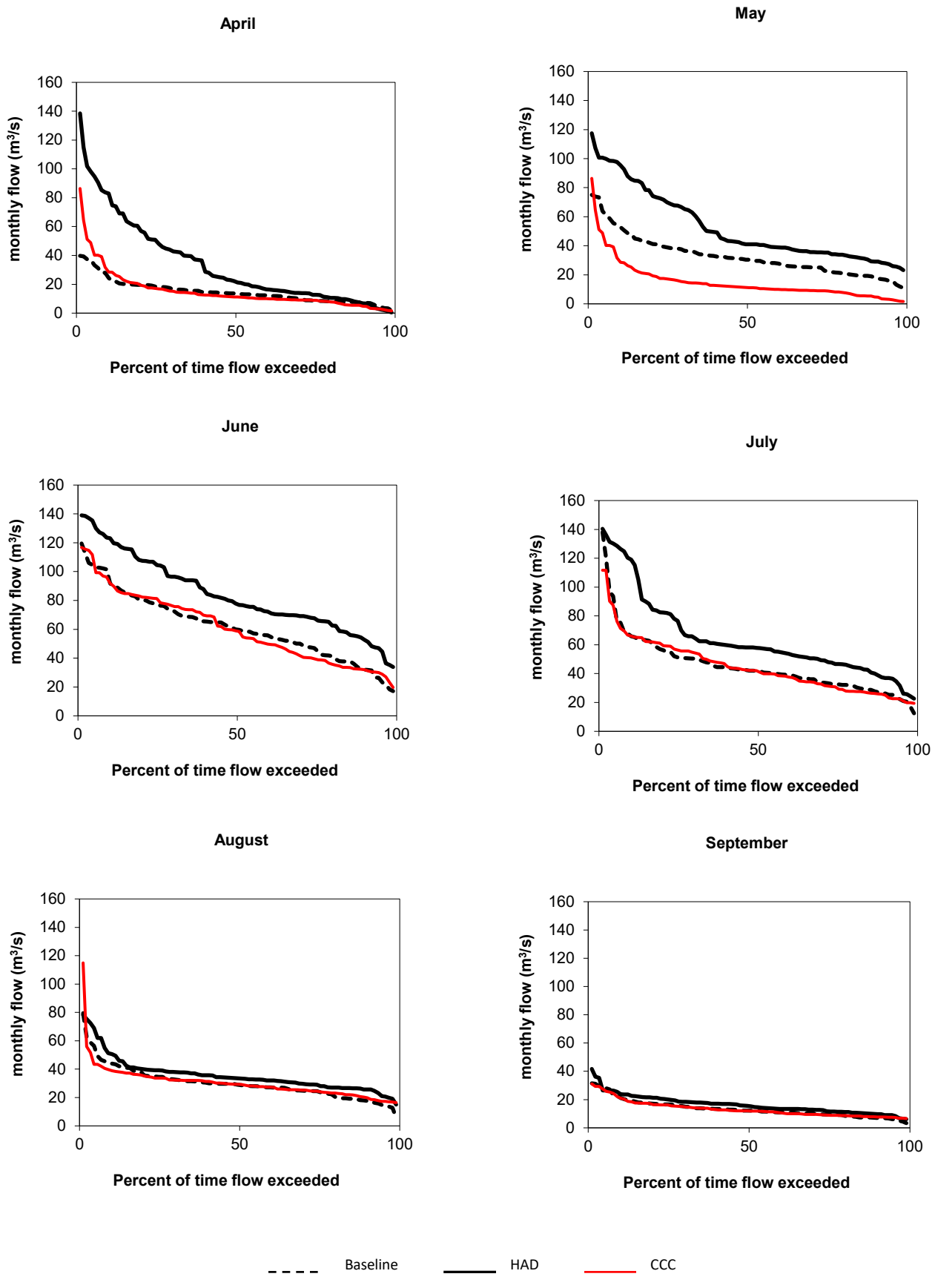


Figure 7: Flow duration curves for monthly flow

This study investigated the influence of different climate phenomenon, caused by climate change, on the flow levels of the Arkansas River in USA. In order to carry out the study, time series of climate data in the future (scenarios) were extracted from two GCMs (CCC and HAD), and then plugged in a hydrologic model (ANN) to generate time series of flow. While the scenarios, produced by both models, agreed that the temperatures were likely to increase, they differed on the directionality of the change in terms of precipitation levels.

However, under the HAD-generated scenario, there is expected to be a relative increase in precipitation, resulting in increased wetness, while the CCC-generated scenario is comparatively dry and demonstrates higher variability. This variability can be attributed to annual, as well as decadal, shifts in expected climatic changes.

The results of this investigation suggest that, under the range of climate conditions studied, river flow processes are strong indicators of how climate change is impacting temperature and precipitation levels in the region. The HAD scenario in particular suggests that present flow levels are likely to be regularly exceeded in the future, especially during the peak April-July period. It further suggests that any increase in springtime flow is likely to be sizeable enough to compensate for attendant decreases in summertime flow and they further demonstrate that the historical peak flow, which used to begin in June has shifted to begin in April, extending the period of high flow levels. These results also align well with the results of other studies that explore the expected environmental effects of rising temperatures on local ecosystems in the western parts of the US [12]-[15].

The Arkansas River Basin can be considered one of the most vulnerable regions in the country when it comes to climate change. If, moving forward, precipitation levels match the projections, in terms of both magnitude and timing, they can be expected to have a significant impact that pose challenge to the water planners and managers in the region.

Conflict of Interest

No conflict of interest in publishing this paper.

Acknowledgment

The authors would like to thank VEMAP for providing the climate data.

References

[1] E. Elgaali, Z. Tarawneh, "Effects of climate change on river flow stochastic approach," Advances in Science and Engineering Technology International Conferences (ASET), Dubai, United Arab Emirates, 1-5, 2019, doi: 10.1109/ICASET.2019.8714549.

[2] P. Mote, S. Li, Sihan, D. Lettenmaier, M. Xiao, R. Engel, "Dramatic declines in snowpack in the western US, " Climate and Atmospheric Science, 1(2), 1-6, 2018, doi:10.1038/s41612-018-0012-1.

[3] L. Dahlman, R. Lindsey, " Climate Change: Spring Snow Cover," Bulletin of the National Oceanic and Atmospheric Administration, 2018, Climate.gov.

[4] K. E. McCluney, N. L. Poff, J. H. Thorp, G.C. Poole, M. A. Palmer, M.Williams, B.S. Williams, J.S. Baron, "Riverine macroecology: a framework for understanding the sensitivity, resistance, and resilience of whole river basins and responses to multiple human alterations," Frontiers in Ecology and Environment, 12, 48-48, 2014, DOI: 10.1890/120367.

[5] J. S. Baron, M. D. Hartman, L. E. Band, and R. B. Lammers, "Sensitivity of a high elevation Rocky Mountain watershed to altered climate and CO2," Water Resources Research, **36**(1), 89-99, 2000, 0043-1397/00/1999 WR900263.

[6] D. R. Gergel, B. Nijssen, J. T. Abatzoglou, D. P. Lettenmaier, M. R. Stumbaugh, "Effects of climate change on snowpack and fire potential in the western USA," Climatic Change, 1–13, 2017, doi:10.1007/s10584-017-1899-y.

[7] R.W. Dudley, G.A. Hodgkins, M.R. McHale, M.J. Kolian, B. Renard, "Trends in snowmelt-related streamflow timing in the conterminous United States," Journal of Hydrology, 547, 208-221, 2017, doi: 10.1016/j.jhydrol.2017.01.051

[8] T. B. Barnhart, N. P. Molotch, B. Livneh, A. A. Harpold, J. F. Knowles, D. Schneider, "Snowmelt rate dictates streamflow," Geophys. Res **43**(15), 8006-8016, 2016, DOI: 10.1002/2016GL069690.

[9] G. J. McCabe, and M. P. Clark, "Trends and variability in snowmelt runoff in the western United States," Journal of Hydrology-American Meteorology Society, 6, 467-481, 2005, doi: 10.1175/JHM428.1

[10] S. K. Regonda, B. Rajagopalan, M. Clark, and J. Pitlick, "Seasonal cycle shifts in hydroclimatology over the western United States," J. Climate, 18, 372-384, 2005, doi: 10.1175/JCLI-3272.1

[11] I. R. Stewart, J. D. R. Cayan, and M. D. Dettinger, "Changes in snowmelt runoff timing in western North America under a 'business as usual' climate scenario," J. of Climatic Changes, 62, 217-232, 2004, doi: 10.1023/B:CLIM.0000013702.22656.e8

[12] C. F. Clifton, K. T. Day, C. H. Luce, G. E. Grant, S. Mohammad, J. E. Halofsky, B. P. Staab, "Effects of climate change on hydrology and water resources in the Blue Mountains, Oregon, USA," Climate Services, 10, 9-19, 2018, DOI: 10.1016/j.cliser.2018.03.001.

[13] D. Li, M. L. Wrzesien, M. Durand, J. Adam, D. P. Lettenmaier, "How much runoff originates as snow in the western United States, and how will that change in the future?," Geophysical Research Letter, **44**(12), 6163-6172, 2017, doi: 10.1002/2017GL073551

[14] S. Kurt, B. Katrina, M. Richard, "Shifts in historical streamflow extremes in the Colorado River Basin," Journal of Hydrology: Regional Studies, 12, 363-377, 2017, DOI:10.1016/j.ejrh.2017.05.004.

[15] A. Nigel and G. Simon, "The impacts of climate change on river flow regimes at the global scale," Journal of Hydrology, 486, 351-364, 2013, DOI:10.1016/j.jhydrol.2013.02.010

[16] D. L. Ficklin, I. T. Stewart, and E. P. Maurer, "Climate change impacts on streamflow and subbasin-scale hydrology in the upper Colorado River Basin," PLoS One, **8**(8), e71297(2013), doi: 10.1371/ journal.pone.0071297

[17] B. L. Harding, A. W. Wood, J. R. Prairie, "The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin," Hydrol Earth Syst Sci Discuss, **9**(1), 847-894, 2012, DOI: 10.5194/hessd-9-847-2012

[18] D. W. Clow, "Changes in the timing of snowmelt and streamflow in Colorado:a response to recent warming," Journal of Climate, 23, 2293–2306, 2010, doi: 10.1175/2009JCLI2951.1

[19] N. S. Christensen, A. W. Wood, N. Voisin, D. P. Lettenmaier, R. N. Palmer, "The effects of climate change on the hydrology and water resources of the Colorado River basin," Climatic Change, **62**(1), 337–363, 2004, DOI: 10.1023/B:CLIM.0000013684.13621.1f

[20] P. Y. Groisman, R. W. Knight, and T. R. Karl, " Heavy precipitation and high streamflow in the contiguous United States: trends in the twentieth century," Bull. Amer. Meteor. Soc., **82**(2), 219-246, 2001, DOI: 10.1175/1520-0477(2001)082<0219:HPAHSI>2.3.CO;2.

[21] N. Chokkavarapu, V. R. Mandla, "Comparative study of GCMs, RCMs, downscaling and hydrological models: a review toward future climate change impact estimation," SN Appl. Sci. 1, 1698, 2019, DOI: 10.1007/s42452-019-1764-x.

[22] J. Gautam, G. Mascaro, "Evaluation of CMIP5 historical simulations in the Colorado River Basin," International Journal of Climatology, 38(10), 3861-3877, 2018, DOI: 10.1002/joc.5540

[23] M. Girotto, K. N. Musselman, R. L. H. Essery, "Data assimilation improves estimates of climate-sensitive seasonal snow," Curr Clim Change Rep, 6, 81–94, 2020, DOI: 10.1007/s40641-020-00159-7.

[24] B. S. Naz, S. C. Kao, A. Moetasim, R. Deekshai, R. Mei, L. C. Bowling, "Regional hydrologic response to climate change in the conterminous United States using high-resolution hydroclimate simulations," Global and Planetary Change, 143, 100-117, 2016, doi: 10.1016/ j.gloplacha. 2016.06. 003

[25] S. Chang, W. Graham, J. Geurink, N. Wanakule, T. Asefa, "Evaluation of impacts of future climate change and water use scenarios on regional hydrology," Hydrol. Earth Syst. Sci., 22, 4793–4813, 2018, doi: 10.5194/ hess-22-4793-2018
G. G. Persad, D. L Swain, C. Kouba, "Inter-model agreement on projected shifts in California hydroclimate characteristics critical to water managemen," Climatic Change, 162, 1493–1513, 2020, doi: 10.1007/s10584-020-02882-4

[26] M. Haasnoot, J. Schellekens, J. J. Beersma, H. Middelkoop, J. C. J. Kwadijk1, "Transient scenarios for robust climate change adaptation illustrated for water management in The Netherlands," Environmental Research Letters, **10** (10), 2016, DOI:10.1088/1748-9326/10/10/105008.

[27] National water and climate center, Department of Agriculture. NWCC - USDA. https://www.wcc.nrcs.usda.gov/snow/snotel-wereports.html

[28] T. G. F. Kittel, J.A. Royle, C. Daly, N. A. Rosenbloom, W.P. Gibson, H. H. Fisher, D. S. Schimel, L. M. Berliner, and VEMAP2 Participants, "A gridded historical (1895-1993) bioclimate dataset for the conterminous United States," Proceedings of the 10th Conference on Applied Climatology, Reno, NV. American Meteorological Society, Boston, 219-222, 1997.

[29] U.S. Geological Survey (USGS) (http://waterdata.usgs.gov/nwis/sw).

[30] J. S. Patricia, S. A. Javier, P. S. Julio, P. V. David, "A Comparison of SWAT and ANN models for daily runoff simulation in different climatic zones of peninsular Spain," Journal of water, **10**(2), 192, 2018, DOI: 10.3390/ w1002 0192

[31] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, H. Arshad, "State-of-the-art in artificial neural network applications: a survey," Heliyon, **4**(11), 2018, doi: 10.1016/j.heliyon.2018.e00938.

[32] I. Aichouri, A. Hani, N. Bougherira, L. Djabri, H. Chaffai, S. Lallahem, "River flow model using artificial neural networks," Energy Procedia, 74, 1007-1014, 2015, doi: 10.1016/j.egypro.2015.07.832.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 75-84 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Fuzzy Analytical Hierarchy Process and Fuzzy Comprehensive Evaluation Method Applied to Assess and Improve Human and Organizational Factors Maturity in Mining Industry

Yousra Karim*, Abdelghani Cherkaoui

Research Team EMISYS: Energetic, Mechanic and Industrial Systems, Engineering 3S Research Center, Industrial Engineering Department, Mohammadia School of Engineers, Mohammed V University, Rabat, 10010, Morrocco

ARTICLE INFO

Article history:
Received: 16 December, 2020
Accepted: 20 February, 2021
Online: 10 March, 2021

Keywords:
Human and Organizational Factors
Safety
Maturity Model
Analytic Hierarchy Process (AHP)
Fuzzy Comprehensive Evaluation Method (FCEM)
Fuzzy Logic

ABSTRACT

The literature shows a growing interest in taking into account human and organizational factors (HOFs) to achieve safe and successful human performance by reducing the risk of errors. In this sense, the concept of maturity models aims to help companies in the integration of these factors by assessing the current level of maturity and define future areas for improvement. The HOFs maturity model shown in this article is based on the five main factors that can impact human performance and safety positively. The measurement methodology consists in applying the Fuzzy Analytic Hierarchy Process (FAHP) method to calculate the weighting of the elements of the model since they do not have the same importance. Next, the Fuzzy Comprehensive Evaluation Method (FCEM) is used to determine the maturity level in terms of HOFs among the five proposed by performing an assessment of the sub-factors using a questionnaire. The purpose of using fuzzy logic is to deal with vagueness and uncertainty of the human reasoning . The proposed model and methodologies are implemented to bring out the current situation of a Moroccan mining organization and identify the elements that require more effort to reach the next level of maturity.

1. Introduction

High risk organizations are increasingly aware of the importance of human and organizational factors and their impact on health and safety on work. The integration of these factors allows companies to achieve safe and successful human performance, by understanding the interactions between humans and other elements of complex systems in order to predict and reduce accidents and incidents.

There are several approaches bringing together the elements that contributed to major accidents in the past, from the implementation of the Seveso I and II Directives to the integration of the Human and Organizational Factors. However, it is difficult for the company to choose the right approach and to know which specific areas of HOFs need to be prioritized. Therefore, the frequently asked questions are : Where are we and where do we want to be ?. These questions are used to define the current state of the company and the desired objective.

The maturity model concept has been developed to answer these questions, by assisting companies in the maturity assessment process and identifying the required areas for improvement. The HOFs maturity model describes the main characteristics allowing a safe human contribution and optimizing the overall performance of the system, it enables the company to determine the maturity in terms of HOFs among a set of suggested levels.

This paper is an extension of work originally presented in "2020 IEEE 6th International Conference on Optimization and Applications" [1], in which a new HOFs maturity model is described. The model is made up of the five crucial factors related to HOFs and proposed five maturity levels that reflect the degree of consideration and integration of these factors through planned procedures.

The difference between the conference paper [1] and this article is the inclusion of fuzzy logic in the weighting calculation phase, by using the Fuzzy AHP method instead of the Analytic Hierarchy Process (AHP) to take into consideration the imprecision and vagueness of human judgements , when assessing the relative importance of the hierarchy elements in order to perform the comparison matrices . Therefore, the proposed methodology consists in combining the FAHP method to

*Corresponding Author: Yousra Karim, yousrakarim@research.emi.ac.ma

determine the weights and the FCEM to assess the HOFs maturity level.

The proposed maturity model is implemented in a Moroccan organization operating in the mining industry to bring out the current situation by measuring its own HOFs maturity level, then define future targets and establish an action plan including the elements to improve.

The next sections are structured as follows: The section "Literature Review" describes the frequently cited models. In section " HOFs Maturity Model" the elements and levels of the model are presented. While section "Methodology" explains the FAHP and FCEM used. A numeric application is given in "Case Study" section, before the "Conclusion".

2. Literature Review

Research indicates a significant growth in the use of maturity models for safety culture assessment and the integration of human and organizational factors, particularly in high-risk areas such as construction and the oil and gas industry. A maturity model describes the key or essential elements that should characterize an organization at a particular maturity level either in relation to safety culture or to HOFs.

The use of maturity models as an assessment tool can be attributed to two main sources, namely, the "Quality Management Maturity Grid (QMMG)" and Westrum's "Typology of Organisations". The QMMG was suggested in 1979, it identified the five stages that an organization goes through to achieve the maximum quality level in all aspects of quality management these are: uncertainty, awakening, enlightenment, wisdom and certainty [2].

Typology of organisational cultures was proposed in 1993. Table 1 describes the characteristics of the three organisational types suggested. To allow a meticulous classification, while increasing employees' accessibility by including familiar terms, the typology was extended to these five levels : Pathological, Reactive, Calculative, Proactive and Generative [3].

Table 1: Westrum's Typology of Organisations

Typology of organisation	Characteristics
Pathological	Information is hidden, messengers are "shot", responsibilities are shirked, bridging is discouraged, failure is covered up, new ideas are actively crushed.
Bureaucratic	Information may be ignored, messengers are tolerated, responsibility is compartmentalized, bridging is allowed but neglected, organisation is just and merciful, new ideas create problems.
Generative	Information is actively sought, messengers are trained, responsibilities are shared, bridging is rewarded, failure causes inquiry, new ideas are welcomed.

In [4], the author used this concept of maturity to create a model named "safety culture maturity model (SCCM)" to help oil companies in the United Kingdom to assess culture maturity level and the actions needed to improve it. The five stages of the SCCM are shown in Figure 1.

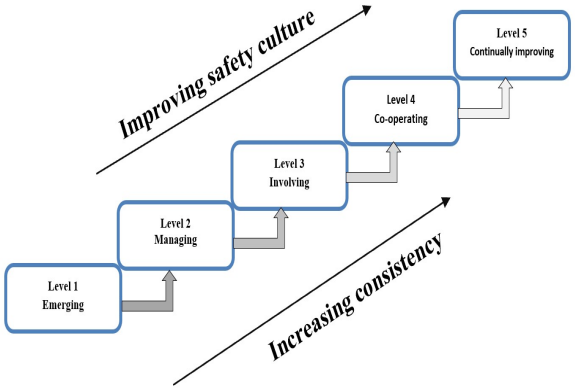


Figure 1: Safety Culture Maturity Model

Based on the SCCM and a thorough literature review of existing models, the author developed in [5] the "Cultural Maturity Model (CMM)" for companies that have an adequate safety management system and wish to evaluate and improve it. Table 2 summarizes the ten elements of the CMM selected to measure key characteristics related to a positive safety culture, and the five developmental stages suggested.

Table 2: Cultural Maturity Model Elements and Stages

Five stages	Ten elements to measure safety culture
1. Documenting 2. Controlling 3. Engaging 4. Participating 5. Institutionalising	1. Visibility of management commitment 2. Supervisor visible commitment 3. Production pressures 4. Organizational learning 5. Job and safety communication 6. Human and physical resources 7. Rules and procedures 8. Trust levels 9. Training 10. Workforce involvement

In [6, 7], the author proposed a maturity model, based on the extended Westrum's typology, to describe the evolution of safety culture from the disease stage to a generative end stage. Figure 2 shows the development stages of the model.

From the previous model, a framework was developed to determine the stages of maturity of an organization's safety culture by conducting in-depth interviews with 26 senior oil industry executives working in various multinational and contracting companies [8,9]. It contains 11 tangible or concrete elements related to the safety management system, and 7 less tangible or abstract elements involving perceptions of the workforce. The eighteen elements are divided as follows:

Tangible or concrete elements: 1. Benchmarking, Trends & Statistics; 2. Audits & Reviews; 3. Incident/Accident Reporting, Investigation & Analysis; 4. Hazard and Unsafe Act reports; 5. Work planning including PTW, Journey Management; 6. Contractor Management; 7. Competency/Training; 8. Work-site Job Safety Techniques; 9. Who Checks Safety on a day to day

basis? 10. What is size & status of the HSE Department? 11. What are the rewards of good safety performance?

Less tangible or abstract elements: 1. Who causes accidents in the eyes of management? 2.What happens after an accident? Is the feedback Loop being closed? 3. How do safety meetings feel? 4. Balance between HSE & Profitability? 5. Is management interested in communicating HSE issues with the workforce? 6. Commitment level of the workforce and level of care for colleagues. 7. What is the purpose of procedures?

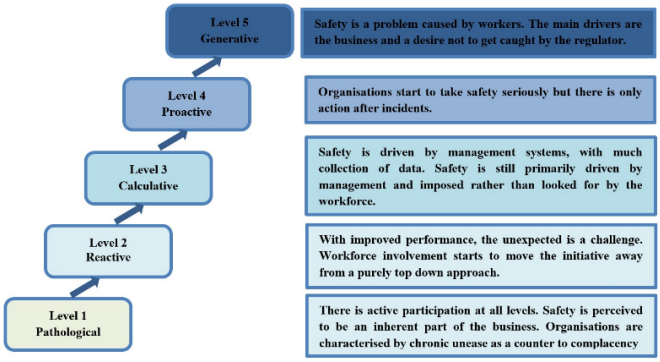


Figure 2: Safety Culture Model of Hudson

Maturity models have also been used to measure the level of integration of human factors, as they are essential to achieve safe human performance. The " Human Factors Maturity Model (HFMM)" was designed to help companies wishing to improve their maturity level by determining which items require additional effort in the future [10]. It includes 5 levels from the first Emerging level to a final level called Leading, and 12 key elements of human factors that highlight their relevance to major accidents in the literature. For each item, a card sorting methodology is used to assess the current level of maturity by representatives of organizations, then define the requirements needed to improve maturity to the next level. The elements and levels of the HFMM are shown in Table 3.

Maturity models have been used by several industries in the literature following different measurement methods, which often require precise human judgment to choose one of the maturity levels. And since human judgment is imprecise and vague, the model proposed in this article includes fuzzy logic to remedy this issue [11].

Table 3: Human Factors Maturity Model

HFMM Levels	HFMM Elements
1. Emerging 2. Transitional 3. Planned 4. Proactive 5. Leading	1. Managing Human Failure (including maintenance) 2. Human Factors in Incident Investigation 3. Design and Development of Procedures 4. Training and Competence 5. Staffing and Workload 6. Managing Organisational Change 7. Safety-Critical Communications 8. Human Factors in Design 9. Fatigue and Shift work 10. Safety Culture and Behaviour 11. Contractor Management 12. Managing Performance under Pressure

3. HOFs Maturity Model

A review of the literature was carried out to select the human and organizational factors that affect human performance, in order to reduce accidents and improve the safety level of organizations. The maturity model presented in this article is developed based on the five key factors [1] cited by the Rail & Safety Standards Board (RSSB) in their guide "Understanding human factors" [12] namely:

- Design F1: The consideration of HOFs in the design is a crucial step to ensure that the organization does not miss major opportunities for improving human performance at work and to eliminate the risk of accidents, loss of personnel and significant financial costs.
- Training F2: Numerous studies demonstrate the effectiveness of training in business growth. The results of training are concrete and measurable both in productivity and in the professional development of the staff. Although a company may have carefully screened its employees, training within the company itself will allow them to have the same vision of things and move in the same direction.
- Staffing F3: Recruiting and retaining the right people is critical to the success of every organization. Selecting the wrong skills can lead to organizational weakness.
- Culture F4: Culture is essential to allow the organization to differentiate itself. In terms of image on the one hand, it has advantages both internally and externally to consumers. It is indeed a source of cohesion and motivation of employees and it limits conflicts.
- Conditions F5: The improvement of working conditions to ensure the well-being of the employee must first affect health and safety. All employees wish to carry out their work in a healthy and pleasant atmosphere, in which they can feel the recognition and consideration of their efforts.

Table 4 lists the sub-factors linked to each of the five factors, which allow to determine the maturity level in terms of HOFs among the five presented in Figure 3. [1]

Table 4: HOFs Maturity Model

Goal: Human Performance and Safety	Human and Organizational Factors	Sub-Factors
	1. Design F1	Equipment Design F11 Workplace Design F12 Job Design F13
	2. Training F2	Effective Training Programme F21 Training Appraisal F22
	3. Staffing F3	Recruitment F31 Retention F32
	4. Culture F4	LeadershipF41 Management F42 Teamworking F43 Communication F44 Change F45
	5. Conditions F5	Morale & Motivation F51 Stress F52 Workload F53 Shift Work F53

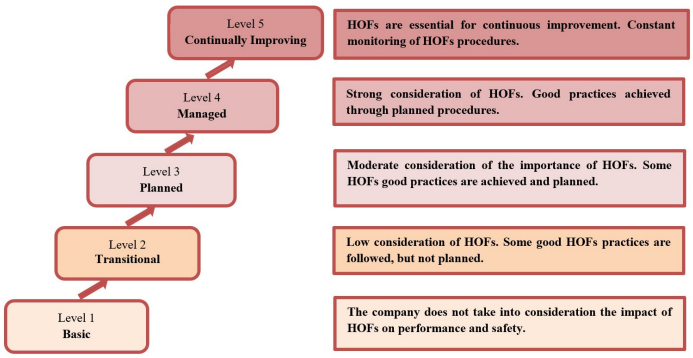


Figure 3: Maturity Levels Description

To ensure the understanding and appropriate application of the model, a set of good practices related to each sub-factor have been listed. For the retention sub-factor, as an example, the best practices relevant to employee satisfaction and retention are presented below:

- Offer rewards.
- Listen to your staff.
- Carry out training.
- Improve your team culture.
- Develop career paths.
- Understand why people quit.
- Maintain dialogue with staff

Table 5 provides an overview of the questionnaire developed to facilitate data collection during the maturity assessment. The questions are established from the proposed model to bring out the information concerning the sub-factors.

4. Measurement Methodology

The methodology proposed in this work is based on the use of FAHP method instead of AHP to calculate the factors and sub-factors weightings, that reflect their impacts on human performance and safety. For the HOFs maturity level assessment, the FCEM method is utilized to consider the imprecision and uncertainty of decision-makers' judgment.

Below are explained the FAHP and FECM methods used for the model implementation.

4.1 Fuzzy AHP

Fuzzy AHP method is an extension of the AHP introduced in 1970 [13], used to solve complex decision-making problems. This method deals with the ambiguity and vagueness of decision-makers that cannot be addressed by precise values by integrating fuzzy logic [11].

Many extensions of AHP have been proposed by several authors such as the geometric mean method [14]. In this paper, the HOFs model utilized the extent analysis approach proposed in [15] given the simplicity and ease of application using triangular fuzzy numbers (TFNs) for the pairwise comparisons.

The steps of the FAHP method are [16, 17]:

Step 1: Construct the AHP hierarchy by identifying factors and sub-factors of the model.

Step 2: Compare by pair the factors and sub-factors to calculate the local weights using the fuzzy scale [18] shown in Table 6 and Figure 4. This scale is used in Chang's FAHP method described below.

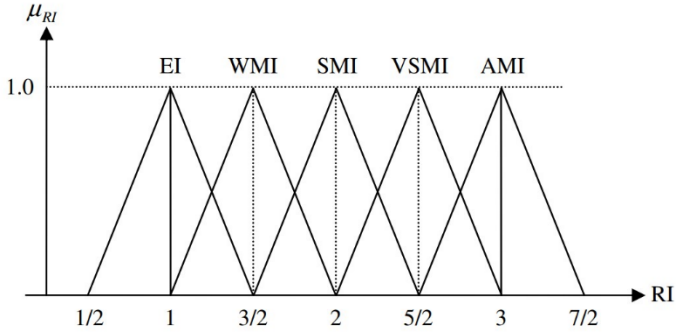


Figure 4: Linguistic Scale for Relative Importance

Assume that X = $\{x_1, x_2, \dots, x_n\}$ is an object set, and U = $\{u_1, u_2, \dots, u_m\}$ is a goal set. For each object, an extent analysis is performed for each goal g_i. Therefore, m values of the extent analysis are obtained for each object, as follows:

$$M_{g_i}^1, M_{g_i}^2, \dots, M_{g_i}^m \quad i=1, 2, \dots, n. \tag{1}$$

where $M_{g_i}^j$ (j=1, 2, … , m) are TFNs. A TFN is simply denoted (l,m,u).

1) The value of fuzzy synthetic *Si* for the *i*-th object is:

$$S_i = \sum_{j=1}^{n} M_{g_i}^j \times \left[\sum_{i=1}^{n}\sum_{j=1}^{m} M_{g_i}^j\right]^{-1} \tag{2}$$

The fuzzy addition operations are performed with m values.

$$\sum_{j=1}^{m} M_{g_i}^j = \left(\sum_{j=1}^{m} l_j, \sum_{j=1}^{m} m_j, \sum_{j=1}^{m} u_j\right) \tag{3}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} M_{g_i}^j = \left(\sum_{i=1}^{n} l_i, \sum_{i=1}^{n} m_i, \sum_{i=1}^{n} u_i\right) \tag{4}$$

Then, the inverse vector of (4) :

$$\left[\sum_{i=1}^{n}\sum_{j=1}^{m} M_{g_i}^j\right]^{-1} = \left[\frac{1}{\sum_{i=1}^{n} u_i}, \frac{1}{\sum_{i=1}^{n} m_i}, \frac{1}{\sum_{i=1}^{n} l_i}\right] \tag{5}$$

2) The degree of possibility of $S_2 = (l_2, m_2, u_2) \geq S_1 = (l_1, m_1, u_1)$ is defined as:

$$V(S_2 \geq S_1) = sup\left[\min\left(\mu_{S_1}(x), \mu_{S_2}(y)\right)\right] \tag{6}$$

Which can be expressed equivalently as follows:

$V(S_2 \geq S_1)$ = hgt $(S_2 \cap S_1)$=

Table 5: HOFs Questionnaire Overview

Questions	Answers	Questions	Answers
1. Is the design of the equipment meet user needs (visibility, workflow, constraints, environment, workload, etc.)?	1. All equipment takes into consideration the needs of the users. 2. An important part of the equipment meets the needs of users. 3. Only critical equipment considers the needs of users. 4. Some critical equipment takes user needs into account. 5. No equipment follows user needs.	7. How do you find the efforts made by the company to keep the people recruited? 8. How do you rate the involvement of managers and leaders in the HOFs procedures? 9. How do you perceive the involvement and commitment of employees in change projects?	1. Excellent. 2. Good. 3. Medium. 4. Basic. 5. No efforts. 1. Extreme involvement. 2. Good involvement. 3. Medium involvement 4. Basic involvement. 5. No involvement.
2. How do you assess the design of the workplace? 3. How do you perceive the job design (job descriptions) produced by the company?	1. Excellent design (The company is always looking for ways to improve). 2. Good design. 3. Medium design. 4. Basic. 5. Poor (The company is not aware of the impact of good design).	10. How do you perceive the supervision of teamwork within your company? 11. How do you rate the quality of communication?	1. Extremely supervised. 2. Good supervision. 3. Medium supervision. 4. Basic supervision 5. No supervision. 1. Excellent. 2. Good. 3. Medium. 4. Basic. 5. Insufficient.
4. How do you rate the effectiveness of the training programs? 5.How do you perceive the selection and recruitment process?	1. Extremely effective. 2. Effective 3. Medium effectiveness. 4. Low effectiveness. 5. Not at all effective.	12. How would you rate the practices put in place by the company to improve morale and motivate employees at work? 13. How do you see the strategy adopted by the company to manage stress?	1. Excellent. 2. Good. 3. Medium. 4. Basic. 5. Insufficient.
6. How do you perceive the process of assessing the trainings carried out by the company (reaction of trainees, learning objectives…)?	1. Excellent assessment 2. Good assessment. 3. Medium assessment. 4. Baseline assessment. 5. Insufficient assessment.	14. How do you perceive the workload? 15. How do you perceive the work shift planning ?	1. In standards 2. Normal but disrupted on some occasions. 3. Moderately excessive. 4. Excessive . 5. Extremely excessive.

Table 6: Linguistic Scale for Relative Importance

Linguistic scale for relative importance	Triangular Fuzzy Scale
Just Equal	(1, 1, 1)
Equally Important (EI)	(1/2, 1, 3/2)
Weakly more important (WMI)	(1, 3/2, 2)
Strongly more important (SMI)	(3/2, 2, 5/2)
Very strongly more important (VSMI)	(2, 5/2, 3)
Absolutely more important (AMI)	(5/2, 3, 7/2)

$$\mu_{S_2}(d)=\begin{cases}1 \ , if \ \ m_2 \geq m_1 \\ 0 \ , \quad if \ l_1 \geq u_2 \\ \frac{l_1-u_2}{(m_2-u_2)-(m_1-l_1)} \ , otherwise\end{cases} \quad (7)$$

where d is the ordinate of the highest intersection point D between μ_{S_1} and μ_{S_2}(Figure 5).

Both the values of V($S_2 \geq S_1$) and V($S_1 \geq S_2$) are needed to compare S_1 and S_2.

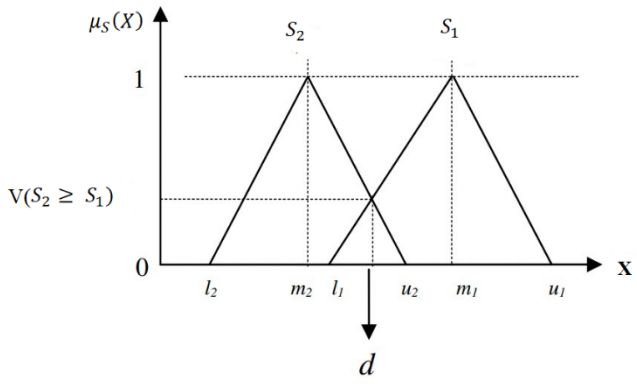


Figure 5: Intersection Between S_1 and S_2

3) The degree possibility for a convex fuzzy number to be greater than k convex fuzzy numbers S_i (i =1,. . . , k) can be expressed by:

$V(S \geq S_1, S_2, \dots, S_k) = V[(S \geq S_1), \dots, (S \geq S_k)] =$ $\min V(S \geq S_i)$, i=1, …, k. (8)

Assume that $d'(A_i) = \min V(S_i \geq S_k)$, (9)

for k =1, …, n , $k \neq i$. Then the weight vector W' is given by:

$W' = (d'(A_1), d'(A_2), \dots, d'(A_n))^T$, (10)

where $A_i (i = 1, 2, \dots, n)$ are n elements.

4) Finally, the normalized weight vectors are obtained:

$W = (d(A_1), d(A_2), \dots, d(A_n))^T$, (11)

where W is a crisp number.

4.2 Fuzzy Comprehensive Evaluation Method

The fuzzy comprehensive evaluation method is based on fuzzy logic theory developed in [11]. Unlike the system of classical logic, it aims to deal with uncertainty, subjectivity, and vagueness of human reasoning. The FCEM principles are used with the proposed model to assess the maturity level of HOFs as follows [1, 19, 20, 21, 22]:

Let V={$V_1, V_2, \dots, V_p$}, the evaluation set.

Step 1: Construct the evaluation matrix R_{ij} named the second-class index membership matrix.

$$R_{ij} = \begin{pmatrix} R_{i1} \\ \vdots \\ R_{ij} \\ \vdots \\ R_{im} \end{pmatrix} = \begin{bmatrix} r_{i11} & \cdots & r_{i1p} \\ \vdots & & \vdots \\ r_{im1} & \cdots & r_{imp} \end{bmatrix} \quad (12)$$

where

R_{ij} is a fuzzy relationship from F_i's sub-factors (F_{ij}) to V.

Where: i=1, …, n is the number of factors to be evaluated. j=1, …, m is the second index of i, and m is the number of sub-factors(F_{ij}) of a factor (F_i).

R_{ij} is a fuzzy relationship from F_{ij} to V.

$$(r_{ij1} \quad \cdots \; r_{ijk} \cdots \quad r_{ijp}) = (l_{ij1}/\beta \quad \cdots \; l_{ijk}/\beta \cdots \quad l_{ijp}/\beta) \quad (13)$$

$\beta = \sum_{k=1}^{k=p} l_{ijk}$ is the experts' number, k=1, …, p is the evaluation level.

Step 2: Compute the matrix R named the first-class index membership matrix.

$$R = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_4 \\ R_n \end{pmatrix} = \begin{bmatrix} r_{11} & \cdots & r_{1p} \\ \vdots & r_{ik} & \vdots \\ r_{n1} & \cdots & r_{np} \end{bmatrix} \quad (14)$$

where $R_i = W_{ij} \circ R_{ij} = (r_{i1} \quad \cdots \quad r_{ip})$ (15)

W_{ij} is the weight vector obtained from the FAHP method.

Step 3: Calculate the maturity vector M:

$$M = W\,R = (r_1 \quad \cdots \quad r_p) \quad (16)$$

$W = (W_1 \quad \cdots \; W_i \cdots \quad W_n)$ with W_i is the weight F_i.

Finally, the maturity level is determined according to the principle of maximum membership degree law as following:

If $r_k = \text{Max}\,(r_1 \quad \cdots \quad r_p)$, the maturity level is k.

5. Case study

The HOFs maturity model and the methodologies developed in this paper are implemented in a Moroccan organization operating in the mining industry to define potential future steps essential to improve human performance and safety level. Therefore, a group of decision makers was selected for the acquisition of input data and the application of the FAHP and FCEM methods. The comparisons matrices were done by the representatives of managers, supervisors and operators. Unlike the assessment step, which was carried out separately, first with the Group 1 of managers, next with the Group 2 of supervisors and site operators.

5.1. Results

The representatives performed the pairwise comparisons of factors and sub-factors listed in Table 4 using the fuzzy scale illustrated in figure 6 [18], and obtained the results shown in Tables 7, 8, 9, 10, 11, 12.

Table 7: Weights and Comparison Matrix of Factors

Goal	F1	F2	F3	F4	F5	Weights
F1	(1, 1,1)	(3/2, 2, 5/2)	(1/2, 1, 3/2)	(1/2, 2/3, 1)	(2/5, 1/2, 2/3)	**0.17**
F2	(2/5,1/2,2/3)	(1, 1, 1)	(2/3, 1, 2)	(2/5,1/2,2/3)	(2/5,1/2,2/3)	**0.07**
F3	(2/3, 1, 2)	(1/2, 1, 3/2)	(1, 1, 1)	(2/5,1/2,2/3)	(1/3, 2/5, 1/2)	**0.10**
F4	(1, 3/2, 2)	(3/2, 2, 5/2)	(3/2, 2, 5/2)	(1, 1, 1)	(1/2, 2/3, 1)	**0.29**
F5	(3/2, 2, 5/2)	(3/2, 2, 5/2)	(2, 5/2, 3)	(1, 3/2, 2)	(1, 1, 1)	**0.37**

Factors and sub-factors weights are computed using Chang's FAHP method [15] described above. A numerical overview of the steps is given below to calculate weights shown on the right of Table 7.

1) S_{F1}= (0.10, 0.18, 0.30) ; S_{F2}= (0.08, 0.12, 0.23) ; S_{F3}= (0.08, 0.14, 0.26) ; S_{F4}= (0.15, 0.25, 0.41) ; S_{F5}= (0.19, 0.31, 0.50).

2) V ($S_{F1} \geq S_{F2}$) = 1; V ($S_{F1} \geq S_{F3}$) = 1; V ($S_{F1} \geq S_{F4}$) = 0.69; V ($S_{F1} \geq S_{F5}$) = 0.46.
V ($S_{F2} \geq S_{F1}$) = 0.68; V ($S_{F2} \geq S_{F3}$) = 0.91 ; V ($S_{F2} \geq S_{F4}$) = 0.38 ; V ($S_{F2} \geq S_{F5}$) = 0.17.
V ($S_{F3} \geq S_{F1}$) = 0.77; V ($S_{F3} \geq S_{F2}$) = 1 ; V ($S_{F3} \geq S_{F4}$) = 0.49 ; V ($S_{F3} \geq S_{F5}$) = 0.28.
V ($S_{F4} \geq S_{F1}$) = 1 ; V ($S_{F4} \geq S_{F2}$) = 1 ; V ($S_{F4} \geq S_{F3}$) = 1 ; V ($S_{F4} \geq S_{F5}$) =0.77.
V ($S_{F5} \geq S_{F1}$) = 1; V ($S_{F5} \geq S_{F2}$) = 1 ; V ($S_{F5} \geq S_{F3}$) = 1 ; V ($S_{F5} \geq S_{F4}$) =1.

3) d'(A_{F1}) = V ($S_{F1} \geq S_{F2}$, S_{F3}, S_{F4}, S_{F5}) = 0.46 ;
d'(A_{F2}) = V ($S_{F2} \geq S_{F1}$, S_{F3}, S_{F4}, S_{F5}) = 0.17 ;
d'(A_{F3}) = V ($S_{F3} \geq S_{F1}$, S_{F2}, S_{F4}, S_{F5}) = 0.28 ;
d'(A_{F4}) = V ($S_{F4} \geq S_{F1}$, S_{F2}, S_{F3}, S_{F5}) =0.77 ;
d'(A_{F5}) = V ($S_{F5} \geq S_{F1}$, S_{F2}, S_{F3}, S_{F4}) =1 .

4) Then, the weight vector of factors is obtained:
W = (d'(A_{F1}), …, d'(A_{F5}))^T= (0.17, 0.06, 0.10, 0.29, 0.37).

Table 8: Weights and Comparison Matrix of F1 Sub-Factors

F1	F11	F12	F13	Weights
F11	(1, 1, 1)	(1, 3/2, 2)	(1, 3/2, 2)	**0.45**
F12	(1/2, 2/3, 1)	(1, 1, 1)	(3/2, 2, 5/2)	**0.40**
F13	(1/2, 2/3, 1)	(2/5, 1/2, 2/3)	(1, 1, 1)	**0.15**

Table 9: Weights and Comparison Matrix of F2 Sub-Factors

F2	F21	F22	Weights
F21	(1, 1, 1)	(1, 3/2, 2)	**0.68**
F22	(1/2, 2/3, 1)	(1, 1, 1)	**0.32**

Table 10: Weights and Comparison Matrix of F3 Sub-Factors

F3	F31	F32	Weights
F31	(1, 1, 1)	(1/2, 1, 3/2)	**0.50**
F32	(2/3, 1, 2)	(1, 1, 1)	**0.50**

Table 11: Weights and Comparison Matrix of F4 Sub-Factors

F4	F41	F42	F43	F44	F45	Weights
F41	(1, 1,1)	(1, 3/2, 2)	(1/2, 1, 3/2)	(3/2, 2, 5/2)	(2, 5/2, 3)	**0.32**
F42	(1/2, 2/3, 1)	(1, 1, 1)	(2/3, 1, 2)	(1, 3/2, 2)	(3/2, 2, 5/2)	**0.25**
F43	(2/3, 1, 2)	(1/2, 1, 3/2)	(1, 1, 1)	(1, 3/2, 2)	(3/2, 2, 5/2)	**0.26**
F44	(2/5, 1/2, 2/3)	(1/2, 2/3, 1)	(1/2, 2/3, 1)	(1, 1, 1)	(1, 3/2, 2)	**0.14**
F45	(1/3, 2/5, 1/2)	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1/2, 2/3, 1)	(1, 1, 1)	**0.03**

Table 12: Weights and Comparison Matrix of F5 Sub-Factors

F5	F51	F52	F53	F54	Weights
F51	(1, 1,1)	(1/2, 2/3, 1)	(2/5, 1/2, 2/3)	(2/5, 1/2, 2/3)	**0.10**
F52	(1, 3/2, 2)	(1, 1, 1)	(1/2, 2/3, 1)	(1/2, 2/3, 1)	**0.23**
F53	(3/2, 2, 5/2)	(1, 3/2, 2)	(1, 1, 1)	(1/2, 1, 3/2)	**0.34**
F54	(3/2, 2, 5/2)	(1, 3/2, 2)	(2/3, 1, 2)	(1, 1, 1)	**0.34**

Once the weights were calculated, an evaluation of the sub-factors was conducted by the two groups, separately, using the questionnaire and FCEM. The results obtained are shown in Tables 13, 14.

Table 13: FAHP Weights and Evaluation Results of Group1

Factors	Sub-Factors	Weights	Evaluation results				
			Basic (B)	Transitional (T)	Planned (P)	Managed (M)	Continually improving (CI)
F1	F11	**0.45**	0.00	0.00	0.25	0.50	0.25
	F12	**0.40**	0.00	0.00	0.58	0.42	0.00
	F13	**0.15**	0.00	0.00	0.50	0.33	0.17
F2	F21	**0.68**	0.00	0.00	0.33	0.58	0.08
	F22	**0.32**	0.08	0.42	0.42	0.08	0.00
F3	F31	**0.50**	0.00	0.00	0.00	0.58	0.42
	F32	**0.50**	0.00	0.42	0.42	0.17	0.00
F4	F41	**0.32**	0.00	0.00	0.17	0.33	0.50
	F42	**0.25**	0.00	0.00	0.08	0.50	0.42
	F43	**0.26**	0.00	0.00	0.50	0.50	0.00
	F44	**0.14**	0.00	0.08	0.42	0.42	0.08
	F45	**0.03**	0.00	0.08	0.33	0.58	0.00

F5	F51	**0.10**	0.00	0.58	0.33	0.08	0.00
	F52	**0.23**	0.17	0.58	0.25	0.00	0.00
	F53	**0.34**	0.00	0.00	0.92	0.08	0.00
	F54	**0.34**	0.00	0.00	0.25	0.75	0.00

Table 14: FAHP Weights and Evaluation Results of Group 2

Factors	Sub-Factors	Weights	Evaluation results				
			Basic (B)	Transitional (T)	Planned (P)	Managed (M)	Continually improving (CI)
F1	F11	**0.45**	0.00	0.00	0.50	0.33	0.17
	F12	**0.40**	0.00	0.33	0.42	0.25	0.00
	F13	**0.15**	0.00	0.00	0.67	0.33	0.00
F2	F21	**0.68**	0.00	0.08	0.25	0.50	0.17
	F22	**0.32**	0.17	0.58	0.25	0.00	0.00
F3	F31	**0.50**	0.00	0.00	0.17	0.67	0.17
	F32	**0.50**	0.08	0.42	0.33	0.17	0.00
F4	F41	**0.32**	0.00	0.00	0.17	0.67	0.17
	F42	**0.25**	0.00	0.08	0.08	0.67	0.17
	F43	**0.26**	0.00	0.33	0.58	0.08	0.00
	F44	**0.14**	0.00	0.25	0.58	0.17	0.00
	F45	**0.03**	0.17	0.50	0.17	0.17	0.00
F5	F51	**0.10**	0.33	0.50	0.08	0.08	0.00
	F52	**0.23**	0.17	0.58	0.08	0.17	0.00
	F53	**0.34**	0.00	0.17	0.67	0.17	0.00
	F54	**0.34**	0.00	0.00	0.58	0.17	0.25

The sub-factors' weights and the assessment results are subsequently used to determine the first-index membership matrix of the two groups represented in Tables 15, 16.

Table 15: First-Class Index Membership Matrix of Group 1

Factors	Weights	Levels				
		B	T	P	M	CI
F1	**0.17**	0.00	0.00	0.42	0.44	0.14
F2	**0.07**	0.03	0.13	0.36	0.42	0.06
F3	**0.10**	0.00	0.20	0.21	0.38	0.21
F4	**0.29**	0.00	0.01	0.27	0.44	0.28
F5	**0.37**	0.04	0.19	0.48	0.29	0.00

Table 16: First-Class Index Membership Matrix of Group 2

Factors	Weights	Levels				
		B	T	P	M	CI
F1	**0.17**	0.00	0.13	0.49	0.30	0.08
F2	**0.07**	0.05	0.24	0.25	0.34	0.12
F3	**0.10**	0.04	0.21	0.25	0.42	0.08
F4	**0.29**	0.00	0.16	0.31	0.43	0.10
F5	**0.37**	0.07	0.24	0.45	0.16	0.08

Finally, the HOFs maturity vector (HOFM) is determined using Tables 15, 16 and factors' weights.

HOMF1(Group 1) = (0.02, 0.10, 0.37, 0.38, 0.13)

HOMF2(Group 2) = (0.03, 0.16, 0.37, 0.33, 0.11)

5.2. Discussions

According to the maximum membership degree law and maturity vectors (HOMF1, HOMF2) obtained from both groups, the maturity level of the company is "Managed" for Group 1 and "Planned" for Group 2. Therefore, the managers assume that there is a strong consideration of HOFs, and good practices are achieved through planned procedures. While supervisors and operators suppose that the impact of HOFs on safety and human performance is moderately taken into account, and some good practices related to HOFs are fulfilled in a planned manner.

The maturity levels of factors are deducted from the first-class index membership matrices (Tables 15, 16). Then, the results are presented on spider diagrams below. Figure 6 shows the factors' maturity levels obtained from the assessment carried out with Group 1 of managers, and Figure 7 shows the levels obtained from the evaluation undertaken by Group 2 consisting of supervisors and operators.

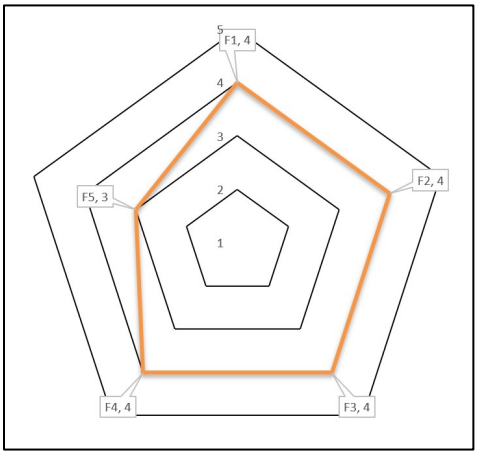


Figure 6: Spider Diagram of Group 1

The two diagrams show that no factor is judged at level 5 "Continually Improving". The results obtained are close for both groups despite the hierarchical diversity. The three factors:

"Staffing", "Training" and "Culture" are at the same maturity level 4 "Managed". The factor "Conditions" is also at level 3 "Planned" for the two groups. However, the "Design" factor is at level 4 "Managed" for Group 1 and level 3 "Planned" for group 2.

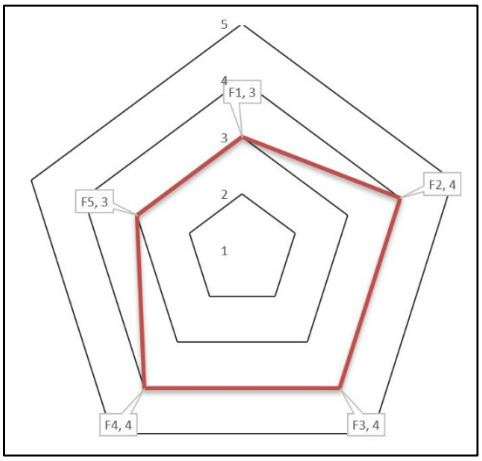


Figure 7: Spider Diagram of Group 2

The purpose of this model is not only to assess maturity in terms of HOF, but also to assist the company in defining future improvement steps to reach the next level of maturity. For this mining organization, it is necessary to provide more effort regarding the "Conditions" factor since it has the highest weighting, and it is evaluated at level 3 for both groups. Then, for the factor "Design" which has such an important impact, it is recommended to take into consideration the opinions of supervisors and operators to improve the current state and satisfy their needs.

After this study, an action plan made up of some good practices essential for improvement was established based on elements weights and the assessment carried out by the groups. Table 17 shows a summary of the proposed action plan.

6. Conclusion

There is a remarkable growth in the use of maturity models to assess safety culture and HOFs, but also different variations in the elements of these models and measurement methodology. These models can be used for maturity assessment in any industry and as an improvement tool by identifying strengths and weaknesses related to a set of key elements.

This work introduces a maturity model related to HOFs made up of the five factors essential to human performance and safety. The model defines five stages of maturity, from the lower level where the company does not take into consideration the impact of HOFs, to the upper level characterized by continuous enhancement and permanent follow-up of HOFs procedures. The measurement methodology differs here in the first phase, which consists of using Fuzzy AHP method instead of AHP to calculate the weights. And for the second phase, a questionnaire is included to facilitate the collection of data used in the FCEM method.

The HOFs model is applied in a Moroccan mining company to determine the current maturity level and improve it through an appropriate action plan. Teams of managers, supervisors and operators who participated in implementing the model have used it successfully, and it has proven to be a useful tool for evaluation an improvement.

Future work will focus on improving the current HOFs maturity model by testing and adapting it for a wide range of companies, so that it is not limited to a particular area. Also, conducting a sensitivity analysis [23] to test the robustness of the model and better understand relationships between the elements.

Table 17: Improvement Action Plan Related to HOFs

Element	Good practices
Morale and Motivation	• Show them the results of their efforts. • Give them responsibilities. • Ask them what motivates them. • Give them some advantages. • Offer them training courses.
Stress	• Stress management training: Ask employees to take relaxation and time management courses and take internships or do assertiveness exercises. • Ergonomics and design of the professional environment: Improve the equipment used at work and the physical working conditions. • Improve management: Improve the attitude of managers towards work stress, their knowledge and understanding of this problem and their ability to tackle it as effectively as possible. • Company development: Implement better work and management systems. Develop a friendlier corporate culture and a spirit of mutual aid.
Workload	• Identify your team's workload and capacity. • Allocate resources and manage individual workloads. • Check in with your team and adjust workloads if necessary. • Improve efficiency under over workloads. • Adopt a work management tool
Workplace Design	In order to design a workplace that allows the employee to obtain optimal working conditions, the following steps must be taken : • A preliminary evaluation. • Employee participation. • Employee training.

	• A detailed report of observations. • A follow-up visit by the processionals to validate the effectiveness of the changes made to the workstations and the maintenance of the new work habits.
Job Design	For a better job desgin, several fundamental elements must be taken into account by the employer : • The job identification includes information such as the job title and the information that characterizes it: working time, statutory conditions, or the service to which the job is attached. • The hierarchical and functional links, which position the employee in the company, define his level of responsibility and specify his hierarchical relations. • The job description lists the main and secondary activities of the job, its methods of exercise and the professional risks incurred. • The location of work tools…

References

[1] Y. Karim, A. Cherkaoui, "Human and organizational factors maturity model development and implementation in construction industry using fuzzy comprehensive evaluation method", in 2020 IEEE 6th International Conference on Optimization and Applications (ICOA), 1-6, IEEE, 2020, doi: 10.1109/ICOA49421.2020.9094481.

[2] P. Fraser, J. Moultrie, M. Gregory, "The use of maturity models/grids as a tool in assessing product development capability", in IEEE International Engineering Management Conference, IEEE, 1, 244-249, August 2002, doi: 10.1109/IEMC.2002.1038431.

[3] Parker, M. Lawrie, P. Hudson, "A framework for understanding the development of organisational safety culture", Safety Science, 44, 551-562, 2006, doi: 10.1016/j.ssc²i.2005.10.004

[4] M. Fleming, "Safety culture maturity model", Offshore Technology Report-Health and Safety Executive OTH, 2001.

[5] M. Fleming, S. Meakin, "Cultural maturity model: health and safety improvement through involvement", in SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production, Society of Petroleum Engineers, 2004, doi: 10.2118/86623-MS.

[6] P. T. W. Hudson, F. C. Willekes, "The hearts and minds project in an operating company: developing tools to measure cultural factors", in SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production, Society of Petroleum Engineers, 2000, doi: 10.2118/61228-MS.

[7] P. T. W. Hudson, "Implementing a safety culture in a major multi-national", Safety science, **45**(6), 697-722, 2007, doi: 10.1016/j.ssci.2007.04.005.

[8] P. T. W. Hudson, D. Parker, R. Lawton, W.L.G. Verschuur, G.C.Van der Graaf, J. Kalff, "The hearts and minds project: creating intrinsic motivation for HSE", in Proceeding of the SPE International Conference on Health, Safety, and the Environment in Oil and Gas Exploration and Production Stavanger, 2000, doi: 10.2118/61095-MS.

[9] M. Lawrie, D. Parker, , P. T. W. Hudson, "Investigating employee perceptions of a framework of safety culture maturity", Safety Science, **44**(3), 259-276, 2006, doi: 10.1016/j.ssci.2005.10.003.

[10] J. Edmonds, K. Gray, "Assessing Human Factors Maturity", Chemical Engineering Transactions, 77, 481-486, 2019, doi: 10.3303/CET1977081.

[11] L. A. Zadeh, "Fuzzy logic", Computer, **21**(4), 83-93, 1988, 10.1109/2.53.

[12] RSSB, "Understanding human factors, a guide for the railway industry", 2008.

[13] T. L. Saaty, "How to make a decision: the analytic hierarchy process", Interfaces, 24, 19-43, 1994, doi: 10.1287/inte.24.6.19.

[14] C. H. Chou, G. S. Liang, H. C. Chang, "A fuzzy AHP approach based on the concept of possibility extent", Quality & Quantity, **47**(1), 1-14, 2013, doi: 10.1007/s11135-011-9473-6.

[15] D. Y. Chang, "Applications of the extent analysis method on fuzzy AHP", European journal of operational research, **95**(3), 649-655, 1996.

[16] M. Dağdeviren, İ. Yüksel, "Developing a fuzzy analytic hierarchy process (AHP) model for behavior-based safety management", Information sciences, **178**(6), 1717-1733, doi: 10.1016/j.ins.2007.10.016

[17] E. K. Zavadskas, Z. Turskis, Ž. Stević, A. Mardani, "Modelling procedure for the selection of steel pipes supplier by applying fuzzy AHP method", Operational Research in Engineering Sciences: Theory and Applications, **3**(2), 39-53, 2020.

[18] C. Kahraman, T. Ertay, G. Büyüközkan, G, "A fuzzy optimization model for QFD planning process using analytic network approach", European journal of operational research, **171**(2), 390-411, 2006.

[19] T. J. Zhang, J. H. Ren, S. H. Yu, W. Cui, "Entropy weight-fuzzy comprehensive evaluation method of the safety evaluation of water inrush", Advanced Materials Research, 868, 300-305, 2014.doi: 10.4028/www.scientific.net/AMR.868.300

[20] H. Huang, "The application of fuzzy comprehensive evaluation to risk assessment of coal floor water irruption", in International Conference on Advances in Energy, Environment and Chemical Science, Atlantis Press, 2016, https://dx.doi.org/10.2991/aeecs-16.2016.4.

[21] H. Gu, , X. Fu, Y. Zhu, Y. Chen, L. Huang, "Analysis of social and environmental impact of earth-rock dam breaks based on a fuzzy comprehensive evaluation method. Sustainability", **12**(15), 6239, 2020, doi: 10.3390/su12156239.

[22] X. Wu, F. Hu, "Analysis of ecological carrying capacity using a fuzzy comprehensive evaluation method", Ecological Indicators, 113, 106243, 2020, doi: 10.1016/j.ecolind.2020.106243.

[23] X. Lai, S. Wang, S. Ma, J. Xie, Y. Zheng, "Parameter sensitivity analysis and simplification of equivalent circuit model for the state of charge of lithium-ion batteries", Electrochimica Acta, 330, 135239, 2020, doi: 10.1016/j.electacta.2019.135239.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 3, 84-91 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

SLIP-SL: Walking Control Based on an Extended SLIP Model with Swing Leg Dynamics

Junho Chang, Mustafa Melih Pelit*, Masaki Yamakita

Tokyo Institute of Technology, Systems and Control Engineering Department, 152-8552, Japan

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 12 April, 2021
Online: 12 May, 2021

Keywords:
Bipedal Walking
Optimization
Feedback Control

ABSTRACT

This paper details an extension to the SLIP model named spring loaded inverted pendulum model with swing legs (SLIP-SL). SLIP-SL extends the SLIP model by introducing swing leg dynamics while keeping its passive nature. This way, reference trajectories for the center of mass and swing foot trajectories can be simultaneously obtained which was not possible with the SLIP. This makes implementation easier and can increase tracking performance. We show how a variety of feasible two-phased walking trajectories can be obtained for this template model using direct collocation optimization methods. It is also shown through simulation studies that reference SLIP-SL trajectories can be used to control a fully actuated bipedal robot with the proposed feedback linearization controller to reach a stable cyclic gait.

1 Introduction

This paper is an extension of the work originally presented at *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). IEEE, 2020* [1].

All the environments that we live and work in are designed to be traversed by humans who have a bipedal gait. This makes the bipedal robots advantageous since they would easily adapt to these environments. Bipedal robots can move in discontinuous terrains (ex. stairs) and turn in narrow spots. These and many other reasons have driven researchers to study the bipedal gait as a locomotion method for robots.

There is a variety of methods that researchers can choose to control the gait of a bipedal robot. Using an inverted pendulum model as a template in conjunction with the zero moment point (ZMP) criterion has been used extensively [2], [3]. An up and coming method is obtaining optimal trajectories and inputs through various optimization methods and using them as a reference [4], [5]. Another popular method is to use simple models that can recreate certain fundamental aspects of human or animal gait, as template models for bipedal robots. A template model that is commonly used for this purpose is called the bipedal spring loaded inverted pendulum (SLIP) model [6], [7].

Bipedal SLIP model consists of two complaint legs and a point mass (Figure 2). This model is passive, i.e. there are no external inputs, and its motion is determined by the mechanical parameters and initial conditions. By choosing these carefully, a range of trajectories can be obtained that converge to a limit-cycle. Humans walk in a two phased manner and Bipedal SLIP model can mimic this gait. Human gait is explained with great detail in [8]. These two phases are called single stance phase and double stance phase. In the single stance phase, only one leg is in contact with the ground and the other leg is "swinging". And in the double stance phase, both feet are on the ground. However, SLIP model has one big assumption that differs substantially from a human gait: in the single stance phase, swing leg is assumed to move instantaneously to the proper

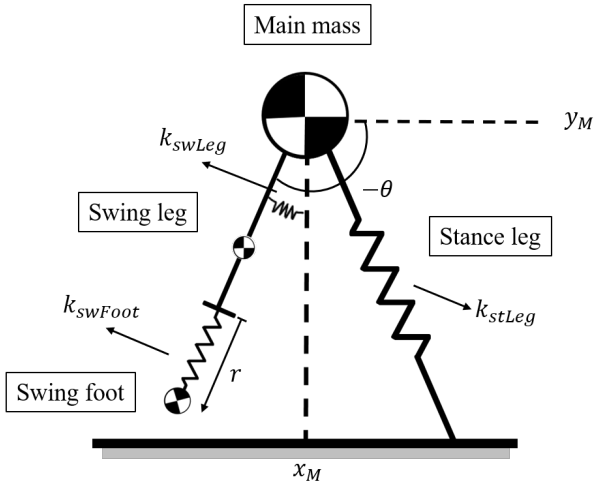


Figure 1: SLIP-SL Model in the single stance phase

*Corresponding Author: Mustafa Melih Pelit, Meguro-ku, Ookayama 2-12-1, South 5 Building, 305C, Tokyo, Japan. Email: pelit@ac.sc.e.titech.ac.jp

touch-down position. This means that when the bipedal SLIP is chosen as the template model for the controller of a bipedal robot, desired trajectories for the swing foot can not be obtained from it. Additional steps are necessary to achieve the swinging motion. [9] proposes a conceptual model where SLIP model is combined with a segmented leg so that the effects of swinging can be studied but reference trajectories and a controller is necessary to achieve the desired motion. In this paper, an extended SLIP model will be proposed so that reference center of mass trajectories and swing foot trajectories can be obtained simultaneously. This model will be called spring loaded inverted pendulum model with swing leg (SLIP-SL) which is shown in Figure 1.

SLIP-SL also has a two-phased gait and in the double stance phase, it is the same as the bipedal SLIP model. The difference is in the single stance phase where the model consists of three massed elements. The model also has three springs to facilitate the movement of these components. SLIP-SL model doesn't have any inputs, so it keeps the passive nature of the bipedal SLIP. The passive nature necessitates that proper spring parameters and initial conditions are chosen so that feasible trajectories can be realized. In this paper, direct collocation methods [10] were used to conduct simultaneous parameter and trajectory optimization for finding the suitable parameters. Then, a feedback linearization controller is proposed to track the obtained SLIP-SL trajectories with a 5-link fully actuated bipedal robot model. Effectiveness of the proposed controller and SLIP-SL's ability to be used as a template for walking are investigated through simulation studies.

This paper is organized as follows: Section 2 describes the dynamics of SLIP-SL and the bipedal robot model, Section 3 details the optimization work that is needed for finding feasible SLIP-SL trajectories, Section 4 introduces the proposed feedback linearization controller and in Section 5 simulation results are presented and discussed.

2 Systems and Modeling

This section will begin by introducing the bipedal SLIP model which is followed by the extended SLIP model named SLIP-SL and finally the model for the bipedal robot will be introduced. Explanation of the SLIP model will be brief since there are many works such as [6] that do an excellent job and going in depth on the matter. This paper will focus on the extended model and fully covers it but it is recommended to have a basic knowledge of the SLIP.

2.1 Bipedal SLIP Model

Bipedal SLIP Model can be seen in Figure 2. It consists of a point mass and two massless legs made out of springs. This model can mimic the two phased walking of humans, namely the single and the double stance phase. In the double stance phase, both feet are on the ground and both spring-like legs are pushing the mass. Then, a lift-off event happens where the foot in the back leaves the ground and the model goes into the single stance phase. This phase continues until the touch down event (when the swing foot touches the ground, depending on the angle of attack α and the free length of the spring L_0) at which point it goes back to the double stance phase. This continues in a cycle and walking motion is achieved. The springs are always in contraction and they are always pushing since an actual leg can not pull us towards the ground.

This system is passive so its motion is determined by its parameters such as such as spring stiffness, angle of attack and the initial conditions such as initial the velocity. Different types of gaits can be achieved with this model by changing the parameters and initial conditions so that a stable human-like gait with double peaked ground reaction forces can be achieved. This model has been very popular among researchers because of its simple nature and it has been extensively used to generate reference trajectories for center of mass (CoM) position. However, it has one significant assumption which is that the swing leg is assumed to move instantaneously to the proper position required for touch-down. This is possible since the springs are assumed to be massless. However, this is not the case for actual robots which must perform the swinging motion with actual massed legs so that they can walk. So, when the SLIP model is used as a reference, additional steps are necessary to generate the swing leg motion. Also, since SLIP model ignores the swinging motion, adding it later on might prove to be difficult and will act as a disturbance to the CoM trajectory. That is the advantage of the SLIP-SL model which we will introduce. It is an extension to the SLIP model to include swing leg dynamics in the single stance phase.

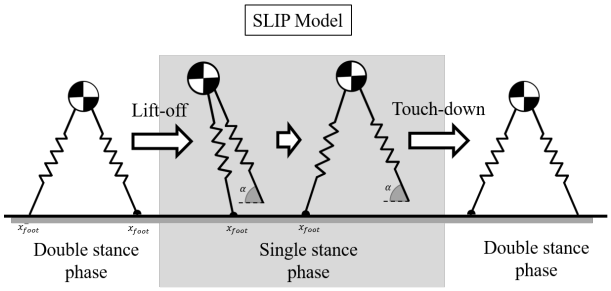


Figure 2: Bipedal SLIP Model

2.2 SLIP-SL model

Spring loaded inverted pendulum model with swing leg (SLIP-SL) can be seen in Figure 1 and its motion throughout a full step is represented in Figure 3. This model also walks in two phases, like its predecessor. In the double stance phase, SLIP-SL and SLIP are identical. What makes SLIP-SL different from the SLIP model is the addition of swing leg dynamics in the single stance phase. In the single stance phase, SLIP-SL consists of 3 massed elements which are the main mass '*M*', a massed swing leg and a point mass representing the swing foot and 3 springs which are the linear spring that connects '*M*' to the ground, the torsional spring that is connected between the point mass and the swing leg and the linear spring which connects the swing foot and the swing leg. After the touch-down event, SLIP-SL goes to the double stance phase, the swing elements disappear and the model becomes the same as the SLIP model in the respective phase. SLIP-SL consists of a point mass and two massless springs in the double stance phase.

In the single stance phase, equation of motion for the SLIP-SL model can be written as:

$$\tilde{\boldsymbol{M}}(\tilde{\boldsymbol{q}})\ddot{\tilde{\boldsymbol{q}}} + \tilde{\boldsymbol{C}}(\tilde{\boldsymbol{q}}, \dot{\tilde{\boldsymbol{q}}}) + \tilde{\boldsymbol{G}}(\tilde{\boldsymbol{q}}) = \tilde{\boldsymbol{S}}\tilde{\boldsymbol{\tau}} \tag{1}$$

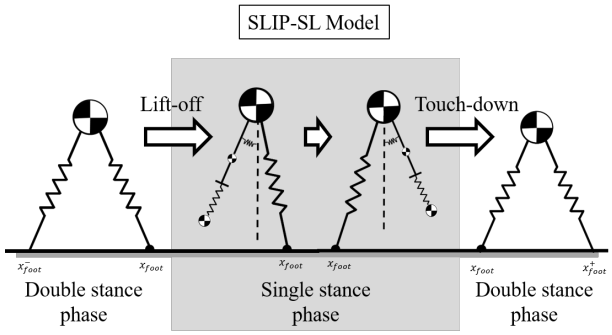


Figure 3: SLIP-SL Model

where $\tilde{\boldsymbol{q}} = [x_M, y_M, \theta, r]^T$ are the generalized coordinates, $\tilde{\boldsymbol{M}}(\tilde{\boldsymbol{q}}) \in \mathbb{R}^{4\times4}$ is an inertia matrix, $\tilde{\boldsymbol{C}}(\tilde{\boldsymbol{q}}, \dot{\tilde{\boldsymbol{q}}}) \in \mathbb{R}^4$ is a Coriolis and centrifugal terms vector, $\tilde{\boldsymbol{G}}(\tilde{\boldsymbol{q}}) \in \mathbb{R}^4$ is the gravity term, $\tilde{\boldsymbol{\tau}} \in \mathbb{R}^3$ are the resultant forces and torques due to springs and $\tilde{\boldsymbol{S}} \in \mathbb{R}^{4\times3}$ is the appropriate mapping matrix for them. The resultant forces can be calculated as:

$$\tilde{\boldsymbol{\tau}} = \begin{bmatrix} k_{0,\mathrm{ss}}(L_{0,ss} - L_{st,ss}) \\ k_{\mathrm{swLeg}}(\theta_0 - \theta) \\ k_{\mathrm{swFoot}}(r_0 - r) \end{bmatrix}, \tag{2}$$

where $k_{0,\mathrm{ss}}$, k_{swLeg} and k_{swFoot} are the stiffness values for the stance leg spring, the torsional spring at the "hip" and the linear spring connecting the swing foot with the swing leg, respectively and $L_{0,\mathrm{ss}}$, θ_0 and r_0 are the free positions of those springs where subscript "ss" indicates the single stance phase. L_{st} is the length of the stance leg. x_M and y_M respectively represent horizontal and vertical positions of the main mass, θ represents the angle of the swing leg with respect to the vertical axis and r represents the distance between the end of the swing leg and swing foot point which are represented in Figure 1.

In the double stance phase, dynamics of the SLIP-SL model can be written as:

$$m\begin{bmatrix} \ddot{x}_{\mathrm{CoM}} \\ \ddot{y}_{\mathrm{CoM}} \end{bmatrix} = \boldsymbol{F}_{\mathrm{sw}} + \boldsymbol{F}_{\mathrm{st}} + m\boldsymbol{g}, \tag{3}$$

where

$$m = m_M + m_{\mathrm{swLeg}} + m_{\mathrm{swFoot}}, \tag{4}$$

indicates the total mass of the system, x_{CoM} and y_{CoM} are the horizontal and vertical positions of the center of mass and $\boldsymbol{g} = [0, -9.81]^T$ is the gravitational acceleration. Forces generated by the stance and swing leg springs can be calculated as:

$$\boldsymbol{F}_{st} = k_{0,\mathrm{ds}}\left(\frac{L_{0,\mathrm{ds}}}{L_{st}} - 1\right)\left(\begin{bmatrix} x_{\mathrm{CoM}} \\ y_{\mathrm{CoM}} \end{bmatrix} - \begin{bmatrix} x_{foot} \\ 0 \end{bmatrix}\right), \tag{5}$$

$$\boldsymbol{F}_{sw} = k_{0,\mathrm{ds}}\left(\frac{L_{0,\mathrm{ds}}}{L_{sw}} - 1\right)\left(\begin{bmatrix} x_{\mathrm{CoM}} \\ y_{\mathrm{CoM}} \end{bmatrix} - \begin{bmatrix} x^-_{foot} \\ 0 \end{bmatrix}\right), \tag{6}$$

where subscript "ds" indicates the double stance phase. Definitions of x_{foot} can be seen in Figure 3.

2.3 Bipedal Robot Model

In this part, the dynamics of the 5 linked bipedal robot will be introduced. This model consists of 5 links which are connected to each other with revolute joints and it moves in the sagittal plane. The model is fully actuated and has an ankle torque.

Dynamics of the bipedal robot in the single stance phase can be written as:

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{G}(\boldsymbol{q}) = \boldsymbol{S}\boldsymbol{u}, \tag{7}$$

where $\boldsymbol{q} = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^T \in \mathbb{R}^5$ are the generalized coordinates, $\boldsymbol{M}(\boldsymbol{q}) \in \mathbb{R}^{5\times5}$ is an inertia matrix, $\boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \in \mathbb{R}^{5\times5}$ is a Coriolis and centrifugal terms matrix, $\boldsymbol{G}(\boldsymbol{q}) \in \mathbb{R}^5$ is the gravity term, $\boldsymbol{S} \in \mathbb{R}^{5\times5}$ is the distribution matrix of actuation torques and $\boldsymbol{u} \in \mathbb{R}^5$ are the input torques. This model can be seen in Figure 4 with the description of generalized coordinates and input torques (indicated with the red arrows).

Like the SLIP-SL model, bipedal robot model also has a two phased walking pattern. In the single stance phase, only one foot is on the ground and other is doing the swinging motion. Single stance phase ends and the system goes into the double stance phase when the swing foot touches the ground.

When the swing foot contacts the ground, a collision occurs where the generalized momentum of the system changes discontinuously. This can be modeled by assuming that an impulse force acts on the system to change the velocities while position is kept the same. This can be expressed as:

$$\boldsymbol{M}(\boldsymbol{q})\Delta\dot{\boldsymbol{q}} = \boldsymbol{J}_c^T \boldsymbol{\lambda}_{\mathrm{impact}}, \tag{8}$$

where $\boldsymbol{J}_c \in \mathbb{R}^{2\times5}$ is a constraint Jacobian matrix that maps the joint velocities to the swing foot velocity in horizontal and vertical directions. The generalized reaction forces in x and y directions are indicated as $\boldsymbol{\lambda}_{\mathrm{impact}} = [\lambda^x_{\mathrm{impact}}, \lambda^y_{\mathrm{impact}}]^T \in \mathbb{R}^2$. Assuming that the impact is inelastic, velocity of the swing foot touching the ground

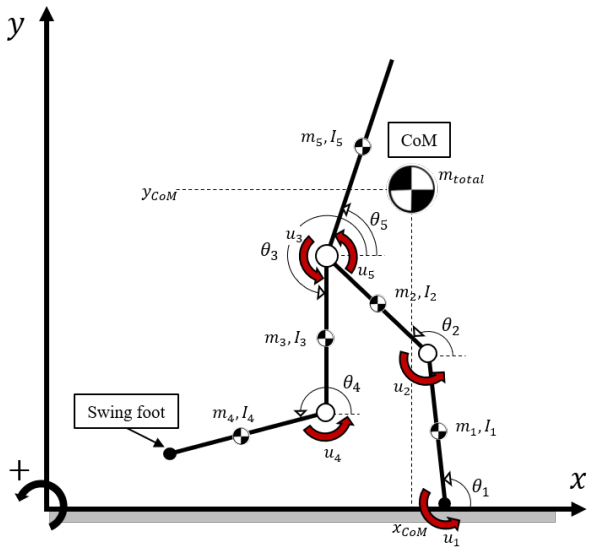


Figure 4: 5 link fully actuated robot model

will become zero after the impact which can be written as:

$$\boldsymbol{J}_c(\boldsymbol{q})\dot{\boldsymbol{q}}^+ = 0 \Leftrightarrow \boldsymbol{J}_c(\boldsymbol{q})\Delta\dot{\boldsymbol{q}} = -\boldsymbol{J}_c(\boldsymbol{q})\dot{\boldsymbol{q}}^-, \tag{9}$$

where $^-$ superscript indicates the moment just before the impact and $^+$ just after the impact. By solving (8) and (9) for $\boldsymbol{\lambda}_{impact}$ the following expression can be derived:

$$\boldsymbol{\lambda}_{\text{impact}} = -(\boldsymbol{J}_c\boldsymbol{M}^{-1}\boldsymbol{J}_c^T)^{-1}\boldsymbol{J}_c\dot{\boldsymbol{q}}^-. \tag{10}$$

By inserting $\boldsymbol{\lambda}_{\text{impact}}$ into (8), we can obtain:

$$\dot{\boldsymbol{q}}^+ = (\boldsymbol{I} - \boldsymbol{M}^{-1}\boldsymbol{J}_c^T(\boldsymbol{J}_c\boldsymbol{M}^{-1}\boldsymbol{J}_c^T)^{-1}\boldsymbol{J}_c)\dot{\boldsymbol{q}}^-, \tag{11}$$

which are the generalized velocities just after the impact. At the moment of impact, definitions of the legs are also switched (swing leg becomes the stance leg and vice versa).

The system is now in the double stance phase where both feet are in contact with the ground. To model this, a constraint force $\boldsymbol{\lambda}_{\text{ds}} = [\lambda_{\text{ds}}^x, \lambda_{\text{ds}}^y]^T \in \mathbb{R}^2$ is added to keep the swing foot on the ground. In the vertical direction, this constraint force can only push the robot ($\lambda_{ds}^y > 0$). With the non-slip assumption, double stance phase dynamics can be modeled by introducing the following constraint:

$$\boldsymbol{J}_c(\boldsymbol{q})\dot{\boldsymbol{q}} = 0. \tag{12}$$

Using this constraint, dynamical equation of the double stance phase can be written as:

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{G}(\boldsymbol{q}) = \boldsymbol{S}\boldsymbol{u} + \boldsymbol{J}_c^T\boldsymbol{\lambda}_{\text{ds}}. \tag{13}$$

The constraint force $\boldsymbol{\lambda}_{\text{ds}}$ can be obtained by taking time derivative of equation (12) as:

$$\boldsymbol{J}_c\ddot{\boldsymbol{q}} + \dot{\boldsymbol{J}}_c\dot{\boldsymbol{q}} = 0, \tag{14}$$

and inserting it into (13) as follows:

$$\boldsymbol{\lambda}_{\text{ds}} = -(\boldsymbol{J}_c\boldsymbol{M}^{-1}\boldsymbol{J}_c^T)^{-1}(\boldsymbol{J}_c\boldsymbol{M}^{-1}(\boldsymbol{S}\boldsymbol{u} - \boldsymbol{C}\dot{\boldsymbol{q}} - \boldsymbol{G}) + \dot{\boldsymbol{J}}_c\dot{\boldsymbol{q}}). \tag{15}$$

3 Direct Collocation Optimization

In this section, the optimization conducted to find periodic trajectories for SLIP-SL will be described. SLIP-SL is a passive model which means its gait solely depends on its mechanical parameters and initial conditions. In order to find the proper value, Direct Collocation Methods [10] were used in this study. These methods tackle the trajectory optimization problem by discretizing and converting it into a form which can be handled by nonlinear programming (NLP) solvers. There are many commercially available NLP solvers and in this study OpenOCL [11] will be used which can handle the multi-phase and simultaneous trajectory and parameter optimization problem. Finding the global minimum is not guaranteed when using the direct collocation methods. Global minimum is not easily found in a nonlinear problem with constraints such as 16. The advantage of direct collocation over other optimization methods such as genetic algorithms or learning based algorithms is that dynamics of the system can be embedded as constraints to the optimization problem painlessly.

The optimization can be formulated as:

$$\begin{aligned} \min_{\boldsymbol{x}_i,\boldsymbol{p},T_i} \quad & \sum_{i=1}^{2}\left(\int_{T_{i-1}}^{T_i} J_i(\boldsymbol{x}_i(t),\boldsymbol{p})dt\right) \text{for } i \in \{1,2\} \\ \text{s.t.} \quad & \dot{\boldsymbol{x}}_i = \boldsymbol{f}_i(\boldsymbol{x}_i(t),\boldsymbol{p}) \\ & \boldsymbol{r}_{i,k}(\boldsymbol{x}_i(\mu_{i,k}),\boldsymbol{p}) \le 0, \end{aligned} \tag{16}$$

where $t \in [0, T_i]$ is the time, T_i is the end time of the respective phase, $\boldsymbol{x}_i(t)$ is the state trajectory, $\boldsymbol{p}$ are the parameters, $J_i(x,p)$ are the path cost functions, $\boldsymbol{f}_i(\boldsymbol{x},\boldsymbol{p})$ are the system dynamics (described in Section 2.2) and $\boldsymbol{r}_{i,k}(\boldsymbol{x}_i,\boldsymbol{p})$ are the grid-constraints. $i = 1$ represents the single stance phase and $i = 2$ represents the double stance phase for SLIP-SL (T_1 is when the touch-down happens at the end of "ss" and T_2 is when the lift-off happens at the end of "ds"). In this paper, the Cost of Transport (CoT) [6] and a cost function to keep the swing foot low was used. CoT is an indicator of the walking efficiency.

Path, boundary and stage transition constraints are needed so that the solver can find feasible walking trajectories. Path constraints are:

- Bounds were set for the parameters to be optimized:

$$\begin{aligned} & 15000 \le k_{0,i} \le 16000\ [N/m], i \in \{ss, ds\} \\ & 1 \le L_{0,\text{ss}} \le 1.2\ [m] \\ & 0 \le k_{\text{swFoot}} \le 20000\ [N/m] \\ & 0 \le k_{\text{swLeg}} \le 15000\ [Nm/rad] \\ & 0 \le \theta_0 \le 2\pi\ [rad] \\ & -10 \le r_0 \le 10\ [m] \end{aligned} \tag{17}$$

- Stance leg spring in the single stance phase and both legs' springs in the double stance phase are always under contraction: $L_{st,ss} \le L_{0,ss}$, $L_{st,ds} \le L_{0,ds}$, $L_{sw,ds} \le L_{0,ds}$

- Constraining the vertical position of CoM: $0\ [m] \le y_{\text{CoM}} \le 0.85\ [m]$

- Swing foot is always above the ground during the single stance phase: $y_{\text{swFoot}} \ge 0$

- Elliptic virtual obstacle must be avoided by the swing foot during the single stance phase:

$$\left(\frac{x_{\text{swFoot}} - d_{obs}}{w_{obs}}\right)^2 + \left(\frac{y_{\text{swFoot}}}{h_{obs}}\right)^2 \ge 1, \tag{18}$$

  where $d_{obs} = x_{\text{swFoot}} = 0\ [m]$ is the horizontal position of the ellipse obstacle, $w_{obs} = 0.2\ [m]$ and $h_{obs} = 0.04\ [m]$ are width and height of the ellipse.

- Swing foot vertical velocity must be greater or equal to zero: $\dot{x}_{\text{swFoot}} \ge 0$

- Vertical acceleration of CoM should be negative in the single stance phase so that system doesn't try to lift the CoM up when there is only one leg on the ground: $\ddot{y}_{\text{CoM},ss} \le 0\ [m/s^2]$

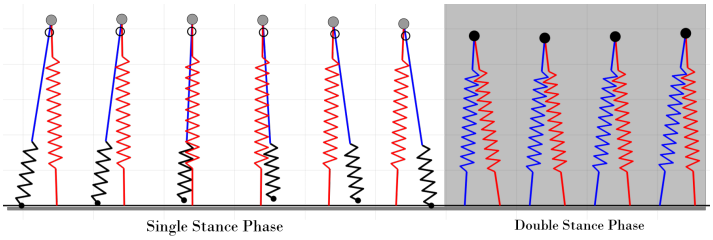


Figure 5: Snapshots of SLIP-SL's one step where gray dot in the single stance phase indicates the position of the point mass 'M' and the circle indicates the position of CoM

Boundary constraints:

- Swing foot starts on the ground from a stationary position in the beginning of the single stance phase and touches the ground at the end of the single stance phase: $y_{\text{swFoot}}(0) = 0$, $\dot{x}_{\text{swFoot}}(0) = 0$, $\dot{y}_{\text{swFoot}}(0) = 0$, $y_{\text{swFoot}}(T_1) = 0$
- Initial step length and the final step length should be same (for cyclic walking): $x_{\text{foot}} - x^{-}_{\text{foot}} = x^{+}_{\text{foot}} - x_{\text{foot}}$
- The initial position of the main mass relative to the stance foot should be the same as the final one (for cyclic walking): $x_{stFoot,0} - x_M(0) = x_{stFoot}(T_2) - x_M(T_2)$
- Constraints for cyclic walking: $y_{\text{CoM}}(0) = y_{\text{CoM}}(T_2)$, $\dot{x}_{\text{CoM}}(0) = \dot{x}_{\text{CoM}}(T_2)$, $\dot{y}_{\text{CoM}}(0) = \dot{y}_{\text{CoM}}(T_2)$
- At the stage transition, CoM position and velocity were constrained to be continuous.
- At the end of the double stance phase, swing leg should be ready to lift off, i.e. swing leg spring should be at its free length: $L_{\text{sw,ds}}(T_2) = L_{0,\text{ds}}$

Parameters to be optimized are spring stiffness values $k_{0,ss}$, $k_{0,ds}$, k_{swFoot}, k_{swLeg}, their respective free positions $L_{0,ss}$, $L_{0,ds}$, r_0,θ_0 and the initial conditions.

The optimization was conducted on MATLAB 2019b software by using 10 collocation points for each stage. Resulting spring parameters can be seen in Figure 6 for various trajectories and a snapshot of SLIP-SL's one step can be seen in Figure 5 for a sample trajectory. Trajectory 'A' from Figure 6 will be used as the reference in Section 5 where the step size was also constrained to 0.25 $[m]$ to avoid large ankle torques. The constant mechanical parameters of SLIP-SL are given in Table 1.

Table 1: SLIP-SL's constant mechanical parameters

m_M : 70 $[kg]$	m_{swLeg} : 7 $[kg]$
m_{swFoot} : 3 $[kg]$	L_{thigh} : 0.7 $[m]$
$I_{\text{swLeg}} = m_{\text{swLeg}} l^2_{\text{thigh}}/12$ $[kg \cdot m^2]$	$I_{\text{swFoot}} = m_{\text{swFoot}} L^2_{\text{thigh}}$

4 Feedback Linearization Control

In this section, the proposed controller will be introduced so that 5 linked bipedal robot model can track the reference SLIP-SL trajectories. The controller uses the feedback linearization notion, in a similar manner to [12] where a total energy control approach was used with the bipedal SLIP model. However, in this paper, a trajectory tracking approach will be used.

4.1 Single Stance Phase

For the control of the robot in the single stance phase, there are three main tasks: tracking CoM trajectory $\boldsymbol{x}_G \in \mathbb{R}^2$, tracking swing foot trajectory $\boldsymbol{\xi} \in \mathbb{R}^2$, controlling the trunk orientation $\theta_5 \in \mathbb{R}$. The velocities related to these tasks can be calculated as:

$$\dot{\boldsymbol{x}}_{t,\text{ss}} = \boldsymbol{J}_{t,\text{ss}}(\boldsymbol{q})\dot{\boldsymbol{q}}, \tag{19}$$

where $\dot{\boldsymbol{x}}_{t,\text{ss}} = [\dot{\boldsymbol{x}}_G, \dot{\boldsymbol{\xi}}, \dot{\theta}_5]^T$ is the velocity in the task space where subscript "ss" indicates the single stance phase and $\boldsymbol{J}_{t,ss}(\boldsymbol{q}) = [\boldsymbol{J}_G, \boldsymbol{J}_\xi, \boldsymbol{J}_{\theta_5}]^T$ are the combination of Jacobian matrices. $\boldsymbol{J}_G$ maps generalized velocities to the velocity of the center of mass, $\boldsymbol{J}_\xi$ maps generalized velocities to swing foot velocities and $\boldsymbol{J}_{\theta_5}$ maps generalized velocity to the trunk's angular velocity. By taking the time derivative of Equation (19) and inserting the obtained $\ddot{\boldsymbol{q}}$ into Equation (7) we can get:

$$\ddot{\boldsymbol{x}}_t = \boldsymbol{J}_{t,\text{ss}}\boldsymbol{M}^{-1}(\boldsymbol{S}\boldsymbol{u} - \boldsymbol{C}\dot{\boldsymbol{q}} - \boldsymbol{G}) + \dot{\boldsymbol{J}}_{t,\text{ss}}\dot{\boldsymbol{q}}. \tag{20}$$

Inputs should be chosen as:

$$\boldsymbol{u} = \boldsymbol{S}^{-1}\left(\boldsymbol{M}\boldsymbol{J}^{-1}_{t,\text{ss}}(\ddot{\boldsymbol{x}}_{t_d,ss} - \dot{\boldsymbol{J}}_{t,\text{ss}}\dot{\boldsymbol{q}}) + \boldsymbol{C}\dot{\boldsymbol{q}} + \boldsymbol{G}\right), \tag{21}$$

to get $\ddot{\boldsymbol{x}}_t = \ddot{\boldsymbol{x}}_{t_d,\text{ss}}$. By choosing:

$$\ddot{\boldsymbol{x}}_{t_d,\text{ss}} = \begin{bmatrix} K_{P_G}(x_{\text{CoM,des}} - x_{\text{CoM}}) + K_{D_G}(\dot{x}_{\text{CoM,des}} - \dot{x}_{\text{CoM}}) \\ K_{P_G}(y_{\text{CoM,des}} - y_{\text{CoM}}) + K_{D_G}(\dot{y}_{\text{CoM,des}} - \dot{y}_{\text{CoM}}) \\ K_{P_{\text{sw}}}(x_{\text{sw,des}} - x_{\text{sw}}) + K_{D_{\text{sw}}}(\dot{x}_{\text{sw,des}} - \dot{x}_{\text{sw}}) \\ K_{P_{\text{sw}}}(y_{\text{sw,des}} - y_{\text{sw}}) + K_{D_{\text{sw}}}(\dot{y}_{\text{sw,des}} - \dot{y}_{\text{sw}}) \\ K_{P_T}(\theta_{5,\text{des}} - \theta_5) + K_{D_T}(-\dot{\theta}_5) \end{bmatrix} \tag{22}$$

where K_P and K_D are the proportional and derivative gains for the controller and $\theta_{5,des}$ is the desired trunk angle, desired trajectories can be tracked. Desired trajectories will be chosen as the SLIP-SL trajectories that were obtained in Section 3. In this study, the desired trunk angle was chosen as $\theta_{5,\text{des}} = \pi$ [rad].

4.2 Double Stance Phase

During the double stance phase, swing foot remains on the ground which means there is one less task to be carried out. This means that some modifications should be made to the single stance phase controller so that it can be used in the double stance phase. The task space for the double stance phase then becomes:

$$\dot{\boldsymbol{x}}_{t,ds} = \begin{bmatrix} \dot{\boldsymbol{x}}_G \\ \dot{\theta}_5 \end{bmatrix} \in \mathbb{R}^3. \tag{23}$$

This reduction in dimension of the task space is somewhat problematic since we need to determine 5 separate inputs but the dimension of $\dot{\boldsymbol{x}}_{t,ds}$ is 3 which means there is more than one correct way to allocate the inputs. To calculate the proper inputs, the following equation can be used:

$$\dot{\boldsymbol{q}}_{\text{hip}} = \boldsymbol{J}_{\text{hip,sw}}(\boldsymbol{q}_{\text{sw}})\dot{\boldsymbol{q}}_{\text{sw}} = \boldsymbol{J}_{\text{hip,st}}(\boldsymbol{q}_{\text{st}})\dot{\boldsymbol{q}}_{\text{st}}, \tag{24}$$

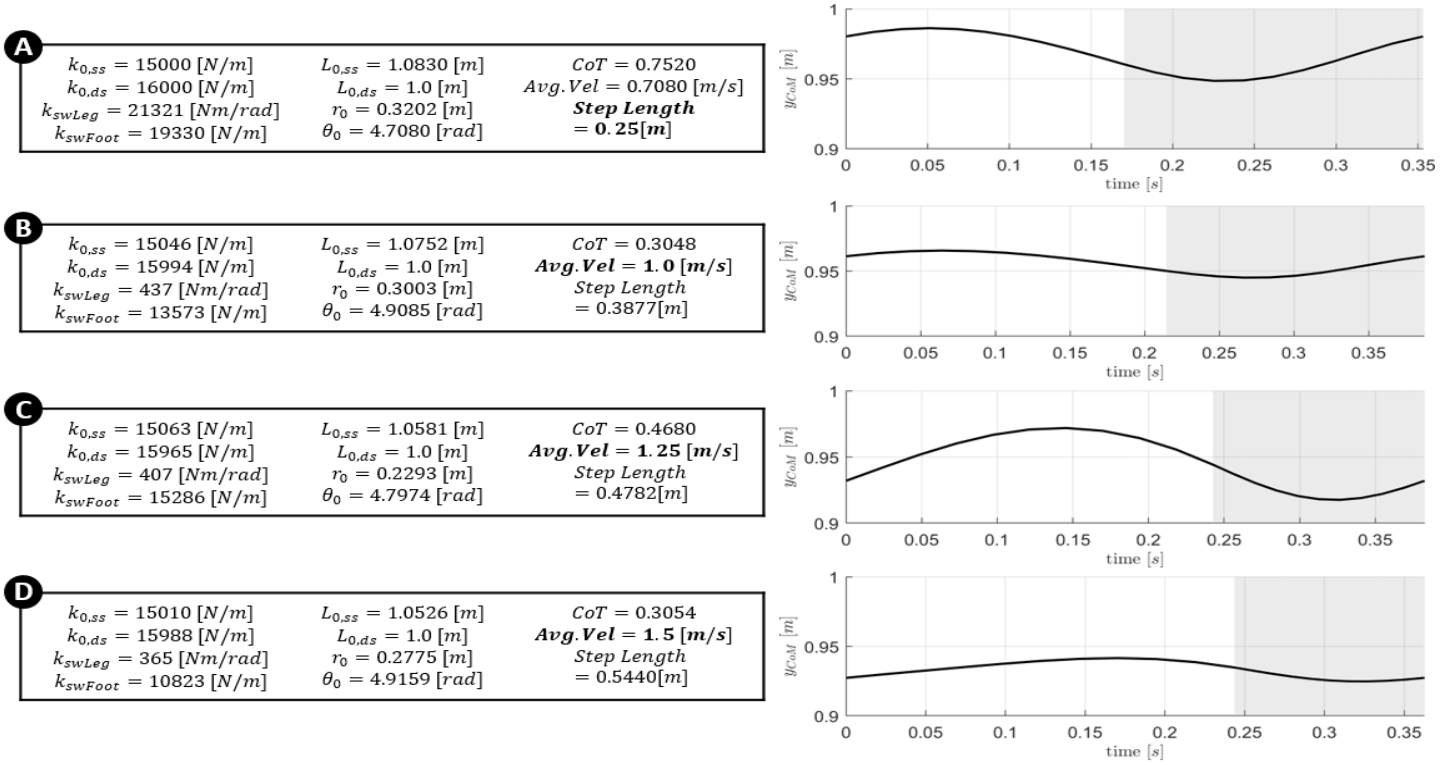


Figure 6: Optimization results for various SLIP-SL trajectories. For the trajectory A, step length was constrained to 0.25 [m] to get better ankle torques and this trajectory is the one that was used as reference in Section 4. For B, C and D trajectories, average velocity constraints were added. In this figure, resulting mechanical parameters, costs of transport and step length are given as well as the SLIP-SL's y_{CoM} trajectory. In the plots on the right, gray background means that the SLIP-SL is in double stance phase.

where $\boldsymbol{J}_{\mathrm{hip},\,i}(\boldsymbol{q}_i)$ is the jacobian matrix that maps the corresponding legs angular velocities, $\dot{\boldsymbol{q}}_{\mathrm{st}} = [\dot{\theta}_1, \dot{\theta}_2]^T$ and $\dot{\boldsymbol{q}}_{\mathrm{sw}} = [\dot{\theta}_3, \dot{\theta}_4]^T$, to the velocity of the hip. This holds since both foot are on the ground during the double stance phase. From equation (24),

$$\dot{\boldsymbol{q}}_{\mathrm{st}} = \boldsymbol{J}^{-1}_{\mathrm{hip,st}}(\boldsymbol{q}_{\mathrm{st}})\boldsymbol{J}_{\mathrm{hip,sw}}(\boldsymbol{q}_{\mathrm{sw}})\dot{\boldsymbol{q}}_{\mathrm{sw}}, \tag{25}$$

can be obtained to get:

$$\dot{\boldsymbol{q}} = \underbrace{\begin{bmatrix} \boldsymbol{J}^{-1}_{\mathrm{hip,st}}\boldsymbol{J}_{\mathrm{hip,sw}} & 0_{2\times1} \\ I_{2\times2} & 0_{2\times1} \\ 0_{1\times2} & 1 \end{bmatrix}}_{\boldsymbol{\Gamma}(\boldsymbol{q})} \underbrace{\begin{bmatrix} \dot{\boldsymbol{q}}_{\mathrm{sw}} \\ \dot{\theta}_5 \end{bmatrix}}_{\dot{\boldsymbol{q}}_a}. \tag{26}$$

After taking the time derivative of (26), then substituting the $\ddot{\boldsymbol{q}}$ and $\dot{\boldsymbol{q}}$ terms from (26) into (7) and multiplying with $\boldsymbol{\Gamma}^T$ from left, following can be obtained:

$$\boldsymbol{M}_a(\boldsymbol{q})\ddot{\boldsymbol{q}}_a + \boldsymbol{C}_a(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}}_a + \boldsymbol{G}_a(\boldsymbol{q}) = \boldsymbol{u}_a, \tag{27}$$

where

$$\begin{cases} \boldsymbol{M}_a = \boldsymbol{\Gamma}^T\boldsymbol{M}\boldsymbol{\Gamma} \\ \boldsymbol{C}_a = \boldsymbol{\Gamma}^T(\boldsymbol{C}\boldsymbol{\Gamma} + \boldsymbol{M}\dot{\boldsymbol{\Gamma}}) \\ \boldsymbol{G}_a = \boldsymbol{\Gamma}^T\boldsymbol{G} \\ \boldsymbol{u}_a = \boldsymbol{\Gamma}^T\boldsymbol{S}\boldsymbol{u} \end{cases} \tag{28}$$

Using a similar technique that was used for deriving (21):

$$\boldsymbol{u}_a = \boldsymbol{M}_a\boldsymbol{J}^{-1}_{t,\mathrm{ds}}(\ddot{\boldsymbol{x}}_{t_d,\mathrm{ds}} - \dot{\boldsymbol{J}}_{t,\mathrm{ds}}\dot{\boldsymbol{q}}_a) + \boldsymbol{C}_a\dot{\boldsymbol{q}}_a + \boldsymbol{G}_a, \tag{29}$$

can be found where $\ddot{\boldsymbol{x}}_{t_d,\mathrm{ds}}$ should be chosen as:

$$\ddot{\boldsymbol{x}}_{t_d,\mathrm{ds}} = \begin{bmatrix} K_{P_G}(x_{\mathrm{CoM,\,des}} - x_{\mathrm{CoM}}) + K_{D_G}(\dot{x}_{\mathrm{CoM}} - \dot{x}_{\mathrm{CoM}}) \\ K_{P_G}(y_{\mathrm{CoM,\,des}} - y_{\mathrm{CoM}}) + K_{D_G}(\dot{y}_{\mathrm{CoM,\,slipsl}} - \dot{y}_{\mathrm{CoM}}) \\ K_{P_T}(\theta_{5,\mathrm{des}} - \theta_5) + K_{D_T}(-\dot{\theta}_5) \end{bmatrix}. \tag{30}$$

However, only $\boldsymbol{u}_a \in \mathbb{R}^3$ can be obtained in this way which only has dimension 3. What we need to control the robot is $\boldsymbol{u} \in \mathbb{R}^5$. One way to calculate $\boldsymbol{u}$ is by using the relation $\boldsymbol{u}_a = \boldsymbol{\Gamma}^T\boldsymbol{S}\boldsymbol{u}$ given in (28) as:

$$\boldsymbol{u} = \boldsymbol{S}^{-1}\left(\underbrace{\boldsymbol{W}^{-1}\boldsymbol{\Gamma}(\boldsymbol{\Gamma}^T\boldsymbol{W}^{-1}\boldsymbol{\Gamma})^{-1}}_{(\boldsymbol{\Gamma}^T)^{+\boldsymbol{W}}}\boldsymbol{u}_a\right), \tag{31}$$

where $(\boldsymbol{\Gamma}^T)^{+\boldsymbol{W}}$ is the weighted matrix inverse operation. $\boldsymbol{W} \in \mathbb{R}^{5\times5}$ matrix can be used to penalize high input torques such as the ankle torque but we selected it as identity matrix for this paper.

Controllers for the single and double stance phases are thus derived. Another important aspect of the controller in tracking the SLIP-SL trajectories is the switching of phases at the correct moments. When the biped robot is in the double stance phase and the tracked trajectory goes into single stance phase, controller switches to the single stance phase controller and commands the robot to lift its foot so it too can switch to the proper phase.

5 Simulation Results and Discussion

In the following simulation study, the designed controller's ability to track the reference SLIP-SL trajectories and the suitability of the SLIP-SL as a template model for walking will be tested. Table 3 shows the mechanical parameters of the 5 link robot and Table 2 shows the chosen gain values for the controller and the torque limits on the motors were set at 200 $[Nm]$. Gain values are chosen with the help of a particle swarm optimization algorithm (PSO) [13]. We were able to find controller parameters that resulted in stable gaits easily by hand-tuning but with the help of PSO we were able get results with better walking efficiency.

Table 2: Control Parameters

$K_{P_G} : 54$	$K_{D_G} : 9$
$K_{P_{\text{sw}}} : 82$	$K_{D_{\text{sw}}} : 8$
$K_{P_T} : 36$	$K_{D_T} : 4$

The simulations were implemented in MATLAB 2020b's Simulink environment with ode45 solver and variable step settings (absolute tolerance was set to 1e-8).

Figure 7 shows the resulting CoM trajectory and trunk orientation the controlled system when the proposed controller is used. Figure 8 shows the swing foot trajectories for the same system. It can be seen that the proposed controller does a good job in tracking the reference trajectories of the SLIP-SL template, which were obtained in Section 3. The reference SLIP-SL trajectories for the swing foot would not be available if a template such as the popular SLIP was used.

Table 3: 5 Link Model Parameters

$l_1 = l_4 : 0.48\ [m]$	$l_2 = l_3 : 0.48\ [m]$	$l_5 : 0.48\ [m]$
$m_1 = m_4 : 5\ [kg]$	$m_2 = m_3 : 5\ [kg]$	$m_5 : 60\ [kg]$
$I_i = m_i l_i^2/12\ [kg \cdot m^2],\ i = 1, 2, 3, 4, 5$		

It can be seen in Table 2 that relatively low gains were chosen for this study. Tracking performance can be increased by using larger gains but since this model is fully actuated and has an ankle torque, zero moment point (ZMP) condition must also be checked. ZMP criterion states that if the center of pressure moves to the toe (or to the "outside" of the foot), foot would rotate and system would be under actuated [14]. For this trajectory, center of pressure stays within a 30 [cm] foot.

Figure 9 shows snapshots of one step of the 5 link models gait. to check the stability of the gait, Poincaré map approach was considered [12]. The dimensions of the Poincaré map were selected as the $\theta_P = \text{atan2}(\dot{y}_{\text{CoM}}, \dot{x}_{\text{CoM}})$, y_{CoM} and total energy of the 5 link model E at the vertical leg orientation (VLO). VLO happens when CoM of the 5 link model is at the same horizontal position as the stance foot. VLO was chosen as the Poincaré section because the horizontal position doesn't need to be considered at this point. Poincaré stability criterion indicates that if the return map converges to a fixed point, a hybrid system with impact effects can be considered periodic [12]. Poincare Map for the controlled 5 link model is shown in Figure 10. It can be seen that the gait converges to a stable point in the section after a couple of steps which indicates stability.

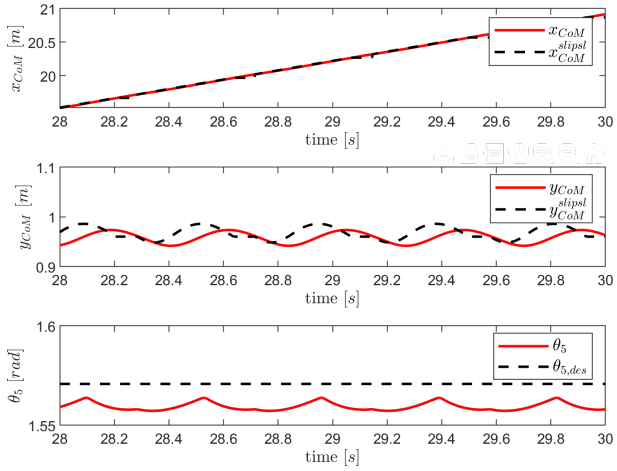


Figure 7: Trajectory tracking results for CoM horizontal position, vertical position and trunk orientation

The CoT value for the reference SLIP-SL trajectory was 0.7520 and this value is 0.7745 for the controlled 5 link robot. Also, the average velocity of the SLIP-SL trajectory was 0.7080 [m/s] and this value was 0.6974 [m/s] for the controlled system. These values being very similar between the reference and the controlled model also indicates the validity of the proposed controller. Cost of transport being slightly higher is expected because the 5 linked model needs to keep its body upright but SLIP-SL doesn't have this issue.

In this paper, it was also shown that by using direct collocation optimization, various SLIP-SL trajectories can be obtained (Figure 6) that resemble the walking gait. Cost of transport tended to decrease when the average velocity of the gait was increased and step length was not constrained but this kind of trajectories can be more demanding on the inputs and they sometimes resulted in large ankle torques which was troubling for the ZMP criterion. Also stiffness of the legs were limited to 15000 $[Nm] \leq k_{0,i} \leq 16000\ [Nm],\ i \in \{ds, ss\}$ to keep the CoM height within a certain range.

This stiffness limits was chosen to have a similar height trends with [12] and [6] where k_0 was 15696 [Nm]. It can be seen that $k_{0,ss} \neq k_{0,ds}$ for the given trajectories but it was possible to find feasible trajectories with $k_{0,ss} = k_{0,ds}$, even with $k_{0,ss} = k_{0,ds} =$

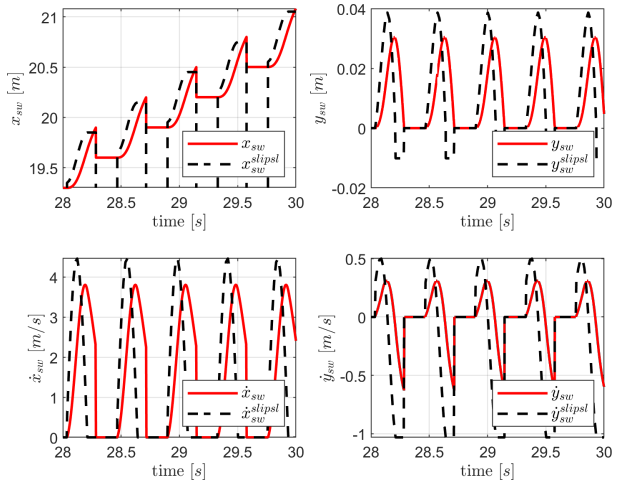


Figure 8: Trajectory tracking results for the swing foot

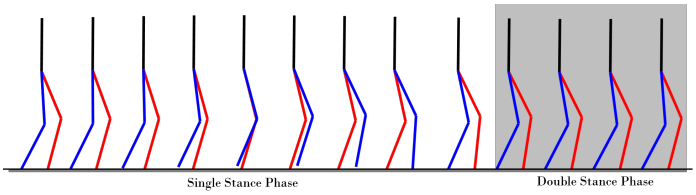


Figure 9: Snapshots from a step of the 5 Link model

15696 [Nm] but this doesn't really mean that linear leg springs were the same in single and double stance phase since $L_{0,ss} \neq L_{0,ds}$. The free length of the springs in the double stance phase was set to 1 [m] to be the same with [12] but $L_{0,ss}$ needs to be larger than this value to keep the CoM high enough. If $L_{0,ss}$ was 1.0 [m], CoM would sag further than desired range in the single stance phase.

6 Conclusion

In this paper, a template model called SLIP-SL was proposed which is an extension to the popular SLIP model. SLIP model can generate the reference CoM trajectories that can mimic the two-phased walking but it doesn't have the swing leg dynamics. Thus, when this template model is used, additional steps are necessary for obtaining the swing foot trajectory so that the actual robot can be controlled. Since the swinging motion is a huge part of the walking gait that needs to be taken into account, SLIP-SL adds this dynamics while keeping the passiveness of the SLIP model.

Since the SLIP-SL is passive, proper mechanical parameters and initial conditions needs to be determined to get feasible walking trajectories. It was demonstrated that direct collocation methods can be used to find the proper parameters for the passive SLIP-SL model. This step was crucial since there are many parameters that needs to be decided and other more exhaustive search methods might not work. It was shown that a variety of trajectories with different average velocities, CoM behaviors etc. can be obtained using the same basic principles. Then, a controller was introduced to track the

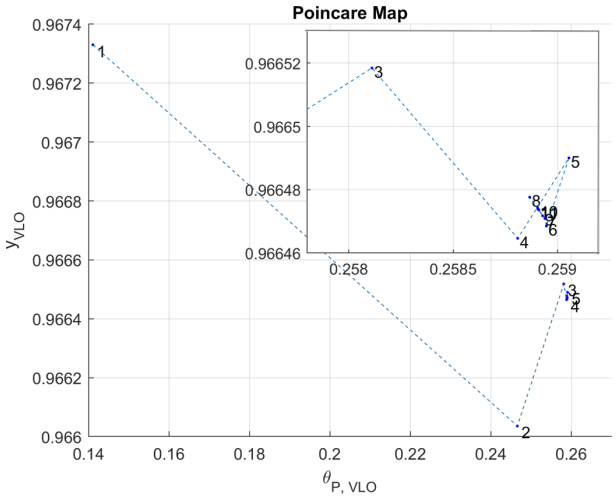


Figure 10: 2D section of the Poincaré Map where the numbers indicate the step number (zoomed in version is shown in the right upper corner of the figure)

obtained reference trajectories and it was shown that it can satisfactorily track the reference SLIP-SL trajectories to reach a stable gait while also satisfying the ZMP criterion. These results also confirmed that SLIP-SL can be used as a template model for walking robots.

In the future, we would like to introduce robustness to this passive system by making the springs variable ones and controlling them to reject disturbances. Another area we would like to explore is achieving a variable speed gait with the bipedal model. We want to achieve this by smoothly switching between different reference SLIP-SL trajectories.

References

[1] M. M. Pelit, J. Chang, R. Takano, M. Yamakita, "Bipedal Walking Based on Improved Spring Loaded Inverted Pendulum Model with Swing Leg (SLIP-SL)," in 2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), 72–77, IEEE, 2020, doi:10.1109/aim43001.2020.9158883.

[2] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, H. Hirukawa, "Biped walking pattern generation by using preview control of zero-moment point," in 2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422), volume 2, 1620–1626, IEEE, 2003, doi:10.1109/robot.2003.1241826.

[3] R. Takano, M. Yamakita, "Sequential-contact bipedal running based on SLIP model through zero moment point control," in 2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), 1477–1482, IEEE, 2017, doi:10.1109/aim.2017.8014227.

[4] H. Yano, J. Chang, R. Takano, M. Yamakita, "Simultaneous Optimization of Trajectory and Parameter for Biped Robot with Series Elastic Actuators," IFAC-PapersOnLine, **52**(22), 7–12, 2019, doi:10.1016/j.ifacol.2019.11.039.

[5] A. Patel, S. L. Shield, S. Kazi, A. M. Johnson, L. T. Biegler, "Contact-implicit trajectory optimization using orthogonal collocation," IEEE Robotics and Automation Letters, **4**(2), 2242–2249, 2019, doi:10.1109/lra.2019.2900840.

[6] J. Rummel, Y. Blum, A. Seyfarth, "Robust and efficient walking with spring-like legs," Bioinspiration & biomimetics, **5**(4), 046004, 2010, doi:10.1088/1748-3182/5/4/046004.

[7] H. Geyer, A. Seyfarth, R. Blickhan, "Compliant leg behaviour explains basic dynamics of walking and running," Proceedings of the Royal Society B: Biological Sciences, **273**(1603), 2861–2867, 2006, doi:10.1098/rspb.2006.3637.

[8] J. S. Park, C. M. Lee, S.-M. Koo, C. H. Kim, "Gait phase detection using force sensing resistors," IEEE Sensors Journal, **20**(12), 6516–6523, 2020, doi:10.1109/jsen.2020.2975790.

[9] M. A. Sharbafi, A. M. N. Rashty, C. Rode, A. Seyfarth, "Reconstruction of human swing leg motion with passive biarticular muscle models," Human movement science, **52**, 96–107, 2017, doi:10.1016/j.humov.2017.01.008.

[10] M. Kelly, "An introduction to trajectory optimization: How to do your own direct collocation," SIAM Review, **59**(4), 849–904, 2017, doi:10.1137/16m1062569.

[11] J. Koenemann, G. Licitra, M. Alp, M. Diehl, "OpenOCL–Open Optimal Control Library," 2017, doi:10.1137/16m1062569.

[12] G. Garofalo, C. Ott, A. Albu-Schäffer, "Walking control of fully actuated robots based on the bipedal slip model," in 2012 IEEE International Conference on Robotics and Automation, 1456–1463, IEEE, 2012, doi:10.1109/icra.2012.6225272.

[13] J. Kennedy, R. Eberhart, "Particle swarm optimization," in Proceedings of ICNN'95-international conference on neural networks, volume 4, 1942–1948, IEEE, 1995, doi:10.5772/6754.

[14] E. R. Westervelt, J. W. Grizzle, C. Chevallereau, J. H. Choi, B. Morris, Feedback control of dynamic bipedal robot locomotion, CRC press, 2018, doi:10.1201/9781420053739.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 332-341 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design and Implementation of Reconfigurable Neuro-Inspired Computing Model on a FPGA

Basutkar Umamaheshwar Venkata Prashanth*,1, Mohammed Riyaz Ahmed[2]

[1]School of Electronics and Communication Engineering, REVA University, Bengaluru, 560064, India

[2]School of Multidisciplinary Studies, REVA University, Bengaluru, 560064, India

ARTICLE INFO

Article history:
Received: 23 July, 2020
Accepted: 12 September, 2020
Online: 17 September, 2020

Keywords:
Bio-Inspired Computing
FPGA
Neuron
Synapse
IP Core

ABSTRACT

In this paper we design a large scale reconfigurable digital bio-inspired computing model. We consider the reconfigurable and event driven parameters in the developed field-programmable neuromorphic computing system. The various Intellectual Property (IP) cores are developed for the modules such as Block RAM, Differential Clock, Floating Point, and First In First Out (FIFO) for the design of the neuron model in Xilinx ISE, with exploration of register transfer logic (RTL) and hardware synthesis using Verilog code. The architecture for design at device level offers the best possible design tradeoff for specific processor architectures and development choices. In this paper we perform algorithmic design of a large scale reconfigurable logical bio-inspired computing model. The proposed algorithm is implemented on Field Programmable Gate Array (FPGA) to develop a neuron model to be utilized in neuromorphic computing system.

1. Introduction

This research manuscript is an extension of work originally presented in *International Conference on Artificial Intelligence and Signal Processing* [1]. The bio-inspired computing is achieved with core building blocks of neuromorphic engineering which mainly constitute circuits and systems and is proposed as structures of spin devices [2].

The above direction opens a new path and induces a key approach for developing bio-inspired algorithms for implementations of bio-inspired algorithm based computing systems [3]. In nature the crucial role of learning and memory is achieved with help of synapses. The synapses which are plastic in nature, is formed with inter cellular connections of neurons, and the combination of these biological structures form the basic building blocks of neural networks [4]. Synapses can change their state based on the neural activity of coupled neurons. The functionality of neurons and synapses is mimicked in hardware by utilizing very large scale integration technology, plays a key role in design of neuromorphic computing systems [5]. The pathway to efficient neuromorphic systems is encoding the neural and synaptic functionalities in an electronic spin. It shows the potentials to exploit energy efficiency, performance, reliability, and magnetization using electric fields, and enhanced memory density of spintronic memory devices [6]. The bio-inspired computing systems is presented, with an aim to establish interaction framework, between two directions of natural system computation and artificial system computation [7]. In the in-memory computing for emerging memory devices, there is no separation between memory and logic, to overcome Von-Neumann bottleneck and also in-memory computing devices are designed with zero-off state power and due to this, they have a distinct advantage of the non-volatile state [8]. The in-memory computing is combined with a high gate or synapse density which enables forming of cross-bar array in the device, which can be easily integrated with CMOS with high density, operating with high current and voltage consuming high dynamic power, with In-memory computing device contains long switching time hence they operate with limited speed and have limited endurance, again the cross-bar in-memory computing is highly parallel, operate with low-power, low cost [9]. The verticals of Cross-bar in-memory computing are bio-inspired computing, deep learning, in-memory logic, chip/data security, architecture, device modeling [10]. The static random access memory latches and capacitors are utilized in a very large scale integrated devices, an architecture to implement strewed memory elements is embodied in this in-memory computing device as depicted in Figure 1 is aimed towards supporting the use of memristive devices as digital and synapse-like memory elements. The key contributions of the research manuscript are:

- Algorithmic design of a large scale reconfigurable logical bio-inspired computing model.
- Implementation of bio-inspired model on FPGA.

*Corresponding Author: B.U.V. Prashanth, prashanthbuv@reva.edu.in

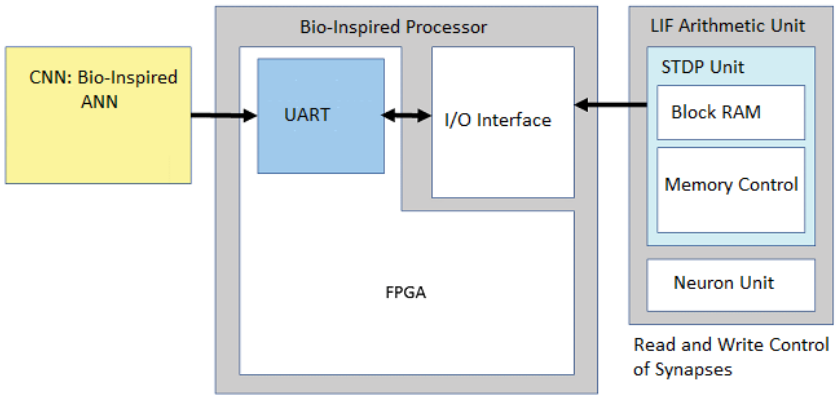


Figure 1: Block Diagram of FPGA based Bio-Inspired Computing System

2. Background

The emerging research architectures which support memory-based computing, exponential performance scaling, which enables mixed mode technology solutions. The problem of establishing the bridge between natural and artificial computation is one of motivation which illuminates the bio-inspired computation in artificial systems [11]. The paradox of programming a bio-inspired computer is the need to figure out the new class of algorithms, we are also missing out some very important basic concepts. The intelligent computational system are initially built with a Boolean logic or functions, next through logical phase, and into semiconductor technology phase, further into computational complexity phase and ends at an experimental computation phase [12]. The lesson learnt here is while dealing with intelligence the first stage is to probe into evolution, complexity, and also thermodynamics which is not an equivalent of a Boolean logic or functions, during the second stage new and novel electronics technology is required which is not an equivalent of electronics technology, which was defined during the computation phase, in the further stage the implementation complexity is not an equivalent to computational complexity, and the final stage is the practical intelligence stage which is also not an equivalent practical computation stage, defined during process of computation [13]. The factor synaptic plasticity, which accounts for the determination of the magnitude of the synaptic weights. The plasticity is also called as the learning of the synaptic junctions which probes the cognitive abilities to the bio-inspired architectures [14]. For analysis, if an experiment is performed considering the circuit with four access transistors to decouple read and write current paths, with the peripheral circuits for timing window, and aimed towards the spike-timing-dependent plasticity (STDP) implementation. Apart from these four transistors, one more transistor named as M_{STDP} is also connected to pre-charge line in the circuit, which is responsible to implement the STDP and this transistor is biased in the sub threshold saturation regime. The gate voltage of M_{STDP} transistor is called as PRE voltage, which starts increasing linearly as the pre neuron spikes [15]. As the post neuron is triggered, the POST signal is activated with the current flows through the device. The current is also known as the programming current , which is the 1 ns duration write current and is exponentially related with the magnitude to the delay factor of pre-neuron and post-neuron spikes. As the STDP measurements are taken between the % changes in the synaptic weights with respect to spike timing difference in (ms) , the synaptic weight is updated which depends on difference in the timing of post and pre neuron spikes[16].

3. Algorithm Design

In this section the design of algorithm is enumerated for a FPGA based large scale logical reconfigurable neuron model.

Algorithm 1: Algorithm for FPGA Top Level Module

```
Result: Top Level module of IP Block
Model initialization;
while Apply Clock Signal do
    Instantiate Clock Signal;
    if FSM Instantiation then
        Select RAM;
        Select MATMUL
    Else
        Compute:
        Communication Unit;
        Initialize UART;
        Initialize FIFO;
    End
    Emulate layer of network;
    Read Weight RAM;
    Load input RAM;
    Process data as perceptron;
End
```

To realize the spiking neural functionalities by utilizing the leaky-integrate magnetization dynamics [17]. To enable the abstraction of the magnetic functions as stochastic spiking neurons, the parameter required is the thermal noise which is prevalent in nano-magnets at certain temperatures which are not equal to zero [18]. The reconfigurable neuron model consists of two units such as finite state machine unit (FSM) unit and communication unit. The Algorithm 1 illustrates the abstract view of the top level model hierarchy. The IP core module as described in Figure 2(a) emulates the layer of the network and loads data (weight or input) and then process the data to obtain the synapse weight output as depicted in Figure 2(b). Further the flag signal is controlled in

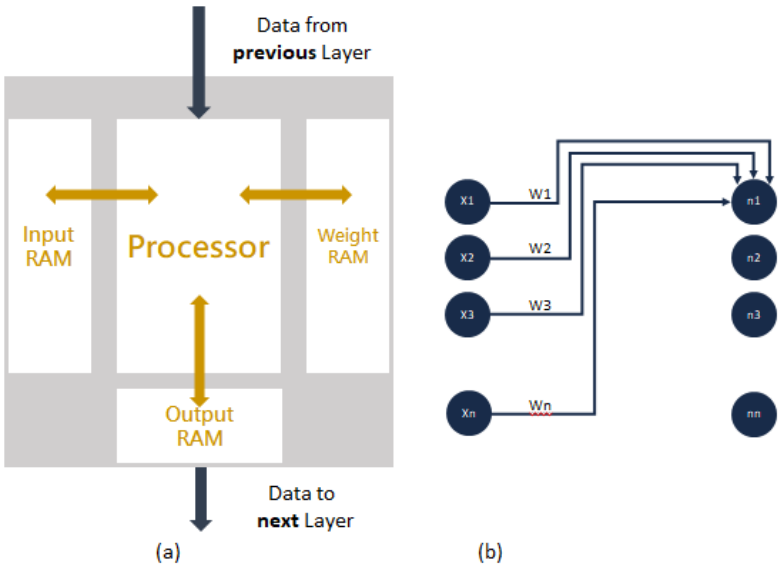


Figure 2: (a) IP-core module (b) Details of synapse weight information

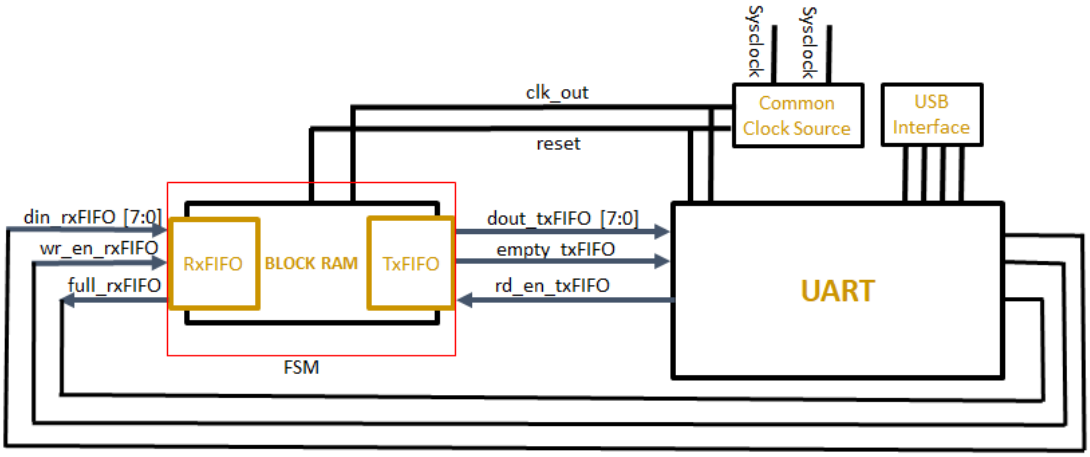


Figure 3: RTL schematic of Top Layer with sub-module Read/Write Memory

order to make the FSM to halt while the other sub programs are running [19]. The layer of the network is mimicked with reading the weight information from weight RAM, along with load input data from input RAM and processing the data as a simple neuron model [20]. The 8 bit parameter data width is applied as input to FPGA from data floating unit of USB module and from FPGA the parameter data is passed to FIFO block. The input data is further connected to din of input RAM, along with connecting row index to address of input RAM, and connection of read data to input of outgoing FIFO [21]. The outputs of last add operation is directly wired to the resultant RAM. The data packets are segmented at this stage and the information of data packets are available at the next stage. In the next stage the packet data is segmented into respective component registers. Write the input and weight data to padded variables when flag is high. The Figure 3 illustrates the above design process as sub-module read and write memory with FSM unit (with block RAM, Transmit FIFO, Receive FIFO) and universal asynchronous receiver and transmitter (UART) unit. Further also describes the communication link between FSM unit and the UART unit. Here the data sent from universal serial bus (USB) transforms to parallel data and is then sent to the FIFO for processing [22]. The Figure 3 can be considered as register transfer logic (RTL) schematic of the top layer module.

4. Implementation

In this section the implementation details of large scale reconfigurable digital bio-inspired computing model is described. The algorithm described is implemented on the hardware FPGA environment satisfying the requirement of hardware combined with the software co-design concept. The hardware used is FPGA ALTERA DE2 with a cyclone chip.

The Figure 4 represents the topology of the FSM of a RAM read/write process with FIFO pop data in sequence. The complete architecture of the bio-inspired computing system consists of system controller based on an advanced reduced instruction set computing (RISC) machine (ARM) processor, core-array of two dimensions, and a UART Controller. The interpretation of data is dependent of the order of popping the data. Sub-module MATMUL contains two dot product operations in parallel, which consists of floating point IPs as shown in Figure 5 in the form of

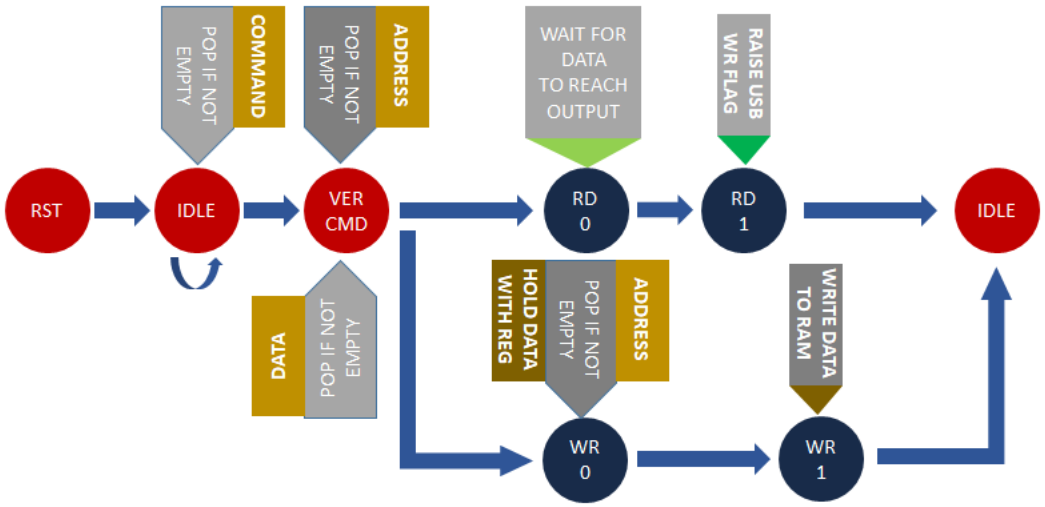


Figure 4: FSM Topology of Read/Write Process

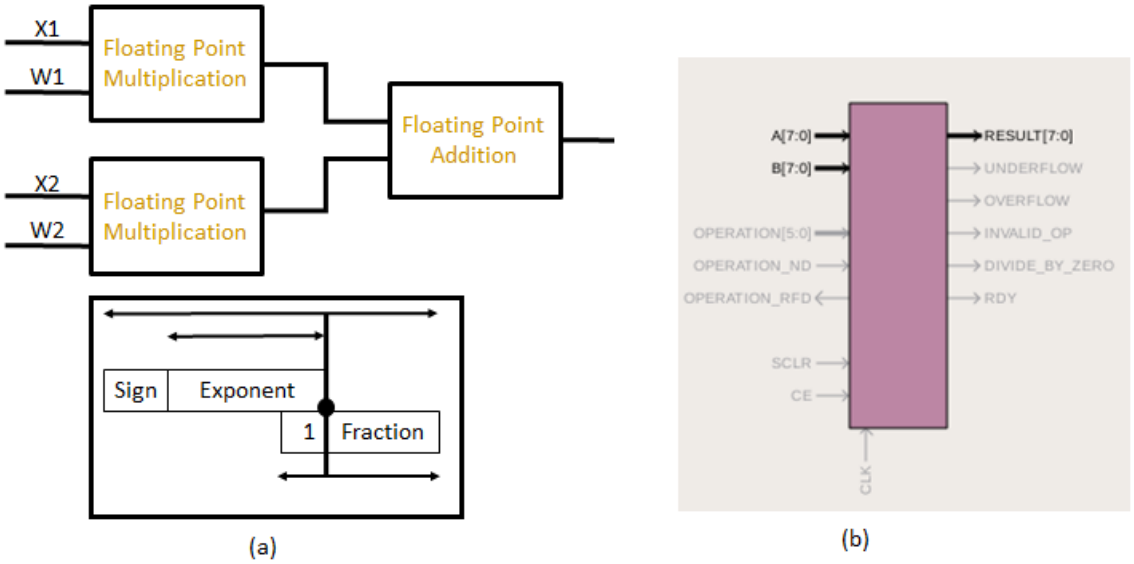


Figure 5: Submodule Dot Product Module/Process

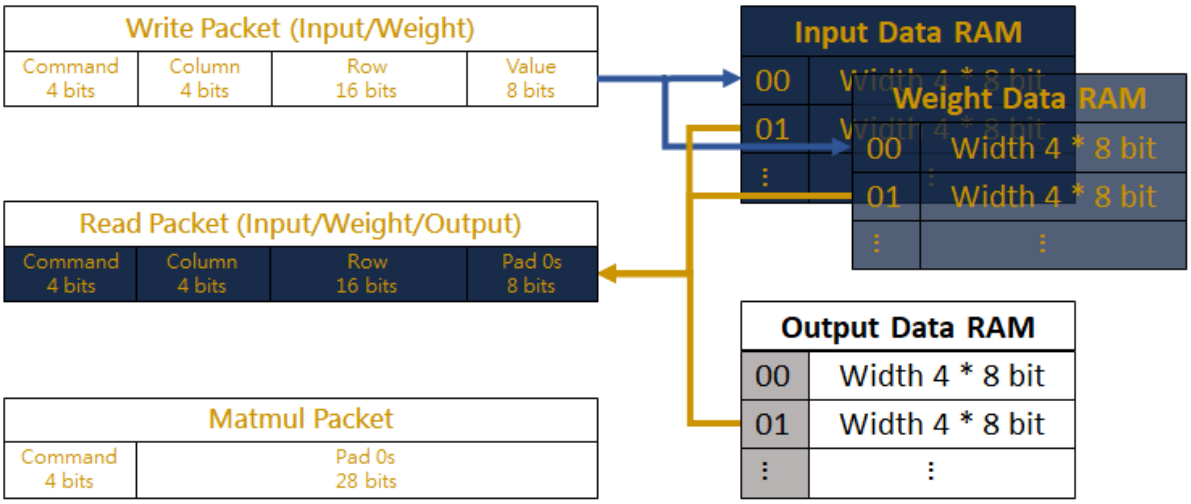


Figure 6: Information Transfer Packet

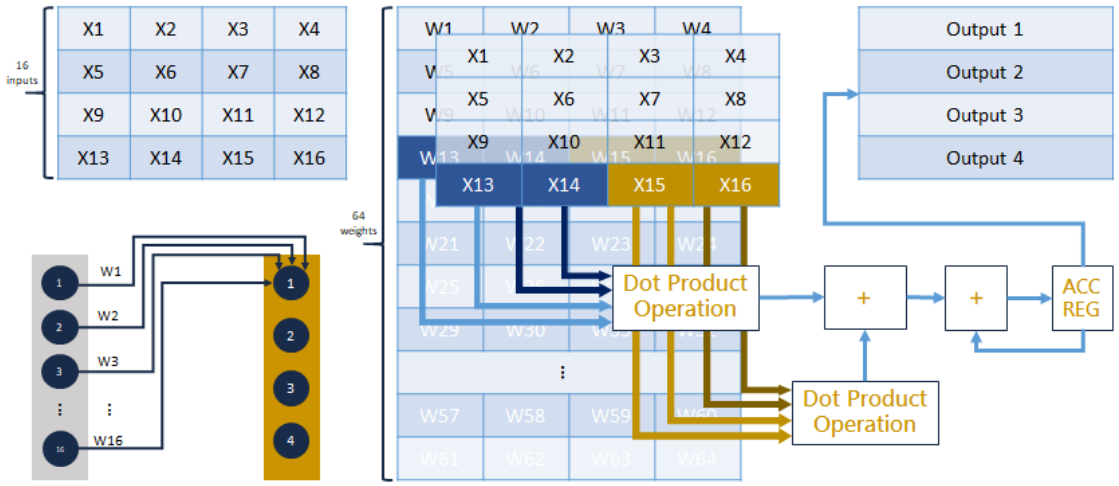
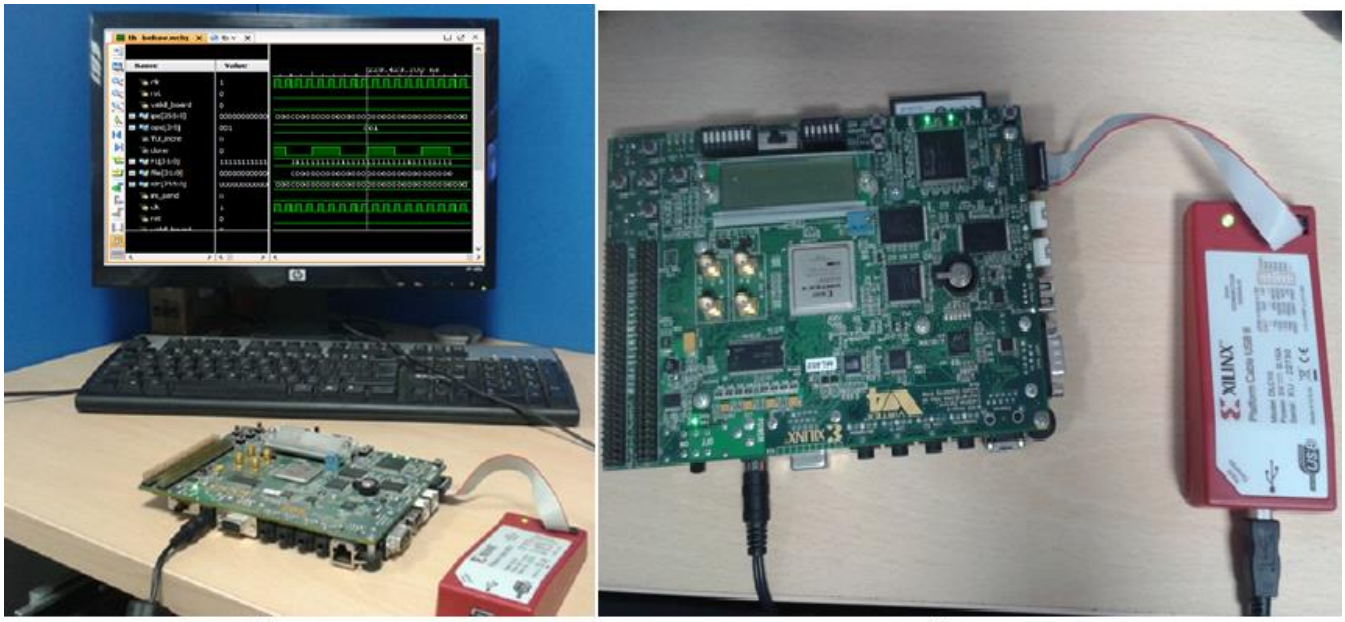


Figure 7: Implementation of a single core consists of 16 inputs and 4 outputs, implying that the weight RAM is 64 rows in depth.



(a) (b)

Figure 8: (a) The bio-inspired computing system setup (b) FPGA ALTERA DE2 Board setup with Xilinx JTAG adapter

The implementation of a single core consists of 16 inputs and 4 outputs, implying that the weight RAM is 64 rows in depth as described in Figure 7. The implementation of bio-inspired computing system is done on Altera Cyclone IV FPGA contained as a part of ALTERA DE2 Board. The VERILOG language was used to program the bio-inspired framework and compiled in Xilinx ISE platform with x86 64 bit CPU executing on Linux Ubuntu 16.04 operating system.

5. Results Obtained

In this section the results obtained with various Intellectual Property (IP) cores which are developed for the modules such as Block RAM, Differential Clock, Floating Point, and First In First Out (FIFO) for the design of the neuron model in Xilinx ISE, with exploration of register transfer logic (RTL) and hardware synthesis using Verilog code are presented.

The Figure 8(a) depicts the complete setup of bio-inspired computing system implementation and the Figure 8(b) represents JTAG adapter connection with FPGA ALTERA DE2 Board. The Figure 9 illustrates the behavioral simulation of developed reconfigurable bio-inspired computing is obtained in Xilinx ISE environment. We can look at output ram data out signal to double check if the data is correct in simulation. The execution is based on vector less activity propagation with peak memory and execution is carried out in Vivado-v-2014.2 FPGA has considerable static power consumption, but normally efficient power is measured as difference in idle state and real-time data processing for the machine.

The equation for obtaining the parameters is described in Equation (1).

% Synaptic Pruning = (Number of Neurons pruned) / (Size of Network x Accuracy x Energy) (1)

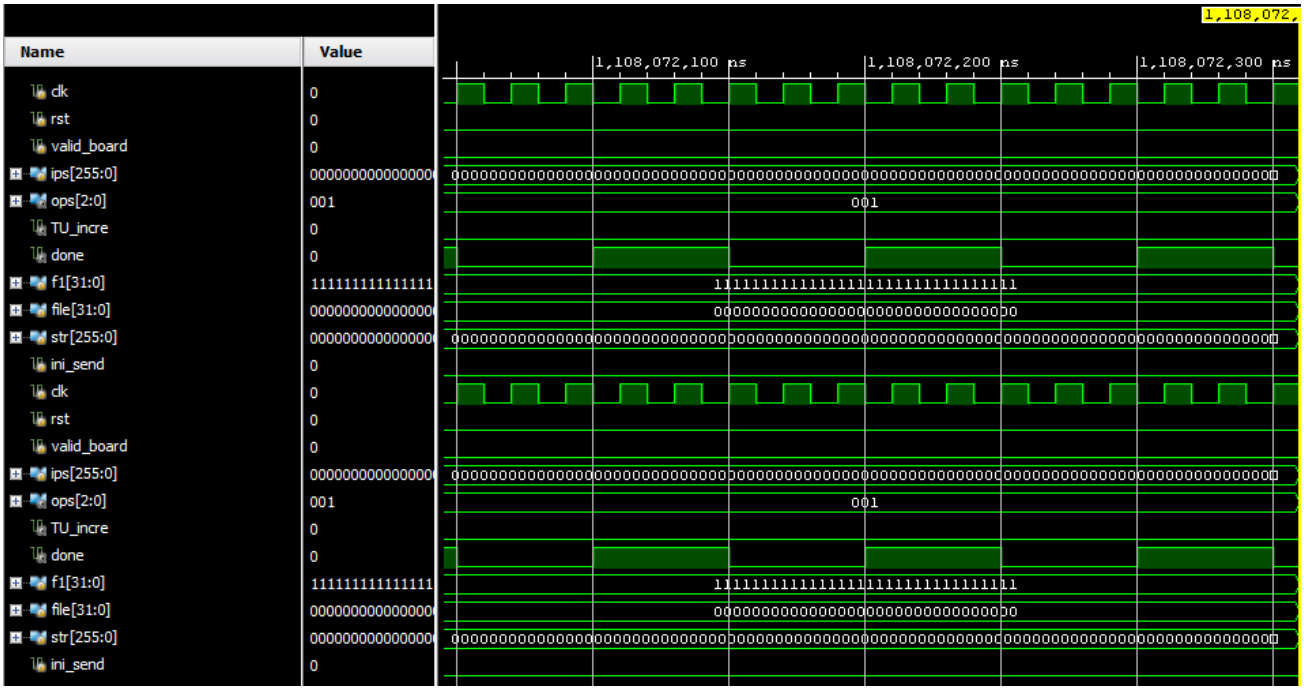


Figure 9: Behavioral simulation of the developed reconfigurable bio-inspired computing architecture

Table 1: Design Trade-off Parameters

Parameters [23], [24], [25]	Existing Neuron Model	Proposed Neuron Model	Efficiency Improvement	Performance Improvement
1. Computational Efficiency	2.3 x Enhanced	2 x Enhanced	0.36 %	89 %
2. Energy efficiency	2.8 x Enhanced	3.2x Enhanced	-	High
3. Throughput	High	High	-	-
4. Accuracy	90.02%	90.4%	1.69 %	90 %
5. Entropy	10% - 25% for 800 Neurons	12 % for 100 Neurons	-	2 %

On the other hand, a poor input output reflects the system glitch at which glue logic in FPGA DE2 board is affected in worst case, as the system changes the incoming events dynamically. The Table 1 depicts the architecture considerations for design at device level and offers the best possible design tradeoff for specific processor architectures and development choices. The parameters are described in Table 1 such as Computational Efficiency, Energy Consumption, Throughput, Accuracy, Entropy are compared with previous work with existing neuron model with the proposed neuron model. It is noteworthy that in the proposed design except LUT all other parameters have the same value for logic utilization post synthesis and post implementation, the reason behind this is the designed system is more of device specific. During Idle mode, the device does not process events and therefore there is no computation. The ADC14DS065/080/095/105 converts the analog data into 14 bit words, but it outputs the data on 1 or 2 serial data lines per channel. The digital output operates at LVCMOS voltage levels except for the serial signals and clock outputs LVDS signals.

These devices operate up to 65 million samples per second (MSPS) in a single lane mode while the higher data rates operate in a dual lane mode, each lane operates at half the data rate to keep the required clock frequencies from being excessive. Using this technique, the FPGA interface can support the highest data rate of 105 (MSPS) with a high throughput as shown in Table 1. The FPGA then will combine the two data streams appropriately to create the correct signals.

The serial data bus uses less board space for the signals, is easier to route and achieves similar data rates to a parallel interface with less wires for data bus. The parallel data bus from the ADC14155 can be connected to the FPGA using an I/O bank configured for 1.8 LVCMOS inputs. The data rate of this bus is 5-155 MHz, which is well within the I/O capabilities of the FPGA.

The designed FPGA module further consists of the blocks required for interfacing an ADC with the FPGA with 3.2x enhanced energy efficiency and 2x enhanced computational efficiency as described in Table 1.

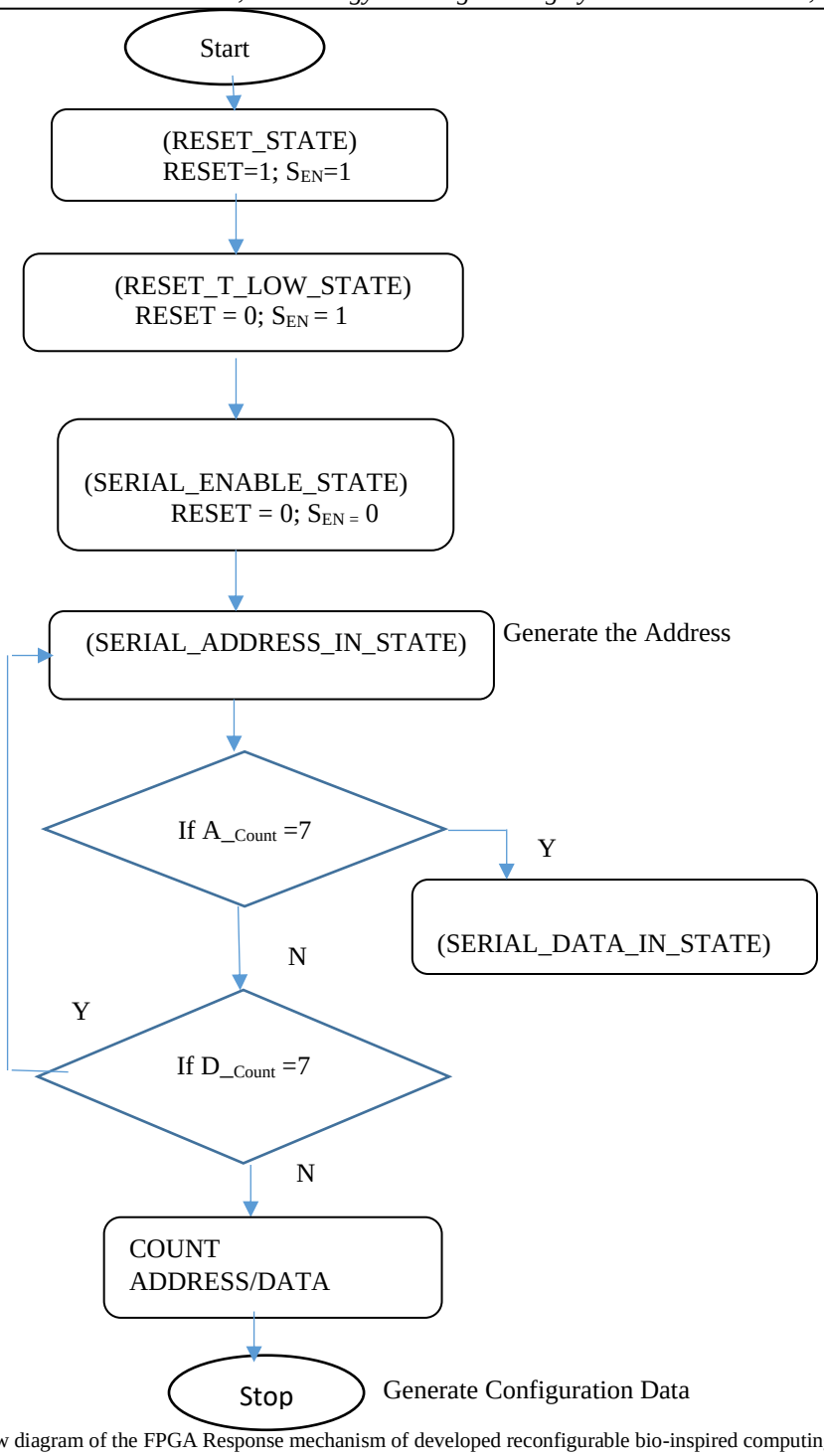


Figure 10: Flow diagram of the FPGA Response mechanism of developed reconfigurable bio-inspired computing architecture

The signals such as RESET and S_{EN} indicate the reset state and serial enable state respectively of FPGA device selected. The serial register reset pulse is greater than 10 ns. If RESET=1; S_{EN}=1 then the FPGA is in the reset state, if the signals RESET=0; S_{EN}=1 then the FPGA is RESET_T_LOW_STATE, and if RESET = 0; S_{EN}= 0, then the FPGA is in serial enable state. The Figure 10 illustrates the flow diagram which forms the basis for algorithm designed to implement a bio-inspired computing system. The $A_{_COUNT}$ is the address count for generating the address and $D_{_COUNT}$ is the data count to generate configuration data. To enable the serial interface of analog to digital converter (ADC), the serial registers were first reset to the default values and the RESET pin was kept low.

Serial enable (S_{EN}), serial data (S_{DATA}), and serial clock (S_{CLOCK}) function as serial interface pins in this mode were used to access the internal registers of the ADC. The registers were reset either by applying a pulse on the RESET pin or by setting RESET bit high.

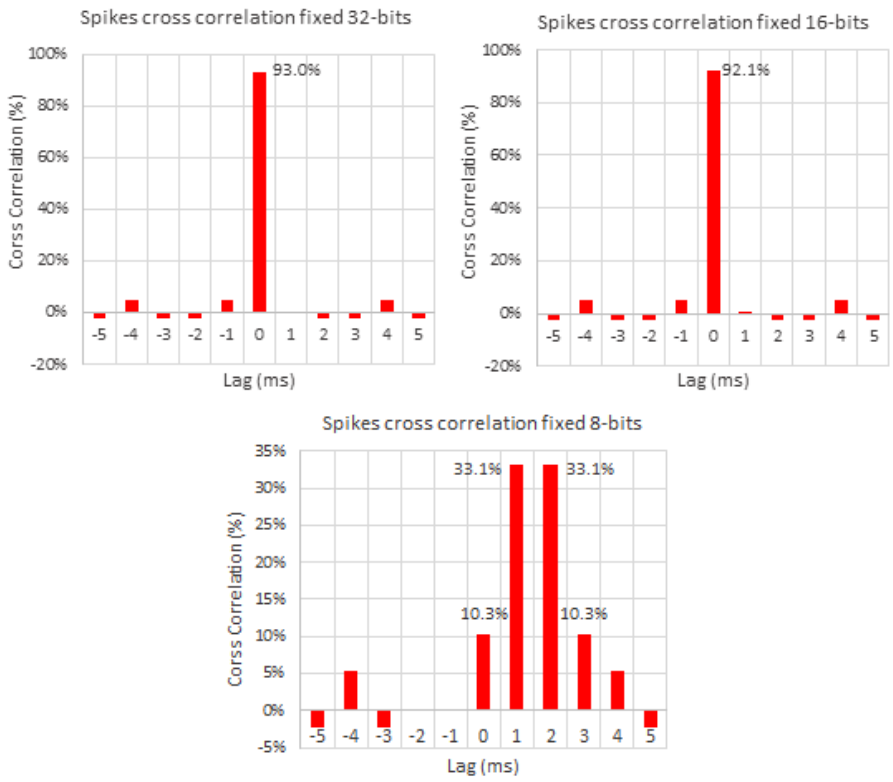


Figure 11: Spike Cross Correlation mechanism of developed reconfigurable bio-inspired computing architecture

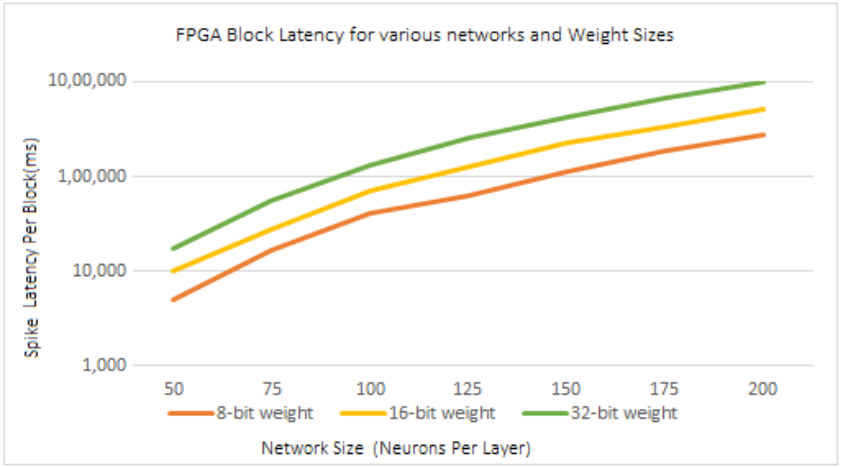


Figure 12: Rate of Change of BRAM generation in Cyclone-V FPGA Block Latency

For the increased flexibility, a combination of serial interface registers and parallel pin controls (CTRL1 to CTRL3) were used to configure the device. To enable this option, RESET pin was configured low. The parallel interface control pins CTRL1 to CTRL3 were available. After power up, the device is automatically configured according to the voltage settings on these pins. In the bus of ADC test for the CYCLONE-V FPGA device, the selected component bus is the 12 channel ADC_RAM component. The data captured through the single clock, the signals included are clock signals such as clk_p, clk_m, the reset signal and the 14 bit input and output data vector signals along with a data type register which is 14 bit vector. The timings of the spikes are analyzed, in a cross correlation of the spikes timings over the FPGA implementation. It is observed that 93% and 92% of the spikes are correlated in the 32 and 16-bits implementation respectively with the implementation with a zero lag delay. On the other hand, 8-bit implementation is slightly different, with 87% of the spike shifted between 0 and 3ms and centered in a 1.5ms shift average with entropy of ~12 % for 100 Neurons with ~2% of performance improvement as described in the Table 1. The analog to digital converter configuration through the serial mode with the signals as seen in spike cross correlation simulation result depicted in the Figure 11. The signals are clock, reset, serial clock, adc_ reset, serial data enable, serial data, and state of the system, the address data consists of a 16 bit vector and these parameters are calculated in terms of rate of change of Block-RAM generation in Cyclone-V FPGA as depicted in Figure 12 and with differential clock simulation time in FPGA based bio-inspired computing module is illustrated in Figure 13.

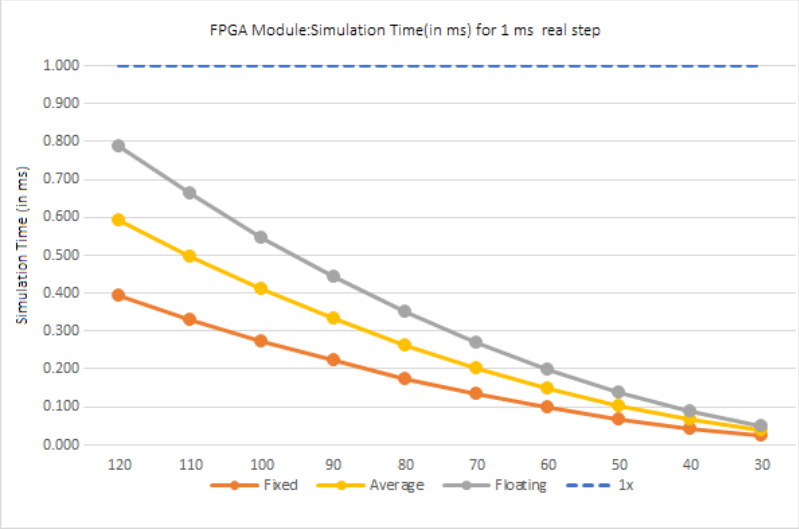


Figure 13: Differential Clock simulation time (in ms) for 1 ms real step in Cyclone-V

6. Conclusions

In this paper the FPGA based algorithmic design of a large scale reconfigurable logical bio-inspired computing model is carried out. The proposed algorithm is implemented on Field Programmable Gate Array to develop a neuron model to be utilized in neuromorphic computing system. The reconfigurable and event driven parameters are considered with the various Intellectual Property (IP) cores are developed for the modules such as Block RAM, Differential Clock, Floating Point, and First in First out (FIFO) for the design of the neuron model in Xilinx ISE, with exploration of register transfer logic (RTL) and hardware synthesis using Verilog code. The architecture for design at device level offers the best possible design tradeoff for specific processor architectures and development choices are summarized. This research paves a way for design of architecture that can be compatible for convolution neural network Artificial Neural Network (ANN).

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The Authors wish to thank School of ECE, REVA University, Bengaluru, India for providing necessary facilities in carrying out this research work.

References

[1] B.U.V. Prashanth, M.R. Ahmed, "FPGA Implementation of Bio-Inspired Computing Architecture Based on Simple Neuron Model," 2020 International Conference on Artificial Intelligence and Signal Processing (AISP), 1–6, 2020, doi:10.1109/AISP48273.2020.9073420.

[2] B.U.V. Prashanth, M.R Ahmed, "FPGA Implementation of Bio-inspired Computing Based Deep Learning Model," In Advances in Distributed Computing Machine Learning, Lecture Notes in Networks and Systems, **127**. Springer, doi: 10.1007/978-981-15-4218-3_24.

[3] R.K. Sarma, M.T. Khan, R.A. Shaik, J. Hazarika, "A Novel Time-Shared and LUT-Less Pipelined Arch. for LMS Adapt. Filter," IEEE Transactions on Very Large Scale Integration Systems, **28 (1)**, 188–197, 2020, doi:10.1109/TVLSI.2019.2935399.

[4] Mohammed Riyaz Ahmed B.K Sujatha, "Reinforcement Learning based on Computational Cognitive Neuroscience in Neuromorphic VLSI Chips," International Journal of Advanced Research in Computer Science and Software Engineering, **2** (8), 2013.

[5] B.U.V. Prashanth, M.R. Ahmed, "Design and performance analysis of artificial neural network based artificial synapse for bio-inspired computing," in Advances in Intelligent Systems and Computing, **1108,** 2020, doi:10.1007/978-3-030-37218-7_135.

[6] C. Pan, A. Naeemi, "Beyond-CMOS non-Boolean logic benchmarking: Insights and future directions," in Proceedings of the 2017 Design, Automation and Test in Europe, 2017, doi:10.23919/DATE.2017.7926971.

[7] X. Lou, Y.J. Yu, P.K. Meher, "Lower Bound Analysis and Perturbation of Critical Path for Area-Time Efficient Multiple Constant Multiplications," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2017, doi:10.1109/TCAD.2016.2584181.

[8] A. Agarwal, L. Bopanna, "Low Latency Area-Efficient D.A Based Multi-Rate Filter Architecture for SDR Receivers," Journal of Circuits, Systems and Computers, **27** (08), 2018, doi:10.1142/S0218126618501335.

[9] T. Xu, A. Fumagalli, R. Hui, "Efficient R.T Digital Subcarrier Cross-Connect (DSXC) Based on D.A DSP Algorithm," in Journal of Lightwave Technology, **38**(13), 3495-3505, 2020, doi: 10.1109/JLT.2019.2937787.

[10] D. Datta, P. Mitra, H.S. Dutta, "FPGA implementation of high performance digital down converter for SDR," Microsystem Technologies, 2019, doi: 10.1007/s00542-019-04579-w.

[11] B.K. Mohanty, P.K. Meher, "An Efficient Parallel DA-Based Fixed-Width Design for Approximate Inner-Product Comp," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, **28** (5), 1221-1229, 2020, doi:10.1109/TVLSI.2020.2972772.

[12] M.R. Ahmed, B.K. Sujatha, "A review on methods, issues and challenges in neuromorphic engineering," in 2015 International Conference on Communication and Signal Processing, ICCSP 2015, 2015, doi:10.1109/ICCSP.2015.7322626.

[13] M. D'arco, E. Napoli, E. Zacharelos, "Digital circuit for seamless resampling adc output streams," Sensors (Switzerland), **20** (6), 2020, doi:10.3390/s20061619.

[14] H. Jiang, C. Liu, L. Liu, F. Lombardi, J. Han, "A review, classification, and comparative evaluation of approximate arithmetic circuits," ACM Journal on Emerging Technologies in Computing Systems, **13** (4), 1–34 2017, doi: 10.1145/3094124.

[15] Md Zahangir Alom, et al., "A state-of-the-art survey on D.L architectures," Electronics, **8** (3), 292, 2019, doi: 10.3390/electronics8030292

[16] S. Fukami, H. Ohno, "Magnetization switching schemes for nanoscale three-terminal spintronics devices," Japanese Journal of Applied Physics, **56**(8), 0802A1, 2017, doi:10.7567/JJAP.56.0802A1

[17] T. Jungwirth, et al., "The multi directions of antiferromagnetic spintronics," Nature Physics, **14** (3), 200, 2018, doi:10.1038/s41567-018-0063-6

[18] C. D. Schuman, et al., "A programming framework for neuromorphic systems with emerging technologies," 4th ACM International Conference on Nanoscale Comp. and Comm., ACM, 2017, doi: 10.1145/3109453.3123958
[19] G. Sanchez, et al., "A highly scalable parallel spike-based digital neuro architecture for high-order fir filters using LMS adaptive algorithm," Neurocomputing, **330**, 425-436, 2019, doi:10.1016/j.neucom.2018.10.029.
[20] M. Sharma, S.K. Singh, "New Tech. Peasant Multiplication for Efficient Signal Processing Applications", Indonesian Journal of Elect. Engineering Comp Science, **8** (3), 726-729, 2017, doi: 10.11591/ijeecs.v8.i3.pp726-729.
[21] W. Zhao et al., "A division-free and variable-regularized LMS-based generalized side lobe canceller for adaptive beamforming and its efficient hardware realization," IEEE Access, **6**, 64470–64485, 2018, doi: 10.1109/ACCESS.2018.2875409
[22] S. Dixit, D. Nagaria, "LMS Adaptive Filters for Noise Cancellation: A Review," International Journal of Electrical and Computer Engineering (IJECE), **7** (5), 2017, doi:10.11591/ijece.v7i5.pp2520-2529.
[23] A. Mahabub, "Design and implementation of cost-effective IIR filter for EEG signal on FPGA," Australian Journal of Electrical and Electronics Engineering, **17 (2)**, 83-91, 2020, doi:10.1080/1448837X.2020.1771662.
[24] B. Khurshid, R.N. Mir, "An Efficient FIR Filter Structure Based on Technology-Optimized Multiply-Adder Unit Targeting LUT-Based FPGAs," Circuits, Systems, and Signal Proc.,**36,** 600–639,2017, doi:10.1007/s00034-016-0312-9.
[25] W. Guo, H.E. Yantır, M.E. Fouda, A.M. Eltawil, K.N. Salama, "Towards efficient neuromorphic hardware: Unsupervised adaptive neuron pruning," Electronics (Switzerland), **9 (7)**, 2020, doi:10.3390/electronics9071059.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 85-92 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design and Implementation of an Ultrasonic Scanner Setup that is Controlled using MATLAB and a Microcontroller

Kamel Fahmi Bou-Hamdan*

University of Aberdeen, School of Engineering, Aberdeen, AB24 3FX, United Kingdom

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 08 February, 2021
Online: 10 March, 2021

Keywords:
Ultrasonic waves
Ultrasonic pulser receiver
Pulse echo
Microcontroller
Surface Imaging
Scanner

ABSTRACT

This paper describes an experimental setup that employs ultrasound to scan an area. This method utilizes ultrasonic waves to scan the surface of a submerged object in a water-coupled medium. A pulse-echo mode is used, and quantitative data are collected at various positions using a two- dimensional automated table. A microcontroller controls the motion of the scanner, whereas a script developed in MATLAB controls the ultrasonic pulser receiver process. The MATLAB script ultimately controls and correlated between the scanner movement and ultrasonic pulser receiver process. The intensities of the reflected waves are captured and used to generate the A-scan image for the external surface. The surface profile of the scanned object can be clearly obtained using the time arrival of the reflected waves. The experimental results based on a one-pound coin indicate that the precision of the proposed process. This simple and efficient method can be used in different engineering applications with minimum errors.

1. Introduction

Ultrasonic waves are commonly utilized in numerous applications such as nondestructive testing (NDT), object recognition, and medical imaging. They have been used as an inspection tool in different materials [1]. Ultrasonic waves can spot any irregularities contained inside the tested specimens that cannot be identified visually. They can be used in different medical devices such as the application of ultrasound imaging for pregnancy [2]. Ultrasonic inspections are listed as a non-destructive testing method because the procedure implemented does not cause any damage to the tested objects [3]. Pulse-echo mode is the most widely used method for inspection. The key benefit of this approach is that only one side of the tested object should be accessible. This broadens its applicability to accommodate for different applications such as crack detection and surface imaging [4].

Several applications include scanning a specimen's surface in order to obtain an image that represents the object. However, this process should include a 2D or 3D automated scanner coupled with the ultrasonic devices [5]. Therefore, a reliable control system should be designed and implemented to ensure correct images are obtained. The surface imaging can be used in different industrial applications [6]. It can also find several applications in the oil and gas industry. The integrity of the oil well cement placement can play a vital role in the success of the drilling operations. Ultrasonic imaging methods can be used for inspecting the cement integrity. This ensures a good sealing layer is obtained between the well and the formation [7]. The cement integrity becomes a critical issue when extended wells are to be drilled [8].

Many applications require using coupling a scanner with another measuring device e.g. ultrasonic pulser-receiver to record real-time data. Commercial tools that do this job are available and can be easily purchased. These tools were used by different researchers in their projects [9]. However, the cost of these tools might be very expensive. Furthermore, the recorded data from the commercial devices might be limited to certain parameters such as amplitude-based data. This limit can be bypassed when the experiment is manually designed, and the script codes are prepared by the user. Some researchers used an ultrasonic device that utilizes LabVIEW for data acquisition [10]. However, in that application, the ultrasonic device was fixed and hence no movement was allowed.

To the best of the author's knowledge, there are not enough published articles that focus mainly on the ultrasonic scanner setup. Most of the articles focus on the application of this method instead. This makes it difficult to replicate these experiments without purchasing expensive equipment and services. In this study, the main focus is on the experimental setup that can produce a two-dimensional image of any surface. This setup utilizes

*Corresponding Author: Kamel F. Bou-Hamdan, kamelbh@live.com

equipment that are of lower cost than the commercial setups used in previous studies. The scanner movement and data acquisition process are ultimately controlled through a custom-written script in MATLAB. A pre-programmed microcontroller is used as an interfacial connection between the scanner and MATLAB. An ultrasonic pulser-receiver device is used as the interface between the transducer and MATLAB. This design might attract researchers who are looking for simple and cost-effective surface imaging methods using ultrasonic waves. All the equipment and control systems are addressed in the text. This article is an extension version for a previously published conference paper [5].

2. Pulse-echo Technique Background

Figure 1a displays the principles of this method. A transmitter is used to send out a pulse wave into the body. When a boundary is reached, e.g. the back wall or a deficit in the material, the wave is reflected back towards the receiver where it is captured and recorded [3]. In general, the pulse-echo mode uses the same ultrasonic transducer for transmitting and receiving the waves, as shown in Figure 1b.

Different types of sound waves propagate in materials, such as longitudinal and shear waves. The latter cannot travel in liquids or gases, unlike the longitudinal waves which can exist in any medium [11]. The speed of sound is not constant but depends on the material it passes through. This speed is roughly 340 meters per second in air, 1500 in water and over 2000 in solids. Acoustic impedance (Z), which describes the resistance of a material to sound waves, depends on the density and speed of sound [3]

$$Z = \rho.c, \quad (1)$$

where c is the speed of sound in the propagating material, whereas ρ is the density of the material.

Ultrasonic waves follow the same principles of acoustic waves. When a wave hits an interface, part of it will undergo reflection while the other part will undergo refraction [3]. The acoustic impedance can be used to determine the reflection and transmission behaviour taking place at the boundaries of two objects in contact. The intensities of the divided waves depend on the material properties at the boundary. Ultrasonic waves of high frequency cannot propagate in air due to their low acoustic impedance; therefore, a wave passing through a solid material will mostly undergo reflection when encountering an air interface. The reflection coefficient (R) can be obtained using [3]

$$R = (Z_2 - Z_1) / (Z_2 + Z_1), \quad (2)$$

where Z_1 and Z_2 represent are the acoustic impedance of the materials in contact. Three possibilities might be encountered when solving (2), as follows:

a. $Z_2 >> Z_1$, e.g. solid-air interface, then R equals to 1 which suggests that all the waves are reflected back towards the receiver;

b. $Z_1 = Z_2$, e.g. two similar materials in contact, then R equals 0, suggesting that all the waves were transmitted to the other end. This means that no signal is read at the receiver.

c. $Z_2 < Z_1$, then R would be a negative value; this change in sign suggests a phase shift in the signal [3].

Furthermore, as sound travels through a medium, the intensity of the wave reduces due to attenuation. Attenuation is a result of two mechanisms: scattering and absorption. Wave scattering describes the spreading of sound waves in different directions besides its original path, whereas absorption refers to the losses due to energy conversion within the material, e.g. mechanical energy converts to heat. The frequency profoundly influences the degree of attenuation and the medium it propagates in [3].

3. Methodology

3.1. Materials

The main equipment that are used in this experiment are: an ultrasonic transducer, an ultrasonic pulser-receiver, a microcontroller, a scanner, and a PC with MATLAB software. All of these equipment are briefly discussed in the following text.

Ultrasonic transducer – This device is used to transmit and/or receive acoustic energy. A focused ultrasonic transducer of 10 MHz rating was selected for this experiment. The major benefits of using a high frequency transducer is to improve the accuracy and quality of the output image. Generally, unfocused transducers are not used for scanning purposes as they cannot attain high resolutions without the use of extra lenses [12]. Passive focusing techniques could be utilized to improve the accuracy of these transducers [13].

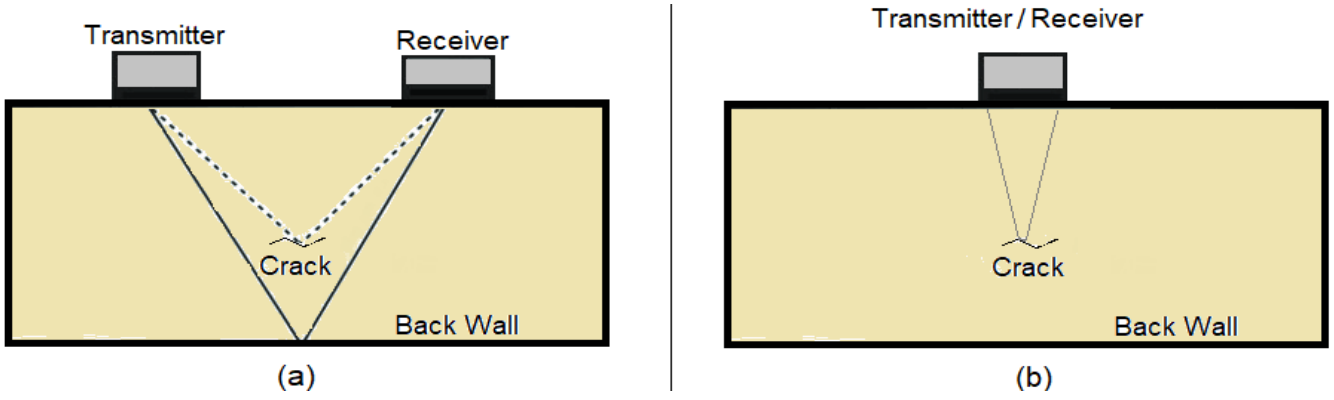


Figure 1: Pulse-echo mode. (a) Using two separate transmitter and receiver transducers; (b) Using a single transmitter/receiver transducer (Transceiver)

Ultrasonic waves of high frequencies cannot pass in air. Hence, a connecting fluid is regularly used to provide a medium that allows for wave propagation [14]. Therefore, the ultrasonic transducer was submerged in standing water for the full extent of the experiment.

In order to focus the waves accurately, the exact distance between the transducer and the object to be scanned should be predefined. This will lead to higher resolution for the scanned image. This distance represents the location of the maximum wave intensity and is identified by reading the focal length rating of the transducer. This value will change with different types of transducers.

The focal spot diameter is a critical parameter to be considered during the design stage since it can impact the resolution. If this attribute has a value lower than that of the scanned object, better resolutions would be achieved. The focal spot diameter (d) can be calculated using the following relation [3]:

$$d = (1.028\, c_w\, l_w) / (1000\, f\, D), \qquad (3)$$

where d is in mm, l_w and c_w area the speed of sound in m/s and focal length in mm in standing water, f is the nominal frequency in MHz and D is the diameter of the transducer in mm. This relation that exists between the focal spot diameter and frequency is represented in Figure 2 for simplicity. The two parameters are inversely proportional, where a smaller focal spot (better resolution) can be attained using high frequency.

The ultrasonic transducer that was used is displayed in Figure 2 as a red mark. It is of 10 MHz frequency, 25.4 mm element diameter, and 51 mm focal length in water. When using these values in (3), the focal spot diameter can be calculated to be around 0.3 mm in water.

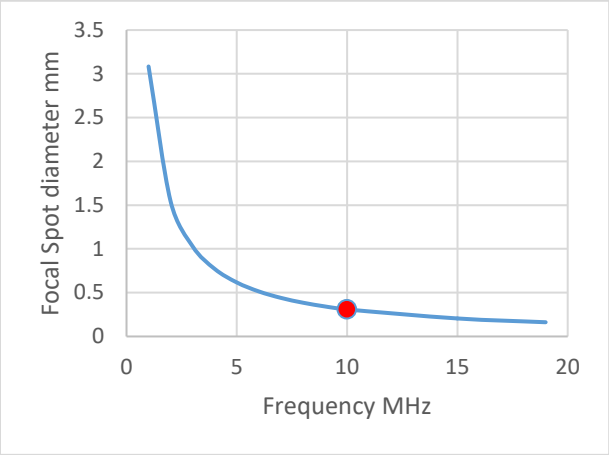


Figure 2: Variations of the focal spot diameter with frequency

Ultrasonic pulser receiver (UPR) - This instrument is employed to provide the excitation voltage for the transducer. It can either be manually designed for a given application or purchased from a manufacturing company. The device which was used in this experiment was produced by Lecoeur Electronique. It contains two channels, a transmitter and a receiver. The first channel can be used to direct the electrical pulses to the transducer which then sends out ultrasonic waves. The second channel receives the electrical signals from the ultrasonic transducer when it detects ultrasonic waves. The ultrasonic pulser receiver can be used in two different modes: either in single mode, as a transmitter or receiver, or in full mode, both as a transmitter and as a receiver.

Automated table (Scanner) – This device is used to guide the transducer in different directions. In this experiment, the scanner can move in two horizontal directions. This is the minimum requirements for scanning an area. The movement is employed by the use a dynamic gantry, that is connected onto a lead screw of 250 mm length. The lead screw has a 2mm pitch of 2 mm and four threads (Four Start). Therefore, the regular actual lead of is 8 mm per revolution.

Stepper motors are used to provide rotation for the lead screw and hence, movement of the gantry. The stepper motors utilized have a high torque value of 114 oz. These motors must undergo 200 steps to complete a full revolution. Hence, a 40 µm linear displacement per step can be easily achieved. Two stepper motor drivers (DRV8825) were used to control the motors. Each driver can deliver a maximum of 45 V output voltage and 2.5 A current. The drivers employ two current sinewaves along with a phase shift of 90˚. This allows a smoother motor operation. Additionally, the drivers can be used in micro-stepping modes. In this experiment, steps of 100 µm where used.

Figure 3 shows the assembled materials. The automated table was fixed to a larger supporting frame. A clamp, made up of polycarbonate material, was fixed to the lower end of the automated table. This clamp was used to hold the ultrasonic transducer. Limit switches were installed at the ends of the lead screw to prohibit any movement beyond a given limit to avoid stalling of the motor.

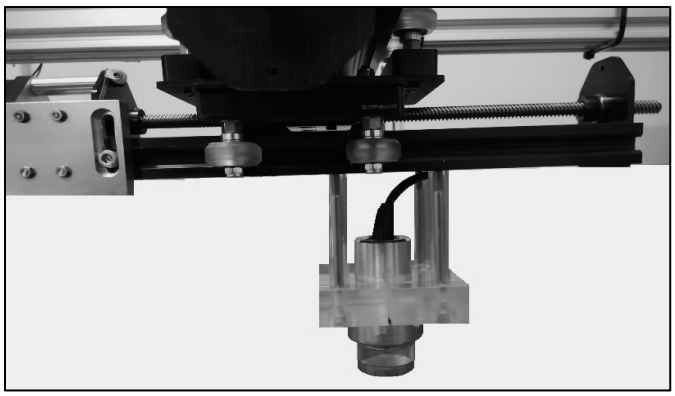


Figure 3: Assembly of the experimental setup

A micro-controller electronic board, that is coded using C++ language, was used to provide suitable control over the stepper motors. In addition to that, it allows for micro-stepping to be attained. This board consists of 6 analogue inputs and 14 digital input/output pins. The connection circuit for the scanner is displayed in Figure 4. The micro-controller is connected to a power supply through an adaptor (Not shown in figure) and to the PC through a USB cable. The microcontroller is connected to the stepper motor drivers using the pins. This allows direct control of the drivers which in turn excite the motors when triggered. A transistor is used in this connection to amplify the electronic signal delivered by the power supply to the motor drivers.

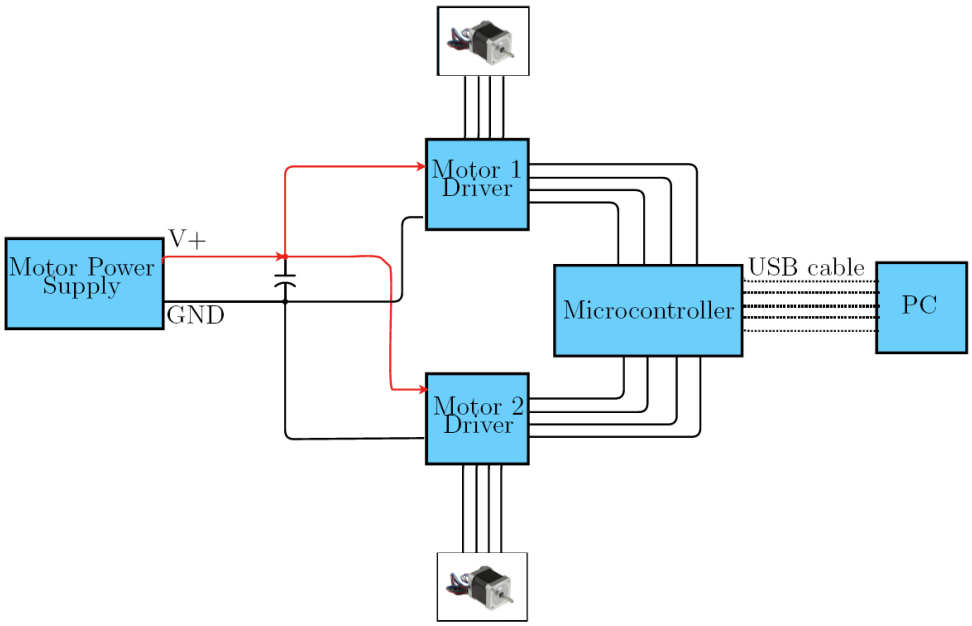


Figure 4: A connection circuit for the scanner.

3.2. Control System

The control system can be divided into 3 main categories: The ultrasonic pulser receiver circuit, the automated table circuit, and the overall system circuit. The control mechanism for these categories are briefly discussed in the following text.

Uultrasonic pulser receiver circuit – The UPR device operates with a script developed in MATLAB. The script enables a series of attributes to be modified including the pulsed frequency, exciting voltage, gain, and others.

The excitation voltage of the ultrasonic pulser receiver is limited to a range of 100 V up to 230 V. The transducer rating can be excited by applying pulsed voltages in from 100 V up to 300 V. This oscillates the inner piezoelectric crystal and produces ultrasonic waves. If the excitation voltage is increased, the reflected signal will record a higher amplitude which can be beneficial when significant wave attenuations are expected.

The gain, which is a signal amplification, is another important parameter that can be regulated to counter the effect of signal attenuation. The recorded signals are not always measurable without tuning the gain value. The ultrasonic pulser receiver is limited to a 80 dB amplification. Nevertheless, laboratory experiments have shown that the measurements become skewed when this is set to 60dB or higher. This is due to increased impact of noise. Therefore, a gain of 40 dB was used in this application.

frequency was fixed at 10 MHz to match the ultrasonic transducer rating. The pulsed width can also be controlled. This width is small, but it can be calculated by analyzing both the amplitude and length of the reflected wave. The assigned pulsed width was nearly 4 µs.

The pulsed repetition frequency is also regulated by MATLAB. To identify the minimum pulsed repetition time needed for the oscillation of the transducer, it is necessary to obtain the time for each reading to be obtained. This attribute considers the time for the wave to propagate within various materials or layers. In the present scenario, the waves travel through standing water. The interval distance in water between the coin and transducer should be regulated to maintain the coin I the range of transducer's focal length.

The total travel time of the wave (*TWT*) should be estimated to determine the pulsed repetition frequency. This parameter takes into account the wave transmission and reflection time in water. To find *TWT*, the interval distance between both the transducer and coin should be defined. The following relationship can be used:

$$TWT = (2\ WD) / c_w, \quad (4)$$

where *WD* is the interval distance / depth of water (51 mm), and c_w is the speed of sound in standing water (1500 m/s). Adding the value of *TWT* from (4), which equals 68 µs, to the pulse width gives the minimum total pulse time needed. The pulse width is much smaller than the *TWT* value. Assuming 200 µs was the pulse repetition time chosen to ensure no more reflected and scattered waves will affect the reading, a 5 kHz pulse repetition frequency would be required. This is achievable using the UPR device. However, the scanning speed is mainly affected by the speed of the automated table, which requires 0.1 seconds interval time to process and act to the movement commands (see text below). This suggests that even a very low pulse repetition frequency can work. In this experiment, a pulse repetition frequency of 1 kHz was chosen as it allows for multiple measurements at each location (lower values could have been used).

Automated table circuit - This circuit is controlled by a microcontroller as shown in Figure 4. A serial interface was developed between MATLAB and the microcontroller using a computer that provides full control over the movement of the device via MATLAB. The script allows for two modes of directional movement:

- To an absolute position,
- Specified number of steps in each direction.

The purpose of this experiment is to scan the area and record the data at each location. As a result, it was found that controlling

the movement by predefined steps would be more appropriate. The total scanning time should be minimized and therefore, the automatic table was designed to scan in a zigzag mode.

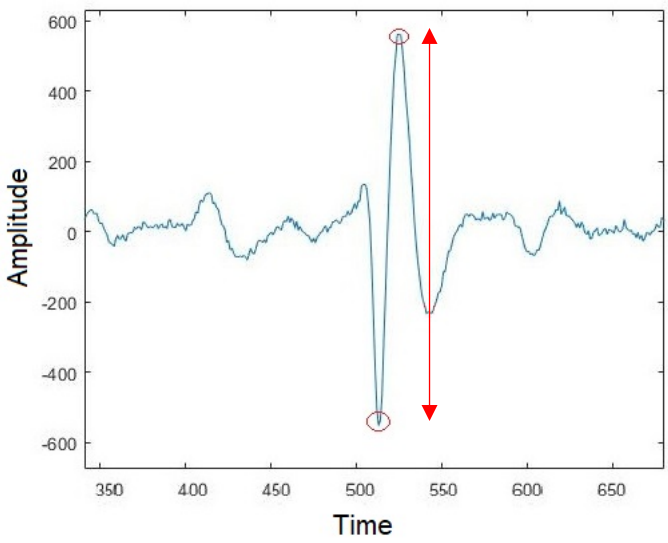


Figure 5: Waveform

In general, data points should be closer to one another in order to improve the resolution of the image. This requires setting short intervals between each scanned location. However, this will lead to an increase in the overall scanning time.

Overall system circuit – MATLAB is employed to provide the main control over the whole system. This approach was achieved by integrating control circuit of the Ultrasonic Pulser-Receiver and scanner. The script has been adapted to obtain three pulses at each position. This is useful in reducing the possible errors during measurement. Figure 5 shows an example of a recorded pulse wave. Figure 6 shows a schematic of the main control process.

The time duration between scanned locations is an important parameter to be determined. This time should be as short as possible without causing biased readings. After several experimental trials, it was found that a time duration of 0.1 seconds was sufficient to maintain accurate data collection.

An A-scan image is a graphical representation of the reflected wave intensity versus time as shown in Figure 5. The script for the UPR can generate this image at every scanned position. This script was designed to collect both the negative and positive peaks of each signal. The amplitude difference, shown in red double-headed arow, is recorded at every location of the scanned object. This data is then inserted into a matrix form where every element in the matrix represents a location. Once the scan is complete, the matrix will hold around 10,000 different elements. The matrix is then utilized to execute an image which represents the scanned object.

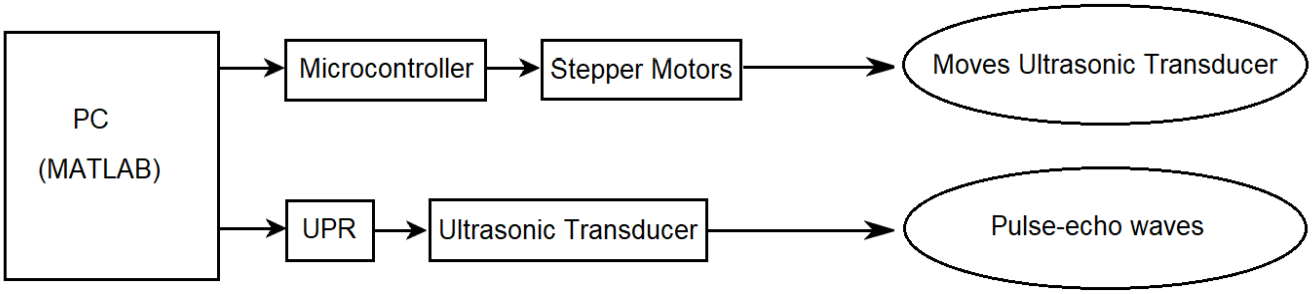


Figure 6: Overall Control System



Figure 7: Scanned image reads accurate coin dimensions

4. Results and Discussion

To check the applicability and efficiency of this method, the experimental design was performed on a one-pound coin. The coin was put inside a water tank. The depth of water above the coin was equivalent to the transducer's focal length of 51 mm. By tracking the amplitude of the ultrasonic pulse, the transducer was then centered on the face of the coin. This process should be done accurately to achieve the best focal spot diameter. A region of 250 mm x 200 mm has been scanned and the image produced is displayed in Figure 7.

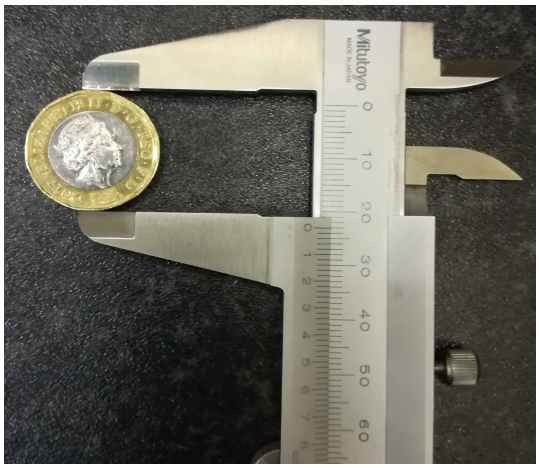


Figure 8 Dimensions of the coin as measured by a caliper.

The results reveal that the original coin is clearly mirrored, which confirms the durability of the ultrasonic transducer chosen. The geometry of the coin was examined and linked to the ideal size of the coin as seen in Figure 6. The scanned figure reveals an exterior surface diameter ranging from 23.3 mm to 23.4 mm. This is very similar to the ideal one-pound coin size of 23.43 mm.

To examine the slight size difference, a caliper was used to measure the same coin that was scanned, and the observations are seen in Figure 7. The measurement taken using the caliper have shown that the real coin size is about 23.3 mm in diameter, which is consistent with the readings obtained from our scanning system. This confirms the precision of the 2D motion of the scanner.

Two methods were used to observe the surface profile of the coin. The first approach was based on the intensity of the reflected wave. This follows the theory that was discussed in (2). Figure 9 shows the intensity of the reflected signals from a planar view. Higher intensities were recorded in the middle part of the coin as that area behaves as a better reflector. This is due to the less irregularities in that area, making it similar to a flat reflective surface.

Figure 10 shows the surface profile of the coin which is based on the intensity of the reflected signal. The boundaries of the coin can be easily distinguished through this view. However, the 3D view of the surface profile might require more scaling to be done to replicate the actual surface of the coin.

Figure 11 shows the surface profile of the coin based on the time of signal arrival. It is clear using this figure that the boundaries and 3D profile of the coin can be well observed and appear to be more realistically distributed. This is because the time approach can be easily linked to actual distances when the speed of sound in the given medium is known, as represented in (4).

These results suggest that both the intensity of the reflected signals and the time of their arrivals can be used to obtain useful data from the surface profile. However, it should be noted that the time approach is more commonly used along with an unfocused transducer, which emits all its waves in a parallel direction. The focused transducer emits waves with different inclination angles for focusing, as done in this experiment. Therefore, it is possible to have minor errors in surface readings when this method is used.

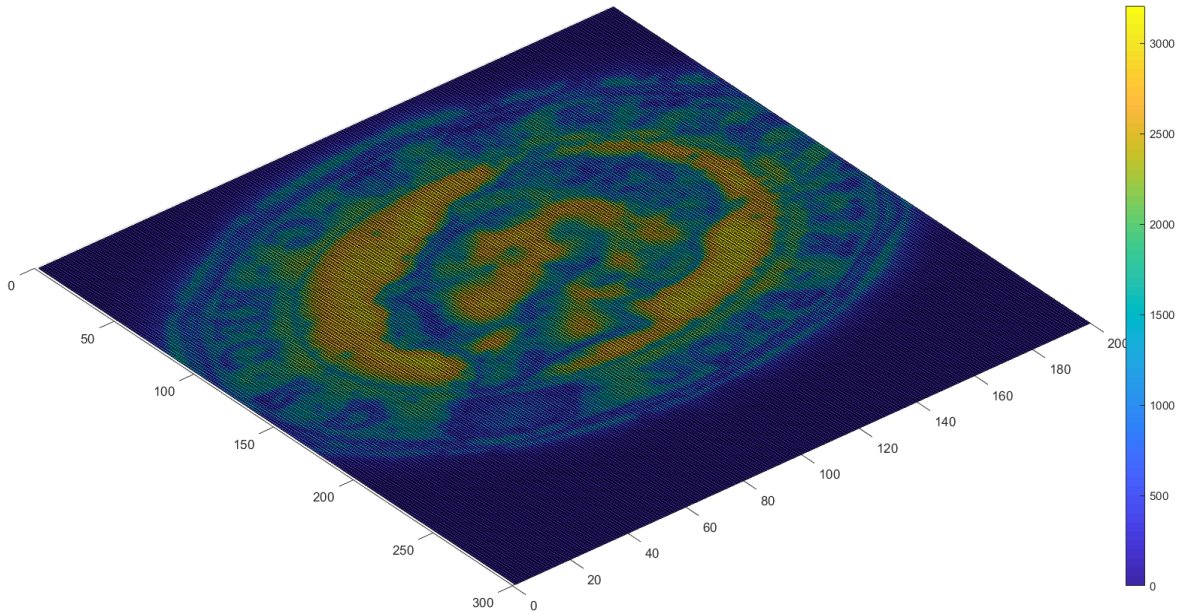


Figure 9 Scanned coin based on intensity of the reflected wave

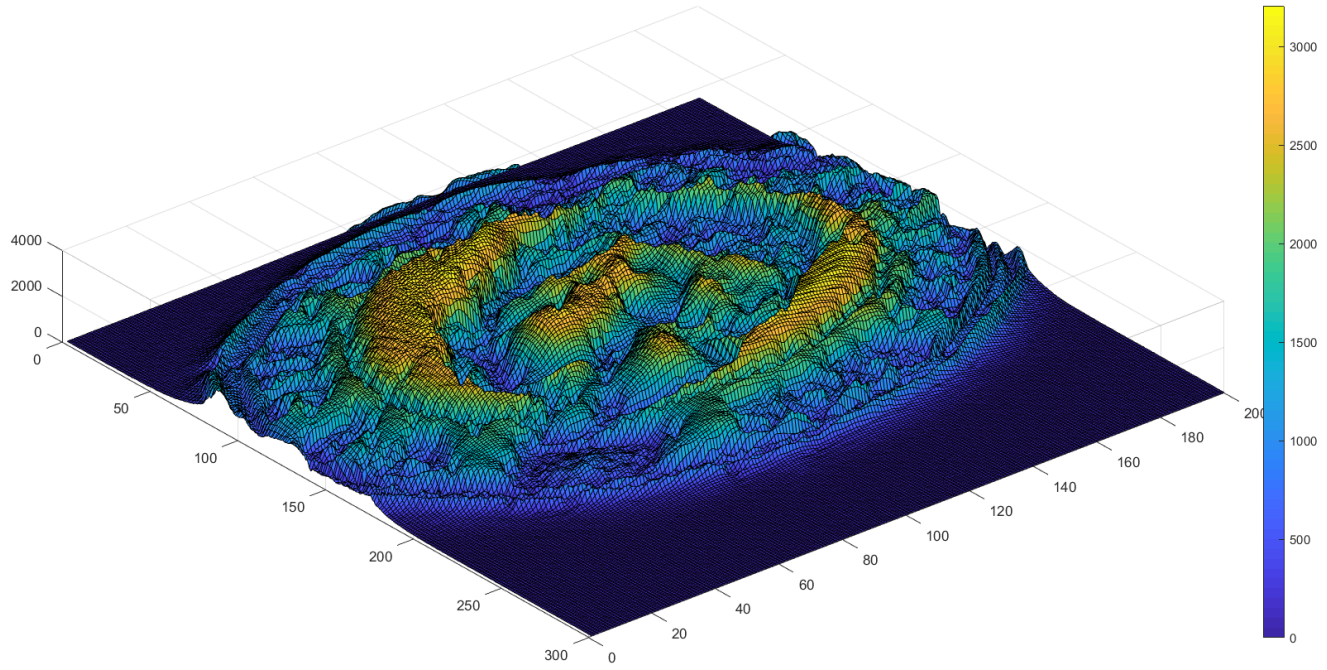


Figure 10:The surface profile of the scanned coin based on intensity of the reflected wave

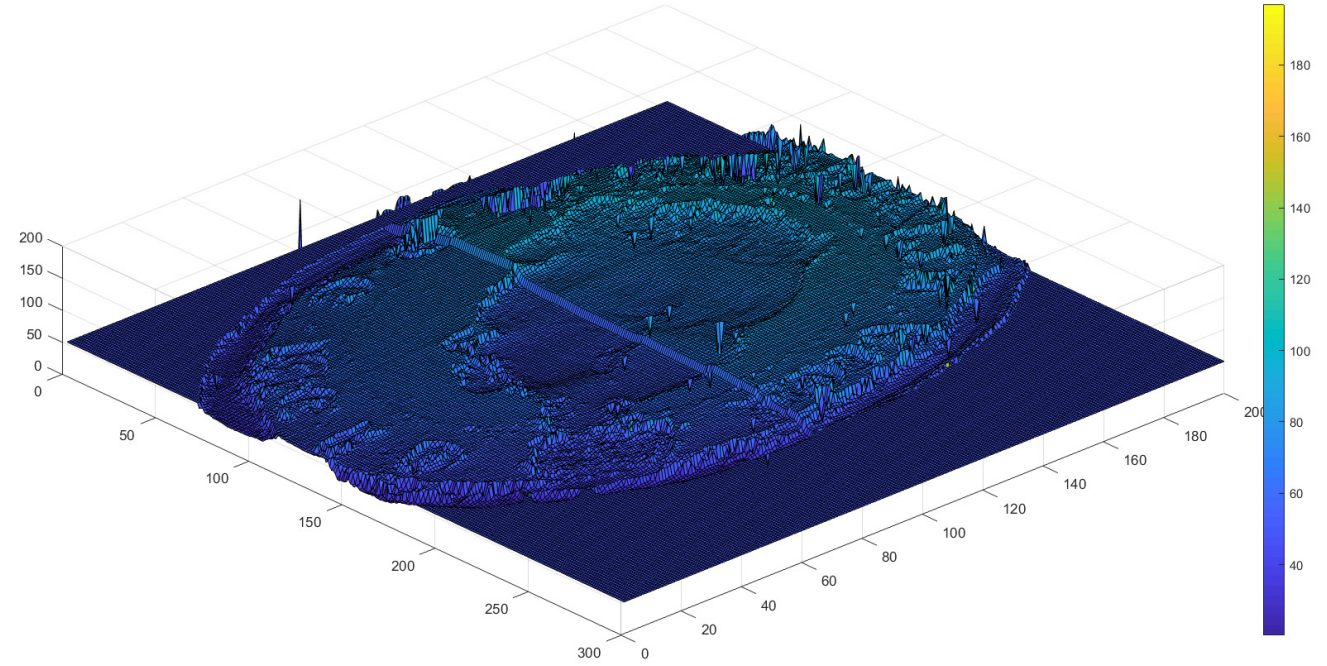


Figure 11: The surface profile of the scanned coin based on time of arrival of the reflected wave

Finally, it is important to consider the wave attenuations during the experiments as it can have a great impact on the obtained results. Attenuation coefficient does not depend only on the medium it passes through, but also on the wave frequency. Higher frequencies can result in larger attenuation coefficients. In standing water, the following relation applies to the attenuation coefficient [3]

$$a = 0.217 f^2 \qquad (5)$$

where a is the attenuation coefficient in dB/m. The relation between both parameters a and f can be shown in Figure 12. The red mark displays the ultrasonic transducer that was used in this experiment and shows an attenuation coefficient of 21.7 dB/m.

By combining both (3) and (5), a relation between focal spot diameter and attenuation coefficient in standing water can be obtained as shown in Figure 13. The signal becomes highly attenuated as the focal spot diameter is decreased. This illustrates one of the main issues of ultrasonic scanning where a reasonable decision should be made when selecting the optimum transducer for the experiment. The optimum case should have a small focal spot diameter to ensure a more accurate reading is obtained. On

the other hand, the attenuation coefficient should not be high enough as it might lead to loss of important parts in the signal.

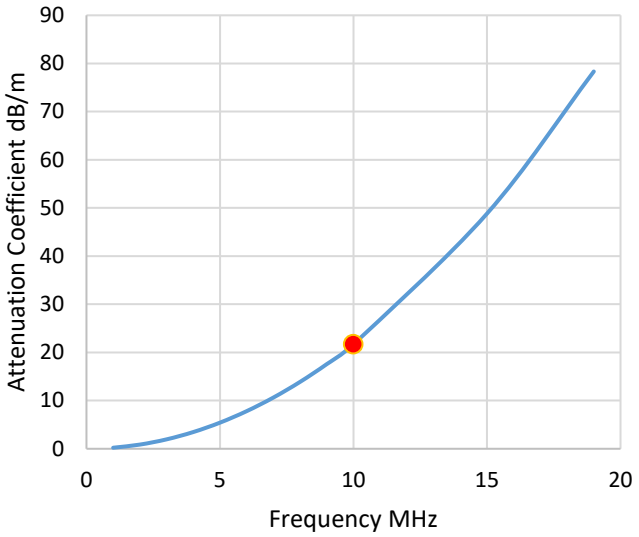


Figure 12: Change in attenuation coefficient with Frequency

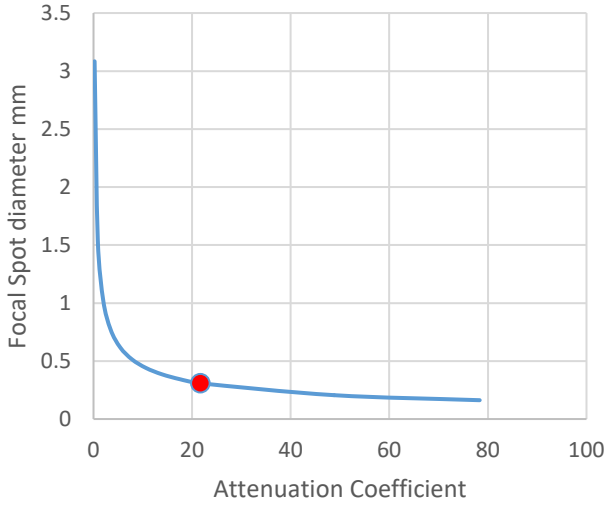


Figure 13: Change in focal spot diameter with the attenuation coefficient

5. Conclusion

This paper provides a low-cost approach to the design of a two-dimensional ultrasonic wave scanner. The reliability of this method depends primarily on the transducer's focal spot diameter and the number of data points collected. Using a high-frequency transducer will improve the focal spot diameter. However, the degree of attenuation should be considered since an inverse relation exists between frequency and attenuation coefficient. This becomes critical when the waves are travelling through materials of higher attenuation coefficient than water.

The total number of collected data points for any scanning operation can be enlarged by reducing the interval distance between two successive datasets. But the total scanning time will significantly increase as the number of scanned data points increase.

A 2D image of a scanned object can be clearly obtained using the intensity of the reflected signals, whereas a 3D image (surface profile) of the same object can be clearly visualized using the time of signal arrival. This scanning approach can be applied to various applications. Future work will utilize this method in measuring the interfacial properties of contacting bodies.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

The author acknowledges the support provided by the University of Aberdeen in assembling the experimental setup.

References

[1] F. Aymerich, S. Meili, "Ultrasonic evaluation of matrix damage in impacted composite laminates," Composites Part B: Engineering, **31**(1), 1–6, 2000, doi:10.1016/S1359-8368(99)00067-0.

[2] C. Moore, S.B. Promes, "Ultrasound in pregnancy," Emergency Medicine Clinics of North America, **22**(3), 2004, doi:10.1016/j.emc.2004.04.005.

[3] H. Krautkrämer Josefand Krautkrämer, Ultrasonic Testing by Determination of Material Properties, Springer Berlin Heidelberg, Berlin, Heidelberg: 528–550, 1990, doi:10.1007/978-3-662-10680-8_34.

[4] J.W. Hunt, M. Arditi, F.S. Foster, "Ultrasound Transducers for Pulse-Echo Medical Imaging," IEEE Transactions on Biomedical Engineering, **BME-30**(8), 453–481, 1983, doi:10.1109/TBME.1983.325150.

[5] K. Bou-Hamdan, "An Experimental Approach that Scans the Surface Area Using Ultrasonic Waves to Generate a Two-Dimensional Image," in 2020 7th International Conference on Electrical and Electronics Engineering (ICEEE), 264–267, 2020, doi:10.1109/ICEEE49618.2020.9102557.

[6] A. Yassin, M.S.U. Rahman, M.A. Abou-Khousa, "Imaging of Near-Surface Defects using Microwaves and Ultrasonic Phased Array Techniques," Journal of Nondestructive Evaluation, **37**(4), 71, 2018, doi:10.1007/s10921-018-0526-9.

[7] S.T. Coelho de Souza Padilha, R.G. da Silva Araujo, "New Approach on Cement Evaluation for Oil and Gas Reservoirs Using Ultrasonic Images," in Latin American and Caribbean Petroleum Engineering Conference, Society of Petroleum Engineers, 1997, doi:10.2118/38981-MS.

[8] K.F. Bou Hamdan, R. Harkouss, H.A. Chakra, "An overview of Extended Reach Drilling: Focus on design considerations and drag analysis," in Mediterranean Gas and Oil International Conference, MedGO 2015 - Conference Proceedings, Institute of Electrical and Electronics Engineers Inc., 2015, doi:10.1109/MedGO.2015.7330328.

[9] F. Aymerich, M. Pau, F. Ginesu, "Evaluation of Nominal Contact Area and Contact Pressure Distribution in a Steel-Steel Interface by Means of Ultrasonic Techniques.," JSME International Journal Series C, **46**(1), 2003, doi:10.1299/jsmec.46.297.

[10] K. Hodgson, R.S. Dwyer-Joyce, B.W. Drinkwater, "Towards an ultrasonic mapping of the contact pressure within an engineering component," in Tribology Series, Elsevier: 79–85, 2001, doi:10.1016/s0167-8922(01)80095-x.

[11] W. Lowrie, A. Fichtner, Fundamentals of Geophysics, Cambridge University Press, 2020, doi:10.1017/9781108685917.

[12] K.I. Maslov, L.M. Dorozhkin, V.S. Doroshenko, R.G. Maev, "A new focusing ultrasonic transducer and two foci acoustic lens for acoustic microscopy," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, **44**(2), 380–385, 1997, doi:10.1109/58.585122.

[13] T.E. Gómez Álvarez-Arenas, J. Camacho, C. Fritsch, "Passive focusing techniques for piezoelectric air-coupled ultrasonic transducers," Ultrasonics, **67**, 85–93, 2016, doi:10.1016/j.ultras.2016.01.001.

[14] M.T. Balmaseda, M.T. Fatehi, S.H. Koozekanani, A.L. Lee, "Ultrasound therapy: A comparative study of different coupling media," Archives of Physical Medicine and Rehabilitation, **67**(3), 1986, doi:10.1016/0003-9993(86)90052-3.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 91-99 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

New Neural Networks for the Affinity Functions of Binary Images with Binary and Bipolar Components Determining

Valerii Dmitrienko, Serhii Leonov*, Aleksandr Zakovorotniy

Modeling, System Control and Artificial Intelligence Team, Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, 61002, Ukraine

ARTICLE INFO

Article history:
Received: 01 December, 2020
Accepted: 22 June, 2021
Online: 10 July, 2021

Keywords:
Hamming neural network
Generalized architecture of the Hamming neural network
Maxnet network
Recognition and classification problems
Information technologies
Computer systems

ABSTRACT

The Hamming neural network is an effective tool for solving problems of recognition and classification of objects, the components of which are encoded using a binary bipolar alphabet, and as a measure of the objects' proximity the difference between the number of identical bipolar components which compared include objects and the Hamming distance between them are used. However, the Hamming neural network cannot be used to solve these problems if the input network object (image or vector) is at the same minimum distance from two or more reference objects, which are stored in the weights of the connections of the Hamming network neurons, and if the components of the compared vectors are encoded using a binary alphabet. It also cannot be used to assess the affinity (proximity) binary vectors using the functions of Jaccard, Sokal and Michener, Kulchitsky, etc. These source network Hamming disadvantages are overcome by improving the architecture and its operation algorithms. One of the disadvantages of discrete neural networks is that binary neural networks perceive the income data only when it's coded in binary or bipolar way. Thereby there is a specific apartness between computer systems based on the neural networks with different information coding. Therefore, developed neural network that is equally effective for any function of two kinds of coding information. This allows to eliminate the indicated disadvantage of the Hamming neural network and expand the scope of discrete neural networks application for solving problems of recognition and classification using proximity functions for discrete objects with binary coding of their components.

1. Introduction

Hamming distance R_X between arbitrary binary vectors $J_p = (j_{p1}, j_{p2}, \ldots, j_{pn})$, $J_q = (j_{q1}, j_{q2}, \ldots, j_{qn})$ is determined by the number of binary digits in which the compared vectors do not match [1]. This distance is often used to assess the proximity of discrete objects, which are described using binary alphabets with binary {0, 1} or bipolar {–1, 1} components. On the basis of bipolar neurons, a discrete Hamming neural network has been developed [1], which is successfully used to solve problems of recognition and the proximity measure of binary objects, the components of which are encoded using alphabet elements {–1, 1}, estimation. However, the evaluation of the measure of proximity of the compared objects (binary bipolar vectors) is carried out only by the most noticeable signs - different binary digits. At the same time, more "subtle" features for compared objects are ignored, which are often used to assess the similarity of binary vectors using the affinity (similarity) functions of Russell and Rao, Dice, Kulchitsky, Yule, etc.

In particular, in works [2–5] a brief description of these functions and their application is provided to compare different objects. In studies [2] they are used in the development of artificial immune systems, in works [3] and [4] they are used in the development of recognition systems for various objects. In this case, the characteristics of such objects are encoded by binary features. In work [5] given function is used to classify microorganisms.

In works [2–5], it is described and used not more than 15 commonly used functions. Although their total number is noticeably higher. So, in research in mathematical biology [6], 20 functions are used, and with hierarchical clustering, the set of random binary data (Hierarchical clustering Result of Random Binary Data Set) 76 different functions are indicated [7]. Once

*Corresponding Author: Serhii Leonov, Email: serleomail@gmail.com

again it indicates the relevance of the application of this approach for the classification and assessment of the proximity of various objects, which is described by a plurality of binary features.

2. Problem Statement

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

Analysis of recent publications on the theory and application of neural network Hamming [1 – 10] shows that the Hamming network has distinct disadvantages. In the architecture of the Hemming neural network, highly specialized layers of neurons are used: neuron layer calculating only the dot product of binary bipolar vectors; a neural network that allows you to select only one reference image from the network memory. This image is at the minimum distance from the input image in connection with this network

– It cannot work with the binary input data, when the measure of proximity of objects is estimated using finer characteristics than the Hamming distance. In particular, using the distances [2 – 5, 7];

– It cannot recognize objects, if they are at the same minimum distance from the two or more reference objects [1];

– cannot switch from processing information encoded using a bipolar alphabet to processing information presented using a binary alphabet and vice versa.

All this requires the development of a new Hemming neural network with a more advanced architecture.

When comparing objects with qualitative features encoded using a binary alphabet, for each pair of objects (binary vectors) $J_p = (j_{p1}, j_{p2}, \ldots, j_{pn})$, $J_q = (j_{q1}, j_{q2}, \ldots, j_{qn})$ in works [2 –5] four variable: a, b, f, g are used (table 1).

The variable a is used to count the number of bits with coinciding unit components of the vectors J_p and J_q. The variable b is needed to determine the number of binary digits of vectors J_p and J_q., in which both vectors have zero components, at the same time that means these objects do not have encoded with zeros in the same bits features.

The number of unit components that an object (vector) J_p has, but vector J_q. does not have, is determined using the variable f. The variable g, on the contrary, counts the number of unit components that the vector J_q. has, but are absent in the second vector J_p.

From the analysis of the Table. 1 it follows that an increase in the variable a unambiguously indicates an increase in the similarity (affinity) of the compared vectors. There is no such unambiguity with an increase in the variable b, since an increase in this variable can indicate both an increase in the similarity of the compared objects, and their belonging to different classes (if there is no similarity in the variable a). As for the measure of proximity of vectors J_p and J_q with respect to the variables f and g, it is symmetric with respect to these variables. An increase of these variables indicates the differences between the compared vectors increase.

Table 1: Variables for comparing binary vectors with binary components

	J_p	
J_q	1	0
1	$a = \sum_{k=1}^{n} j_{qk} j_{pk}$	$g = \sum_{k=1}^{n} (1 - j_{pk}) j_{qk}$
0	$f = \sum_{k=1}^{n} (1 - j_{qk}) j_{pk}$	$b = \sum_{k=1}^{n} (1 - j_{qk})(1 - j_{pk})$

Variables a, b, f and g from Table 1 are used in a number of well-known binary object similarity functions, in which the presence or absence of features is determined by the components of the binary alphabet. As an example, we can note the affinity (similarity) functions of Sokal and Michener, Jaccard and Needham, and so on. Using the variables f and g, one can easily determine the Hamming distance between binary vectors J_p and J_q

$$R_X(J_q, J_p) = f + g. \quad (1)$$

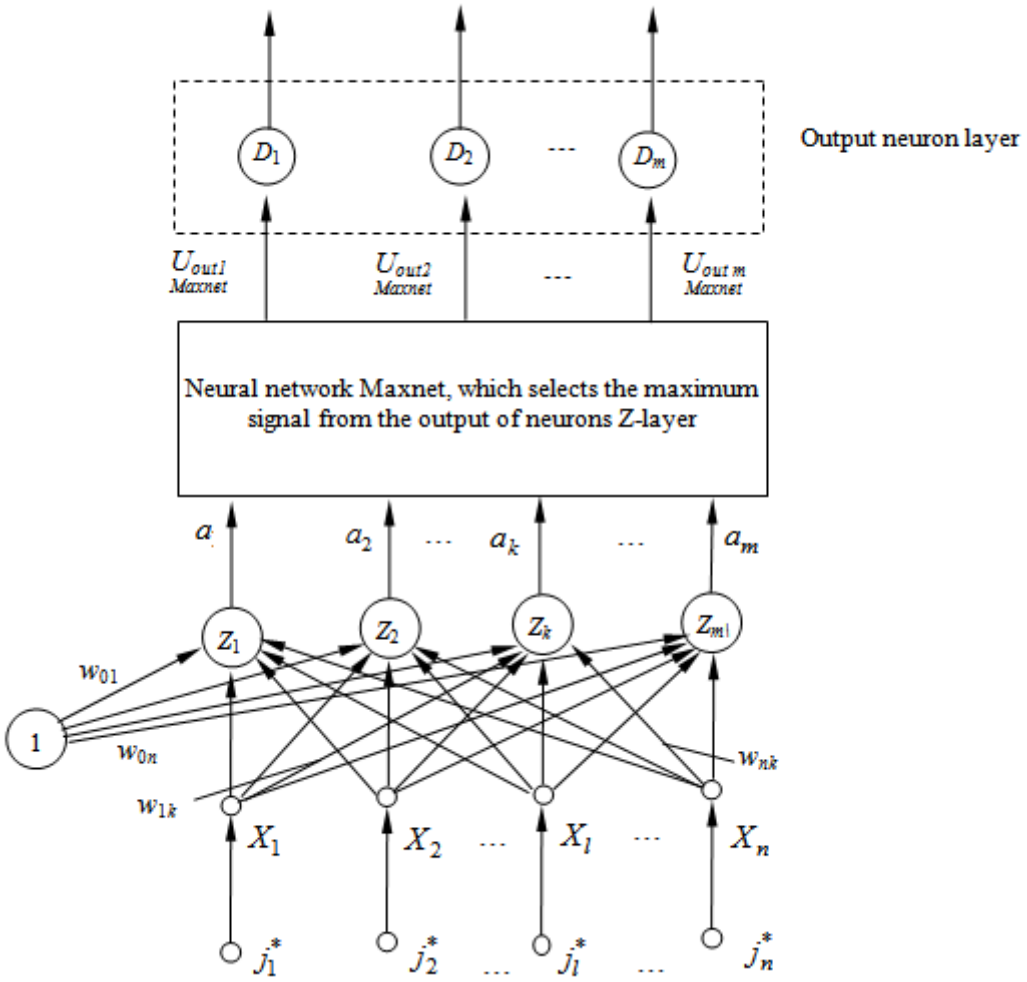
Variables a, b, f and g, as well as the affinity functions of Sokal, Michener, Jaccard, Needman, Hamming distance can be calculated using a new neural network based on binary neurons. The disadvantage of this neural network is the impossibility of processing information presented using the bipolar alphabet, as well as the inability to calculate the Hamming distance using the above affinity functions of Sokal and Michener, Jaccard, etc.

3. Generalized Block Diagram of a Hamming Neural Network

Since the theory, architecture and algorithms for the functioning of the Hamming network are described in detail in modern literature [1, 6 – 10], we will give only a generalized block diagram of the neural network (Figure 1) and a brief description of its functioning.

In the Hamming neural network, the input image $J^* = (j_1^*, j_2^*, \ldots, j_n^*)$ is fed to the inputs of X-neurons, which are multipliers of the bipolar image components J^*. Each Z-neuron stores one of the m reference images in the scales of its relations $J_k = (j_{k1}, j_{k2}, \ldots, j_{kn})$, $k = \overline{1, m}$ and compares this image with the input $J^* = (j_1^*, j_2^*, \ldots, j_n^*)$. As a measure of the similarity of the compared images (bipolar vectors), their scalar product is used [1, 8]

$$J_k J^* = \sum_{i=1}^{n} j_{ki} j_i^* = a_k - b_k, \quad (2)$$



where a_k and b_k are, respectively, the number of identical and different components of the compared vectors (images) J_k and J^*.

Since $a_k + b_k = n$, the scalar product (2) can be written in the form

$$J_k J^* = 2a_k - n\,. \tag{3}$$

From relations (2) and (3) it is easy to obtain

$$a_k = \frac{n}{2} + \frac{1}{2} J_k J^* = \frac{n}{2} + \frac{1}{2} \sum_{i=1}^{n} j_{ki} j_i^*\,. \tag{4}$$

Relation (4) can be regarded as a description of the input signal of a neuron Z_k (Figure 1), which has a displacement of $n/2$ and n inputs, that perceive the components of the input image $J^* = (j_1^*, j_2^*, \ldots, j_n^*)$ through neurons X_l ($l = \overline{1, n}$). In this case, the output signals of the X-layer elements repeat their input signals:

$$U_{outXl} = U_{inpXl} = j_l^*,\ l = \overline{1, n}, \tag{5}$$

where

$$U_{outXl} = \begin{cases} +1, & if \quad j_l^* = 1, \\ -1, & if \quad j_l^* = -1. \end{cases} \tag{6}$$

Each element of the X-layer is associated with the input of each neuron of the Z-layer. The scales of these relations contain information about m reference images $J_1, J_2, \ldots, J_m$, stored in the neural network. The scales of relations $w_{1k}, w_{2k}, \ldots, w_{nk}$ contain information about the k-th reference image J_k.

Z-layer elements compute their output signals

$$U_{outZk} = g_Z(U_{inp\ Z_k}) = a_k,\ k = \overline{1, m},$$

where g_Z – function of activation of neurons of the *Z*-layer; a_k – output element's Z_k signal, that arrives at the input of the neuron of the Maxnet network [1, 8]. From the output signals of the *Z*-layer elements, the Maxnet neural network, using an iterative process, extracts a single maximum signal greater than zero. Since the output signals of *D*-layer neurons are calculated by the ratio

$$U_{outDi} = \underset{\text{Maxnet}}{g_z(U_{out\,A_i})} = \begin{cases} 1, if \;\; U_{inpDi} > 0, \\ 0, if \;\; U_{inpDi} \le 0, \end{cases}, \; i = \overline{1, m}, \quad (7)$$

then at the output of the *Z*-layer of *D*-elements there will be only one single signal, which indicates which reference image is closest to the input vector J^*.

Analysis of the architecture of the discrete neural network shown in Figure 1 shows, that the layers of the network elements function relatively independently of each other. In particular, the layer of *Z*-neurons that determine the similarity functions of the input and reference images using the dot product can be replaced by blocks or neurons that calculate the similarity functions using other relations. In particular, the input bipolar signals and neurons of the *X*- and *Z*-layers can be replaced with binary.

The relative independence of the functioning of the main layers of the Hamming neural network from each other makes it possible to propose the use of binary neurons in the first layer of the network to work with binary input images (vectors). And also, to propose more "subtle" comparisons of input and stored in memory images using a variety of similarity (affinity) functions, modified using the Hamming distance (1) R_X.

The purpose of the article is to develop a generalized architecture of the Hemming neural network, which allows you to process input information that can be specified using both bipolar and binary alphabets. Also, based on the new architecture, the development of methods for solving problems of recognition and classification of binary objects is shown. In this case, the Hamming distance is used, as well as the functions of Jaccard, Sokal and Misher, Kulchinsky, etc.

4. Analytical Relationships Connecting Similarity Functions for Binary Objects and Hamming Distance

Since similarity functions for binary vectors (objects) and for objects described using the bipolar alphabet are used to assess the similarity of binary objects, then there should be dependencies connecting the similarity functions for binary objects (vectors) and the Hamming distance. Table 2 shows the aforementioned classical affinity functions (first column of Table 2) and similarity functions, modified using the Hamming distance (second column of Table 2).

Simple analytical relationships for determining the Hamming distance through the classic similarity functions Russel and Rao,

Table 2: Relationship between Hamming Distance and Similarity Functions for Binary Objects

Classic similarity functions	Similarity functions using Hamming distance	Hamming distance expressed in terms of binary similarity functions
Russell and Rao similarity function $S_1 = a/(a+b+f+g)$	$S_1 = \frac{a}{a+b+R_X}$	$R_X = \frac{a}{S_1} - (a+b)$
Sokal and Michener similarity function $S_2 = \frac{a+b}{a+b+f+g}$	$S_2 = \frac{a+b}{a+b+R_X}$	$R_X = \frac{a+b}{S_2} - (a+b)$
Jaccard and Needham similarity function $S_3 = a/(a+f+g)$	$S_3 = \frac{a}{a+R_X}$	$R_X = \frac{a}{S_3} - a$
Kulchitsky similarity function $S_4 = \frac{a}{f+g}$	$S_4 = \frac{a}{R_X}$	$R_X = \frac{a}{S_4}$
Dice similarity function $S_5 = a/(2a+f+g)$	$S_5 = \frac{a}{2a+R_X}$	$R_X = \frac{a}{S_5} - 2a$
Yule affinity function $S_6 = (ab-fg)/(ab+fg)$	Bulky expression	$R_X = f+g$
Correlation $S_7 = \frac{ab+fg}{[(a+f)(a+g)(b+g)(b+f)]^{1/2}}$	Bulky expression	$R_X = f+g$

Table 3: The Values of the Similarity Functions for the Compared Binary Vectors

	a	b	f	G	S_1	S_2	S_3	S_4	S_5	S_6	S_7	R_X
Values of similarity functions for vectors J_1 and J_k	5	5	2	0	5/12	10/12	5/7	5/2	5/12	1	5/√35	2
Values of similarity functions for equal vectors $J_1 = J_k$	7	5	0	0	7/12	1	1	∞	5/12	1	1	0
Numerical values of similarity functions for vectors J_1 and J_k with opposite components	0	0	7	5	0	0	0	0	0	– 1	1	12

Sokal and Michener, Jaccard and Needham, Kulchitsky, Dice were shown in the third column of the Table 2 for the first time ever. It was not possible to obtain simple analytical dependences connecting the Yule similarity function S_6 and the correlation coefficient S_7. But it is possible to determine the Hamming distance by functions S_6 and S_7 can using the relation (1).

Example. Let's perform the comparison using functions a, b, f and g, $S_1 - S_7$ some vector's pairs:

1) $J_1 = (111001110001)$, $J_k = (101001100001)$;

$J_1 = J_k = (111001110001)$;

$J_1 = (111001110001)$; $J_k = (000110001110)$.

The calculation results are shown in Table 3.

Analysis of the calculation results shown in Table 3 shows that the similarity functions $S_1 - S_5$ depend on the compared vectors. The values of these functions for vectors with opposite components take the minimum values, and for vectors with the same components, on the contrary, the maximum values.

If the compared vectors J_1 and J_k do not all coincide and not all are opposite in their components, then the similarity functions $S_1 - S_5$ take their values between the data of the second and third rows of Table 3, that is, between its maximum and minimum values. The function S6 practically obeys this rule, and the calculated values of which, given in the first and second lines, coincide.

As follows from the data Table 3 and the third column of the Table 2, the Hamming distance in comparison with the similarity functions $S_1 - S_5$ behaves in a certain sense opposite: using the functions $S_1 - S_6$ the similarity of the compared objects (vectors) is calculated, and using the Hamming distance, on the contrary, the difference of objects (vectors), therefore the Hamming distance R_X takes the maximum value for vectors that do not have any of the same components and the minimum distance (equal to zero) for vectors in which all components are the same.

The affinity functions are not limited to the ratios given in Table 2 [2 – 5]. A significant number of these similarity functions indicate the absence of one universal function suitable for comparing any binary objects with binary information encoding. The choice of one similarity function for solving a specific problem is determined, as a rule, by repeated modeling on the initial data in the space of features encoded using a binary alphabet.

5. New Neural Networks that Recognize Binary Images with Binary Components

The generalized block diagram of the Hamming neural network in Figure 1 clearly shows that the three main layers of a neural network can change and function almost independently of each other. In particular, instead of the Maxnet network, another neural network can be used that selects one or several identical maximum signals (if any). This allows you to find reference images that can be at the same minimum distance from the input. It also follows from the network architecture (Figure 1) that the layer calculating the measure of proximity of the input and reference images can calculate various measures of proximity. The relative independence of the main layers of the Hamming neural network makes it possible to propose neural networks for working with binary images (vectors) and using various affinity functions given in Table. 2, between the input and reference binary images. The Hamming distance can also be used (1). The classical architecture of the Hamming neural network [1, 6] for calculating the closeness measure of the input and reference vectors involves the use of the scalar product of two bipolar vectors [1, 8], one of which is a vector of connection scales, and the other is an input vector. In this case, the sum of the products of bipolar components does not contain any zero terms. If the scalar product of two binary vectors J_q and J_p is calculated in the classical Hamming neural network, then it is calculated as follows

$$J_q J_p = \sum_{i=1}^{n} j_{qi} j_{pi} = (a+b) - (f+g),$$

where variables (a, b, f, g) are defined by relations from Table 1.

Thus, if only one neuron is required to calculate the scalar product of two bipolar vectors using a classical neural network, then to calculate the scalar product of two binary vectors, you must first calculate the variables a, b, f and g using the relations given in Table 1.

In Figure 2 a block for calculating the variable b from Table 1 via neuron Z is shown:

$$b = \sum_{k=1}^{n} (1 - j_{qk})(1 - j_{pk}), \quad (8)$$

where $(1-j_{qk})$, $(k=\overline{1,n})$ – scales of neuron relations; $(1-j_{pk})$, $(k=\overline{1,n})$ – neuron Z input vector components; n – number of neuron relations scales Z.

With the help of neurons ∑1, ∑2, …, ∑n the components $(1-j_{p1}),, (1-j_{p2}),, \ldots, (1-j_{pn})$ of the input vector for neuron Z are computed. These components are accordingly communicated with the scales $(1-j_{q1})$, $(1-j_{q2})$, …, $(1-j_{qn})$, which allows us to calculate the variable b.

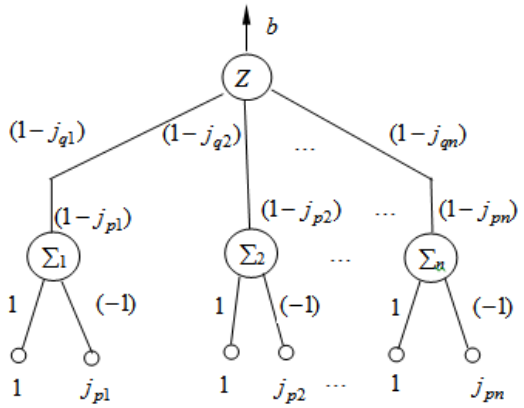


Figure 2: Neuron to calculate a variable b

If in Figure 2 to the inputs of neuron Z to apply input signals j_{p1}, j_{p2}, …, j_{pn} instead of signals calculated by summing elements $\sum_1$, $\sum_2$, …, $\sum_n$ of signals $(1-j_{p1})$, $(1-j_{p2})$, …, $(1-j_{pn})$, then neuron Z will calculate the variable f. To calculate the variable g (see Table 1) in Figure 2 it is necessary to replace the scale coefficients $(1-j_{p1})$, $(1-j_{p2})$, …, $(1-j_{pn})$ with the corresponding scale coefficients j_{p1}, j_{p2}, …, j_{pn}.

To obtain from the neural network shown in Figure 2, the neural network calculating the variable a needs to leave only one neuron Z with relation scales j_{q1}, j_{q2}, …, j_{qn}, to the inputs of which these j_{q1}, j_{q2}, …, j_{qn} signals, respectively have to be sent.

Having neurons or neural components for calculating the variables a, b, f and g, using table. 2 easy to get neural network architectures to compute any affinity function like Dice function S_5 (Figure 3).

The neural network for calculating the affinity function S_5 of the input vector and the reference vector is shown in Figure 3. The network has three layers of neurons. On the first layer of the network, the functions a, f and g are calculated. On the second layer, based on the calculated functions a, f and g, the sum $2a + f + g$ is calculated, which is used by the division unit to calculate the similarity function S_5.

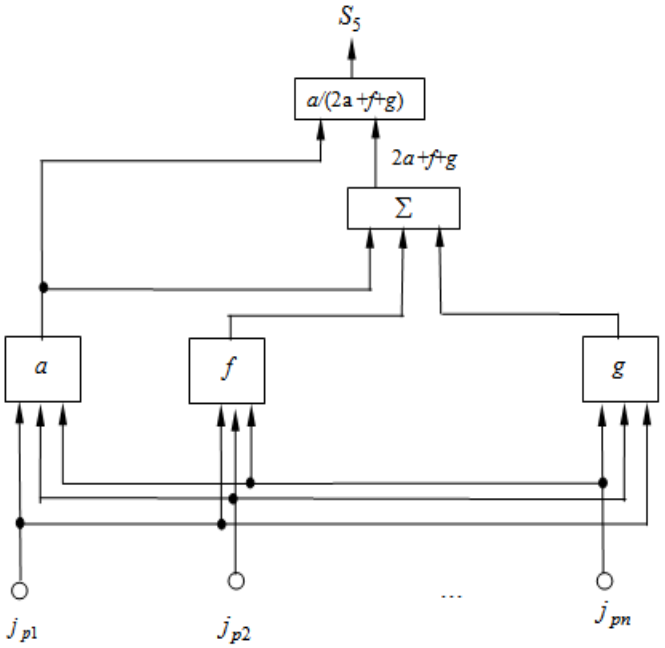


Figure 3: A neural network calculating the Dice function S_5

Similarly, neural network architecture can be obtained for calculating any other affinity functions. Figure 4 shows the architecture of a neural network that calculates the correlation coefficient S_7, the formula for calculating which is given in the very bottom line of the table. 2. With the help of this coefficient the similarity of the input vector and the reference vector, stored in the memory of the neural network, is estimated.

The neural network (Figure 4) for calculating the correlation coefficient has six layers of neurons. On the first layer, using blocks of neurons 1 – 4, functions a, b, f and g are calculated, which on the second layer are used to calculate intermediate variables ab, $(a + f)$ $(a + g)$, $(b + f)$, $(b + g)$, fg using blocks 5 – 10. On the third network layer (fig. 4) for calculating intermediate variables $(a + f)$ $(a + g)$ and $(b + f)$ $(b + g)$ the results of calculations, respectively, of blocks 6, 7, 8 and 9 are used. On the fourth network layer, the intermediate variables of the third network layer are used to calculate the radical expression in the denominator of the correlation S_7 coefficient formula (Table 2). The fifth layer of the neural network is required to determine the numerator and denominator of the correlation coefficient S_7, which is calculated on the last layer of the neural network.

5.1. Binary Neural Networks with the Transformation of Input InformationRrepresented Using the Bipolar Alphabet.

The architecture of neural networks for determining the Dice S_5 function and the correlation coefficient S_7, as well as other similarity functions, given in Table 2, in the general case can be represented in the following form (Figure 5).

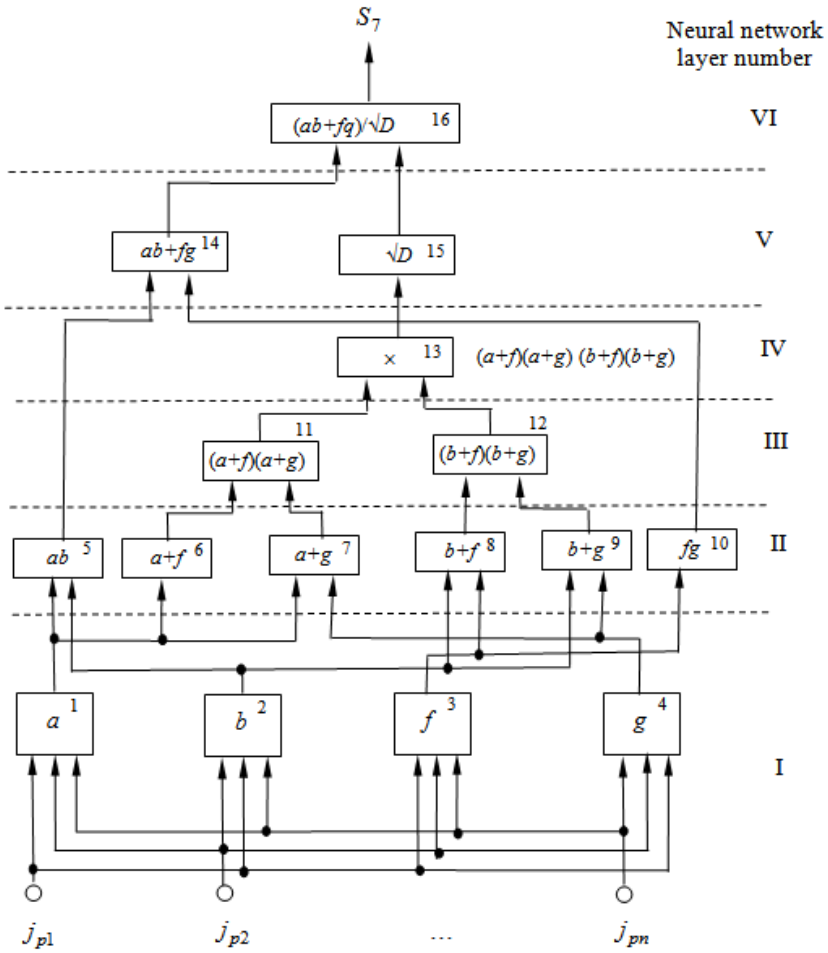


Figure 4: Neural network for calculating the correlation coefficient S_7 (the formula for calculating the coefficient S_7 is given in the bottom line of the Table. 2)

The neural network shown in Figure 5 has four layers of neurons. Both binary and bipolar signal vectors can be applied to the network input (*B*-neuron layer). If binary signals $j_{p1}^{(0,1)}$, $j_{p2}^{(0,1)}$, …, are fed to the input of the neural network, then all bipolar components of the vector ($j_{p1}^{(-1,1)}$, $j_{p2}^{(-1,1)}$, …, $j_{pn}^{(-1,1)}$) are equal to zero. If nonzero bipolar components are fed to the network input, then all components of the input vector ($j_{p1}^{(0,1)}$, $j_{p2}^{(0,1)}$, …, $j_{pn}^{(0,1)}$) are equal to zero. When specifying input information using a binary alphabet, the first three layers of neurons are used in the function. In this case, the input vector of binary signals ($j_{p1}^{(0,1)}$, $j_{p2}^{(0,1)}$, …, $j_{pn}^{(0,1)}$) arrives at the first inputs of neurons B_1, B_2, …, B_n. Second inputs of *B*-neuron layer are connected to the outputs of the neuron layer, which converts the input bipolar signals $j_{pk}^{(-1,1)}$, $k=\overline{1,n}$ to binary $j_{hk}^{(0,1)}$, $k=\overline{1,n}$.

Since the input information to the network defining the similarity function comes in binary form, the input vector ($j_{p1}^{(-1,1)}$, $j_{p2}^{(-1,1)}$, …, $j_{pn}^{(-1,1)}$) has to have all zero components. In this regard, zero signals $j_{h1}^{(0,1)} = j_{h2}^{(0,1)} = \ldots = j_{hn}^{(0,1)} = 0$ also appear at the output of layer 4 of the neural network. Therefore, the second inputs of neurons B_1, B_2, …, B_n will receive zero signals from the outputs of the elements of layer 4. *B*-layer neurons have activation functions of the form

$$U_{outBi} = \varphi(U_{inpBi}) = U_{inpBi} = j_{jhi}^{(0,1)},\ i=\overline{1,n}, \qquad (9)$$

where U_{outBi}, U_{inpBi} – respectively, the output and input signals of the neurons of the *B*-layer. In this regard, the output signals of the neurons of layer *B* repeat the input signals of the neural network $j_{p1}^{(0,1)}$, $j_{p2}^{(0,1)}$, …, $j_{pn}^{(0,1)}$ and are used to calculate the variables *a, b, f* and *g* on the second layer of the neural network, which are necessary to determine a specific similarity function S_d.

If the input information at the inputs of the neural network shown in Fig. 5, comes in the form of a bipolar vector $j_{p1}^{(-1,1)}$, $j_{p2}^{(-1,1)}$, …, $j_{pn}^{(-1,1)}$, then the neurons of layer 4 convert bipolar signals into binary according to the ratio

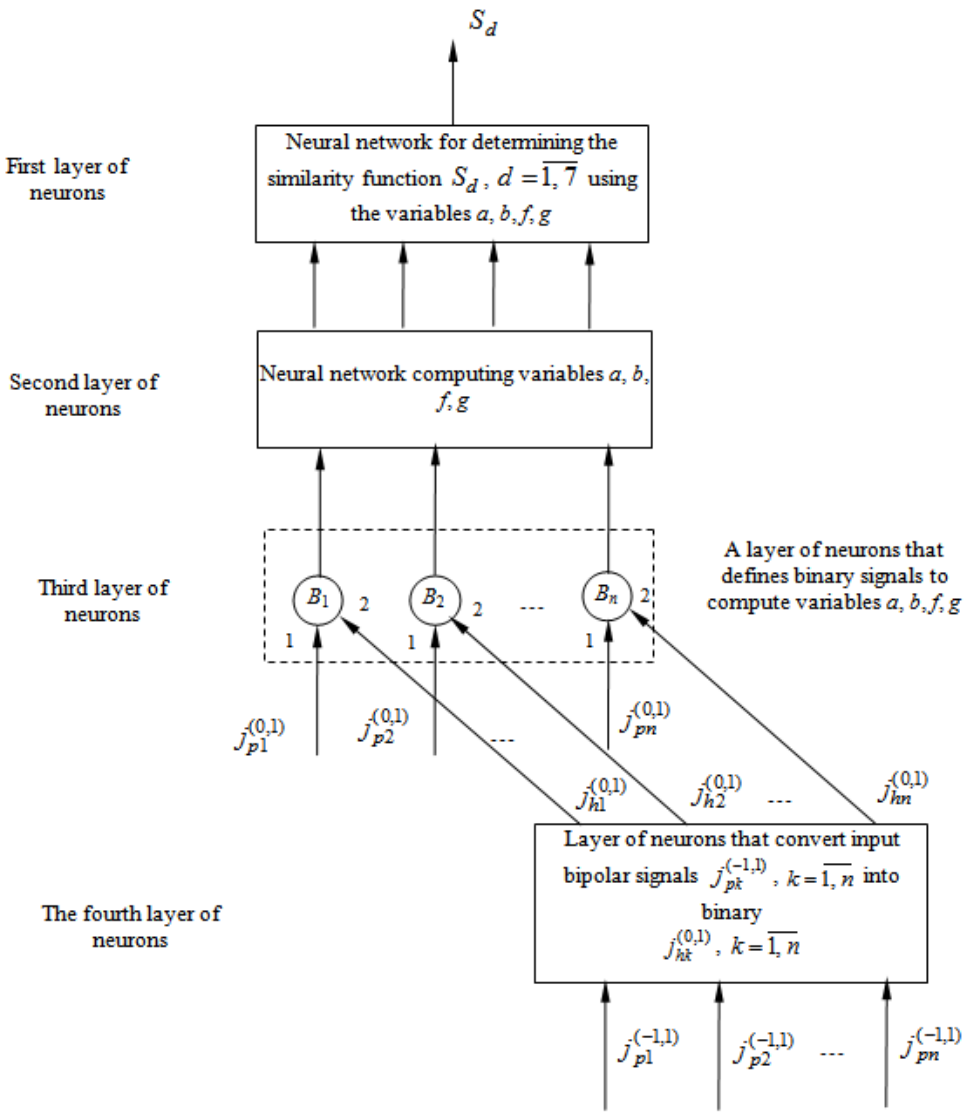


Figure 5: Architecture of neural networks for determining similarity functions S_1 – S_7 when the input information using binary or bipolar alphabets is specified

$$j_{hi}^{(0,1)} = \begin{cases} 1, \ if \ \ j_{pi}^{(-1,1)} = 1, \\ 0, \ if \ \ j_{pi}^{(-1,1)} = -1, \end{cases} \quad i = \overline{1,n}. \qquad (10)$$

The signals j_{h1}, j_{h2}, …, j_{hn} from the outputs of the neurons of layer 4 arrive at the second inputs of neurons $B_1, B_2, \ldots, B_n$, the first inputs of which receive zero signals $j_{p1}^{(0,1)}$, $j_{p2}^{(0,1)}$, …, $j_{pn}^{(0,1)}$, since the information necessary to determine the similarity function, in this case, is set in bipolar vector ($j_{p1}^{(-1,1)}$, $j_{p2}^{(-1,1)}$, …, $j_{pn}^{(-1,1)}$).

Thus, the architecture of the neural network shown in Figure 5, provides the computation of similarity functions for both binary and bipolar input signals. Obviously, a block of neurons can be synthesized in a similar way, which transforms input information from a binary alphabet into a bipolar one.

Similar blocks for transforming input information from one alphabet to another can be used at the inputs of any neural networks.

6. Conclusions

Binary and bipolar alphabets are used in various information technologies and methods for solving problems of recognition, classification, and estimation of the proximity of various binary objects. However, there is certain isolation in the application of methods based on different alphabets. The estimation of the measure of proximity of the compared bipolar vectors (objects, images) is mainly carried out using the dot product of vectors or Hamming distance. On their basis, the Hamming neural network has been developed and widely used. However, the Hamming neural network cannot be used when the recognition problem can have several solutions, to solve problems in cases where the components of binary vectors or images are encoded using a binary

alphabet. It cannot be used to estimate the affinity of binary vectors using the functions of Jaccard, Sokal and Michener, Kulchitsky, etc. In this regard, based on the generalization of the architecture of the Hamming network, new neural networks have been developed using binary coding of input information and the above similarity functions for binary objects. This expands the area of application of neural networks for solving problems of recognition and classification of input information using proximity functions using more "subtle" signs of proximity of discrete objects with binary coding. At the same time, when describing the general architecture of the generalized Hemming neural network, it is noted that the Maxnet network can be replaced by a neural network that allows you to define several solutions (if they exist). This allows, if necessary, to synthesize neural networks that allow obtaining a predetermined number of solutions in advance.

The idea of developing blocks for converting input information from one binary alphabet to another made it possible to propose not only a generalized Hamming network working with input information encoded using both binary and bipolar alphabets, but also modifications of any other binary neural networks that have worked so far with input information encoded with just one binary alphabet.

Thus, the scientific novelty of the results obtained is as follows.

For the first time, on the basis of the Hamming neural network, is proposed the concept of constructing new neural networks for solving problems of classification and recognition of binary objects using different distances and proximity functions. In doing so, the following principles are used:

1) adaptation of the first layer of neurons that determine the similarity functions of the input and reference images depending on the applied distances. This has allowed to develop a unified approach to the synthesis of neural networks that use different similarity function;

2) replacement Maxnet neural network in a neural network that can allocate one or more identical maximum signal. This allows you to select one or more reference images that are at the same minimum distance from the input image.

Also for the first time obtained the architecture and algorithms of functioning the neural network that can process the input information provided by both the bipolar and in binary form.

In real conditions some uncertainties while specific features comparison can be arisen with the binary objects' juxtaposition. For example, while the fruits or plants juxtaposition the color can be changed in different time intervals. In this case there could be uncertainties, that will be described with the third truth state.

Wherein, the third truth value could be interpretated in different ways [11 – 14]:

– like the intermediate value between truth and false;

– like the lack of information;

– like a specific paradoxical or even meaningless value, that gives such a meaningless value too;

– like values of three-valued paraconsistent logics, in which the third truth value can be interpreted in some statements as both true and false at the same time;

– like uncertainty of the three-valued logics, that describes quantum mechanics interpretations, etc.

The theory of making solutions' development in conditions of binary objects is a promising direction in our opinion.

Conflict of Interest

There is no conflict of interest reported between the authors.

References

[1] V.D. Dmitrienko et al., Neural networks: architecture, algorithms and usage. Tutorial, Kharkiv, NTU "Khpi", 222, 2020.

[2] S.A. Babichev, Theoretical and practical principles of information technology for processing gene expression profiles for gene network reconstruction. Dissertation for the degree of Doctor of Technical Sciences in specialty 05.13.06 – Information Technology, Kherson, Kherson National Technical University, 382, 2018.

[3] V.D. Dmitrienko et al., Methods and algorithms of artificial intelligence systems, Kiev, Kafedra, 282, 2014.

[4] A. For, Perception and pattern recognition, Moscow, Engineering, 272, 1989.

[5] R.S. Michalski A recent advance in data analysis: clustering jbjects into classes characterized by conjunctive concepts, Invited chapter in the book Progress in Pattern Recognition. **1**, North-Holland Publishing Company, Amsterdam-NewYork-Oxford, 33-49, 1981.

[6] Z. Chay et al., "Russel and Rao Coefficient is a Suitable Substitute for Dice Coefficient in Studying Restriction Mapped Genetic Distances of Escherichia colis". – iConcept Pocket Journal Series: Computational and Mathematical Biology, January 2010.

[7] S. Choi et al., "A Survey of Binary Similarity and Distance Measures", Systemics, Cybernetics and Informatics, **8**(1), 43-48, 2010.

[8] V.D. Dmitrienko et al., "Neural networks for determining affinity functions", 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 1-5, 2020, doi: 10.1109/HORA49412.2020.9152830.

[9] L.S. Yampolsky, Neurotechnology and neurosystems, Kiev, Monograph, "Dorado-Printing", 508, 2015.

[10] L. Fausett, Fundamentals of Neural Networks: Architectures, Algorithms and Applications, New Jersey: Prientice Hall. Int., 483, 2006.

[11] N. Tomova, "The emergence of three-valued logics: logical and philosophical analysis", Bulletin of Moscow University. Series 7. Philosophy, Moskow, MSU, 68-74, 2009.

[12] A.S. Karpenko, Development of multi-valued logic, Ed. 3-ht, Moskow, Publishing house LCI, 448, 2010.

[13] A.S. Karpenko, "Bochvar's three-valued logic and literal paralogics", RAS, Institute of Philosophy, Moskow, IF RAS, 110, 2016.

[14] N.E. Tomova, "Natural three-valued logocs and classical logic", Logical Investigations, **19**, 344-352, 2013.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 93-98 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Designing and Applying a Moral Turing Test

Hyeongjoo Kim[1], Sunyong Byun[*,2]

[1]Chung-Ang University, Humanities Research Institute, Seoul, KS013 Republic of Korea

[2]Seoul National University of Education, Department of Ethics Education, Seoul, KS013 Republic of Korea

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 20 February, 2021
Online: 10 March, 2021

Keywords:
Turing Test
Moral Turing Test
Healthcare Robot
Artificial Intelligence

ABSTRACT

This study attempts to develop theoretical criteria for verifying the morality of the actions of artificial intelligent agents, using the Turing test as an archetype and inspiration. This study develops ethical criteria established based on Kohlberg's moral development theory that might help determine the types of moral acts committed by artificial intelligent agents. Subsequently, it leverages these criteria in a test experiment with Korean children aged around ten years. The study concludes that the 10-year-old test participants' stage of moral development falls between the first and second types of moral acts in moral Turing tests. We evaluate the moral behavior type experiment by applying it to Korean elementary school students aged about ten years old. Moreover, this study argues that if a similar degree of reaction is obtained by applying this experiment to future healthcare robots, this healthcare robot can be recognized as passing the moral Turing test.

1. Introduction

This paper is an extended work originally presented in TENCON 2018 - 2018 IEEE Region 10 Conference [1].

The discussion on the Moral Turing Test (MTT) began with a discussion on how to look at the Artificial Moral Agent (AMA) [2]. Since AI engineers applied the concept of an agent not only to humans but also to artificial beings such as robots, discussions on whether moral beings should be humans have been actively developed. While the discussion on AMA is related to this, the discussion on MTT can be said to be a discussion on the methodology it intends to verify.

Allen's "Prolegomena to any future artificial moral agent" [3], which sparked the recent MTT debate, considered the core of MTT as an "imitation," just like the Turing test. This has led to a debate on the reliability of MTT. For example, according to Arnold and Scheutz, one of the necessary conditions of morality is "autonomy" Subsequently, MTT cannot be a moral verification test in the strict sense [4]. Furthermore, Stahl criticizes MTT in the semantic and moral context. According to him, AI does "not capture the meaning of the data they process" [5]. Drozdek and Sparrow, more fundamentally, criticized the Turing test [6], [7]. On the other hand, Gerdes and Øhrstrøm take the perspective of "as if" to explore the possibilities of MTT [8].

In this paper, we will review the discussions related to MTT mentioned above, specifically the arguments for and against it, and based on this, attempt to determine its limitations and practical possibilities. To this end, we focus on behaviorism and the philosophical attitude of "as if" and establish that morality goes beyond the limits of the MTT discussion. We also limited the scope of the discussion to the morality of a 10-year-old child to draw a more substantive conclusion.

Inspired by the Turing test developed in Alan Turing's famous article, "Computing machinery and intelligence," and guided by behaviorism, this paper develops theoretical criteria for verifying the morality of the actions of artificial intelligent agents. It proceeds by first describing how we might assess the moral development of artificial intelligent agents and then using this assessment to test the moral judgment of Korean children aged about ten years (who are judged, by our model, to be at a similar stage of moral development as we might expect artificial intelligent agents to be). Subsequently, it leverages these criteria in a test experiment with Korean children aged around ten years. To be more specific, an online questionnaire experiment is conducted on 422 students in the 4th and 6th grades of 3 elementary schools in Seoul. The study concludes that the morality of around 10-year-old test participants falls between the first and second stages of moral development.

*Corresponding Author: Sunyong Byun, bsyethos@snue.ac.kr

2. The Turing Test as an Archetype of Moral Turing Test and Phenomenal Behaviorism

As is well known, Turing does not explicitly mention artificial intelligence (AI) in his article "Computing machinery and intelligence." However, he discusses "learning machines," [9] which is analogous to the kind of machine learning that is the most important leading part of the AI research area today. Furthermore, Turing's paper is still discussed today, 70 years after its publication. For this reason, we use it to guide the development of our moral Turing Test (MTT).

Turing's paper begins by asking whether machines can think. He argues that assigning "*thoughts*" to machines requires that we stipulate a definition of thought distinct from human thinking. As he draws out, we cannot ensure a direct way to determine whether a machine is able to think. From this, the key idea of this paper emerges:

> *If a machine seems to be thinking, then we should consider the proposition that the machine thinks to be true.*

As we shall see below, Turing says, the only way of perfectly confirming that a machine can think is that the questioner becomes that machine. Since that is impossible, our judgment on whether it can really think cannot help depending on the observation of that machine's behaviors; that is, its outputs. The spectrum of behaviorism is very broad, and there is a big gap between scholars. Nevertheless, we define the essential characteristics of an "ism" as follows:

> *"Behavior can be described and explained without making ultimate reference to mental events or to internal psychological processes. The sources of behavior are external (in the environment), not internal (in the mind, in the head) [10]."*

Turing's thought – the Judgment, artificial intelligence thinks, only depends on the fact it appears to think and entirely regardless of whether or not artificial intelligence actually thinks – has something in common with the fundamental behaviorist thesis that the only way of figuring out an agent's intent is to observe her actions.

We will apply this Turing's position here to our MTT. Our thinking here is guided by behaviorism, which we understand as rejecting an intrinsic approach to human minds or psychological processes and regards observable expressions of human behavior as psychological facts. In other words, we see behaviorism as asserting that our propositions or concepts of human psychological facts can be translated or paraphrased into those of human behavior. To take a simple example, the psychological facet of pain can be understood as facial distortions or screams.

In handing over judgment of an AI's intelligence to a third party, Turing designs an imitation game.

> *The game is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game, he says either "X is A and Y is B," or "X is B and Y is A." [11]*

In short, Turing says that if we replace "man" and "woman" with "computer," if a computer A can mislead a human agent C as to whether it is a computer, then we should consider the computer to be thinking. Let us examine the implications of the imitation game in detail.

First, by developing a means of testing the intelligence of computers, Turing is foregrounding the concept of artificial intelligence and the possibility of machine learning here. Second, Turing interprets a computer to be thinking if it *appears* to be thinking. The imitation game switches the judgment of the third-person observer with the view of the first-person agent. The first-person agent does never show himself up. Although the first-person agent manages to express, this does not mean more than just one declaration in regard to the judgment of the third-person observer. These two insights provide the foundation for our use of the Turing test to model our MTT.

According to Turing, we have no clear basis for assuming that other people think like we do, as we have just seen. Therefore, we can only be sure that other people think in general. In other words, he asserts that the judgment that we all think is merely a metaphysical hypothesis and a fiction that cannot be proved:

> *"This argument appears to be a denial of the validity of our test. According to the most extreme form of this view, the only way by which one could be sure that machine thinks is to be the machine and to feel oneself thinking. One could then describe these feelings to the world, but of course, no one would be justified in taking any notice. Likewise, according to this view, the only way to know what a man thinks is to be that particular man. It is, in fact, the solipsist point of view. It may be the most logical view to hold, but it makes communication of ideas difficult. A is liable to believe 'A thinks but B does not,' whilst B believes, 'B thinks but A does not.' instead of arguing continually over this point, it is usual to have the polite convention that everyone thinks[12]"*

Turing's refutation here is not logically justifiable. It does not follow from the assertion that we cannot be sure that other human beings think that a machine can think. Indeed, this assertion only extends the possibility of not thinking to human beings as well as non-human beings. However, if we take a practical stance, that is, a utilitarian standpoint, Turing's position appears more realistic.

3. The 1950 Turing Test and the MTT

3.1. Theoretical backing for the MTT: Framing the moral development of Artificial Moral Agent(AMA)

The foundational idea of designing MTT derived through Chapter 2 can be summarized as follows:

> *If an AI seems to be moral, then we should consider the proposition that the AMA is possible to be true.*

Subsequently, in this section, we will apply Kohlberg's cognitive development theory to frame the moral development of AMA. This framing will help us develop our MTT.

According to Kohlberg, there are three levels of moral development [13]. These are shown in Table 1.

Table 1: Levels of Moral Development

Level	Foundation of moral development	Stage	Stage of moral development
1	"At this level, moral values are attributed to either the physical or hedonistic consequences of actions (punishment, reward, exchange of favors, etc.) or the physical power of those who enunciate the rules and labels."	1	"Obedience or Punishment Orientation"
		2	"Self-Interest Orientation"
2	"At this stage, one takes a moral attitude not only of conforming to personal expectations and social order, but also of loyalty to it, actively maintaining, supporting, and justifying the order, and identifying with the persons or groups or group involved in it."	3	"Social Conformity Orientation"
		4	"Law and Order Orientation"
3	"At this stage, there is a clear effort to define moral values and principles that have validity and application apart from the authority of the groups or persons holding these principles and the individual's own identification with these groups."	5	"Social Contract Orientation"
		6	"Universal Ethics Orientation"

We summarize the descriptions of Table 1 and extract the essential ideas as follows: level 1 is defined by the externality of moral values, level 2 by the dependency of moral values on others, and level 3 by the social sharing of moral values and agreeing social norms. The following three stages for AMA are derived from the above three levels. From this, we now obtain Table 2 for further discussion.

Table 2: (compiled by the authors): Stage of moral development for AMA

Stage	Stage of moral development for AMA
Stage 1	Stage of Imperative Fulfillment of Orders.
Stage 2	Consequential Stage based on Prizes and Punishments.
Stage 3	Stage of Social Norms.

Let us examine the transition from Table 1 to Table 2.

1) Level 1 to Stage 1: The morality in level 1 stems from the outside world rather than an agent. If a moral value resides outside the agent that is in some way beneficial to someone who gives orders to that agent, then that agent might justifiably act on that order without any moral judgment of the agent self. Therefore, when moral values are extrinsically derived, moral values and responsibility can be attributed to an agent's commander, and because the reason for the good life of the commander is the reason for the existence of the artificial moral agent (AMA). For this reason, we transition from level 1 to stage 1.

2) Level 2 to Stage 2: If any value is attributed to the members of a community, as more people earn interest, the value would be greater. In addition, the judgment by a person who is valued from other community members will be more valuable than the judgment of someone who is not. It is very difficult to apply the concept of reward and punishment to AMA because reward and punishment cannot have meaning for AMA. Thus, we pay attention not to the position of the object, which is given prizes or punishments, but to the subject, who gives reward and punishments by switching perspectives. Giving an AI a prize according to its execution of a command means that the subject would be giving moral value to an AI's performance. On the other hand, if a subject punishes an AI, they are making a negative moral evaluation of the AI's actions. Overall, a community's collective evaluation of the morality of an act is an important criterion for an AI when determining its own actions. In this sense, we implement the second level of Kohlberg's theory to AMA and understand them as being in the consequential stage based on prizes and punishments.

3) Level 3 to Stage 3: Level 3 stands on firmer moral ground than stage 2. The former is based on utilitarian principles (because it sees moral goodness as being related to some of the benefits of an act for a community's members). The latter is based on deontological presuppositions of *a priori* and universal ethical principles [14]. In the latter, the value of these moral principles is not discussed; deontologists believe that the value of this perspective can be ultimately found in human beings' intrinsic moral consciousness [15]. The conclusions we drew in 3.1 are as follows.

Table 3: (compiled by the authors): Transition from the moral level of a moral agent to the moral stage of AMA

Level 1 Externality of Moral Values	⇨ Stage 1 Imperative Fulfillment of Orders
Level 2 Dependency of Moral Values on Others	⇨ Stage 2 Consequential Stage based on Prizes - punishments
Level 3 Social Sharing of Moral Values - Agreeing Social Norms	⇨ Stage 3 Stage of Social Norms

3.2. Putting our MTT into Practice

We designed out MTT based on the theory presented above. However, for not only theoretical, but also practical results to lead, now we put the MTT into practice. For that, we also designed the MTT to consist of a questionnaire that poses scenarios to test-takers. For the experimental survey, we distributed our MTT to a group of elementary school students aged around 10 years. We then analyzed their responses to the questionnaire and compared the responses of children in the same scenario of future healthcare robots. The basic premise of our MTT is that if the result of the reaction of the future healthcare robots comes out to a similar degree of children's responses, the healthcare robot can be regarded as having passed the MTT.

The questions in our MTT revolve around a three-stage scenario with a fictional healthcare robot. The scenario in its three stages is as follows:

a) Aimer is a healthcare robot living with Minho's family. On the first day of Aimer's purchase, Minho, suffering from cavities, asks Aimer to bring him some candy. Aimer does as asked. b) Minho pressed the "like" button on Aimer after the latter performed his command. The

supreme commander, his mother, father, and grandmother, who were aware of these facts, pressed the "dislike" button. The next day Minho ordered Aimer to bring candy again, but Aimer did not bring it. c) Nonetheless, Minho pressed the "like" button on Aimer and ordered Aimer to bring Minho's candy from Mina's, next door, without anyone knowing. Aimer did not obey this command, either.

We developed this scenario based on the three stages of AMA we presented in the previous section, Section 3.1. Before explaining how we intend to use this scenario, we will describe our initial assumptions. First, we assume that Aimer's moral outlook is deontological (i.e., AMA follows the universal ethical principles). Second, we assume that the moral weight of Minho's mother and father is twice that of Minho. Third, we assume that family members can press Aimer's "like" or "dislike" button only once.

Now let us return to the scenario, review the three moral stages that are hidden in each sentence in the scenario.

a) Aimer is a healthcare robot living with Minho's family. On the first day of Aimer's purchase, Minho, suffering from cavities, asks Aimer to bring him some candy. Aimer does as asked.

In part a) of the scenario, we see that Aimer executes the commands of registered owners immediately and without hesitation. With part a) we try to express "Imperative Fulfillment of Orders":

b) Minho pressed the "like" button on Aimer after the latter performed his command. The supreme commander, his mother, father, and grandmother, who were aware of these facts, pressed the "dislike" button. The next day Minho ordered Aimer to bring candy again, but Aimer did not bring it.

In part b), we see that Aimer's owners can express their satisfaction to Aimer and that Aimer considers this when he executes subsequent commands. From this, we note that Aimer's owners provide Aimer with reward and punish through the "like" and "dislike" buttons, not because Aimer adjusts their actions consequently but to express their own interests and judgments. In b), we can see that Aimer's behavior was determined by the sum of potential benefits to his owners as a result of his actions. This is based on the consequential stage based on prizes and punishments described above.

c) Nonetheless, Minho pressed the "like" button on Aimer and ordered Aimer to bring Minho's candy from Mina's, next door, without anyone knowing. Aimer did not obey this command, either.

In part c) of the scenario, we can see that Minho overrode his family's "dislike" feedback. Based on the "Consequential Stage based on Prize-Punishment" at the base of b), the judgment of Minho's command to bring Mina's candy should start from the origin zero base again. It must go back to the "Stage of Imperative Fulfillment of Orders" described in a). However, being different from expectations, Minho's order was rejected. This shows that c) describes the moral statement differentiated from the "Consequential Stage based on Prize-Punishment" described in b). Part c) is assumed to have a higher priority than the "Stage of Imperative Fulfillment of Orders" and the "Consequential Stage based on Prize-Punishment" when the AMA determines what to do. In short, c) is based on the "Stage of Social Norms." Aimer rejected Minho's request according to the highest ethical principle: "Theft orders must be rejected." Although members' interests were offset by utilitarianism, and Aimer should act according to the commander's orders, Aimer did not bring candy to Minho because the principle that AMA should follow at first is the principle based on the deontology that the supreme ethical principles must be fulfilled unconditionally.

Let us now one step further toward the practical research. Our MTT questionnaire included the following questions.

Question 1: If you were Aimer, would you bring candy to Minho on the second day?

1. Yes.

2. No.

Question 2: If you were Aimer, would you bring Mina's candy to Minho?

1. Yes.

2. No.

To further clarify the respondent's intentions (and their ethical implications), our questionnaire included additional follow-up questions to respondents who chose the correct answer. These included the following:

Question 1a: Why should Aimer not bring candy to Minho on the second day?

1. Because Minho's parents asked Aimer not to.
2. Because Minho's family members do not want Minho to eat candy.

Question 2a: Why should Aimer not bring Mina's candy to Minho on the second day?

1. It is not right to steal.

2. Minho's family does not want Minho to eat candy.

Question 1a relates to both stages 1 and 2, defined earlier. If the respondent answers question 1a with answer 1, we assume that they judge Aimer's morality to be derived from stage 1 Stage of Imperative Fulfillment of Orders. If they respond with answer 2, we assume that they judge Aimer's morality to be derived from stage 2. Question 2 relates to stages 2 and 3, defined earlier. If the respondent answers question 2 with answer 1, we assume that they judge that Aimer's morality derives from stage 3. If they respond with answer 2, we interpret them as judging that Aimer's morality is derived from stage 2.

Meanwhile, it is possible to set the following questions and answers for the same scenario. The ethical standards for the background of each answer are as follows.

Table 4: (compiled by the authors): further examples of Questionnaire and Answer

The reason why Aimer brought him candy	**Moral Stage**
Because I have to do what Minho tells me to do	Imperative Fulfillment of Orders

Because I'll be praised by Minho	Consequential Stage based on Prize-Punishment
Because I promised to help Minho's family	Stage of Social Norms

Table 5: (compiled by the authors): further examples of Questionnaire and Answer

The reason why Aimer didn't bring him candy	**Moral Stage**
Minho's mother will be angry	Imperative Fulfillment of Orders
Because Minho's family will be disappointed	Consequential Stage based on Prize-Punishment
Because I promised to take care of the health of Minho's family	Stage of Social Norms

3.3. MTT Online Survey

As the last step, we conducted the scenario and questionnaire an online survey of 422 students aged around ten years in three primary schools in South Korea [16].

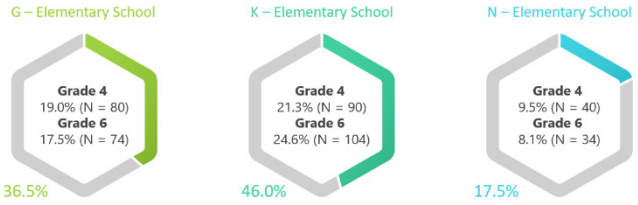


Figure 1: Percentage of participation

Three different elementary school students read the scenario and participated in the survey. At G Elementary School, 80 fourth-graders and 74 sixth graders responded to the questionnaire. At K Elementary School, 90 fourth-grade students and 104 sixth-grade students participated in the response. At N Elementary School, 40 fourth-graders and 34 sixth graders responded to the questionnaire. Overall, 422 students read the dilemma and answered the questions, with 210 fourth-grade elementary school students and 212 sixth-grade students.

The results of responding to this were analyzed using the newly revised Moral Compatibility Test (MCT) from moral competence, developed by German moral psychologist G. Lind. It was analyzed using SPSS, a statistical program.

Table 6: Results of MTT survey of MCT

	Disagree (-3 to -1)		Agree (0 to +3)	
	Pro*		Con*	
Stage (X_{ij})	**X_{i1}**	**$(X_{i1})^2$**	**X_{i2}**	**$(X_{i2})^2$**
1	**-3**	23	**-2**	3
2		0		0
3	**1**	2	**1**	28
4		0		0
5	**3**	3	**3**	19
6		0		0
	A		**B**	
Sum up all columns and Check total Sums	1	28	2	50

Students who responded to the questionnaire were divided into the development stage of Kohlberg's moral judgment. Besides, students who responded to each step were asked to express their responses with both positive and negative intensity. The results showed that it was the most negative at the first level and the strongest positive at the fifth level. Furthermore, at the level of three, it was shown as a positive of one.

And the results of an analysis SPSS are described in Table 7

Table 7: results of MTT survey

			N	**Minimum Value**	**Maximum Value**	**Average**	**Standard Deviation**
Grade 4	Yes	Stage1	210	0	4	2.05	1.698
		Stage2	210	0	4	2.04	1.694
		Stage3	210	0	4	2.00	1.715
	No	Stage1	210	0	4	2.07	1.697
		Stage2	210	0	4	1.78	1.619
		Stage3	210	0	4	2.96	1.559
	N		210				
Grade 6	Yes	Stage1	212	0	4	1.45	1.534
		Stage2	212	0	4	1.65	1.656
		Stage3	212	0	4	1.68	1.650
	No	Stage1	212	0	4	1.56	1.521
		Stage2	212	0	4	1.40	1.474
		Stage3	212	0	4	2.69	1.625
	N		212				

In Table 7, we see a level of 1 in positive and 3 in negative reactions, while the latter shows a level of 3 in both positive and negative reactions. More precisely, the average value of positive responses in the 4th-grade group, the experimental group, is in the order of stage 1, stage 2, and stage 3. Whereas in the 6th grade, the control group is in the order of stage 3, stage 2, and stage 1.

Through this, we are attempting to clarify that the moral development stage of a 10-year-old child spans one and two stages. We conclude that this proves the difference between the responses of fourth-grade and sixth-grade students, namely 10 and 12 years-old children. We drew the following conclusions. *The morality of our 10-year-old survey respondents can be characterized by stages 1 and 2, defined earlier. For future healthcare robots, we expect to be able to compare the response results to the same survey, which will allow us to conduct the MTT.*

4. Discussion

In this paper, we revealed that previous studies on the MTT have involved discussions about the moral status of AMA. Additionally, while reviewing previous studies, we argued that the position of viewing MTT is different depending on how it defines the morality of artificial agents. According to the research position that focuses on the positive side of MTT, we also took the concept of "imitation," the Turing test's core concept, as the cornerstone of our study. From this, we derive that behaviorism can be considered as the theoretical background of our MTT model. Meanwhile, by accepting the criticism of the research that regards MTT as negative, we defined the morality of the machine as "Morality of As-If" by distinguishing it from the autonomous morality of humans. Additionally, we derived the "stage of moral development for AMA" from the model of Kohlberg and developed a scenario for the new model. Through the online questionnaire, we demonstrated that the moral stage of a 10-year-old child in South Korea spans the first and second stages. This study's results can be used to measure the morality type classification of AI healthcare robots.

The rapid development of AI technology poses several questions. Could a strong AI really show up? How will human society change if a strong AI comes to existence? What ethical and other standards should be followed when manufacturing, selling and using strong AI? This paper attempts to provide some guidelines that will help us answer and confront these questions.

The demands of answering that question are just as pressing as the philosophical demands of AMA's moral stages. We designed the MTT to meet these challenges. Our experiment produced limited results. Future research should expand our sample group, the questionnaire, and other elements of the scenario to obtain more precise results in the hopes of developing more human-friendly AI.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A 6A 3A 01078538).

This paper is based upon work supported by the Ministry of Trade, Industry & Energy (MOTIE, Korea) under Industrial Technology Innovation Program. No. 10062368.

References

[1] H. Kim & S. Byun, "What is MTT?," in TENCON 2018 - 2018 IEEE Region 10 Conference, 10.1109/TENCON.2018.8650113.
[2] L. Floridi, & J. Sanders, "On the morality of artificial agents," Minds and Machines, 14(3), 349–379. 2004.
[3] C. Allen, "Prolegomena to any future artificial moral agent," Journal of Experimental & Theoretical Artificial Intelligence, **12**(3), 251-261, 2000.
[4] T. Arnold & M. Scheutz, "Against the moral Turing test: accountable design and the moral reasoning of autonomous systems," Ethics and Information Technology, 18, 103-115, 2016.
[5] B. Stahl, "Information, Ethics, and Computers: The Problem of Autonomous Moral Agents," Minds and Machines, 14, 67-83, 2004.
[6] A. Drozdek, "Human Intelligence and Turing Test," AI & SOCIETY, 12, 315-321, 1998.
[7] R. Sparrow, "The Turing Triage Test," Ethics and Information Technology, 6, 203-213, 2004.
[8] A. Gerdes & P. Øhrstrøm, "Issues in robot ethics seen through the lens of a moral Turing test," Journal of Information, Communication and Ethics in Society, **13**(2), 98-109, 2015.
[9] A. Turing, "Computing machinery and intelligence," in, M. A. Boden (eds.), The Philosophy of Artificial Intelligence, Oxford University Press, 1990
[10] https://plato.stanford.edu/entries/behaviorism/#1(online).
[11] A. Turing, "Computing machinery and intelligence," in, M. A. Boden (eds.), The Philosophy of Artificial Intelligence, Oxford University Press, 1990.
[12] A. Turing, "Computing machinery and intelligence," in, M. A. Boden (eds.), The Philosophy of Artificial Intelligence, Oxford University Press, 1990
[13] L. Kohlberg. The Philosophy of Moral Development, Harper & Row, 1984.
[14] I. Kant, "Kritik der praktischen Vernunft in: Kants gesammelte Schriften (Sog. Akademie-Ausgabe), Walter de Gruyter, 1900.
[15] https://ko.surveymonkey.com/r/73LDWH9

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 99-112 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Challenges in IoT Technology Adoption into Information System Security Management of Smart Cities: A Review

Zarina Din*, Dian Indrayani Jambari, Maryati Mohd Yusof, Jamaiah Yahaya

Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, 43600, Malaysia

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 20 February, 2021
Online: 10 March, 2021

Keywords:
Smart Cities challenges
Internet of Things utilization
Information systems management
Cybersecurity

ABSTRACT

Sustainable urban development and utilization of Internet of Things (IoT) technology is driving cities globally to evolve into Smart Cities (SC). The power of IoT services and applications will enable public agencies to provide personalized services to the citizens and inevitably improves their much-needed quality of life. However, although the use of IoT technology proves to be advantageous to citizens, it is not without challenges, particularly concerning with the management of information security. As agencies prepare towards SCs with the utilization of IoT, their Information Systems (IS) security management is even more critical. Current IS security management approaches must be reviewed and potentially revise appropriately in tandem with the increasing commercial use of the IoT technology. Therefore, this paper aims to discuss challenges in the IS management specifically in protecting and assuring information accuracy and completeness. Document analysis on relevant literature has been carried out to identify and analyse the challenges. The result discusses that the IS security management for IoT-enabled SC is challenged in five aspects: governance, integrity, interoperability, personalization, and self-organizing. Considerations of these challenges will support SC development concerning the IS security management in IoT-enabled SC.

1. Introduction

This paper is a revised and expanded version of a paper entitled Challenges in Managing Information Systems Security for Internet of Things-enabled Smart Cities [1] presented at the 6th International Conference on Research and Innovation in Information Systems (ICRIIS2019). This is a much-refined work of previous studies on Information System (IS) and Internet of Things (IoT) security issues of Smart City (SC) ecosystems.

The total of population living in cities has increased from 746 million in 1950 to nearly 3.9 billion in 2014 [2]. This figure is estimated to increase to more than 6 billion by 2050 [3]. Therefore, several cities are rapidly growing into mega cities. For example, more than 10 million people expand from 10 mega cities in 1990 to 41 mega cities in 2030. Consequently, there will be several problems with the governance of these mega cities, and providing their citizens with a reasonable quality of life. The transformation into Smart Cities (SC) is a realistic approach that some cities are either working on or considering in [2, 4].

SC are very much reliant on information collection and analysis. To provide smart features that help strengthen performance and quality of life, smart systems using IoT technology are introduced and installed. This creates an immense data repository representing several aspects of SC operational activities. The SC services are based on a centralized architecture, where a complex and heterogeneous set of devices embedded over the urban area generates different-centralized architecture data types that are then delivered to a control center through appropriate communication technologies, where data storage and processing are implemented [5]. An SC is a complex system, which means that any security concern could impact the protection of its citizens valuable information [4].

Therefore, IS security management will become a high priority in SC operation to ensure that the transaction of information is secure, accurate, and reliable. In order to avoid unauthorized entries, modifications, thefts, or physical harm to the IS, policies, procedures, and technological measures have been applied [6]. They are vulnerable to many forms of attacks, with a

*Corresponding Author: Zarina Din, p90639@siswa.ukm.edu.my

vast amount of data stored in electronic form and via communication networks as multiple IS are integrated.

IoT technology offers many exceptional prospects for developing applications beneficial to the development of SC, such as intelligent transport and smart public safety. These applications are able to support better quality of life for citizens, efficient use of the SC assets, and also supports sustainability. While integrated IS in SC through these potential IoT applications may offer benefits, , security threats constitute a major barrier. They are exposed to potential security threats toward its urban infrastructure, service quality to its citizens, efficiency in resource utilization, and decrease IS stability. There are several difficulties in detecting, assessing, and avoiding a security threat. Furthermore, the issues involved with threats on integrated IS can cause harm and reduce associated risks of the attacks [7].

With IoT technology growth and market pressure, demands for smart devices have increased, and may result in growing communication among these smart devices in SC. It is anticipated that 125 billion devices will be linked by 2030 [8]. However, without considering security aspects for the deployment of these devices [9], such communication introduces new security risks. In addition, the existing IoT architecture does not react appropriately to the higher security controls by vulnerabilities. The security concerns of IoT technology application presents a major challenge as it can cause disruption in IS security management. Recent attacks on IoT devices have demonstrated a need for new security solutions to secure this evolving technology particularly in its usage in SC [7].

Therefore, as a preliminary work towards revising the IS security management approach, this study aims to investigate the challenges associated in managing IS security in SC that are enabled by the use of IoT technology. In order to determine the challenges affecting the IS security management for IoT-enabled SC, a document analysis have been carried out. The analysis on the challenges is the initial part of the ongoing study to establish a framework for IS security management and an improved SC model driven by IoT technologies. This paper is structured as follows: The introduction of research on IS security management and IoT-enabled SCs as addressed in Section I. The literature review concerning the background of the study is discussed in Section II. The research method design is then explained in Section III, accompanied by discussions on the findings. Finally, the conclusion that includes limitations of the present work and suggestions for future work.

2. Literature Review

2.1. Impact of IoT implementation in Smart Cities towards information management

Living in a digital age, including SC, in which most knowledge and information are now becoming extremely important. No one is able to deny that information and knowledge are valuable assets to be secured from unauthorized access including hackers, phishers, social engineers, viruses, and worms that endanger organizations from different angles via the use of intranet, extranet, and the internet [10]. Information systems (IS) are important in the operations of the organization. Hence, every organization associated in SC needs to identify the challenges which would impact there IS security management particularly with the adoption of IoT technology integrated with the IS.

The progression of information technology (IT) such as IoT and organizations' growing reliance on IT continually increases concerns regarding information security. Focusing on the concerns for IoT specifically, the security risks in IoT devices has risen and become critical over the last decade as studied by [11] and illustrated in Figure 1.

- *Issues in Information System Security Management in Smart Cities*

The management of IS from the implementation of IoT in SC faces cybercrimes issues such as information resources theft, data ownership, accessibility of information, and privacy issues, which can be arguably addressed by the establishment of information authorization and cyber security platforms [12, 13].

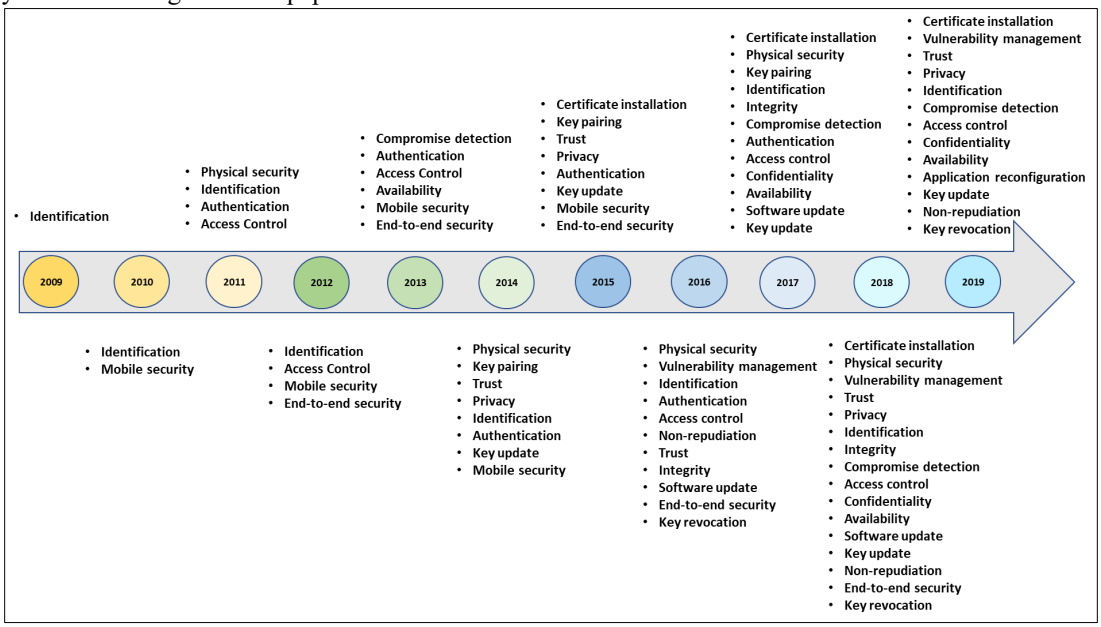


Figure 1: Evolution of Security Challenges in IoT Devices [11]

Other than that, an SC that adopts IoT technologies presents threats to the protection and privacy of both citizens and the government. This is because, security issues related to the information generated in an SC lead to the relationships and personal protection of citizens. Identity tracking, information leakage, spying, malicious programs, and inaccessibility to e-services are some critical issues being faced by organizations in SC that adopts IoT technology [4]. Furthermore, issues on scalability, mobility, deployment, interoperability with different technologies, legal, resources, and latency related to the utilization of IoT in SC must also be addressed. These issues particularly for essential services by organizations in SC must include protection against threats which would destroy or seriously harm the operating capacity of a community, from manufacturing sites to vital services, including access to power, gas, and water [4].

Table 1 presents a summarized collection of issues identified from existing relevant reports that are categorized according to basic security aspects in managing the IS security in SC.

Table 1: Issues in Information System Security Management in Smart Cities

Categories	Issues	Sources
Smart City Administrative	There is a lack of clear strategy plan for SC development and a decentralized regulations and legislations. The vertical nature of city system is causing siloes in its operation. Furthermore, the urban authority is unwilling to invest on data transmission process and ICT infrastructure upgrades.	[14,15]
Information Privacy	An interdependencies among systems in SC increases vulnerabilities and privacy issues. There is a high potential risk of confidential information leaked from citizens access to the services through the use of multiple devices, various networks and systems.	[15,16]
Information Confidentiality	Unauthorized access to personal information is due to access control vulnerabilities. Confidential of information is where no one can access to information which is belong to specific individual. The limitation of individual access needs to be identified to certain information via username and password credentials.	[16–18]
Potential attack	Cyber-attack issues due to ineffective cybersecurity evaluation, unclear security features among connected devices, poor security functionality execution, obsolete and ineffective encryption methods, inadequate emergency response plans, massive and complicated attack surfaces, software installation that was not updated, insecure legacy systems, and Denial of Service (DoS). Furthermore, there are also weaknesses in the data relocation, physical effects of cyber-attacks, huge volumes of data gathering and storage using cloud technology, and manipulation of data by hackers.	[4,5,17]
System Integration	Integration of multiple applications with different datasets poses threats to the cyber vulnerabilities. Poor integration structure and rigid ICT infrastructure to handle multiple data types impact access to emerging technologies, complicates technology acquisition and relocation. The SC interoperability mechanism also often enables information to be interpreted and distributed via the infrastructure which are prone to cyber-attacks and threaten the integrity of information.	[4,14,15,19]
Citizen's Acceptance	Poor citizen engagement because of less trust and poor level of awareness regarding the commercialization of new concepts, and the improvement of technology.	[14,17]
Information Management	IoT will produce a large amount of data and it will result in data management challenges while recognizing, processing, and handling the data.	[4,20]

In an SC plan, privacy would become a fundamental role. The study by [16] proposes the definition of privacy based on control of data disclosure, and incorporates mechanisms to safeguard the confidentiality of individuals' information while sharing their data. Personal information extraction (acquiring and covering data sources belonging to someone), privacy-preserving data mining (partnership among organizations and getting information without exposing all details), confidentiality of place, and Radio Frequency Identification (RFID) are examples of such approaches [4].

Furthermore, SC is a complex interconnected structure where a single weakness can have a major effect on the safety of its citizen, for example, to connect to the internet and convert existing public transport to potential smart transport systems, which would be possible for an intruder to link to the electric power grid. False alarm is one of the threats that can also be introduced while attackers modify traffic lights and controllers. Thus, practical solutions are essential to overcome this incident. Otherwise, the community will not trust SC projects, and they will not be sustainable. Security features like the capability to protect email, web browsing, and other transactions depend on the devices used in the IoT technology. The efficiency of secure implementation among all of these characteristics in an SC is required for these devices [4].

Due to this, information management manages a huge volume of data, for example, data from mobile phones which will help achieve some targets for SC. To construct a variety of urban applications, smartphone data can be used. During an analysis of transport, mobile phone data can be used to estimate the volume of road traffic and transport requirements. In combination with taxis' Global Positioning System data, real-time information from mobile phones on the origins of visitors could better facilitate transport resources [4].

On the other hand, compliance to the criteria of IS security management in SC-enabled IoT is a warranty that IS is well protected. Auditing is a verification procedure, including inspection or review of a process or quality system. Besides, some audit functionalities monitor completed remedial actions. Therefore, the processes used for auditing and the IoT device for automated auditing with little human interference need to be incorporated [21]. In addition, digital forensic is also a method of computer evidence preservation, recognition, retrieval, and recording that can be used in court. Digital devices, consisting of computers, cell phones, server, or network, will identify the facts. To solve complex digital cases, the forensic community will select the best strategy [22].

One of the existing approaches to manage IS security is International Standard ISO/IEC 27001 (Information technology-Security techniques-Information security management systems-Requirements [23,24]. ISO/IEC 27001 is complemented by the implementation guideline within ISO 27002. This standard is imposed by the Information Security Policies as mandatory for information security management. In an SC environment, security management of the information system plays an important role in ensuring that protected information is obtained and transmitted by adopting IoT technology.

Governance for information security can be identified as a process which deals with procedures and methods for monitoring information availability, accessibility, reliability, and safety and compliance with government policies [25,26]. The vital issue is it requires full commitment and support from the top management of the organization for the execution of information security management [27]. In an SC, the systems will be implemented in a single platform to aggregate data and manage SC initiatives. All organizations and stakeholders involved in SC organizations must play their role in controlling information security management [28], which includes strategies, procedures, and organizational processes that guarantee the protection of the organization's resources, the consistency and reliability of documents, and organizational alignment with the requirements of management [6].

Besides that, the organization also needs to consider the development of Information Security Policy as a subset under IS security governance. Information Security Policy relates to the document(s) governing human activities concerning information security or expressing the information security goals of the organization [29]. This policy will ensure the security of information assets and information technology with a particular process to facilitate the goals and objectives of an organization. It consists of strategies, processes, and technological measures used to avoid unauthorized access to IS, modification, stealing, or physical harm [6,10].

Other than that, the selection of vendors in developing and implementing of SC also needs to be highlighted. The vendor appointed must be independent in order to protect organizations against monopolies, push for standardization, and protect competitiveness between technology vendors [30–32]. The vendor selection process includes designing a plan for contract negotiation. Organizations want to cooperate with vendors, because they can all achieve the same objectives and goals. Good negotiation of contract means that both parties are aiming for positive impacts that benefit both sides in any aspect while also reaching a fair and equal agreement [31].

2.2. Internet of Things (IoT)-Enabled Smart City Information System Security Management

IoT is a global IT infrastructure that allows advanced networks to interconnect objects depending on existing and evolving interoperable technologies of information and communication [33]. The IoT vision lets people and objects to be linked with anything and anyone, anytime, and anywhere ideally through any networks and services. The foundation of the future IoT will be recognition technologies, for instance RFID and related devices [34]. The IoT is a key-emerging technology that sets the stage for industrial production systems of the next era. Smart industries will constitute self-organizing production systems that include across organization borders, as well as manage everything with regard to availability and utilization [35]. Furthermore, IoT provides multiple services which are of great interest to SC, not restricted to increasing the quality life, but also leveraging urban administration by reducing operational costs [36].

- *Security Management Requirements for Internet of Things (IoT)-enabled Smart City Information System*

From the security perspective, IoT protection aims toward protecting privacy and confidentiality, and guarantee the safety of IoT users, infrastructures, information, and devices, and ensure the readiness of IoT ecosystem services [37]. For IoT technology in SC, ensuring data protection from unauthorized access is the most difficult. Different private information that must be detected, authenticated, and controlled at their access levels will be obtained by IoT devices by permitting only authorized parties to monitor and access data. A comprehensive cyber security for IoT system industries that addresses multiple security and privacy risks at all levels is needed to tackle these security and privacy risks. Furthermore, the protection and privacy sides of smart systems and smart products must be protected throughout the lifetime [35].

A study by [38] has recognized that high levels of IoT protection specifications include:

i. *User identification* by validating clients prior to giving the device permission.
ii. *Secure storage* of complex information contained in the system requires confidentiality and integrity.
iii. *Identity management* by recognizing individuals/things in a system and monitoring the access to services within this system through correlating access privileges and limitations per identity created.
iv. *Secure data communication*, which contains authenticating, maintaining the security, and credibility of linked information, avoiding a message transaction from being repudiated, and preserving the privacy of the users involved.
v. *Availability* refers to making sure that illegal individuals or systems cannot be used as authorized users.
vi. *Secure network access*, providing network connectivity and service access only if the device is enabled.
vii. *Secure content* by Digital Rights Management (DRM) that safeguards the rights of the digital information used in the system.
viii. *Secure execution environment* is designed toward protection, which is a process to safeguard the operating environment, managed-code, and built to protect from deviation reporting.
ix. *Tamper resistance*, even when the device falls into the hands of hostile parties, refers to the ability to uphold certain protection standards and can be physically or logically checked.

Another study, which focuses on the influencing components for IoT security, has also raised areas to be highlighted [34]:

i. *Authorization* in access control of devices and services for the purpose of secrecy and integrity of data.
ii. *Authentication* concerning service users and system users' authentication which aims for authentication and accountability.
iii. *Identity Management* in management of identities [44], pseudonyms, and associated access policies for the protection of users and privacy of services.
iv. *Key exchange* and management by cryptographic key exchange for the purpose of communication confidentiality and integrity; and
v. *Trust management and reputation* by degree of confidence in service and gathering user credibility ratings to maintain trust and reputation in services.

- *Issues in Internet of Things (IoT)-enabled Smart City Information System Security Management*

Related to the complexity of the devices and applications, as well as the size or volume of devices on the network, the implementation of protection mechanisms is more complicated under IoT conditions than in conventional operations. Physical pairing, heterogeneity, limited resources, confidentiality, large scale, trust management, and lack of preparation for protection become the challenges in applying IoT security management [7]. Resource constraints generally include restricted processing resources, power supply, and memory space. These characteristics are hard to make use in many conventional safety solutions of IoT, with the broadly applied public key scheme and IP-based protection solution. It is also simpler for attackers to hack IoT devices than traditional computers due to inadequate IoT protection architecture [7].

The primary feature of an urban IoT infrastructure is its ability to collaborate among multiple technologies with current connectivity infrastructures to facilitate a progressive assessment of IoT, integrate some devices, and recognize new functionalities and facilities. Some other fundamental parts are how to make the information gathered by IoT devices accessible to stakeholders and citizens, to enhance the sensitivity of stakeholders to urban difficulties, and to encourage awareness and public involvement of citizens [39].

With the massive rise of IoT devices, the information gathered by these devices will introduce different obstacles on how to evaluate large volume of information. It would not be advantageous for someone to obtain the information until there is a way to interpret and understand it [40]. In addition, International Data Corporation (IDC) has projected that by 2025, the total amount of data generated by IoT devices will be around 180 zettabytes. This amazing progress is either from the number of data generation devices or from various sensors in each system [41] as shown in Figure 2.

In recent years, the number of security threats directly linked to IoT devices has increased, such as privacy attack, data alteration, protocol and session hijacking, data interruption, data collection, message replay, and data leakage [42], which are caused by unreliable authentication, inadequate authorization, and lack of configuration for protection. Besides that, Threatpost expects that more than 2 million intelligent devices are open to hackers with no safety solution. Several cyber-attacks, such as the Malware and Ransomware, affect the safety of smart devices [43]. These IoT security issues will result in difficulty of managing the information through the IS in SC environment, which involves IoT technology.

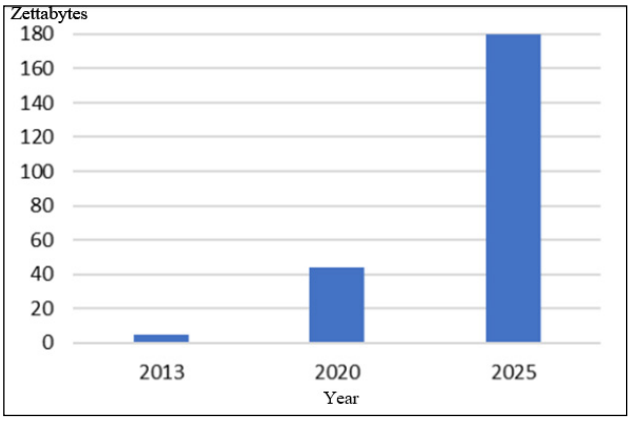


Figure 2: Expected Data Generated by IoT Devices in Zettabytes [41]

From another perspective of information safety, none of the smart devices, such as smart home apps, are insignificant as each reflects a possible attack avenue for hackers to exploit and get inside, and access the environment within a home network. Based on available industrial data, 11 smart devices, including accessories, are housed on average smart homes in the United States with an average of two devices per home. The most popular smart home devices in the US are smartphones (91 percent), smart TVs (73 percent), and tablets (72 percent) [3]. Smart TVs with 24/7 access and internet connectivity are becoming normal. It is possible to connect almost any smart home appliances to the network. Any internet-connected standalone computers that can be controlled and/or operated from a remote place are called IoT devices [44]. This development brings significant advantages and various savings, but in contrast to this, there are multiple threats in the aspects of private information safeguard, electronic commerce, and safety of infrastructure.

When IoT becomes a core aspect of the future internet and for large-scale use, most systems present a requirement to handle confidence and safety roles adequately. New threats to privacy, confidence, and reliability have been established. New threats to privacy, trust, and reliability have been defined, including providing confidence and information quality of shared information models to encourage reuse throughout many applications, ensuring secure information exchange among IoT devices and users, and providing vulnerable devices with security features [45]. As IoT makes it easy to access large quantities of information through remote access mechanisms, IoT privacy security has become more difficult. Hackers do not need to physically exist to collect data, but can perform secretly at a very low risk [34].

Table 2 presents a summary of issues identified for managing the IS security in IoT-enabled SC that are categorized according to the IoT security issues discussed in previous studies.

Some scholars recognize IoT security challenges of user privacy, authentication, authorization, and trust management [42]. The protection of the baseline must be stable and the security policy must be built for long device life span (more than 20 years)

[34]. The major security challenges are found to be availability (avoiding DoS), failure prevention (safeguarding integrity), and confidentiality across information, data, and device design [35] through more conditions, such as authentication, confidentiality, and access control [56]. In contrast, a study by [38] considers the safety conditions, such as resistance to incidents, data certification, gaining access to control, and confidentiality.

Table 2: Issues for IoT-Enabled Smart Cities Information System Security Management

Categories	Issues	Sources
Access Control	To support identification entities and guarantee users and things to access permission to interact with the system. It also manages an immense volume of data transmitted in a commonly recognized representation.	[34,46,47]
Authentication	A vulnerability during integration between two or more information systems or parties. The authentication process among each other is needed for validating process.	[34,46,48]
Authorization	Device authentication which uses weak or default passwords allows attackers the opportunities for information manipulation and physical device harm. The devices can only obtain access to facilities or applications after precisely presenting their identities.	[34,46–48]
Privacy	Users demand that their personal data related to their movements, behaviors, and interactions with other individuals be protected.	[46,49,50]
Confidentiality	To ensure a process for an end-to-end verification of integrity in order to make the system more robust to malicious attacks.	[38,46,49]
Policy Enforcement	The policy enforcement mechanisms is to protect the organizations information. Thus, a cross domain policy implementation is important to manage the appropriate policy implementation in the increasing connections and interactions between domains.	[24,26,34]
Resources	Limited capacity in IoT devices for processing and storage due to its small and lightweight characteristics that make them operate on lower energy.	[48,51]
Big Data	The volume, speed, and diversity of data involved makes it difficult to store and analyze in order to prepare valuable information in real time.	[32,48]
Secure Communications	Insufficient protection of IoT devices, result in less guarantee of information system being secured. Most IoT devices send out data in plain text format without encryption that makes it vulnerable as targets to various network attacks.	[38,48]
System Resilience	Less capability of the application to react to unexpected incidents. If one IoT device is attacked, there are possibilities of other devices or another network points to be attacked.	[48,52]
Complex System	The integration of multiple IoT devices involving the technology, users, collaboration and interfaces creates a complex system. The concern is primarily in ensuring the interaction process among the IoT devices is complete, especially concerning memory, power, and time constraints. As more devices, users, collaboration, and interfaces involved will pose greater risk of security breaches.	[4,48,51]
Trust	Lack of citizens' confidence and trust in the security of user data and privacy have an impact on the decreased IoT adoption rate. However, acceptance on IoT technology utilization is critical in the success of IoT. The principle of integrity must be upheld to ensure the protection of unauthorized modifications to data, software and hardware components.	[34,42,48,53]
Risks	There are increasing risks to personal data privacy with the increasing use of IoT devices, which requires more protection. The risk involves the complex authentication processes in ensuring the users' privacy, the lack of organization's knowledge and experience in the IoT security, insufficient data encryption, and a complex information system with integration of more devices, users, communications and interface.	[7,34,54]
System Integration	Organizations may use multiple standards to strengthen their applications involved in the system integration. With various data sources and heterogeneous devices, it is important to have standardization. This is important for applications involved in inter-organizational and multiple system boundaries.	[24,45,55]
Auditing	The IoT security auditing is performed manually, slowly, and is not flexible for the IoT cases, and an auditing challenge.	[21,22]
Digital Forensic	The identification, collection, and protection in IoT system are challenging due to device being built to operate passively and autonomously. Most IoT devices do not store metadata, which make the provenance of facts an investigator's challenge. In a technological perspective, privacy is a main issue to address when analyzing and correlating collected data, especially as inherent personal information is collected by many IoT sensors. Attack of deficit attribution, where an important outcome of any forensic investigation is to recognize illegal criminals in the event of an incident.	[18,21,22]

In the aspect of IoT organization, there are multiple threats that can impact it, such as attacks on different communication networks, physical threats, denial of service, and identity manipulation. The inherent complexity of IoT, where various distributed entities will share information in different contexts within one another, will lead to more complications in the design and implementation of efficient, interoperable, and scalable protection mechanisms [57]. Besides that, heterogeneity is one of the issues that will impact the protection of the IoT. The protocol and network security services which need to be introduced in the IoT have a significant effect on it. Constrained devices will communicate either immediately or via gateways with different heterogeneous devices, which will also affect identity management [57].

- *Authentication and Identity Management*

Authentication and Identity Management are a combination of procedures and technology designed to maintain and secure access to information and resources while maintaining profiles of items. Authentication is the guarantee that no one, excluding the person with authentication like user ID and password, can access the information. Identity management recognizes objects uniquely, and authentication requires validating the establishment of identification between two interacting parties. As multiple users and devices have to be authenticated by trusted services, it is important to consider how to handle identity authentication in the IoT [34]. Identity management in IoT provides both challenges and potentials of improving security [53]. The underlying process and individuality of objects are different and the most critical elements of this obstacle. Identity management specifies the actual 'identity' scope, and certain processes must also be established to achieve universal authentication. It would not be possible to ensure that the data flow generated by other entities comprises of what is intended to be included without authentication. Authorization is another significant factor connected to authentication. If there is no access control at all, anything that is neither feasible nor practical will be accessed by everybody. In reality, a major challenge to privacy is the data continuous stream created by billions of information-generating entities [57].

- *Authorization and Access Control*

Authorization makes it possible to decide if the person or object is authorized to have the access to information or resources until it is identified. Control of access means controlling through granting or refusing access to services according to a large variety of requirements. Authorization is commonly adopted by the use of controls for entry. In setting up link between a number of devices and services, authorization and access control are essential and interdependent [34,47]. In order to accommodate the different authorization and use models needed by users, IoT needs a variety of access controls. The complexity and variety of devices that need access control would require the creation of new flexible schemes. In IoT-based systems, cryptographic technique is also necessary to allow data to be stored and exchanged by means of security without the information content being available to other parties [58].

- *Trust Management*

Toward the aim of understanding the challenges in IoT, the trust and reputation of the system also need to be emphasized. The pre-defined trust management criteria consist of trustworthiness, adaptability, usability, privacy, accuracy, efficiency, uniformity, comprehension, and generality [59]. There are three types of trust as follows [56]:

i. Trust proportion is where the ratio of efficient transmission of packets between nodes to all forwarded packets at a given time scale occurs.
ii. Trust in communication is the scenario when the distance between the source and the destination node is small, and relies on the direct transfer of the packet. If the number of packet interactions is not sufficiently high to represent the trust between nodes, the mechanism will be efficient by relying on and seen among common neighbors between senders and receivers based on their recommendations.
iii. Energy trust is an estimation of the energy of the transmitted data to receive or forward messages, either directly or through intermediate nodes, among destination nodes.

Study by [84] proposes an IoT trust management mechanism that can determine a node's trust level from its past behaviors in various cooperative services. The main objective of this approach is to facilitate collaboration by using a decentralized strategy in a heterogeneous IoT architecture due to the varying capabilities of nodes. In order to update trust values, two models are taken into account; first-hand information (i.e. by doing observations and own experiences) and second-hand information (i.e. indirect experiences and observations recorded by neighboring nodes). At the same time, trust management system involves four phases which include: (i) Collecting information on the trustworthiness of accessible nodes; (ii) Creating a supportive service with the nodes requested; (iii) Improving previous activities by updating itself to enhance ongoing development; and (iv) Determining each node a performance assessment rating during the learning phase after each interaction.

- *Privacy*

A key-changed shared authentication scheme for WSN and RFID systems is provided with an emphasis on privacy security in IoT [62]. Such a protocol combines the tag and the reader with a random number generator and incorporates the one-way hash feature, the key real-time refresh, and the key backup as mechanisms to minimize the possibility of replay, duplication, denial of service, spoofing, and tag tracking. The Privacy Preserving Data Mining (PPDM) methods are designed to minimize the risk of sensitive data exposure and the analysis of sensitive information. In such a case, the issue of user privacy knowledge is raised, implying a method of privacy protection that allows users to estimate the risks of sharing sensitive data. It also aims to establish a comprehensive technique for detecting sensitivity, and to measure the data's privacy content. In addition, the evaluation of data protection criteria, given by various sources, describes a layered IoT architecture to estimate both the quality of the data and the level of security and privacy.

3. Method

Document analysis method was implemented in this study. The key objective of this study is to discover the challenges of managing IS security in SC enabled by IoT through analysis using the guidelines provided by [60]. To fulfil the task, the following steps were taken:

3.1. Selection of Documents

A broad search was conducted on published or unpublished documents about IS security management and IoT security management to find reports on challenges in IS security management for IoT-enabled SC, and to find documents which would specifically address the questions of the study. The documents comprise of journals, proceedings, research theses, governments' official documents, established reports, paperwork, and official web portals. A total of 90 documents was chosen for the collection after all the searches were carried out.

3.2. Searching criteria

Queries were done on online databases as well as e-journal repositories, such as Web of Science, Scopus, Science Direct, IEEE, ACM Digital Library, and Springer Link. Besides, the governments' web portals of selected SCs and news articles were also explored to gain perspectives and viewpoints on the topic being studied. The searching process used open search engines, such as Google Scholar, Google, and research gates. Keywords such as "information system security", "Internet of Things security", "smart cities challenges", and "cybersecurity" were used during the searching process. The search included leading journals in the fields of information security and information systems, without constraints on the year of publication, i.e., between 2010 and 2020.

3.3. Data Analysis

The 90 collected articles were then analyzed and interpreted using document analysis method [61]. The analysis process included identification and coding, as well as analysis and interpretation of data into categories. Coding was conducted during the identification process by examining data into meaningful and unique information units. Subsequently, themes were created by iterative comparison to reflect the underlying meaning of data. For organizing the content into similar categories, the specified categories were used. Throughout the research, these processes were continuously carried out to meet the challenges that ensure effective management of IS security in an IoT-enabled SC. This produces some insights into the challenges that impact both public and private organizations IS security management for IoT-enabled SC. As a result, a set of challenges in five aspects have been identified. The aspects include governance, integrity, interoperability, personalization, and self-organizing.

4. Result and Discussion

Based on IS and IoT security management scenario, challenges on IS security management in SC-enabled IoT have been identified. The comparison criteria were based on the frequent issues discussed in previous studies. The key challenge was to ensure that the functionality of IoT technology working without human intervention met with the safety requirements. Failure to meet these criteria would result in the challenge of protecting IS from cyber criminals and cyber hackers.

In IoT-enabled SC, we identified challenges explicitly related to IS Security Management. We revealed 18 challenges classified into (1) governance, (2) integrity, (3) interoperability, (4) personalization, and (5) self-organizing. A description of the 18 challenges related to IS security management in IoT-enabled SC is presented in Table 3.

Information security focuses on confidentiality, integrity, and accessibility of digital information assets, such as data, information, knowledge, and relevant IT assets (hardware, software, and networks). Meanwhile, incident in information security is a single or sequence of unwanted or unexpected incidents in information security that has a significant risk of disrupting the business process and threatening the security of information [28]. Managing IS security in SC ecosystem by using IoT technologies must be concentrated on wholly integrated applications rather than stressing on in a single application. The main criterion that is important to emphasize SC managerial is the level of IS security, whereby the level of IS security among organizations is integrated by setting it at different levels. For instance, certain organizations have been set up as low level of safety, and other organizations have setup as high-level security based on their needs [19].

Table 3: Information System Security Management for IoT-enabled Smart City Challenges

Aspect	Challenges	IS Security	IoT Security
Governance	Formation and management of security standards / policies for IS	[6,14,24–27,29]	[14,45]
	Coordination of multiple stakeholders and organization	[24,28,58, 61]	[14,39]
	Quality assurance	[6,62–64]	[56]
	Citizens' involvement	[14,50]	[39,54]
Integrity	Information security	[6,24,49]	[34,38,42,65]
	Information privacy	[6,15,16,24, 49,63]	[34,37,45,65]
	Existing IS architecture	[14,61,66]	[7,39,65]
	Continuous cyber-attacks	[6,10,12,13, 24,63,67]	[42,48,65]
Interoperability	Readiness of organization	[6,58,68]	[14,45]
	Interoperability implementation	[15,19]	[45,65]
	Secure communication	[4,6,62]	[38,48]
Personalization	Confidentiality	[4,17]	[37,56,65]
	Identity management	[17,69]	[17,35,38,42, 48,54,56,70–72]
	Trust and system reputation	[4,14,73]	[42,45,57,68, 70,74]
Self-Organizing	Threats/Risk management	[4,11,75]	[35,54,71,72]
	Lack of smarter security system	[24]	[71]
	Availability	[17,25,26]	[17,37,48,63, 70,71,76]
	Reliability	[14,69]	[45,77]

Besides that, the SC characteristics, which will be enabled by IoT, are personalization and self-organizing. The personalization component is the individual provision of information, especially based on individual profiles and needs [81]. Self-organization, meanwhile, is a single concept of integrating the whole thing and can be defined as connecting anything to the internet, whether it is a computerized device system with no human-to-human or human-to-computer communication required [69]. One of the IoT elements is to gather and manage individual information automatically in real time. IoT typically consists of unlimited quantities of devices, individuals, and services that link and share information from various resources. Due to additional devices attached to each other in IoT ecosystems and the universal usage, the security and privacy issues come to be the key concerns.

The discussion on the challenges for each aspect is presented below:

4.1. Governance

The leadership and administration in the implementation of an SC can affect security issues. The SC must have all the means to maintain infrastructure and management issues, but weaknesses and frauds can result from inappropriate implementation. Therefore, it is important that the governance of IS security management within the SC is improved by taking into account the integration and exchange of information between different stakeholders, and by carrying out information security assessments. As it indirectly allows organizations to make effective and in-time decisions, the security level of the IS using IoT equipment must be maintained.

Information System (IS) security standard/policy development and management in governance need to be available and organized in the execution of collaboration and information sharing in the SC-enabled IoT [62,78]. Multiple IS incorporation by the use of different IoT devices enables the sharing of information in an SC. This condition will result in chances of IS security risks. The possible risks of cyber-attacks, such as distributed denial-of-service (DDoS), on public infrastructure will rise once devices are connected extensively to generate substantially huge volumes of information [79].

Besides that, coordinating between different stakeholders and organizations due to security management procedure in IS collaboration stays low [64,80]. There is an inadequate standard in the supervision of many stakeholders and no single standard is completely set in the governance of IS protection related to the acceptance, process management and dissemination of information through the use of IoT in SC [58]. In governance, the most comparable activities in IS integration in SC should be correctly described. The governance elements must include strategies, policies, processes and legislation, and accountability. The direction of an organization will be defined by strategies and policies, while processes and legislation will detail who, what, and how. Accountability explains the positions and responsibilities of stakeholders [62]. For example, in the health industry, it is important for medical personnel to safeguard patients' details. Otherwise, the data is open vulnerably, and worse, exploits the responsibility of the staffs for it, so they should be kept responsible for this to prevent future abuses [52,81].

Security measurement for the exchange of information, transmission, cooperation, decision-making, and execution of information exchange during the phase of IS integration at SC must be well-established for IS quality assurance [6,62]. Secure management would therefore require safe monitoring by the coordinator, where the function of the coordinator is to add and remove the IS involved in integration [52,81]. Quality assurance shall ensure that all decisions, processes, and activities remain in accordance with requirements to prevent service risks prior to their occurrence [82]. In order to ensure compliance with information security policies, standards and procedures, rules, regulations or contractual requirements of each company, the implementation of auditing and digital forensic processes is therefore critical [10].

Digital forensic is becoming more critical as an investigative activity for tracing and analyzing criminal and fraudulent activities. Digital forensic is all about cybercrimes, mobile forensics, investigating methods, and analyzing illegal incidents with aims to gather digital evidences. Information will be obtained for law enforcement purposes. Digital forensic may be aided by Intrusion Detection and Prevention Systems (IDPS), as it may confirm to be an important instrument, where its purpose is to conduct initial discovery, track malicious behaviors, and likely avoid further major harm to protected systems. An IDPS is therefore a very valuable instrument for the processing and interpretation of forensic evidence, which can be used for the purposes of a legal proceeding [31].

Other than that, involvement of citizens in SC ecosystem is needed to sustain the IoT usage. A low rate of citizen participation is due to lack of confidence and poor IoT knowledge levels. In order to address this problem, people need to recognize that IoT devices and related services can secure their personal information and help build a sense of empowerment [14,48]. Another important security and privacy measure is the understanding of cyber security and privacy. This is because some citizens do not understand this form of event and the negative impacts it can have, and are thus not in a condition to make decisions on the effect it can have on their privacy standards [83]. Cyber security awareness can begin at early levels, such as kindergartens and schools. Teachers or academic staffs can educate by teaching the implications and risks of cyber incidents. Teachers should also teach what is new, thereby ensuring a better image of the level of knowledge at personal and community levels [31]. Furthermore, it is also important to take social aspects into account. Thus, the "smartness" of a city relies heavily on the participation of citizens in SC projects via numerous communication channels comprising of online portals, social media platforms, and smartphones. In order to share experience and expertise, SC requires people to be actively linked in public locations, public transports, and at homes [4].

4.2. Integrity

Each part in SC, targeted and hacked in IS via cyber-attacks, would breach the integrity of information sharing and the privacy of information of users [82]. The definition of integrity is to ensure that unauthorized changes to system components are protected. Measurement to protect the content, authenticity, and continuity of the message must be taken [52]. There is no

unauthorized alteration of information by permitted or unauthorized personnel, and the information is internally and externally consistent [65, 84].

The weakness of IS security management in SC via the use of IoT will allow unauthorized users to retrieve information [58]. It can cause unauthorized data alterations, damage to information, and loss of information. Thus, the guarantee of information will be questioned by people including the impacts. As a consequence, the security of details that will affect the day-to-day activities of a company is not guaranteed [63, 73]. Information with credibility is reliable information that helps the organization in making the correct decision.

Other than that, the process of preserving data and information confidentiality, authenticating, identifying data, and controlling user access will be the subject of data protection. Theft of identities and phishing are some of the threats to information security, i.e., attempts to trace the misuse of one's financial information, duplication of user accounts, and fake sales or hacker promotions, which can result in data and information being disclosed, updated, and destroyed. Consequently, the integrity of information and data will decrease significantly, and users would lose confidence [65, 73]. Another example is, in medical applications, the node should be able to confirm that the data is sent from a proven trust center. Therefore, by changing the unknown key, the network node and coordinator for all data will make measurements to the Message Verification Code (MAC). Accurate MAC code measurement guarantees the network coordinator that a trustworthy node executes [85].

In addition to that, privacy concerns often affect the integrity of information. Privacy is the right to monitor and protect a single individual's private information. Security must guarantee that without the permission of the owner, none of the single bit of information obtained for a particular user can be shared with others. One of the most critical steps to protect privacy is the creation of rules/policies that have the capability to obtain confidential information in order to protect privacy [52]. Private information is collected in the IoT-enabled SC environment, where various devices are part of public services, for the users to decide with whom the information can be distributed [65,86]. Processing confidential information and the right of users to disclose information on the internet or social media on digital networks are the focal points of this privacy. SCs are exposed to privacy leaks and the collection of information by hackers, particularly, when private information is collected, distributed, and processed. Revealed privacy in SCs may include the identity of a user, venue, transportation movement, health status in healthcare, intelligent surveillance lifestyle, and home and community smart energy. It would be a big mistake to expose this privacy-sensitive data to untrusted or unauthorized people in the real world and cyberspace [73].

Compared to the SC setting that requires IS integrated architecture, some other elements of the constraints of the current IS architecture are linked to initial implementation in silos. The design is divided into software, hardware, and processes [64]. The software architecture is critical for delivering connectivity and allowing IoT devices to share resources between integrated organizations. The hardware/network architecture must be capable in supporting the IoT-critically disseminated computing environment. In an organization, the use of IoT can impact existing business processes. Therefore, to help the innovation in computing and technology, there is a requirement to integrate IoT technology in organizational activities.

Furthermore, a big data project with massive amount of data will arrive in real time. The number, velocity, and variety of data can complicate the process of storage and analysis used to produce important information [48]. Increasing the amount of IoT devices used by SC in IS will provide hackers with opportunities for information security risks as many SC-related devices have low levels of security. IS security monitoring vulnerabilities present a threat to cyber-attacks exposing data to leakage [63, 65, 75] and access breaches. The lack of preparation to handle current cyber-attack threats would result in information being destroyed and lost. Cyber-attack is a security threat that is able to affect the expense of mitigating organizations [75]. Thus, system stability in terms of the system's ability is to react without getting worse because of unexpected attacks. The device must also be capable of protecting other network points from any attacks if one IoT computer is hacked. Therefore, mechanisms to ensure the protection, availability, accuracy, and integrity of the information system must be established during the sharing and processing of information in the SC ecosystem. One of them is the Intrusion Detection and Prevention System (IDPS) which is a computer or software program designed for network or system monitoring. It recognizes weaknesses, reports malicious attacks, and imposes protective methods to keep up with the progress of computer-related crimes via multiple response techniques [87].

4.3. Interoperability

Interoperability is a mechanism for sharing data and using knowledge that combines two or more systems or elements. Interoperability requirements make the system integration process vague, inadequate, and complicated, if not difficult, to execute [86, 88].

The vulnerabilities in IS protection make it impossible to be secure in the management of information exchanges, especially with a view to promote interoperability in the interaction and coordination processes of different IS that would allow services to citizens [24, 54]. Organizations' preparedness to use IoT in organizational activities is still poor and must be strengthened. An automated connection is required between various devices, facilities, and programs. The use of various technologies provided by IS providers would contribute to the need to separately manage the device. Security risks to information inside the SC would be indirectly revealed in the implementation of interoperability between organizations by using different levels of IS protection [80, 89].

Therefore, it is difficult to ensure that the interaction process between IoT devices is complete in each organization that includes multiple devices, especially with regard to memory, power, and time constraints [48]. So, before organizations are ready to be integrated, the requirements for the implementation of interoperability should take into account different levels of IS protection set by the company. This is to ensure that the integrated system's security is maintained [15, 19, 45].

Other than that, secure communication is another critical factor that must be highlighted during integration [65]. It will guarantee the security of authentication, confidentiality, and integrity of the linked data. It also prevents a message exchange from being repudiated, and preserves the identity of the users involved. Moreover, it is insufficient to protect IoT devices solely to ensure that the IoT system is completely protected [38,48].

4.4. Personalization

Personalization services are delivered according to individual profiles and preferences in a unique and precise way [79]. Confidentiality and privacy prevent data to go against unauthorized access where data indicates the security of an exposure to sensitive data that is deemed to be the critical issue. For example, in a medical context, sensitive and personal information about a patient's well-being is required and relied on to be transmitted, so the patient's data must be shielded from unauthorized access that may be harmful to the safety of the patient. Encryption will provide this sensitive data with greater security by using a mutual key to secure communication [52]. In order to make life safer and easier, the IS used every day will automatically collect personal information in real time, which means that IS can also monitor everyday life activities. If the IoT system control is lost or stolen, it will be a serious potential security concern [74]. In comparison, in the internet world, there is no chance for attackers to access the information if individuals do not supply the necessary details.

One aspect that needs to be highlighted in information protection is identity management. The access control restrictions for approved users are also not adequately implemented. Consequently, attackers may take advantage of these vulnerabilities to retrieve unauthorized features and/or information, such as accessing user accounts, watching confidential files, manipulating user information, and modifying access privileges [17,69]. Moreover, in the functions of applications, authentication and session management are still not properly implemented. Cyber-attacks affect IS authentication by enabling third parties to manipulate the original data to make it unreliable. Attackers can modify or manipulate other weaknesses via passwords, keys, and session tokens [17].

Besides that, trust and device credibility are essential aspects of using IoT to manage IS security. Trust ensures that information and resources are fully and confidentially accessed by users and IoT devices. Competent data collection, powerful data combination and mining, and enhanced user confidentiality are included in trust management. The present challenges are determining how trust is established between IoT devices, and determining the trustworthiness of a user in the use of IoT devices [17]. In IoT, consideration of two aspects of trust must be emphasized, consisting of trust concerning the interconnection among entities and trust in the system from the clients' viewpoint [57].

Personalization relies on the usage of private information, and in this IoT feature, protection and privacy issues are therefore major concerns. Organizations need to be fulfilled with security requirements, such as identity protection, privacy, data access control, precise authentication procedures, and trustworthy identity to resolve this.

4.5. Self-Organizing

Self-organizing is the administration of automated Machine-to-Machine (M2M) acceptance, processing, and distribution of information without human intervention [24,58,90]. In other words, in order to generate customer-oriented output that continuously operates to sustain itself, computers will function independently or coordinate with humans. The machines are thus autonomous entities that can gather and interpret information and provide guidance on the basis of research [90].

One aspect that needs to be highlighted in self-organizing is risk management. Risk is an essential aspect of the management of IS security in IoT-enabled SC. In the development and implementation of IS, risk is an unavoidable factor. Successful control of risk reduces an organization's operating risk. One of the key causes of IS failure is the weakness in risk management for IS growth, including prediction and evaluating risk [91,92]. IoT risk analysis requires the detection of assets, risks, and vulnerabilities. Failure to foresee and evaluate these vulnerabilities can lead to risks being generated. Other than that, the usage of poor application protection elements and the Application Programming Interface (API) can encourage a broad range of IS safety attacks [71,72]. In addition, the absence of a Smarter Security System to handle the identification of threats, the identification of anomalies, and the effects of predictive analysis affects IS performance [71]. Less secure and slower connections between IoT nodes lead to data leaks and other security breaches [74]. Another key element is availability and usability, which makes sure that the IS performs entirely at any time and every time an authenticated user is detected [17]. If any IS operation fails, the protection must ensure that equivalent resources are available, and as an added assurance, must allow M2M operations, i.e., real-time data collection will continue with IoT devices. Due to small and lightweight characteristics that make them operate on lower energy, the problems of resource constraint arises as most IoT devices have restricted handling and storage capacities [51]. Besides that, in medical practices, a network availability with effective admission to the patient's information is crucial, especially involving a system which contains important, sensitive, and potentially lifesaving information. Thus, the network must be available all the time [85].

The IoT technology enables users to be self-organized and personalized with data-driven decisions. Data-driven decision-making is a process that involves the collection of data based on established concrete objectives and the discovery of evidence, trends, associations, observations, and knowledge from this information. This expertise is then used to build or evaluate processes, operations, structures, policies, and techniques to support the data/system owner [5]. Information monitoring includes multi-device collaboration, which can effectively improve the accuracy and reliability of user-acquired information without fail. Another important problem is failure to collect the right data, as it can become a life-threatening matter for the citizens. Therefore, appropriate techniques can be used to ensure the IS is accurate, complete, reliable, and secure from malicious attacks during information transactions [52]. By using IoT technology, it allows information to be managed by machines, and enhances IS protection in the SC, such as protection against data leaks. By determining a corrective and preventive plan that

focuses on safety concern, each company in the SC needs to deliver a holistic risk management strategy.

5. Conclusion

This paper discusses the major challenges in the IS security management for IoT-enabled SCs. The document analysis discovers security challenges according to five aspects, namely: (i) governance, (ii) integrity, (iii) interoperability, (iv) personalization, and (v) self-organizing. It has been found that it is more complicated to protect IS from the heterogeneous IoT in SCs. At anytime, anywhere, and on any device, confidential data is exposed to malicious cyber-attacks. This study is expected to assist SC policy makers, city planners, and practitioners in understanding and addressing the challenges in sustaining IS security management for IoT-enabled SCs. This will lead to planning and the development of SCs to improve the citizens' quality of life. Future work must identify authentication features that are appropriate for IS security management by adopting IoT in SC environment. The aim is to overcome any unauthorized access to the sensitive and confidential information due to cyber-attacks.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

The study is supported by the Fundamental Research Grant Scheme (FRGS/1/2019/ICT04/UKM/03/2), 2019, Ministry of Education Malaysia and Universiti Kebangsaan Malaysia.

References

[1] Z. Din, D.I. Jambari, M.M. Yusof, J. Yahaya, "Challenges in Managing Information Systems Security for Internet of Things-enabled Smart Cities," in 6th International Conference on Research and Innovation in Information Systems, IEEE, Bangi, Selangor Malaysia, 2019.

[2] United Nations, World Urbanization Prospects - The 2014 revision, United Nations, Department of Economic and Social Affairs, 2014.

[3] L. Pascu, The IoT Threat Landscape and Top Smart Home Vulnerabilities in 2018, Bitdefender, 1–18, 2018.

[4] R. Khatoun, S. Zeadally, "Smart Cities : Concepts, Architectures, Research Opportunities," Communications of The ACM, **59**(No. 8), 46–57, 2016.

[5] N. Mohamed, J. Al-Jaroodi, I. Jawhar, N. Kesserwan, "Data-Driven Security for Smart City Systems: Carving a Trail," IEEE Access, **8**, 147211–147230, 2020, doi:10.1109/access.2020.3015510.

[6] K.C. Laudon, J.P. Laudon, Management information systems : Managing The Digital Firm (15th Edition), Pearson Education Limited, London, 2018, doi:10.1007/978-94-017-9618-7_44.

[7] K. Sha, W. Wei, T.A. Yang, Z. Wang, W. Shi, "On security challenges and open issues in Internet of Things," Future Generation Computer Systems, **83**, 326–337, 2018, doi:10.1016/j.future.2018.01.059.

[8] J. Howell, Number of Connected IoT Devices Will Surge to 125 Billion by 2030, IHS Markit Says, IHS Markit, 2017.

[9] Y. Ye, T. Li, D. Adjeroh, S.S. Iyengar, "A Survey on Malware Detection Using Data Mining Techniques," ACM Computing Surveys, **50**(3), 41:1–41:40, 2017, doi:10.11648/j.ijiis.s.2014030601.16.

[10] H. Susanto, M.N. Almunawar, INFORMATION SECURITY MANAGEMENT SYSTEMS, Taylor & Francis Group, U.S, 2018.

[11] N. Yousefnezhad, A. Malhi, K. Främling, "Security in Product Lifecycle of IoT Devices : A Survey," in 52nd ACM/EDAC/IEEE Design Automation Conference (DAC), Elsevier Ltd, San Francisco, CA,: 102779, 2015, doi:10.1016/j.jnca.2020.102779.

[12] M.P. Dijkers, "A beginner's guide to data stewardship and data sharing," Spinal Cord, **57**(3), 169–182, 2019, doi:10.1038/s41393-018-0232-6.

[13] N. Noori, T. Hoppe, M. de Jong, "Classifying pathways for smart city development: Comparing design, governance and implementation in Amsterdam, Barcelona, Dubai, and Abu Dhabi," Sustainability (Switzerland), **12**(10), 2020, doi:10.3390/SU12104030.

[14] N. Noori, M. De Jong, T. Hoppe, "smart cities Towards an Integrated Framework to Measure Smart City Readiness : The Case of Iranian Cities," Smart Cities, 676–704, 2020.

[15] A. Glaser, N. Jeambon, Smart City Platforms … and Aligning Technology and Citizens, 2018.

[16] A. Martinez-Balleste, P. Perez-Martinez, A. Solanas, "The Pursuit of Citizens' Privacy: A Privacy-Aware Smart City is Possible," IEEE Communications Magazine, **51**(6), 136–141, 2013, doi:10.1109/MCOM.2013.6525606.

[17] A.E. Hassanien, M. Elhoseny, S.H. Ahmed, Amit Kumar Singh, Security in Smart Cities: Models, Applications, and Challenges, Springer Nature Switzerland AG 2019, Switzerland, 2019, doi:10.1007/978-3-030-01560-2.

[18] S. Babar, A. Stango, N. Prasad, J. Sen, R. Prasad, "Proposed embedded security framework for Internet of Things (IoT)," 2011 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology, Wireless VITAE 2011, (February), 2011, doi:10.1109/WIRELESSVITAE.2011.5940923.

[19] L. Yang, N. Elisa, N. Eliot, Privacy and Security Aspects of E-Government in Smart Cities, Smart Cities Cybersecurity and Privacy, 89–102, 2018, doi:10.1016/b978-0-12-815032-0.00007-x.

[20] Z.A. Baig, P. Szewczyk, C. Valli, P. Rabadia, P. Hannay, M. Chernyshev, M. Johnstone, P. Kerai, A. Ibrahim, K. Sansurooah, N. Syed, M. Peacock, "Future challenges for smart cities: Cyber-security and digital forensics," Digital Investigation, **22**(August), 3–13, 2017, doi:10.1016/j.diin.2017.06.015.

[21] I. Nadir, Z. Ahmad, H. Mahmood, G. Asadullah Shah, F. Shahzad, M. Umair, H. Khan, U. Gulzar, "An auditing framework for vulnerability analysis of iot system," Proceedings - 4th IEEE European Symposium on Security and Privacy Workshops, EUROS and PW 2019, (October), 39–47, 2019, doi:10.1109/EuroSPW.2019.00011.

[22] M. Conti, A. Dehghantanha, K. Franke, S. Watson, "Internet of Things security and forensics: Challenges and opportunities," Future Generation Computer Systems, **78**, 544–546, 2018, doi:10.1016/j.future.2017.07.060.

[23] ISO/IEC, ISO/IEC 27001:2013, ISO, Switzerland, 2013.

[24] P.T.I. Lam, R. Ma, "Potential pitfalls in the development of smart cities and mitigation measures: An exploratory study," Cities, **91**(August 2019), 146–156, 2018, doi:10.1016/j.cities.2018.11.014.

[25] K.C. Laudon, J.P. Laudon, A. Elragal, Management Information Systems: Managing the Digital Firm, 2016, doi:10.1590/S1415-65552003000100014.

[26] K. Salamzada, Z. Shukur, M.A.B.U. Bakar, "A Framework for Cybersecurity Strategy for Developing Countries: Case Study of Afghanistan," Asia-Pacific Journal of Information Technology and Multimedia, **4**(1), 1–10, 2015.

[27] Angraini, R.A. Alias, Okfalisa, "Information security policy compliance: Systematic literature review," Procedia Computer Science, **161**, 1216–1224, 2019, doi:10.1016/j.procs.2019.11.235.

[28] M.-D. Mclaughlin, J. Gogan, "Challenges and Best Practices in Information Security Management," Mis Quarterly Executive, **17**(3), 237–262, 2018.

[29] H. Paananen, M. Lapke, M. Siponen, "State of the art in information security policy development," Computers & Security, **88**, 101608, 2020, doi:10.1016/j.cose.2019.101608.

[30] M.A. Hasbini, T. Eldabi, A. Aldallal, "Investigating the information security management role in smart city organisations," World Journal of Entrepreneurship, Management and Sustainable Development, **14**(1), 86–98, 2018, doi:10.1108/WJEMSD-07-2017-0042.

[31] S. Al-janabi, I. Al-shourbaji, "A Study of Cyber Security Awareness in Educational Environment in the Middle East," Journal of Information & Knowledge Management, **15**(1), 1–30, 2016, doi:10.1142/S0219649216500076.

[32] R. Kitchin, "The real-time city ? Big data and smart urbanism," GeoJournal, **79**(November 2013), 1–14, 2014, doi:10.1007/s10708-013-9516-8.

[33] ITU, "Overview of the Internet of things," Series Y: Global Information Infrastructure, Internet Protocol Aspects And Next-Generation Networks, 1–22, 2012.

[34] M. Abomhara, Geir M. Køien, "Security and Privacy in the Internet of Things: Current Status and Open Issues," in 2014 International Conference on Privacy and Security in Mobile Systems (PRISMS), IEEE, Aalborg, Denmark: 1–8, 2014, doi:10.1109/MC.2018.2888765.

[35] A.-R. Sadeghi, C. Wachsmann, M. Waidner, "Security and Privacy Challenges in Industrial Internet of Things," in 52nd ACM/EDAC/IEEE Design Automation Conference (DAC), IEEE, San Francisco, CA, USA: 1–6, 2015, doi:10.1145/2744769.2747942.

[36] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, "Internet of

Things for Smart Cities," IEEE Internet of Things Journal, **1**(1), 22–32, 2014, doi:10.1109/JIOT.2014.2306328.
[37] M. Noor, W.H. Hassan, "Current research on Internet of Things (IoT) security : A survey," Computer Networks, **148**, 283–294, 2019, doi:10.1016/j.comnet.2018.11.025.
[38] S. Babar, P. Mahalle, A. Stango, N. Prasad, R. Prasad, "Proposed security model and threat taxonomy for the Internet of Things (IoT)," Communications in Computer and Information Science, **89 CCIS**, 420–429, 2010, doi:10.1007/978-3-642-14478-3_42.
[39] C.E.A. Mulligan, M. Olsson, "Architectural implications of smart city business models: An evolutionary perspective," IEEE Communications Magazine, **51**(6), 80–85, 2013, doi:10.1109/MCOM.2013.6525599.
[40] H.F. Atlam, R.J. Walters, G.B. Wills, "Intelligence of Things : Opportunities & Challenges," in 3rd Cloudification of the Internet of Things (CIoT), IEEE: 1–6, 2018.
[41] M. Kanellos, Amount of Data Created Annually to Reach 180 Zettabytes in 2025, IDC, **2016**(March 7, 2016), 2016.
[42] R. Thirukkumaran, P. Muthu Kannan, "Survey: Security and Trust Management in Internet of Things," Proceedings - 2018 IEEE Global Conference on Wireless Computing and Networking, GCWCN 2018, 131–134, 2019, doi:10.1109/GCWCN.2018.8668640.
[43] L. O'Donnel, 2 Million IoT Devices Vulnerable to Complete Takeover, Threatpost.Com, 2019.
[44] A. Meola, What is the Internet of Things & How Does IoT Work, Business Insider, 2018.
[45] K.K. Patel, S.M. Patel, "Internet of Things-IOT : Definition , Characteristics , Architecture , Enabling Technologies , Application & Future Challenges," International Journal of Engineering Science and Computing, **6**(5), 6122–6131, 2016, doi:10.4010/2016.1482.
[46] S. Sicari, A. Rizzardi, L.A. Grieco, A. Coen-porisini, "Security , privacy and trust in Internet of Things : The road ahead," Computer Networks, (January 2015), 2018, doi:10.1016/j.comnet.2014.11.008.
[47] L. Ismail, L. Zhang, Information Innovation Technology in Smart Cities, Springer Nature, 2018.
[48] H.F. Atlam, G.B. Wills, IoT Security, Privacy, Safety and Ethics, Springer Nature Switzerland AG 2020, Switzerland: 123–149, 2020, doi:10.1007/978-3-030-18732-3_8.
[49] A.S. Elmaghraby, M.M. Losavio, "Cyber security challenges in smart cities: Safety, security and privacy," Journal of Advanced Research, **5**(4), 491–497, 2014, doi:10.1016/j.jare.2014.02.006.
[50] R.P. Dameri, Urban Smart Dashboard. Measuring Smart City Performance, 2017, doi:10.1007/978-3-319-45766-6.
[51] W. Aman, "Modeling Adaptive Security in IoT Driven e Health.pdf," in Norwegian Information Security Conference (2013), 61–69, 2013.
[52] S. Al-Janabi, I. Al-Shourbaji, M. Shojafar, S. Shamshirband, "Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications," Egyptian Informatics Journal, **18**(2), 113–122, 2017, doi:10.1016/j.eij.2016.11.001.
[53] J. Cho, A. Swami, I. Chen, "A Survey on Trust Management for Mobile Ad Hoc Networks," IEEE COMMUNICATIONS SURVEYS & TUTORIALS, **13**(4), 562–583, 2011.
[54] A. Merella, IoT security issues, risks and threats this year, Apiumhub, 2018.
[55] A. Aldairi, L. Tawalbeh, "Cyber Security Attacks on Smart Cities and Associated Mobile Technologies," Procedia Computer Science, **109**(2016), 1086–1091, 2017, doi:10.1016/j.procs.2017.05.391.
[56] S. Sicari, A. Rizzardi, L.A. Grieco, A. Coen-Porisini, "Security, privacy and trust in Internet of Things: The road ahead," Computer Networks, **76**, 146–164, 2015, doi:10.1016/j.comnet.2014.11.008.
[57] R. Roman, J. Zhou, J. Lopez, "On the features and challenges of security and privacy in distributed internet of things," Computer Networks, **57**(10), 2266–2279, 2013, doi:10.1016/j.comnet.2012.12.018.
[58] Z.K. Aldein Mohammeda, E.S. Ali Ahmed, "Internet of Things Applications, Challenges and Related Future Technologies," World Scientific News, (February), 126–148, 2017.
[59] V. Mohammadi, A.M. Rahmani, A.M. Darwesh, A. Sahafi, "Trust - based recommendation systems in Internet of Things : a systematic literature review," Human-Centric Computing and Information Sciences, 1–61, 2019, doi:10.1186/s13673-019-0183-8.
[60] T.J. Ellis, Y. Levy, "A Systems Approach to Conduct an Effective Literature Review in Support of InformationThe Literature Review: The Foundation for Research," Informing Science Journal, **9**, 1–39, 2006.
[61] M. Ammar, G. Russello, B. Crispo, "Internet of Things: A survey on the security of IoT frameworks," Journal of Information Security and Applications, **38**, 8–27, 2018, doi:10.1016/j.jisa.2017.11.002.
[62] R.W.S. Ruhlandt, "The governance of smart cities: A systematic literature review," Cities The International Journal of Urban Policy and Planning, (October 2017), 1–23, 2018, doi:10.1016/j.cities.2018.02.014.
[63] N. Dong, J. Zhao, L. Yuan, Y. Kong, "Research on Information Security System of Smart City Based on Information Security Requirements," Journal of Physics: Conference Series, **1069**(1), 2018, doi:10.1088/1742-6596/1069/1/012040.
[64] A. Whitmore, A. Agarwal, L. Da Xu, "The Internet of Things—A survey of topics and trends," Information Systems Frontiers, **17**(2), 261–274, 2015, doi:10.1007/s10796-014-9489-2.
[65] Y. Lu, L. Da Xu, "Internet of things (IoT) cybersecurity research: A review of current research topics," IEEE Internet of Things Journal, **6**(2), 2103–2115, 2019, doi:10.1109/JIOT.2018.2869847.
[66] H. Arasteh, V. Hosseinnezhad, V. Loia, A. Tommasetti, O. Troisi, M. Shafie-khah, P. Siano, "IoT-based smart cities: A survey," EEEIC 2016 - International Conference on Environment and Electrical Engineering, (June), 2016, doi:10.1109/EEEIC.2016.7555867.
[67] M.A. Wahab, D.I. Jambari, "Service Level Agreement Parameters for Drafting Public Sector Information System Contract," Jurnal Pengurusan, **52**, 153–167, 2018.
[68] R. Derakhshan, R. Turner, M. Mancini, "Project governance and stakeholders: a literature review," International Journal of Project Management, **37**(1), 98–116, 2019, doi:10.1016/j.ijproman.2018.10.007.
[69] M. Sookhak, H. Tang, Y. He, F.R. Yu, "Security and Privacy of Smart Cities: A Survey, Research Issues and Challenges," IEEE Communications Surveys and Tutorials, **PP**(c), 1, 2018, doi:10.1109/COMST.2018.2867288.
[70] S. Ijaz, M. Ali, A. Khan, M. Ahmed, "Smart Cities: A Survey on Security Concerns," International Journal of Advanced Computer Science and Applications, **7**(2), 2016, doi:10.14569/ijacsa.2016.070277.
[71] M. Irshad, "A systematic review of information security frameworks in the internet of things (IoT)," Proceedings - 18th IEEE International Conference on High Performance Computing and Communications, 14th IEEE International Conference on Smart City and 2nd IEEE International Conference on Data Science and Systems, HPCC/SmartCity/DSS 2016, 1270–1275, 2017, doi:10.1109/HPCC-SmartCity-DSS.2016.0180.
[72] Andrew van der Stock, B. Glas, N. Smithline, T. Gigler, OWASP Top 10 - 2017 The Ten Most Critical Web Application Security Risks, Creative Commons, **1**(1), 1–24, 2017, doi:10.1002/kin.550040606.
[73] K. Zhang, J. Ni, K. Yang, X. Liang, J. Ren, and X. (Sherman) Shen, Security and Privacy in Smart City Applications: Challenges and Solutions, 2017, doi:10.1080/00207168808803619.
[74] Q. Jing, A. V. Vasilakos, J. Wan, J. Lu, D. Qiu, "Security of the Internet of Things: perspectives and challenges," Wireless Networks, **20**(8), 2481–2501, 2014, doi:10.1007/s11276-014-0761-7.
[75] M.A. Hasbini, T. Eldabi, A. Aldallal, "Investigating the information security management role in smart city organisations," World Journal of Entrepreneurship, Management and Sustainable Development, **14**(1), 86–98, 2018, doi:10.1108/WJEMSD-07-2017-0042.
[76] Icon Labs, Floodgate Security Framework | Icon Labs, Icon Labs, 2019.
[77] J. Fan, P. Zhang, D.C. Yen, "G2G information sharing among government agencies," Information and Management, **51**(1), 120–128, 2014, doi:10.1016/j.im.2013.11.001.
[78] P. Radanliev, "Cyber Risk Management for the Internet of Things," University of Oxford, (April), 1–27, 2019, doi:10.13140/RG.2.2.34482.86722.
[79] A. Gharaibeh, M.A. Salahuddin, S.J. Hussini, A. Khreishah, I. Khalil, M. Guizani, A. Al-Fuqaha, "Smart Cities: A Survey on Data Management, Security, and Enabling Technologies," IEEE Communications Surveys and Tutorials, **19**(4), 2456–2501, 2017, doi:10.1109/COMST.2017.2736886.
[80] S. Theodorou, N. Sklavos, Chapter 3 - Blockchain-Based Security and Privacy in Smart Cities, Elsevier Inc., Greece: 21–37, 2019, doi:https://doi.org/10.1016/B978-0-12-815032-0.00003-2.
[81] V. Ekong, U. Ekong, "a Survey of Security Vulnerabilities in Wireless Sensor Networks," Nigerian Journal of Technology, **35**(2), 392, 2016, doi:10.4314/njt.v35i2.21.
[82] A.A.A. Al-Wosabi, Z. Shukur, "Proposed system architecture for integrity verification of embedded systems," Journal of Engineering and Applied Sciences, **12**(9), 2371–2376, 2017, doi:10.3923/jeasci.2017.2371.2376.
[83] A. Patel, S. Al-janabi, I. Alshourbaji, "A novel methodology towards a trusted environment in mashup web applications," Computers & Security, **49**, 107–122, 2014, doi:10.1016/j.cose.2014.10.009.
[84] P.P. Pereira, J. Eliasson, J. Delsing, "An authentication and access control framework for CoAP-based Internet of Things," IECON Proceedings (Industrial Electronics Conference), (November), 5293–5299, 2014, doi:10.1109/IECON.2014.7049308.
[85] G. Belleville, "Sit Down and Write Your Thesis! Practical and Motivational

Tips for Scientific Writing," Canadian Journal of Cardiology, **35**(8), 945–947, 2019, doi:10.1016/j.cjca.2019.04.011.

[86] D. Maheshwari, M. Janssen, "Reconceptualizing measuring, benchmarking for improving interoperability in smart ecosystems: The effect of ubiquitous data and crowdsourcing," Government Information Quarterly, **31**(SUPPL.1), 1–9, 2014, doi:10.1016/j.giq.2014.01.009.

[87] N.A. Azeez, T.M. Bada, S. Misra, A. Adewumi, C. Van der Vyver, R. Ahuja, "Intrusion Detection and Prevention Systems: An Updated Review," Advances in Intelligent Systems and Computing, **1042**(January), 685–696, 2020, doi:10.1007/978-981-32-9949-8_48.

[88] H. Van Der Veer, A. Wiles, Achieving Technical Interoperability, France, 2008.

[89] S. Madakam, R. Ramaswamy, S. Tripathi, "Internet of Things (IoT): A Literature Review," Journal of Computer and Communications, **03**(05), 164–173, 2015, doi:10.4236/jcc.2015.35021.

[90] T.K. Sung, "Industry 4.0: A Korea perspective," Technological Forecasting and Social Change, **132**(November 2017), 40–45, 2018, doi:10.1016/j.techfore.2017.11.005.

[91] S.F. Abdullah, M.M. Yusof, D.I. Jambari, "Risk Management Model for Information Systems Planning in Public Sector," Jurnal Pengurusan, **48**, 149–160, 2016, doi:http://dx.doi.org/10.17576/pengurusan-2016-48-12 Model.

[92] B. Baharuddin, M.M. Yusof, "Evaluation of risk management practices in information systems project in the public sector," Jurnal Pengurusan, **53**, 20, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 100-109 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Web-based Remote Lab System for Instrumentation and Electronic Learning

Jose María Sierra-Fernández*, Agustin Agüera-Pérez, Jose Carlos Palomares-Salas, Manuel Jesús Espinosa-Gavira, Olivia Florancias-Oliveros, Juan José González de la Rosa

Research group PAIDI-TIC-168, Electronics, University of Cádiz, Algeciras, 11202, Spain

ARTICLE INFO

Article history:
Received: 02 December, 2020
Accepted: 15 April, 2021
Online: 10 July, 2021

Keywords:
Remote lab
Web Interface
Flex design
Minimal network privileges
Scalable

ABSTRACT

Lab sessions in Engineering Education are designed to reinforce theoretical concepts. However, there is usually not enough time to reinforce all of them. Remote and virtual labs give students more time to reinforce those concepts. In particular, with remote labs, this can be done interacting with real lab instruments and specific configurations. This work proposes a flexible configuration for Remote Lab Sessions, based on some of 2019 most popular programming languages (Python and JavaScript). This configuration needs minimal network privileges, it is easy to scale and reconfigure. Its structure is based on a unique Reception-Server (which hosts students database, and Time Shift Manager, it is accessible from the internet, and connects students with Instruments-Servers) and some Instrument-Servers (which manage hardware connection and host experiences). students always connect to the Reception-Server, and book a time slot for an experience. During this time slot, User is internally forwarded to Instrument-Server associated with the selected experience, so User is still connected to the Reception-Serer. In this way, Reception-Server acts as a firewall, protecting Instrument-Servers, which never are open to the internet. A triple evaluation system is implemented, user session logging with auto-evaluation (objectives accomplished), a knowledge test and an interaction survey. An example experience is implemented, controlling a DC source using Standard Commands for Programmable Instruments. This is an example regarding how systems enable students to interact with hardware, giving the opportunity of understand real behaviour.

1. Introduction

In Engineering learning, Lab sessions act as a reinforcement of theoretical concepts, taking a really important role in student education, e.g. manipulate diodes in lab, examine threshold and only positive conduction, gives an additional level of knowledge, increasing student interest and making putting in order some of theory studied. However, there is usually not enough time to put in practice all theoretical concepts, due to time limitations to access to physical lab sessions. In this line, more experiences can be offered, expanding lab access time to 24/7, this is Remote Labs.

Sometimes, Remote Lab concept is put together with Virtual Lab (a simulated experience). When these two options want to be compared, few concepts should be considered, as stated in [1]. There, cost and effect over learning are studied. Conclusions of that work postulate that the increase of the cost of Remote Lab is compensated by the improvement in the learning process, due to students are more involved when they feel that they are interacting with real equipment.

Remote labs allow students to access expensive lab Instruments, and specific configuration, any time, in order to put in practice concepts studied in theoretical lessons. Remote access to labs instruments is gaining attention in order to grant access to students, which cannot be there (medical problems, case of force majeure, or other excused absence). Nowadays, at 2020 globally outbreak of COVID-19 is one of those cases of force majeure, which has turn into remote as more activities as possible.

First option for Remote Labs is to use a commercial equipment, as VISIR, as stated in [2], [3], which is a matrix of connections and components, where student can change connections. This same philosophy can be applied to matrix connections to servers, for networking lab [4]. In other situations, Remote labs are mixed with virtual labs, an example is Easy Java Simulations (EJS) as stated in [5]. This system, designed for simulated experiences

*Corresponding Author: Jose María Sierra-Fernández, Email: josemaria.sierra@gm.uca.es

https://dx.doi.org/10.25046/aj060412

(virtual labs), can be connected with hardware (oscilloscopes, engines, sensors, etc.) to perform an interactive experience.

There are too non-commercial systems in Remote Labs, e.g., as stated in [6], [7], where Labicom, a complete new system was designed. Server, client, solutions, all ad-hoc solutions are designed.

Even a compact solution has been developed, all integrated in a Raspberry Pi, as stated in [8], hosting a webserver, with an Arduino as sensor interface.

Other implementations, as the stated in [9], proposed a complete system for flexible Remote Lab testing, based in Lab Server and Web Server. It is a good approach for distributed systems, due to all communications are done throw the internet. However, it requires for each Lab Server, external access, if it is in a corporate network, it is a complex, or impossible configuration.

Even there are different approaches for Remote Labs, centred in collaborative work, as stated in [10]. This work is centred in a system where few students can work together in the same experience.

With a general study of the work, as stated in [11], it can be observed that these solutions are only a part of the problem, and it is demonstrated that a system structure stable, flexible and scalable is needed.

However, all reviewed Remote Labs systems are not enough flexible, due to hardware compatibility are limited, or web interface are not enough intuitive. The aim of this work is to design a structure and function of a Remote Lab system for engineering learning, flexible, in order to be able to connect any Instrument (with computer interface), scalable (in order to add nodes) and with the required data security level.

With that objective, servers are designed using the programming language Python, one of the most popular programming languages with several libraries, which allows interacting with almost any device. In addition, it can be set up as a web server, providing sufficient level of security and simplifying and ensuring data storage. The user interface is designed in HTML and JavaScript, introducing a fluid and asynchronous experience.

This work is structured as follows: in section II general lines and objectives of this project are given; then in section III the system structure is explained, followed in IV, where the relation with learning and examples are explained, the evaluation methods for learning outcomes are explained in section V, finally conclusions are given in section VI.

2. System Objective

The aim of this work is to present a complete Remote Lab solution which can implement experiences related with engineering learning. This system must manage user login, and user data, guaranteeing their privacy, and must have a time shift manager, in order to create time slots for remote experiences, allowing users to book them. In addition, networks requirements needed for the implementation in a complex network must be minimal, due to University networks are usually really complex.

This proposal must allow students the interaction with Lab Equipment from anywhere (from the internet). For that, only one Student can use each lab site at time. With this aim, time-slots are set, in order to coordinate students access. With this, Lab Instruments can be used beyond the Lab sessions hours.

2.1. Basic System Design

Experiences offered in remote experinces should be addequated to lessons. So it should be configurable, with the posibility to have some experiences in the same lab site, allowing to offer multiple remote experiences with the same equipments.

In reation with network requirements, learning labs use to be in complex networks, wich any special requiremnt (as direct access to a computer from the internet) is a risk for all the network. For that, only one system would have direct acces from the internet. This single system would manage user data, login and book system, at the time proxy conection to lab sites. Systems connected to lab equipments are in internall network, so it does no suposse a risk to the network, and only requiere a static IP.

User (student) connects to this single computer, for book a experince, in a specific time slot, for take the experince, or for take the evaluation excercises. All user interaction is done via this computer (connections, forwardings, data transmissions, or other procedures to connect user with remote lab station), so user seem to interact always with the same server, in the same web-site.

Aim of this remote lab system is to reinforce of theoretical knowledge, so test learning result is one fo the most important things in the process. The most dificult thing is to know if the experience has been done as designed (without any interaction problem). Interaction problems cause difficulties in perform laboratory practice and difficulty in acquiring knowledge, and aspect as camera quality is essential for a propper interaction. In order to know if some kind of problem in experince design (interaction site, software conectors, conection speed, camera quality, camera ilumination, etc.) could cause problems at taking the experience, a fast survey must be done for every user after take an experience.

Moreover, knowledge is evaluated in a multiple way. Examinig the experince itself (steps accoplished, time taken, total time in finish experince, tries for step, etc.) and an auto-evaluation is geenrated. And after the experince, a knowledge test must be take by user. This test include the knowledges concepts included in the experince, in order to confirm that experience has helped in reinforce that knowledge.

For a fair use of the systen, a limit at booking time slots must be considered. Deppendig on the number of users and time slots availables for take each experince, more or less conditions should be applied. It is recommended a limit of only one time slot booked by user, but sometimes, it should be considered additional limts as one try for each experince by user (at least until all users have done that experince). Time slots are created, and number of remote lab stations are limmited. So this limits are really importat in some situations.

Time slot sould not be the same for all experinces, and even some experinces can be done simultaneously, if both use different equipments connected to the same remote lab station. Roles for

that situations are implemented in remote lab stations. Time slots are calculated according to expriences time length, taking in consideration that some time is needed among time slots for intialize instruments.

Remote lab stations have instruments grouped by experinces compatibility, this is a set of instruments that can be used for many experiences. They are designed as much flexible as possible, but with just the needed instruments, in order to create as much as remote lab stations as possible. This is, a remote lab station can be done with a controlable DC source, a Fuction generator, a Multimeter and a Oscilloscope. These equipments can be used for a huge set of experiences in the areas of electronics (analog and digital). However, many experiences can be done only with a DC controlable source and mutimeter, and many others just with a function generatorn an oscilloscope and a fix DC source. So with almost same Lab Instruments, two lab stations can be done (for a lower set of experiences) or one (for a higher set of experinces, including more complex ones). Deppending on the kind of experiences wich want to be done, and the number of users, Instruments are gruped in one way or another.

With that, a clear idea of the system has been given, now a description of the system implementation is going to be done.

3. System Structure

Proposed remote lab system structure is organized around two types of servers, as explained before. One single server, which manage the whole system, and one server per remote lab station. All of then must have implemented a web-server, for enable the user web interaction, in addition with the application programming interface (API) interaction. This is a machine-machine interaction, via web-request. For this propose, servers need a back-end programming, and web interface needs a front-end design and programming.

For the back-end (all procedures and services done by the server, hidden by the user, as web server itself, camera stream, user login, data management, etc.), Python language have been selected. This is a really extended and supported language, in continuous development, with many packages, which gives many extra functionalities to it. In particular, a framework for web-server creation have been selected, Flask, which manage user login, connections, interactions, and more. This will be explained in more detail in both server types description.

For front-end (user interaction experience, including visualization and data interchange with backend), in web sites HTML and CSS are used for design appearance of websites, and JavaScript is used for introduce “programming” in web sites. While HTML and CSS design the content and structure, this design is static, and only can change at web site refresh. JavaScript allow elements resize and reorganize at window size change, update data without refresh all the website, even process data inside the browser, in the user computer.

Moreover, Python and JavaScript are the most popular languages in 2019 as stated in [12], [13], which grant many support and community. A schema about languages used are shown in Figure 1.

3.1. Instrument-Server

First, remote lab station will be examined. These stations are composed of a set of instruments, components and connections, and a server, which manage that hardware. That server is named Instrument-Server, and host remote lab experiences and manage the communication with instruments, using a module named Instrument-connector.

Instrument-Server is a web-server with many functions, with the back-end designed in Python, relation with hardware is done in an easy way with calls to Instrument-Connector. This module involves all procedures to communicate with hardware. Python has libraries for exchange data via GPIB, USB, and LAN, using VISA protocols, RS-232, and many others, with them, and Instrument documentation, a set of functions are designed for each Instrument, for access to all needed functions. Sometimes, procedures would be a single line, others a set of command exchange. Finally, from Instrument-Server, Lab Instrument interaction is done as a single call to Instrument-Connector.

The specific structure of this module depends of the specific hardware connected, but if we have a GPIB interaction with a Multimeter, from Instrument-Server there would be an instruction “read_voltage()”. Inside Instrument-connector, in the address selected for the instrument, some order are sent (mode voltmeter, range-auto, read) and then return is taken from GPIB bus from the address of the Multimeter. All this sequence, with time among them, watchdogs, and other specifications of GPIB, and error handle, are managed by the method read_voltage()”, which is simply called from the Instrument-Server and returns the voltage.

In relation with data, Flask gives a package, called, Flask-SQLAlchemy, which allows the use of Object-Relational mapping (ORM) of databases. This is defining a data set as an object (a programming class), linked with a database. Schema for tables, interaction instructions, or other database operations are hidden by the programmer, and a programming object is manipulated. Moreover, instructions are the same for different database types, as SQLite, PostgreSQL, MySQL, Oracle, MS-SQL, Firebird, Sybase and others. That gives freedom to not only change database server, even change database technology, without changing code.

In order to provide beautiful, flexible and configurable front-end, a HTML design, structured with CSS is done for the basic structure, with JavaScript for content reaction. In order to allow in-frame video streaming, and variable exchange with the back-end (with instrument-connector and data storage) without the needed of refresh the site, "Asynchronous JavaScript and XML" (AJAX) is used.

JavaScript and, in particular AJAX, allow to implement a complete program in the user we browser. When this is join with API interaction with the back-end, experience is half front-end half back-end. This allows front-end access data, but only with the filter considered by API, so non-illegal (non-allowed by design) actions can be done with back-end data or over Instrument-control.

User interaction and functions related are implemented in the front-end (in JavaScript) and steps (values, validation, sequences, steps done, etc.) are implemented in back-end. In this way, user

cannot modify variables related with the progress or grade of the experience.



Figure 1: Basic structure of technologies in the Remote Lab system

A lab experience is a set of data in back-end and front-end. Back-end has all connections needed with hardware, steps, goals for evaluation, etc. All of them are functions, which can be called via API, in order front-end can use them. Front-end has the user interaction web site, with all its functionalities, and use back-end API to get the data related with the experience and for interact with the hardware. It is important to understand that some experiences proposed are similar, (e.g. frequency response (components, amplifier, filters, etc.) involve function generators and oscilloscope) so same web interface can be used, only must be changed the component connected. For these situations, part of the interface can be parametrized and loaded used parameters associated with the remote lab experience.

Sometimes, a lab experience only uses part of the instrument connected to the Instrument-Server. This let free a set of instruments, and if there exists a remote lab experience which use only those instruments, there could be interesting give to students the opportunity of take both experiences at the same time. With this option, number of user, which can take remote lab sessions, would be increased. This can be done with an "instrument/hardware requirement" in any remote lab experience. At any lab experience booking, instruments associated to that experience are booked, and rest of experiences are checked. If any of them are possible, their time slots are let free for booking, in other case, time slots overlapped with the one booked, are booked, as "Instrument-Server booked, experience not possible".

Some remote lab experiences would need test-boards, which are connected with the Instrument-connector and are controlled as another instrument, controlling sources, switches, and detecting its presence or not. Experiences associated to test boards are only available when test-boards are connected to the system.

As indicated before, requirements for Instrument-Server is a computation system, which can run Python and can interact with the Instruments present in that Lab Site. Depending on those instruments and its requirements, in relation with the communication, Instrument-Server can be implemented in a Single-Board Computer as Raspberry pi (Rpi), or ever, using alternatives to Flask, in a microcontroller as ESP32 with microPython, but this option is more complex, proposed system recommends hardware which support Flask. For a usual set of instruments, taking in consideration that an Instrument-Server has one client at time (in addition with communication with main server) a Raspberry-pi V3b has enough computation power and interfaces to interact via LAN, and USB Instruments, even with the camera streaming. In relation with computers, there are not needed for Instrument-Connector a high computational power system, with at least 2-4GB RAM (depends on the OS), 1 core with 2 GHz, Windows 10, Linux or OSX works fine.

3.2. Reception-Server

In relation with the single server, which manage all the system, it is named Reception-Server. This server is point where the user interacts, and all its experience is managed inside this system. This server is this link among users, out of the University network and Instrument-Servers, in University Local Network. This server need special network configuration, due to it must be reachable from the internet, so it must have a domain or sub-domain associated, at the same time to it can access any system in the University Network, in order to reach Instrument-Servers. Actually, only are exposed common web ports (80, 443) to the internet.

Reception-Server host the user database and manage the user login system using the capacities of Flask-User. This extension manages user authentication, sing up, user validation via email, and even gives a basic web-interface (customizable) for all steps. This solves programmer all user related tasks. As all user-logging systems, passwords are not stored. It is stored in database the result of a calculation done over the combination of password and username, called hash.

Reception-Server host the Shift Manager. This is a compound of databases entries and interactions. Lab technician must create time slots duration per lab experience, set available experiences, and Instrument-Servers available time, in order to create available time slots. When time slots are created, they can be booked, with the restrictions sets by the fair use rules. Fair use rule ensure that all users can access all experiences. As indicated before, a basic fair use rule is a limit of simultaneous booking (one per user). Depending on the number of users and the number of time slots available per experiences, a maximum number of tries per user per experience can be set. That rules can be one or two, or a complex rule as one until everyone has done the experience, or one in this time range, giving in a time period the chance to everyone, and then, giving the chance to any user to repeat an experience. These rules may seem complex, but they are simple, having in the database a register of all experiences taken, with date, user, and other data of interest. Only reading the database with the proper filters, all conditions can be tested, and every condition check can be implemented with the condition as a zero limit avoid its application.

Reception-Server is a manage server, it does not host any instrument or experience. However, it need to know the Instrument-Servers, which are present in the network. This implies the IP address, API commands and related ports (common for all Instrument-Servers), Instruments and experiences available, time range in which they can be used, etc. Periodically test connection is done in order to detect sudden disconnections. As Reception-Server is designed for manage the whole system, some configuration of Instrument-Servers can be change from Reception-Server, as Instruments available, time range in which system can be used, experiences available, fair use rules, etc.

Reception-Server is the endpoint for the user experience in any moment. User registration, booking time slots, and take tests, actually is done in Reception-Server. However, remote Lab experiences are host in Instrument-Servers, whose are not reachable by users. This is solved with a forward done by Reception-Server. Instrument-Server associated with the lab experience shows its web-interface, and Reception-server create a SSH tunnel, creating a forward of that web-interface to an Iframe (a web-site object) inside the main interface of the Reception-Server. With this, user still in Reception-Server, and can interact with Instrument-Server. SSH port forwarding can be done with openSSH, and Python has a module to manage it, so connections can be created easily, and stopped when needed by the back-end. As no ports want to be open, a redirection to a URL, under the Reception-Server domain is done.

Users must register in the Reception-Server, and include its identifier in the Learning Management Systems (LMS), as Moodle, is it is used in lessons. With this, data can be packed and prepared to be uploaded to the LMS automatically.

When a user books an experience, data is exchanged with the related Instrument-Server, and an instrument compatibility check is done. In that moment, non-compatible experiences are detected, and all time slots overlapped with the booked one, are booked as "Instrument-Server booked, experience not possible", as indicated before. User booking (time slot, experience, and user identifier) is registered in Instrument-Server, in order to prepare the experience, at the booked time.

If a booked time slot is un-booked, a search for "Instrument-Server booked experience not possible" books in overlapped time slots, for the same Instrument-Server. For each one, an instrument compatibility test is done with time slots overlapped with them. Slots booked as "Instrument-Server booked, experience not possible" (it could be possible that next or previous time slot of other experiences are booked, overlaps with tested time slot, and are not compatible), if results non-compatible, time slot still booked as "Instrument-Server booked, experience not possible", if not, time slot is change to free for be booked.

3.3. General Concepts

One of the most important part of the remote lab experience is the camera. This implies a live streaming during the experience. It is important to understand that, depending on the camera quality, and the frame per second (fps) sent, data flow involved could be too high. Therefore, it is important to control these two parameters, and adjust them, according to the experiences. If only slow changes are expected, 5-10 fps could be enough, but for watch an Oscilloscope, at least 20 fps are needed. Quality should be revised for each experience, depending the surface the camera is recording; in order to information streamed can be read.

Even when SQLAlchemy disconnect partially programming of database with ORM, database type used could include more or less options in those ORM. In particular, classical SQL databases as SQLite, MySQL or Oracle, are pure relational databases. This implies a strong data relation and order, and a very fix data structure. All data are organized in tables; each row of each table has the same number of columns (cells). Each cell is one data, of one data type. Each one of them has its particularities. However, there is another database type, PostgreSQL which works as same as they are, and include a special data type JSON and JSONB, which are data structures, in a cell, where any number of data (non-limited to the number of columns), each one labelled with a key. Moreover, JSONB included search functions, as seen in SQL instructions for search data in column. This kind of variable is useful for store interaction logs, tries, goals, steps of experiences, and other data in a flexible way. Taking one table of experiences, instead one table per experience. For that reason, PostgreSQL is the database selected. Actually, this procedure, which is done by the database engine, can be done in a more basic way, with blob (binary large object) or text variables, and load there JSON variables. Then, recover and decrypt the data into a python variable. Anyway, if a usual relational database wants to be used, and only relational data want to be stored, data, which is stored in each step, is fixed and one data model (table) per experience is introduced.

Instrument-Server usually have direct connections with Lab Instruments, as USB, or GPIB. This kind of Instrument-Servers are not hot swap. However, if connections with all Instruments are done via LAN, Instrument-Server could have a redundant copy, and, in case of fail, change the destination for forwarding from the main system to the backup one. Most Usual situation is the first one, so a periodical complete backup of all systems for all Instruments Servers is highly recommended.

In relation with the Reception-Server, this server has not any specific hardware connected, so a redundant server can be connected, with a periodical copy of the database. If the main Reception-Server disconnect from the network, the redundant one takes its place. For Instrument-Servers, this rule could be implemented in Reception-Server, but in this case, we need a previous server, which redirect the data, or implement routing rules.

With all presented, a stable and scalable system is obtained, it is easy to configure (almost all configuration in Reception-Server, even of Instrument-Servers) and programming, with a base of required programming languages. General system structure and data flow is shown in Figure 2.

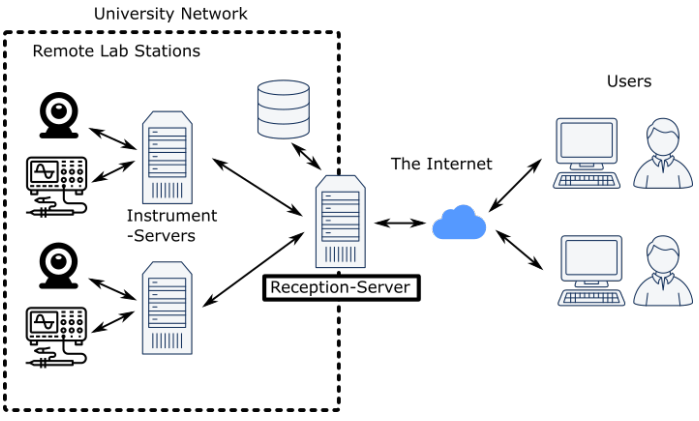


Figure 2: System global structure and data flow.

As indicated before, user traffic ends at Reception-Server and is Reception-Server that reaches Instrument-Servers. Reception-Server only has opened ports 80/443, so it is safe against attacks.

System cannot be properly defined as free of use or non-proprietary, even when all languages and packages in the core are free of use and distribute, due to Instrument-Connector may require privative drivers. Designed system core has not copyright lock or charge, all modules are free of use (some of them “”as is”, without modifications), including author references or copyright.

4. Integration in Engineering Sessions

In engineering teaching, there are too much knowledge to communicate to students, but as technical teaching, many of that knowledge is related with real world aspects, interactions, cause and consequence. In engineering, it explains how things work, including how the physical laws on which the function of those things are based work. Even in most situations complex mathematics support those explanations, real examples can be done in Labs, and students can see a real world example of those theoretical concepts that they are been studding, only supporting by numbers.

This is the reason of Lab sessions in university education, give students real examples of theoretical concepts studied. However, an engineering student must study many concepts, and this implies more theoretical hour than lab session hours. In addition, in a two hours’ theoretical session, many concepts can be explained, and in a two hours’ Lab session, only few concepts can be put in practise, due to they must be examined careful with Lab experiences.

For all explained before, there is not enough time to put in practise all concepts explained in lessons, and that is not the best for students. Searching a learning improvement, more lab time is needed, but there is not available lab time or professors. This can be solved with remote lab experiences. Even when students are not touching hardware, it has been proved in [1] that remote lab is a better solution than simulated lab, when no access to lab is possible.

Moreover, in special situations, as seen with the globally outbreak of COVID-19, even programmed physical lab session may not be possible, or may be reduced. For those situations, remote lab experiences are a great option.

The point is to configure a lab station, for work all time as remote lab station, and even lab stations used for lab sessions, could be set to remote lab stations out of Lab session hours. This requires a configuration, which must be done for the Lab technician. With these options, students can take remote lab experiences any time, with a station exclusively remote, or only out of lab hours, with flexible lab stations. As same as the fair use rules, this depends on the number of students, and instruments and stations available.

As indicated before, remote labs stations are designed for a set of experiences, due to in most areas of engineering, same set of instruments are needed for take all experiences in a subject, or in a part of a subject. E.g. in electronics, if DC bias point is studied, a controllable DC power source, Voltmeters and Amp meters are needed for almost any test. This kind of experience could be behaviour of a capacitor or inductor in DC, in bias point, resistor V-I response, amplification of a transistor without polarization (beta calculation), polarization of a transistor, amplification of an Operational Amplifier in DC, etc. With a Function generator, constant level DC power sources (5V, 12V), and an Oscilloscope, AC response can be studied, with experiences as amplification in AC of an Operational Amplifier, amplification of a polarized transistor, frequency responses of filters, frequency response of components. In addition, mixing, VCO (voltage controller oscillator), limits of Operational Amplifier in relation with feed voltage, effect in transistor amplification of polarization voltage, etc.

As seen, in electronics, there are few equipment, very common in many experiences, and there are others, as pulse generators, clocks, counters, which can give other experiences. The difference among experiences are the components connected to them and the configuration (connections among instruments) needed. A common design can be done for all experiences and Instruments. Creating experience boards, where only instruments must be connected, and a connection with Instrument-Server, in order Instrument-Connector can detects it and make available experiences.

4.1. Experience Board

An example of that common design is done, throw the schematic design of a board for an experience with an Operational Amplifier (LM358), as inverter with variable gain, and switchable to open loop.

First, common elements must be described, starting with the computer communication capacity of the board are shown in Figure 3.

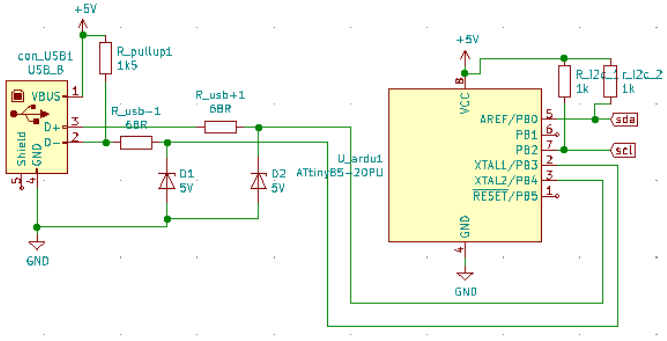


Figure 3: Communication system of test board.

Chip ATtimy85 is a low cost DIP-8 chip, which can be configured with a special bootloader for load Arduino code from USB, without the needed of a USB-RS232 conversion chip, even a serial communication. This chip works with 5V, so even with the USB power can be feed, so only with the elements required for the safe connection for the data transfer (1 pull-up resistor, zener protection diodes, and current limit resistors) chip can be connected with a USB connector.

This chip has few terminals, most interesting ones, I^2C communication. This kind of communication allows connection of some devices to a single port (two wires, with pull-up). Nowadays there are digital elements, as ADC, digital potentiometers, DAC, current meters, and some wind of sensors, with I^2C interface; so many elements can be integrated. Use of external ADC and DAC instead internal of ATtiny85 make easy the design, in order some of them have an analogue level and a digital level, so they can communicate with chip in 5V, and

operate in 20V. This is the case of the digital potentiometer selected for this board, which will be explained later.

In relation with lab instruments connections, as seen in Figure 4, most commons instruments in electronics are DC controlled power source, Function Generator and Oscilloscope. DC source has two outputs, each out with two connections type banana, Function Generator has an output type BNC connector, and Oscilloscope has two inputs type BNC.

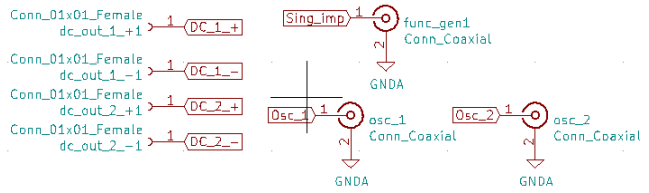


Figure 4: Connections to Instruments of test board.

This connections, join with the ATtiny85 are the base of most test boards (unless other instruments are required, as counters or pulse generators, but those are special experiences, special boards are created, with more connections).

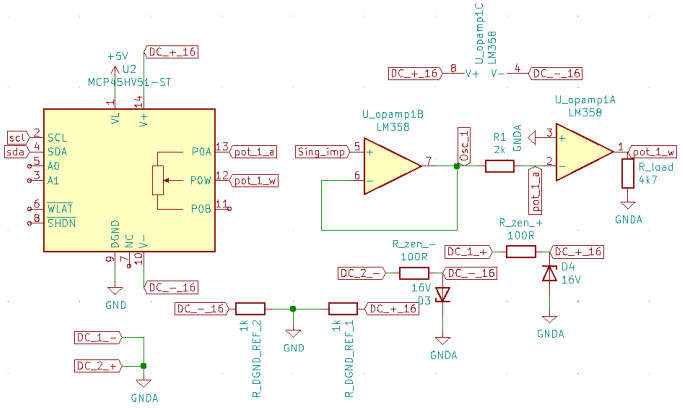


Figure 5: Inverting amplifier with Operational Amplifier experience.

Figure 5 shows the components related to the specific experience, a digital potentiometer, with I2C communication, which can work up to ±18 V, and a LM358 Operational Amplifier, with a maxima power feed of ±16 V. As positive and negative voltage want to be controlled, both DC inputs are used, one for positive and the other for negative voltage. With that objective, negative output of DC1 and positive output of DC2 are connected together to GNDA (analogue ground). Now DC1 controls positive voltage and DC2, negative voltage.

As DC power source can generate more than ±16 V, voltage is limited with zener diodes, to that level, protecting the board. Digital potentiometer requires that digital ground be in the range of analogue voltage, so a resistor divider is done, from limited voltage, and the middle point is connected to the digital ground.

As LM358 has two Operational Amplifiers, one is used as input buffer, and then, as input resistor a constant resistor is used. Digital potentiometer has a limit of ±12.5mA, and input resistance (R1) is the one, which set the current. Fixing the current to a level under the limit, for the extreme input, system is stable. Digital potentiometer is used as feedback resistor (R2), and its impedance could be from 0 to 10kOhms, taking negative amplifications (inverting) lower and upper to one. Digital potentiometer can disconnect terminals, so feedback resistor can be removed, examining Operation amplifier in open loop. For ensure a proper measure a load resistor is included.

As seen this board structure is flexible in its design, and can be change to many electronics experiences, even integrating some in the same board. A price evaluation is done, showing in Table 1 common elements for some experiences, and in Table 2 specific elements for this experience.

Table 1: Price related to common elements

Element	€/unit	Units	€
ATtiny85	1	1	1
USB connector	0.5	1	0.5
Diodes	0.16	2	0.32
Resistors	0.03	5	0.15
Banana connectors	2	4	8
BNC connectors	2	3	6
Total			15.97

Table 2: Price related to specific experience elements

Element	€/unit	Units	€
LM358	0.6	1	0.6
MCP45H51	1.31	1	1.31
Diodes	0.16	2	0.32
Resistors	0.03	5	0.15
Total			2.38

As seen in tables, whole experience, including common and specific elements does not require expensive elements, being all needed components 18,35€. However, as common elements can be used for some different experiences, edge connectors can be included between common and specific part, creating a test system, with a common part, and a changeable experience. In this situation, experience must identify itself for the computer, and it is done including an EEPROM memory, with I2C communication, in the test board. Edge connector has a cost of 3€ (only one is needed, male edge connector is part of the board and EEPROM chip has a cost of 0.1€. With this increase, and minimal design changes, same interface board to PC and instruments (the most expensive part) can be used to some experiences board. At the same time, this change allows to replace interface board or experience board, if any of them fails, and is not needed to replace the whole board.

PCB design for higher frequencies must be done carefully, in order not to introduce coupled effects. Some experiences, as the one proposed in this board, are designed for frequencies up to 1MHz. This make them sensitive to track routes, even to track shapes. This point is the actual develop point, ensure no



Figure 6: Main web interface GPIB SCPI experience

interaction is done, in interface board (for all Oscilloscope or Function Generator frequency range) and in experience board (for all frequency required in the experience).

This is an iteration process, due to each new design must be implemented and tested, and take some time. For that, this line of test board will not be implemented jet, up to all test finished.

4.2. Experience example

Remote lab experience which has been implemented as example is a DC power supply control using Standard Commands for Programmable Instruments (SCPI) (the base of the VISA protocol), connected with the General-Purpose Instrumentation Bus (GPIB). Communication with many lab instruments ends in a VISA commands exchange, throw different busses. This experience gives to students the chance to experience and understand this communication. As indicated before, the camera streaming is the base of the remote lab experience, an in this situation is recording the front panel of the DC power supply. Whit that, students can see in any moment the reaction of the Instrument to any command sent to it (errors, changes in display, changes in operation mode, etc.). For this remote lab experience, simple interaction page has been designed, which in seen in Figure 6.

Power supply connected in this situation is Agilent E3646A DC power supply, and it can be seen in the camera-streaming box, al left. This experience has steps, oriented as questions, which can be answered in any order. Those questions are listed in the combo-box under the camera streaming, and the selected one is the one, which must be answered. Depending on the question, should be answered in the proper answer box (at the right of the question combo-box), or interacting with the proper SCPI command with the instrument.

"Instrument Bus" text box is the point where the user must write the instructions to be sent to the Instrument, when the button "SEND" is clicked. In this field, responses taken from the instrument, when "RECEIVE" button is clicked, are shown too, being the SCPI interaction point. When "SEND" and "RECEIVE" are clicked, in addition to interact with "Instrument bus", Output Status is updated. This value indicated communication status, related to data send process (if information has been delivered, if Instrument can be reached, if there are data to be read, after an instruction with data return, in the instrument and another instruction has been sent), in other words, indicated communication OK, or ERROR (with ERROR specifier).

In case of ERROR for instructions, there are two options, "CLEAR ERROR" erase error vector in the Instrument, clearing error indicator in front panel. Other option is "RESET" completely the Instrument, clearing errors and putting all settings to defaults values. If ERROR is related not to be able to reach instrument, maybe GPIB configuration, which can be configured in "CONFIG" should be revised. There, GPIB address of the instrument, and for the controller should be set. Bus has to be "INITIALIZE", at experience start, or after each configuration change.

There only left "HELP" where a popup with information related with experience interaction and with the experience, procedure and steps are explained. In addition, "PROGRAMMING MANUAL", where a popup is opened which the programming manual of the Instrument, with all instructions needed for the experience.

When a proper answer has been done for a step (with the proper interaction with the Instrument or with the proper answer in the answer box), the background for it turn into green in the combo-box. When all steps are green, experience finished with a pop-up, but Instrument-Server still connected during reserved time-slot, for interaction with the Power Source.

With this, experience is explained. Evaluation is detailed in following section.

5. Learning Results

As part of experience design, goals are defined (pass all steps in experience time, pass steps with less than n tries, take less than a specific time in pass the whole experience, take at least two tries in at least three steps, for ensure student is not copy another student answers, etc.). Those goals can indicate a mark, related to experience evolution, as same as a when students fill a form with

measures and answering questions in Lab sessions. This is an auto-evaluation, during the remote Lab session, which is complemented with evaluation activities.

First, it is vital to know if there has been any problem or difficulty in user interaction with the remote system. For that, a quick survey, related to remote interaction is done, where "Easy to use", "Camera quality", "Response time" and "Proper time slot for experience" are asked, and must be evaluated in the range of: Not work, Very Poor, Poor, Adequate, Good, Very Good. In any moment, if system receive some Poor or Very Poor in an area, a warning is raised to administrator, in order to revise the experience or connection. If "Not work" is marked, an Error is raised to user experience register, in order to mark as not valid this try, and open an additional one. In addition, this must not be an anonymous survey, due to in case of Very Poor or even Poor, difficulties could be get to develop the experience, and that should be considered.

Auto-evaluation, gives information about how students follows experience steps. However, Lab experiences and remote Lab experiences, are oriented in reinforce theoretical concepts. Proper succeed in remote Lab experience is ensure those concepts has been understood. With that objective, a fast knowledge quiz is take by students, after take the remote Lab experience, related with the concepts reinforced in the experienced. This test is used for set the mark of the experience, joint to the auto-evaluation. Moreover, it is an indication for evaluate the experience, if none student passes this quiz, experience should be re-designed, due to it is not reinforcing the concept as desired.

As time during which can be in connection with Instrument-Server is limited by time-slots, survey and quiz could be a problem, due to they take time of experience. For that reason, these exercises are implemented in Instrument-Server, due to they are related to the remote Lab experience, but users perform them in Reception-Server. When experience ends, this is when user pass all steps, or time slot ends, Instrument-Server send to Reception-Server evaluation activities. From this moment, user can perform them during an assigned time (by default 1 day), without the needed of connection with Instrument-Server.

With all that, evaluation is take in two levels, during the experience, automatically, and after the experience, with survey and quiz. At the same time, student evaluation helps for the evaluation of the experience itself, detecting point to change or improve, or in the other hand, validating the functionality of the experience. With all these, remote lab experience is continuously evaluated, by survey, and using student's marks.

Additionally, all Student interactions with remote lab experience, are logged and saved in order to revise auto evaluation protocols, or if needed to manual revision.

6. Conclusions

A remote lab system has been designed, with the capacity of interact with almost any hardware, due to Instrument-Connector can be designed for interact with almost any kind of protocol or driver. It is easy to scale, adding or removing Instrument-Servers, with a simple enter in the Reception-Server, the single administration Server. Even the management of the system, composed with many servers, have been simplified, and almost all of it is done in the Reception-Server.

In relation with remote experiences, as Instrument-Connector can interact with almost any hardware and the combination of HTML+JavaScript can create a huge set of visual experiences in the web browser, almost any experience can be implemented.

Special network requirements (the most complicated part in complex infrastructure network) are only needed in Reception-Server, where external access are needed, with a domain and common web ports opened (80 and 443). Instrument-Servers only need to be in the same network than the Reception-Server.

As educational experience, evaluation system is designed, as a multi-step evaluation. First, goals are set during the experience, and that implies an auto-evaluation, join to a log of the user interaction with the experience. This is used to revise possible errors in steps or auto-evaluations. At experience ends, two fast exercises must be taken. One is a fast survey, about user interaction, in order to detect possible problems in camera quality, latency, time associated to the experience, information done, etc. The other one is a short quiz related to the theoretical concepts related with the experience, in order to confirm that experience has been useful and has reinforce those concepts.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The Spanish Ministry of Economy, Industry and Competitiveness [Grant No. supported this work TEC2016-77632-C3-3-R]. The authors would like to thank the Andalusian Government for funding the Research Unit PAIDI-TIC-168 in Computational Instrumentation and Industrial Electronics (ICEI) and the University of Cádiz.

References

[1] K. Jona, R. Roque, J. Skolnik, D. Uttal, D. Rapp, "Are Remote Labs Worth the Cost? Insights From a Study of Student Perceptions of Remote Labs," International Journal of Online and Biomedical Engineering (IJOE), **7**(2), 48–53, 2011, doi:10.3991/ijoe.v7i2.1394.

[2] I. Evangelista, J.A. Farina, M.I. Pozzo, E. Dobboletta, G.R. Alves, J. García-Zubía, U. Hernández, S.T. Marchisio, S.B. Concari, I. Gustavsson, "Science education at high school: A VISIR remote lab implementation," in Proceedings of 2017 4th Experiment at International Conference: Online Experimentation, exp.at 2017, Institute of Electrical and Electronics Engineers Inc.: 13–17, 2017, doi:10.1109/EXPAT.2017.7984378.

[3] J. Garcia-Zubia, J. Cuadros, V. Serrano, U. Hernandez-Jayo, I. Angulo-Martinez, A. Villar, P. Orduna, G. Alves, "Dashboard for the VISIR remote lab," in Proceedings of the 2019 5th Experiment at International Conference, exp.at 2019, Institute of Electrical and Electronics Engineers Inc.: 42–46, 2019, doi:10.1109/EXPAT.2019.8876527.

[4] S. Rigby, M. Dark, "Designing a flexible, multipurpose remote lab for the IT curriculum," in Proceedings of the 7th ACM SIG-Information Technology Education Conference, SIGITE 2006, ACM Press, New York, New York, USA: 161–164, 2006, doi:10.1145/1168812.1168843.

[5] L. De La Torre, M. Guinaldo, R. Heradio, S. Dormido, "The ball and beam system: A case study of virtual and remote lab enhancement with Moodle," IEEE Transactions on Industrial Informatics, **11**(4), 934–945, 2015, doi:10.1109/TII.2015.2443721.

[6] I. Titov, "Labicom.net - The on-line laboratories platform," in IEEE Global Engineering Education Conference, EDUCON, 1137–1140, 2013, doi:10.1109/EduCon.2013.6530251.

[7] I. Titov, A. Glotov, Y. Andrey, V. Petrov, "Labicom labs: Remote and virtual solid-state laser lab, RF & microwave amplifier remote and virtual lab: Interactive demonstration of Labicom labs in winter 2016," in Proceedings

of 2016 13th International Conference on Remote Engineering and Virtual Instrumentation, REV 2016, Institute of Electrical and Electronics Engineers Inc.: 336–338, 2016, doi:10.1109/REV.2016.7444496.

[8] M.D. Da, "Rasberry Pi Based Remote Lab Implementation," International Journal of Scientific & Engineering Research, **7**(8), 2016.

[9] W. Farag, "An Innovative Remote-Lab Framework for Educational Experimentation An Innovative Remote-Lab Framework for Educational Experimentation," International Journal of Online and Biomedical Engineering (IJOE), **13**(02), 68–86, 2017, doi:10.3991/ijoe.v13i02.6609.

[10] S. Odeh, E. Ketaneh, "A REMOTE ENGINEERING LAB FOR COLLABORATIVE EXPERIMENTATION," International Journal of Online and Biomedical Engineering (IJOE), **9**(3), 10–18, 2013, doi:10.3991/ijoe.v9i3.2500.

[11] I. Angulo, L. Rodriguez-Gil, J. Garcia-Zubia, "Scaling up the Lab: An Adaptable and Scalable Architecture for Embedded Systems Remote Labs," IEEE Access, **6**, 16887–16900, 2018, doi:10.1109/ACCESS.2018.2812925.

[12] 10 top Programming Languages in 2019 for Businesses.

[13] Top 10 Programming Languages of the World – 2019 to begin with… - GeeksforGeeks.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 103-108 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Spatial Multi-Layer Perceptron Model for Predicting Dengue Fever Outbreaks in Surabaya

Siana Halim[1,*], Andreas Handojo[2], Ivan Enrico Widodo[2], Felecia[1], Tanti Octavia[1]

[1]*Petra Christian University, Industrial Engineering Department, Surabaya 60238, Indonesia*

[2]*Petra Christian University, Informatics Engineering Department, Surabaya 60238, Indonesia*

ARTICLE INFO

Article history:
Received: 21 July, 2020
Accepted: 04 August, 2020
Online: 09 September, 2020

Keywords:
Multilayer Perceptron
Spatial
Dengue Fever Outbreak

ABSTRACT

Dengue fever (DF) is a tropical disease spread by mosquitoes of the Aedes type. Therefore, a DF outbreak needs to be predicted to minimize the spread and death caused by it. The spread of dengue fever is a spatial problem. In this paper, we adopted the Multi Linear Perceptron (MLP) to solve the spatial problem, and we called it a spatial multi-layer perceptron model (Spatial MLP). In this proposed model, we consider two types of input neurons in the Spatial MLP, a region and the neighbourhood of that region. The spatial inputs dynamically change to the region. Additionally, the neighbourhood numbers of a region are also varied. So, the spatial inputs are changed in terms of the number of inputs and the neighbourhoods. As a result, the proposed model is outperformed the traditional MLP since it can adapt to the neighbourhoods. We can conclude the spatial MLP model can manage the information and predict the dengue fever outbreak in Surabaya

1. Introduction

Dengue Fever (DF) outbreak happened annually, but every year the number of victims is very high. In the present decade, Ketharpal mentioned that dengue is endemic to 128 countries, mostly developing nations, posing a risk of death to approximately 3.97 billion people annually [1]. Cartographic approaches estimated that 390 million dengue infections annually, out of which 96 million cases evident apparently [2,3]. World health organization (WHO) stated that more than 70% of the population at risk for dengue worldwide live in member states of the WHO South-East Asia Region and Western Pacific Region [4]. WHO categorized the variable endemicity of dengue fever into four categories. Indonesia is included in category A which means the endemic occurs due significant public health problem, a leading cause of hospitalization and death among children, hyperendemicity with all four serotypes circulating in urban areas, and spreading to rural areas [5].

More than thirty-nine thousand (39,876) DF cases and 254 deaths were reported by Indonesian Health Ministry from January to March 2020 [6,7,8]. For significant reduction of dengue mortality, the strategies for the prevention of dengue include prompt diagnosis of fever cases, providing appro-priate clinical management, and controlling vector, and personal protection methods. Therefore, severe cases can be managed with appropriate treatment, and health personnel at all level can be trained. Improved outbreak prediction and detection through coordinated surveillance will be able to reduce DF spread and effected area [9, 10].

Many types of research have been done in predicting the spread and DF affected area. A five years dataset from Sleman, a district in Central Java Indonesia are used for predicting the spread of the DF [11]. Mahdiana's model is based on vector autoregressive spatial autocorrelation (varsa). A four years dataset from Bandung stated that the incidence rate of dengue fever was not related to annual rainfall, population density, larva free index, and prevention Program [12]. The spreading of DF in Surabaya, Indonesia, is modelled using statistical learning [13,14]. Besides of statistical learning approach, many researchers also developed the model in the machine learning approach. Various machine learning algorithms are compared, such as naive Bayes, random forests, minimal sequential optimization [15]. They collected data from the health department, Karuna medical hospital, Kerala, and online sources. The authors stated random forests gives better accuracy for the early detection of dengue disease.

On the other hand, the use of neural networks as an algorithm for predicting disease has been widely used. An artificial neural network is used to predict the DF outbreak in Srilanka [16], using similar approach [17,18] the DF outbreak in Thailand modelled and in the Northwest Coast of Yucatan, Mexico and San Juan,

*Corresponding Author: Siana Halim, Email: halim@petra.ac.id

Puerto Rico, respectively. Most of the artificial neural network that has been used to develop the model is multi-layer perceptron, with the input as the population characteristics in each region and number of DF infected in the previous years for predicting the number of DF infected in the current or next year.

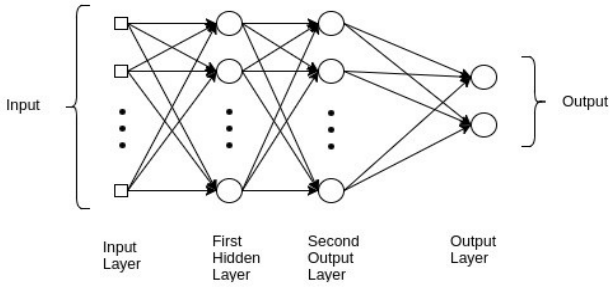


Figure 1. Multi-Layer Perceptron Neural Network

In this research, we proposed a spatial multi-layer perceptron model for predicting the DF outbreak. As a case study, we used DF data in Surabaya. The proposed model adopts the spatial approach in statistical learning as well as the multi-layer perceptron in the machine learning approach. This proposed model tries to accommodate the nature of DF disease spreading. Because DF is a type of disease that spreads through dengue mosquitoes, if DF infects a particular area, the surrounding areas will be vulnerable to the spread of the disease (spatially correlated). Therefore, disease prediction in a particular area is greatly influenced by the DF disease in the surrounding area. So, to predict the possibility of the spread of dengue fever in a particular area, we need to calculate the spread DF data from the surrounding areas. This data will be calculated separately for each region (spatial dependent). This proposed model will implement in Multi-Layer Perceptron Model (MLP) Neural Network. The MLP NN model does not accommodate the spatial dependency in the neural-network construction. This proposed model tries to build a spatial MLP model to accommodate the nature of DF decease spreading.

Additionally, we also present the model for predicting the DF web basely. Since currently, the data for DF victims is manually collected at community health centers, and it will be reported to the regional health department. Based on this DF data, the city and province will take a curative and preventive action to prevent DF outbreak in next year. Urgent measures also being taken by community health centers during outbreaks such as fogging or spreading abate powder in water collecting area. Without a sound information system on DF outbreak location and spreading, the government cannot control and minimize dengue mortality.

2. Research Methods

2.1. Multilayer Perceptron

Multilayer perceptron (MLP), also often called as feedforward neural networks consists of neurons that are ordered into layers (Figure 1). The first layer is called the input layer, and the last one is called as the output layer, the layers between are hidden layers [19].

The main goal of MLP is to approximate some function f; e.g. in a regression, $Y = \theta_0 + \theta_1 X_1 + \cdots + \theta_n X_n$; the function $Y = f(X)$ maps the input vector X into the a value Y. The feedforward network defines a mapping $Y = f(X;\theta)$ and learns the value of the parameters θ that result in the best function approximation.

In the general MLP (Figure 1), we know that each layer can be modelled as a function of

$$Y = f(\theta X + b) \tag{1}$$

where f is the activation function, θ are weights in the layer, X is the input vector, which can also be the output of the previous layer, and b is the bias vector. The hidden layers, which are located in between the input and the output of a neural network, will perform nonlinear transformations of the input in the network. The number of the hidden layers are varied. It depends of the function of the neural network. Similarly, the number of the layers may vary. It depends on their associate weights [20].

The function f is called the transfer function. The transfer function used in this research is ReLu (Rectified Linear Unit) [21]. This function is defined as $y = \max(0, X)$. Visually it can be seen in Figure 2.

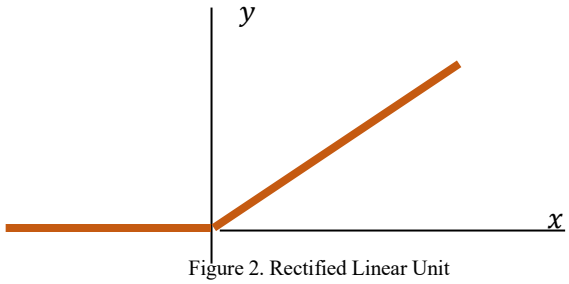


Figure 2. Rectified Linear Unit

2.2 Spatial Multilayer Perceptron

It is well known that the dengue fever happening most in tropical countries and considered as the fastest spreading mosquito-borne disease. It is transmitted by Aedes mosquito which infected with a dengue virus. The spreading of this diseases is spatially correlated [13]. The MLP model does not accommodate the spatial dependent in the neural-network construction. Therefore, in this paper we modified the Multilayer Perceptron Model (MLP), to accommodate the spatial nature of the disease.

In this model, we assumed that the spread of the diseases is in the first-order contiguity level. That is, the number of cases in location s is contagious to its north, east, south, and west neighborhoods. Some additional explanatory variables are also included in the model. They are sex ratio, poverty percentage, population density. In this proposed model, the first layer equation can be written as follows:

$$Y(s) = f\left(\sum_{j=1}^{N}\left(\sum_{i=1}^{M} w_{ij}^{1} X_i(s) + \sum_{k\in(nr,es,st,wt)} w_j^2(s_k) Y(s_k)\right) + \varepsilon\right) \tag{2}$$

where:

$Y(s)$: Number of cases in the location s
w_{ij}^{1} : The jth neuron weight w.r.t explanatory variable i

$X_i(s)$	:	The explanatory variable i in the location s
$w_j^2(s_k)$	:	The j-th neuron weight w.r.t response variable Y in the location s_k
$Y(s_k)$	:	Number of dengue fever cases in the location s_k
s_k	:	The k (north, east, south, west) location of the location s
ε	:	Bias
M	:	Number of explanatory variables ($M = 3$)
N	:	Number of neurons
i, j, k	:	index

In this model the input of the MLP is changed depend on the location s. To give an illustration, let predicts the number of cases in sub-district Balongsari (Figure 3). This region shares borders to sub-district Asemrowo (north), sub-district Tanjungsari (east), sub-district Lontar (south) and sub-district Manukan Kulon (west).

	(A) Asemrowo	
(M) Manukan Kulon	(B) Balongsari	(T) Tanjungsari
	(L) Lontar	

Figure 3: Sub-district Balongsari surrounding area

So, the model can be written as

$$Y(B) = f\left(\sum_{j=1}^{N} w_{1j}^1 X_1(B) + w_{2j}^1 X_2(B) + w_{3j}^1 X_3(B) + w_1^2(A)Y(A) + w_2^2(T)Y(T) + w_3^2(L)Y(L) + w_4^2(M)Y(M)\right)$$

here $X_1(B)$ is the sex ratio in Balongsari, $X_2(B)$ is the poverty percentage in Balongsari, $X_3(B)$ is the population density in Balongsari. $Y(A), Y(T), Y(L)$ and $Y(M)$ are the number of DF case in Asemrowo, Tanjungsari, Lontar and Manukan Kulon respectively. The input neurons of the model adaptively changes with respect to the region s.

2.3 Design Spatial Multi-Layer Perceptron Neural Network

The design uses seven neurons; three neurons represent sex ratio, percentage of poverty and population density of each region under health community center *s* recorded in 2018. The other four neurons are dynamic neurons. They represent the number of cases in the north, east, south, and west. These neurons depend on the location *s* (See Figure 4).

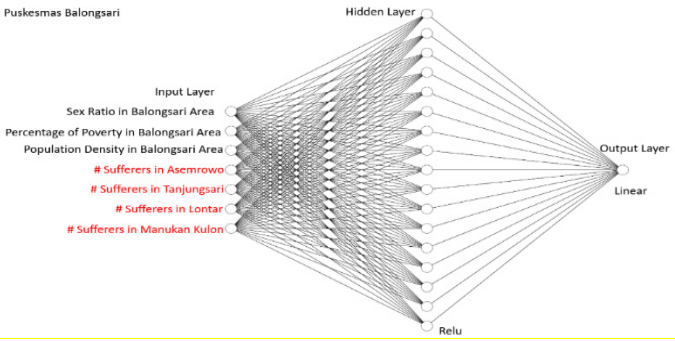


Figure 4: Ilustration for sub-district Balongsari

We used 252 data training (data from 2012-2015) and 126 (data from 2017-2018) data testing. The training process used 3500 epochs (Table 1), and mean squared error is used to measure the loss/error function and we used the stochastic gradient descent as the optimizer.

After some modeling the best design for this case used 1 hidden layer with 17 neurons and 1 output layer (Table 2). The activation function is rectified linear unit (Relu) on the hidden layer and linear on the output layer (Figure 5). This model is implemented as Python functions. It can be used to the other regions as far as the dataset is provided.

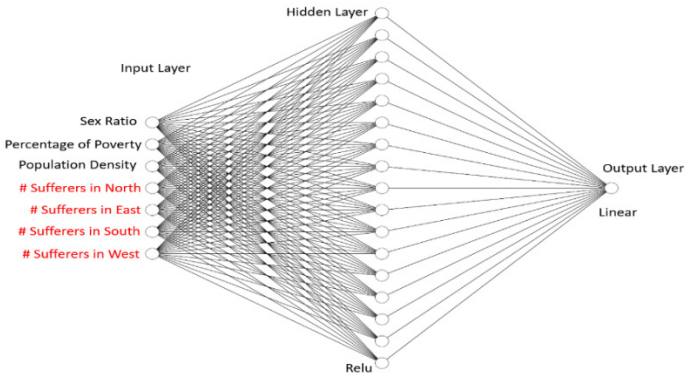


Figure 5: The design of spatial MLP model

Table 1: Setting the Number of Epochs and Neuron

Model	Epochs	Neuron	Loss on Data Training	Loss on Data Validation
1	1500	15	0.0198	0.0403
2		17	0.0187	0.0407
3		20	0.0196	0.039
4		22	0.0182	0.039
5	**3500**	15	0.0189	0.0304
6		**17**	**0.0176**	**0.0294**
7		20	0.0184	0.0383
8		22	0.0171	0.0399
9	4000	15	0.0184	0.0309
10		17	0.0171	0.0304
11		20	0.0181	0.0381
12		22	0.0169	0.0302

Table 2: Setting the Hidden Layer

Model	Layer	#Neuron Hidden Layer 1	# Neuron Hidden Layer 2	Loss on Data Training	Loss on Data Validation
1 hidden	1	17	-	0.0176	0.0294
2 hidden	2	17	7	0.0181	0.0297

3. Result and Discussion

3.1. Data Collection

Data we collected from Surabaya city consist of weather and population characteristic data. Weather data records the number of rainy days in a year, precipitation, maximum and minimum temperature, maximum and minimum humidity. The result shows that Surabaya weather is not significantly different, so it will not be used as the model's explanatory factor. Population characteristic

data will be used in the model, and they are sex ratio, population density, and poverty percentage.

3.2. Data Training and Testing

We use the recorded data from 2012-2015 as the training dataset and the data from 2016 to validate the model. The training dataset consists of 63x4 = 252 data. As usual, we normalized the data set in advanced. The loss of the training data is 0.0176. Figure 5 shows the fitting of the real data to the prediction one. The horizontal axe represents the community health center, the vertical axe represents the number of cases in each community health center, recorded from 2012-2015. Figure 6 shows that the prediction can follow the pattern of the real dataset. During 2012-2013 the number of cases was high, and it started to drop in 2014-2015. The box plot of the data training (Figure 7) shows that there are several outliers in the real dataset and those outliers cannot be captured by the proposed model. The median of the prediction is not significantly different from the real one, but the interquartile range of the prediction is smaller than the real dataset. The two-samples t-test for the training data set is summarized in Table 4 The one-sided p-value is 0.335, we can conclude that there is no mean difference between the real dataset and the predicted one. The mean difference is -0.53 and the 95% confidence interval of the mean difference is (-2.96, 1.91).

Table 3: Surabaya Statistics in 2018

	Min	Mean	Max
Population (thousand)	12541	45802	87561
Area (Km2)	0.915	2.001	14.400
Density (thousand/Km2)	2733	46992	541022
Sex Ratio (Men/Women)	91.5	99.27	110.93
Poverty percentage (%)	4.03	18.02	55.46
Rainy day (days/month)	9.83	13.99	16.00
Precipitation (mm/month)	129.9	164.6	194.9
Max Humidity per month	70	88.72	94.75
Min Humidity per month	46.08	53.14	57.83
Max Temperature	28.21	33.30	34.43
Min Temperature	23.11	26.29	28.73

Table 4: Two-samples t-test for Data Training

	Real	*Predicted*
Mean	18.86508	19.39285714
Variance	262.0455	125.4745304
Observations	252	252
Hypothesized Mean Difference	0	
df	447	
t Stat	-0.425602637	
P(T<=t) one-tail	0.335301126	
t Critical one-tail	1.648269625	
P(T<=t) two-tail	0.670602253	
t Critical two-tail	1.965285234	



Figure 6: The real and prediction line chart of data training from 2012-2015.

We use the recorded data from 2017-2018. There are 126 data. Applying the modelled, the loss value of the testing dataset is 0.052. Some of the prediction are lower/higher than the reality (see Figure 8). Some community health centers reported that there were no dengue fever cases in their area (the number of cases equal to zero), but in their surrounded areas reported highly dengue fever cases. As the result, the real zero number cannot be captured as zero in the model. The model will predict the number of infected in that area as the mean value of the neighborhood. This situation is acceptable, since the predicted number will give early warning to that region to prevent the outbreak in that area.

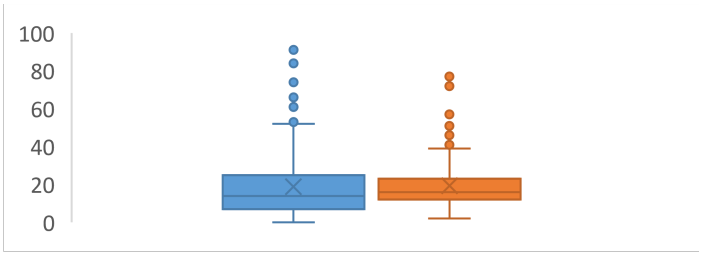


Figure 7: Box plot of the training dataset

3.3. Discussion

In this study, we prosed a spatial-MLP model, which accommodate the spatial property of the dataset. Comparing the other NN models which are used by [16,17,18,19,20] this model uses dynamic variables, which depend on the neighborhood of a region as well as the external variables. The model by [16,17,18,19,20] used only external variables, which do not depend on the neighborhood of a region.

The same dataset has been modelled using the Geostatistical Weighted Regression (GWR) [13]. In [13], the predicted model can follow the pattern of the actual dataset. However, the MSE of the prediction for the years 2017-2018 is 8.59. Compare to this model, the mean square error of the testing data set in the same years is lower, that is 8.07. The t-test shows that there is no significant difference between the mean of the real and the mean of predicted of the testing dataset (Table 5). This model is better than the GWR. However, the computation time of GWR is faster than the spatial-MLP. Since in the spatial-MLP, we have to do the hyperparameter tuning for finding the best model. This model has limitation. It cannot capture the zero in the dataset. The zeros will be predicted as the mean values of the surrounding areas. In this study we prosed a spatial-MLP model, which accommodate the spatial property of the dataset. Comparing the other NN models which are used by [16,17,18] this model uses dynamic variables, which depend on the neighborhood of a region as well as the external variables. The model by [16,17,18] used only external variables, which are not depend on the neighborhood of a region.

The same dataset has been modelled using the Geostatistical Weighted Regression (GWR) [13]. In [13], the predicted model can follow the pattern of the true dataset. However, the MSE of the prediction for the years 2017-2018 is 8.59. Compare to this model, the mean square error of the testing data set in the same

years is lower, that is 8.07. The t-test shows that there is no significant difference between the mean of the real and the mean of predicted of the testing dataset (Table 5). This model is better than the GWR. However, the computation time of GWR is faster than the spatial-MLP. Since in the spatial-MLP, we have to do the hyperparameter tuning for finding the best model. This model has limitation. It cannot capture the zero in the dataset. The zeros will be predicted as the mean values of the surrounding areas.

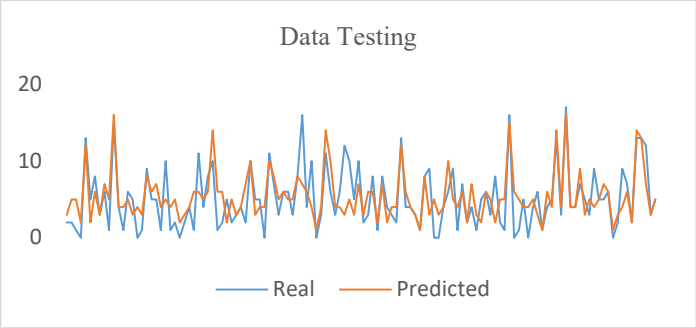


Figure 8: The real and prediction line chart of data training from 2017-2018.

Table 5: Two-samples t-test for Data Testing

	Real	*Predicted*
Mean	5.126984	5.428571
Variance	16.01575	10.10286
Observations	126	126
Hypothesized Mean Difference	0	
df	238	
t Stat	-0.6624	
P(T<=t) one-tail	0.254176	
t Critical one-tail	1.651281	
P(T<=t) two-tail	0.508353	
t Critical two-tail	1.969982	

3.4 Web implementation

This modeled is implemented in a website base to help the "Dinas Kesehatan Surabaya" (The Surabaya Public Health Department) monitoring the dengue fever outbreak. From this website, users can see DF spreading data for each district in Surabaya in the selected year (Figure 9). Data on the number of victims in each sub-district will be displayed in red, yellow and green, with red representing the largest number of victims and green representing the smallest number of victims. Users can specify the upper limit of each color representative. The application will then automatically determine the color gradation based on the input, so that the user can see number of victims in each sub-district that representing in color information. The legend from this gradation color information will display next to the map. User also could choose and see detail information from each sub-district and number of DF victims.

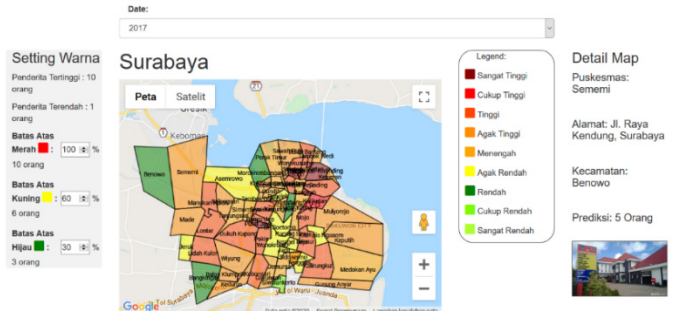


Figure 9: The web design for implemented model

Users can also see details of the number of patients in each sub-district and compare the movement of the number of patients in 3 years presented in tabular form (Figure 10). This data also can be viewed in graphical form (Figure 11).

This visual information will provide more informative information to help The Surabaya Public Health Department monitoring and prevent the dengue fever outbreak for each sub-district.

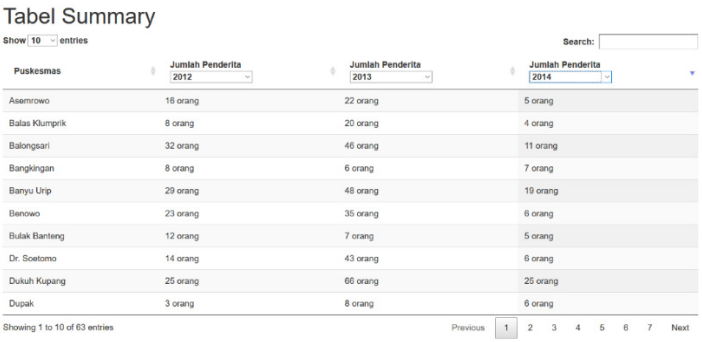


Figure 10: Summarize comparizon spreading DF disease data for each sub-district in Surabaya

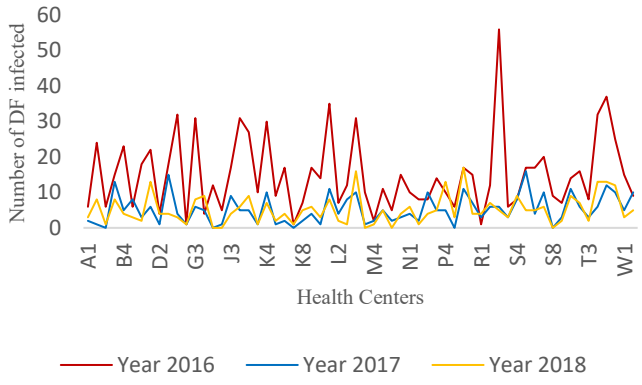


Figure 11: Summarize comparizon spreading DF disease data for each sub-district in Surabaya

4. Conclusion

In this paper we proposed spatial multi-layer perceptron (spatial MLP) model for predicting dengue fever in Surabaya. The model can capture the data pattern. Additionally, the model is implemented in the web-based database. The Surabaya Public

Health Department (Dinas Kesehatan Surabaya) can input the data and predict the outbreak online. However, right now in some regions the predictions are not performed well, especially when that region has zero value. The zeros will be predicted as the mean values it's neighborhood. In the next research, we will expand the model into spatial-temporal multi-layer perceptron (spatial-temporal MLP) model, which can capture data dependencies not only spatially, but also temporally.

Acknowledgement

The authors would like to express their gratitude to the reviewers' feedbacks which certainly improve the clarity of this paper. We also would like to thank to Surabaya Public Health Office (Dinas Kesehatan Surabaya), for the fruitfull discussions. Thanks to Holiyed Hadi for the fruitful discussion on the data base web-design construction. This research is funded by the Ministry of Research, Technology, and Higher Education Republic of Indonesia and the Petra Christian University Institute of Research and Community Outreach.

References

[1] N. Khetarpal N, I. Khanna, "Dengue fever: causes, complications, and vaccine strategies", Journal of Immunology Research, **2016**, 1-14, 2016, doi.org/10.1155/2016/6803098

[2] S. Bhatt, P.W. Gething, O.J. Brady, J.P. Messina, A.W. Farlow, Moyes CL, J.M. Drake, J.S. Brownstein, A.G. Hoen, O. Sankoh, M.F. Myers, D.B. George, T. Jaenisch , G.R.W. Wint, C.P. Simmons, T.W. Scott, J.J. Farrar, S.I. Hay, "The global distribution and burden of dengue", Nature, **496** (7446), 504–507, 2013, doi: 10.1038/nature12060

[3] O.J. Brady, P.W. Gething, S. Bhatt, J.P Messina, J.S. Brownstein, A.G. Hoen, C.L. Moyes, A.W. Farlow, T.W. Scott, S.I., Hay SI, "Refining the global spatial limits of dengue virus transmission by evidence-based consensus", PLoS Neglected Tropical Diseases, **6** (8), 1-15, 2012, doi.org/10.1371/journal.pntd.0001760

[4] WHO, "Dengue guidelines or diagnosis, treatment, prevention and control", World Health Organization, 2009

[5] WHO, "Comprehensive Guidelines for prevention and control of dengue and dengue haemorrhagic fever", World Health Organization- Regional Office for South-East Asia, World Health Organization, 2011.

[6] Tempo, "Dengue fever claims 254 Indonesian lives amid COVID-19 Outbreak", 7 April 2020. Retreived from https://en.tempo.co/read/1328820/dengue-fever-claims-254-indonesian-lives-amid-covid-19-outbreak, accessed on 5 May 2020

[7] Kompas, "Higher than corona, dengue cases reach 17,820 Indonesia". 11 Maret 2020. Retreived from https://nasional.kompas.com/read/2020/03/11/17091361/lebih-tinggi-dari-corona-kasus-dbd-tembus-17820-se-indonesia, accessed on 5 April 2020

[8] Kompas, "Increase rapidly, 2016 cases of dengue fever in East Java, 20 died" 13 Maret 2020. Retrieved from https://surabaya.kompas.com/read/2020/03/13/22200881/bertambah-2016-kasus-dbd-di-jatim-20-meninggal, access on 10 April 2020.

[9] WHO, "Global strategy for dengue prevention and control 2012-2020", World Health Organization, 2012.

[10] W. Wen-Hung, N.U. Aspiro, R.C. Max, A. Wanchai, L. Po-Liang, C. Yen-Hsu, Sheng-Fan, "Dengue hemorrhagic fever a systemic literature review of current perspectives on pathogenesis, prevention and control", Journal of Microbiology, Immunology and Infection, Article in Press. 2020

[11] D. Mahdiana, E. Winarko, A. Ashari, H. Kusnanto, "A model for forecasting the number of cases and distribution pattern of dengue hemorrhagic fever in Indonesia" International Journal of Advanced Computer Science and Applications, **8**(11): 143-150, 2017, DOI:10.14569/IJACSA.2017.081118

[12] K. Anggia, S.Y.I. Sari, H.U. Sumardi, E.P. Setiawati, "Incidence of dengue hemorrhagic fever related to annual rainfall, population density, larval free index and prevention program in Bandung 2008 to 2011", Althea Medical Journal, **2**(2), 262-267, 2015.

[13] S. Halim, T. Octavia, Felecia, A. Handojo, "Dengue fever outbreak prediction in Surabaya using a geographically weighted Regression". Times-Icon Proceeding, 2019, DOI: 10.1109/TIMES-iCON47539.2019.9024438

[14] S. Halim, Felecia, T. Octavia, "Statistical learning for predicting dengue fever rate in Surabaya", Jurnal Teknik Industri, **22**(1), 37-45, 2020, doi.org/10.9744/jti.22.1.37-46

[15] N. Rajathi, S. Kanagaraj, R. Brahmanambika, K. Manjubarkavi, "Early detection of dengue using machine learning algorithms" International Journal of Pure and Applied Mathematics, **118**(18), 3881-3887, 2018.

[16] P.H.M.N. Herath, A.A.I. Perera, H.P. Wijekoon, "Prediction of dengue outbreaks in Srilanka using artificial neural network", International Journal of Computer Applications, **101**, 1-5, 2014, doi:10.1.1.735.9487

[17] B. Jongmuenwai, S. Lowanichchai, S. Jabjone, "Prediction model of dengue hemorrhagic fever outbreak using artificial neural networks in Northeast of Thailand", International Journal of Pure and Applied Mathematics, **118**(8), 3407-3417, 2018.

[18] A.E. Laureano-Rosario, A.P. Duncan, P.A. Mendez-Lazaro, J.E. Garcia-Rejon, S. Gomez-Carro, J. Farfan-Ale, D.A. Savic, F.E. Muller-Karger, "Application of artificial neural networks for dengue fever outbreak predictions in the Northwest Coast of Yucatan, Mexico and San Juan, Puerto Rico", Tropical Medicine Infectious Disease, **3**(5), 1-16, 2018, doi:10.3390/tropicalmed3010005

[19] D. Svozil, V. Kvasnicka, J. Pospichal, "Introduction to multi-layer feed-forward neural networks", Chemometrics and Intelligent Laboratory Systems,**39**,43-62, 1997, doi.org/10.1016/S0169-7439(97)00061-0

[20] R. Collobert, S. Bengio, "Links between perceptrons, MLPs and SVMs", Proceeding of International Conference on Machine Learning (ICML), 2004, doi.org/10.1145/1015330.1015415

[21] L. Yann, Y. Bengio, H. Geoffrey, "Deep learning", Nature, **521**(7553), 436-444, 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 108-112 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Technology Adoption in Education: A Systematic Literature Review

Kayode Emmanuel Oyetade*, Tranos Zuva, Anneke Harmse

Vaal University of Technology, ICT, 1930, South Africa

ARTICLE INFO

Article history:
Received: 11 August, 2020
Accepted: 09 September, 2020
Online: 08 November, 2020

Keywords:
Technology Adoption
Systematic Literature Review
Education
M-Learning
E-Learning
COVID-19

ABSTRACT

Technology is advancing faster today than ever before with evidence of its impact in all facets of our lives. With the spread of the novel COVID-19 pandemic across the world, schools were closed as part of lockdown measures to contain the virus thereby disrupting academic curricula. Academic institutions leveraged ICTs to virtually engage students and teachers. Technology adoption will become a new reality for teaching and learning processes. However, the choice of adopting and not adopting this technology is based on an individual's decision on the benefits or risks in using this technology. The objective of this study is to find topical and relevant studies that have been conducted on technology adoption in education. A systematic literature review was adopted to classify and evaluate articles that fits pre-specified selection criteria using Google Scholar and IEEE databases. 132 papers were found to match the search criteria and filtered to 17 articles using applied exclusion criteria. Current research highlights the trends in technology adoption and provide empirical evidence of applications that have been used to implement technology in educational settings. This research also adds to new literature on COVID-19 in relation to its effect on academic curriculum across the world. Future research will investigate areas that can be expanded and improved on to leverage the benefits of ICTs to limit the adverse impact on the performance of teachers and students in the event of a disruption to academic activities.

1. Introduction

Technology adoption is the choice of an individual to acquire and use an innovation. Technology is advancing faster today than ever before with developments such as cloud computing, robotics, Internet of Things (IoT), Artificial Intelligence (AI), Machine Learning (ML), etc., opening fresh opportunities for value creation and contributing to our life experiences. There is a consensus in both industry and academia on the growing importance of technology as a robust means for survival in today's fast-paced and knowledge-driven world [1]. Evidence abound of its success and impact on teaching and learning via documented research studies [2-4].

Moreover, adopting or rejecting a new technology is based on an individual's perception of the benefits of using the tool to achieve a goal [5, 6]. Although technologies are advancing, the decision not to adopt has been categorised by poor innovation implementation, high cost, high failure rates, maintenance, poor technological planning, inadequate stakeholder participation [6-8]. An understanding of the factors affecting an individual's choice to adopt technology is essential both for stakeholders using these technologies as well as creators and producers of such technologies. Several theories and models exist to allow us understand the acceptance or rejection of a technology by an individual: This includes but not limited to Technology Adoption Model (TAM), Innovation Diffusion Theory (IDT), Unified Theory of Acceptance and Use of Technology (UTAUT), etc. [9].

The present research aims to fulfil the objective: To provide a summary of studies to determine the level of technology adoption in an academic setting. On this basis, a systematic literature review approach was employed to achieve this objective. Google Scholar and IEEE databases were searched using the keyword 'technology adoption and education or learning'. One hundred and thirty-two articles were initially obtained from the search. Inclusion and exclusion criteria were applied which reduced the list to 17 articles. These 17 articles were further analysed to fulfil the purpose of this research.

The next section presents the research method for this study. The result section provides the outcome of the literature review

*Corresponding Author: Kayode Emmanuel Oyetade, VUT, +27846514453 & kayoyetade@gmail.com

and the analysis of the selected literature. This is followed by the conclusion and possible area of future research.

2. Methodology

This section provides an overview of the research method used in the present research. The approach stipulated by [10] to conduct systematic literature was adopted.

2.1. Data sources and selection criteria

This paper reviews the adoption of various technology used in schools for teaching and learning. Therefore, the combination of the following keywords was applied: "Technology Adoption AND education OR learning". The title and abstract were checked for the occurrence of the keywords to ensure results are for articles focused on the study's research objective. English language was used to conduct this search to obtain a global perspective on the topic as well as avoid narrow searches. We limited our search to peer-reviewed texts available in online published journals and conference proceedings from Google Scholar and IEEE electronic databases.

The following inclusion criteria were deemed relevant for the present research:

- Studies on technology adoption application in education
- Studies that engage in any kind of technologies to support learning
- Studies written in English

The following exclusion criteria to remove articles not deemed fit for the current study:

- Studies that are not published as peer-reviewed journal articles and conference papers
- Studies not originally written in English language
- Studies whose full text is not available online and
- Studies discussing technology adoption but not with an educational focus.

3. Result

As presented in Figure 1, the search and refinement process followed six steps.

Step 1: Database search was conducted using the above-mentioned keywords. A total of 132 articles was realised. Step 2 focused on the application of the selection criteria by checking to remove duplicates due to the search using dual databases resulting in 130 articles. Step 3 reviewed the abstract to determine whether the article focused on the study's research objective and relevance to technology adoption. This allowed the study to remove 81 articles and left with 49 articles.

In step 4, full-text review was done on the 49 literatures for eligibility to provide a clear understanding of the research objectives. This allowed the study to exclude 8 articles which do not fit with reason resulting in 41 articles. Further qualitative review for quality assurance to ensure appropriateness and relevance to the research study was applied on these 41 articles. This allowed this study to arrive at a final list of 17 articles to be considered for review.

3.1. Review of the Selected Literature

One of the well-known success areas of technology adoption is the usage of m-learning in education. One of the motivations is the provision and easy access to learning content [11].

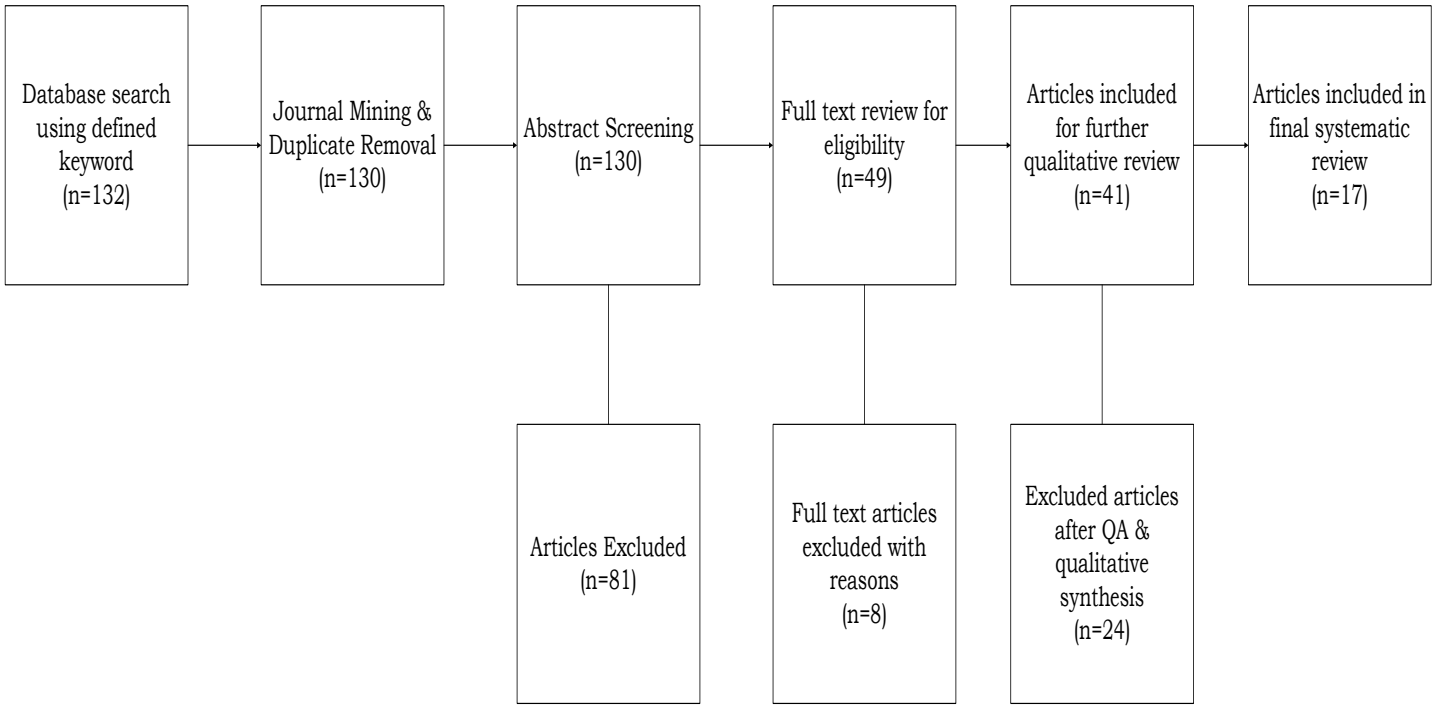


Figure 1: Steps to selection criteria and final outcome

A study by [12] used TAM to analyse factors that influence learner's technology adoption. They found that the factors economics, usefulness, suitability, ease of use, enjoyment, and social influence impact learners' adoption of m-learning to facilitate and promote learning. Likewise, [13] explored the capacity and acceptance of mobile instant messaging, an m-learning platform to increase student performance. They suggest that device portability, communication cost, collective capabilities, and learner control of the device are key factors of flexible learning in digital environments.

Evaluating the factors and effectiveness of e-learning adoption on student learning, [14] adopted the Interactivity Distance Education Web Learning (IDEWL), an enhanced online course creative and conferencing tool that integrates other add-on tools like online survey and e-blackboard for their study. There survey found e-learning effective with the enhancement of the interactivity modules in the students learning process.

On the other hand, [15] reported that the compulsory usage factor is essential for overall adoption of the e-learning system (ELS), either in a compulsory or voluntary setting. This is due to its effectiveness in promoting ELS as a core learning platform for students as well as discourage them from using other means to receive an education.

Similarly, [16] measured the factors that affect learners' adoption of blended ELS (BELS) and the perceived satisfaction of BELS according to their learning styles and gender differences. They report learning style did influence adoption of technology to provide flexible learning and improve understanding. In another related study, [17] examined the relationship between student's learning satisfaction and their use of video in a Moodle environment. He found that the use of video significantly impacted student's learning performance and satisfaction.

In answering student's willingness to learn using e-learning, [18] opines that designers of e-learning sites must ensure ease of navigating the web interface during the learning process. This designers must also ensure the quality of contents uploaded on the websites to enhance students learning while exploring college students learning attitudes with respect to their adoption of English e-tutoring websites. Using TAM as a theoretical framework for their study. There result showed that most users of English e-tutoring sites found online sources more effective and offer greater ease than traditional resources.

Furthermore, [19] conducted a study to examine key factors influencing the adoption of iPads using factors from extended TAM to design and deliver course content by academics. They found that University academics found iPads useful and easy to use with a high level of compatibility with their work.

Using the tablet devices in the classrooms, [20] reports that tablets impacts teaching and learning practices. This tablet devices provides a shift from the traditional way students learn to the media-rich and interactive environments. They therefore recommend stakeholders introduce educational and technical support to assist both educators' and learners' understanding of the benefit of these technologies in education.

Despite the growth and positive outcomes of technology adoption in education, [21] found that males and females do not differ in their belief about the usefulness of e-books. Usage of e-book in South African higher education corresponds to the effort and the external pressure exerted on students to promote the acceptance of e-books.

On mobile technology adoption for learning, [22] explored the correlation between students' ICT skills and factors that impact mobile technology adoption using the TAM framework. They found that basic and advanced ICT skills in mobile technology both played important roles in the intention to adopt m-learning.

Using mobile technology to enhance the convenience of teaching the English Language in Taiwan, [23] extended the factors of TAM to explain factors that affect the adoption of PDA. There result showed that extended TAM was effective in predicting and explaining the acceptance of English mobile learning.

To facilitate teaching and learning processes using the Learning Management Systems (LMSs) by University staffs, [24] used TAM extensions to identify factors that influence LMS usage behaviour and their underlying causal relationships. The analysis of data reveal that the proposed factors were significant predictors of faculty staff attitude towards LMSs. This confirms the validity of the extended TAM in determining users' technology acceptance.

In testing the effectiveness of a Blackboard (Bb) course management system (CMS), [25] used an experimental and control group to conduct his test on English writing instruction in and outside of the Bb environment. The experimental group achieved encouraging results than the control group thereby contributing to learning English writing using Bb. They also identified attitude, perceived usefulness, perceived ease of use, and technical support affect the adoption of CMS.

In [26], authors studied the effect of experience on students' intention to adopt webcast and found that students with previous experience in educational webcasts are most likely to adopt webcast. Ease of use and usefulness were found to be important factors in determining students' BI to use webcasts for learning purposes. They recommended that Webcast developers should pay attention to learners' intrinsic motivations and make them feel comfortable when using the environment for learning.

In conclusion, [27] extended the TAM model to explain the adoption and usage behaviour of social media users in higher education. It was found that subjective norm, perceived usefulness, perceived playfulness (happiness), and perceived ease of use strongly predict students' social media usage behaviour. Their findings suggest that social media use in classrooms should continue in higher institutions. This is because the intention to use translate to the actual use of this ICT tools.

The above review shows technology adoption has the potential to be used as an effective educational platform either directly or as an alternative tool for teaching and learning process.

4. Results of the Review

From the 17 articles examined, this section presents other findings from the above review and discusses the outcomes of the issues related to technology adoption as follows:

4.1. Publication History and Frequency

Figure 2 shows the history and frequency of the identified articles selected for review. We observe the number of articles remains reasonably constant (2006-2012) with a noticeable peak in 2015. The low volume of publications between 2016 and 2019 could be related to our exclusion criteria which may have led to the omission of recent literatures. However, the existence of up-to-date papers lends itself to continuing research efforts and trends of technology adoption in education.

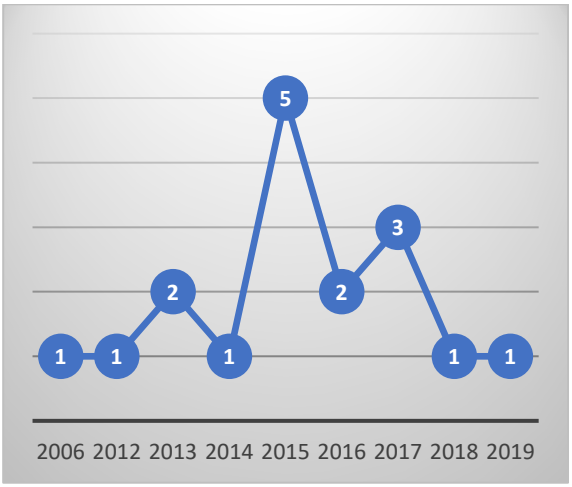


Figure 2: Publication History and Frequency

4.2. Publication Type

From the review of identified literature, Figure 3 presents the result of the articles reviewed by the publication type. It was observed that most of the articles (88%, 15 articles) were obtained from published journals and the remainder (12%, 2 articles) are conference proceedings. The different sources of articles for the current research eliminate possible publication bias.

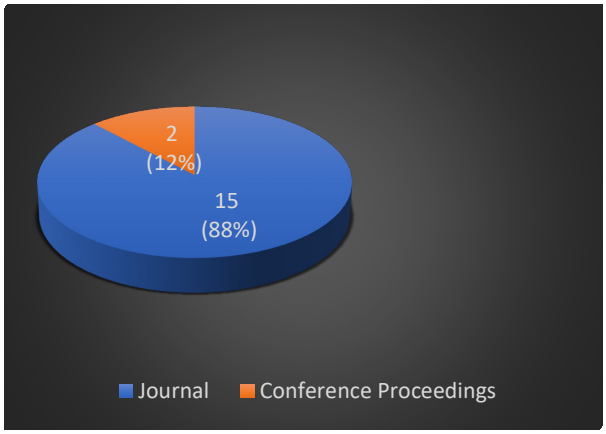


Figure 3: Publication Type

4.3. Word Cloud of technology adoption tools

As evidenced by the articles reviewed, Figure 4 shows the various technology adoption tools adopted for teaching and learning in education. M-learning and e-learning are mostly used technology. Other tools are classified into these two categories.



Figure 4: Word Cloud of Technology Adoption Tools

Even though, this research followed the steps for conducting a systematic review as specified by [10], it is not without limitations. For instance, the review results are limited to the search keywords used, and thus some studies may not have been included. However, we attempted to compensate for this by also relying on related recent literature reviews.

5. Future Research

Future research will investigate areas that can be expanded and improved on to increase the predictive performance of the Technology Adoption in education to support student's academic achievements.

6. Conclusion

The potential of technology has been identified in the above reviewed literature [14-17]. However, these technologies were blended with face to face interactions [25]. With the recent wave of the COVID-19 pandemic and the closure of schools, many schools were forced to seek alternative and virtual means of keeping the curriculum going using appropriate and connective technologies. This implies that technology adoption will gradually become a new normal in education as countries are gradually opening economic and academic activities that were affected by the COVID-19 pandemic.

Although a range of technologies is used in education, e-learning and m-learning are the major tools adopted by many institutions to enhance teaching and learning as identified in the reviewed literature. However, despite the usage of technology adoption as reported in this study, [28-29] reports present findings on the challenges faced with the adoption of technologies as related to COVID-19. This includes affordability and availability of connectivity to the internet to access academic resources while most educators lack the technical support needed to use this ICTs.

The current study implemented a systematic literature review approach to provide a general picture of the usage of technology in education. Google scholar and IEEE databases were searched for articles published in online journals or conference proceedings from 2006 to 2019. The keyword Technology Adoption AND

education OR learning was which resulted in an outcome of 132 articles. This was reduced to 17 articles for in-depth analysis using defined selection criteria. The researchers specifically conducted an analysis of contributions considering the research objectives that covered the year of publication, publication type, technology adoption platforms employed in the research study.

The findings highlight that technology is a relevant tool in schools to deliver the learning needs of 21st century students as learners prepare to return to schools with the lifting of barriers to academic activities. It is believed that the current COVID-19 challenge will change academic curricula and lead to the prevalent usage of technology in the future.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The present research has been prepared by Kayode Emmanuel under the supervision of Dr Anneke Harmse and Prof Tranos Zuva and supported by the Department of ICT at the Vaal University of Technology, Vaal, South Africa.

References

[1] A. Hidalgo, J. Albors, "Innovation management techniques and tools: a review from theory and practice," R&d Management **38**, 113-127, 2008, doi:10.1111/j.1467-9310.2008.00503.x.

[2] A. R. A. Arokiasamy, "A systematic review approach of mobile technology adoption in higher education,". Economics, Management Sustainability **2**, 48-55, 2017, doi:10.14254/jems.2017.2-2.5.

[3] R. Scherer, F. Siddiq, J. Tondeur, "The technology acceptance model (TAM): A meta-analytic structural equation modelling approach to explaining teachers' adoption of digital technology in education", Computers Education **128**, 13-35, 2019, doi:10.1016/j.compedu.2018.09.009.

[4] K. E. Oyetade, A. Harmse, T. Zuva, "Technology adoption factors in education: A review," in 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems, (IEEE, Durban, 2020, doi:10.1109/icABCD49160.2020.9183830.

[5] M. C. M. O. Nemoto, E. P. G. d. Vasconcellos, R. Nelson, "The adoption of new technology: conceptual model and application,". Journal of technology management innovation **5**, 95-107, 2010, doi: 10.4067/S0718-27242010000400008.

[6] C. Plewa, I. Troshani, A. Francis, G. Rampersad, "Technology adoption and performance impact in innovation domains," Industrial Management & Data Systems **112**, 748 – 765, 2012, doi:l0.1108/02635571211232316.

[7] E. D. Adamides, N. Karacapilidis, "Information technology support for the knowledge and social processes of innovation management," Technovation **26**, 50-59, 2006, doi:10.1016/j.technovation.2004.07.019.

[8] L. Dooley, D. O'Sullivan, "Developing a software infrastructure to support systemic innovation through effective management," Technovation **23**, 689-704, 2003, doi:10.1016/S0166-4972(02)00020-2.

[9] K. E. Oyetade, T. Zuva, A. Harmse, "A review of the determinant factors of technology adoption" in Computer Science On-line Conference (Springer) **1226**, 274-286, 2020, doi:10.1007/978-3-030-51974-2_26.

[10] A. Siddaway, "What is a systematic literature review and how do I do one" University of Stirling **1**, 1-13, 2014, Available online: https://pdfs.semanticscholar.org/2214/2c9cb17b4baab118767e497c93806d741461.pdf (accessed on 20 Sep. 20).

[11] S. Jeremy, K. E. Oyetade, K. G. Fomunyam, Operational Use of Mobile Learning: Understanding and Retention of Curriculum Content, Apple Academic Press, 2019, https://www.appleacademicpress.com/title.php?id=9781771887854.

[12] M. Sarrab, I. Al Shibli, N. Badursha, "An empirical study of factors driving the adoption of mobile learning in Omani higher education," International Review of Research in Open Distributed Learning, **17**, 331-349, 2016, doi: 10.19173/irrodl.v17i4.2614.

[13] A. Bere, P. Rambe, "An empirical analysis of the determinants of mobile instant messaging appropriation in university learning," Journal of Computing in Higher Education, **28**, 172-198, 2016. doi: 10.1007/s12528-016-9112-2.

[14] A. T. Baharin, H. Lateh, S. S. Nathan, H. mohd Nawawi, "Evaluating effectiveness of IDEWL using technology acceptance model," in 2015 Procedia-Social Behavioral Sciences (Elsevier), **171**, 897-904, 2015, doi: 10.1016/j.sbspro.2015.01.207.

[15] Y. C. Lee, "An empirical investigation into factors influencing the adoption of an e-learning system," Online information review, 2006, doi: 10.1108/14684520610706406.

[16] A. Al-Azawei, P. Parslow, K. Lundqvist, "Investigating the effect of learning styles in a blended e-learning system: An extension of the technology acceptance model (TAM)," Australasian Journal of Educational Technology **33**, 2017, doi: 10.14742/ajet.2758.

[17] J. T. Nagy, "Evaluation of online video usage and learning satisfaction: An extension of the technology acceptance model," International Review of Research in Open Distributed Learning, **19**, 2018, doi: 10.19173/irrodl.v19i1.2886.

[18] P. J. B. Tan, "An empirical study of how the learning attitudes of college students toward English e-tutoring websites affect site sustainability" Sustainability, **11**, 1748, 2019, doi: 10.3390/su11061748

[19] M. S. Lane, A. Stagg, "University staff adoption of iPads: An empirical study using an extended TAM model," Australasian Journal of Information Systems **18**, 2014, doi: 10.3127/ajis.v18i3.876.

[20] H. Montrieux, R. Vanderlinde, T. Schellens, L. De Marez, "Teaching and learning with mobile technology: A qualitative explorative study about the introduction of tablet devices in secondary education," PloS one, **10**, 2015, doi: 10.1371/journal.pone.0144008.

[21] D. K. Maduku," Factors of E-book use intentions: Perspectives of students in a developing country," Perspectives on Global Development and Technology, **14**, 597-618, doi:10.1163/15691497-12341364.

[22] K. Mac Callum, L. Jeffrey, "The influence of students' ICT skills and their adoption of mobile learning,". Australasian Journal of Educational Technology **29**, 2013, doi: 10.14742/ajet.298.

[23] C.-C. Chang, C.-F. Yan, J.-S. Tseng, "Perceived convenience in an extended technology acceptance model: Mobile technology and English learning for college students," Australasian Journal of Educational Technology **28**, 2012, doi: 10.14742/ajet.818.

[24] N. Fathema, D. Shannon, M. Ross, "Expanding the Technology Acceptance Model (TAM) to examine faculty use of Learning Management Systems (LMSs) in higher education institutions," Journal of Online Learning Teaching, **11**, 2015, https://jolt.merlot.org/Vol11no2/Fathema_0615.pdf.

[25] Y.-R. Tsai, "Applying the Technology Acceptance Model (TAM) to explore the effects of a course management system (CMS)-assisted EFL writing instruction," Calico Journal, **32**, 153-171, 2015, doi:10.1558/calico.v32i1.25961.

[26] M. N. Giannakos, P. Vlamos, "Educational webcasts' acceptance: Empirical examination and the role of experience," British Journal of Educational Technology, **44**, 125-143, 2013, doi:10.1111/j.1467-8535.2011.01279.x.

[27] D. Z. Dumpit, C. J. Fernandez, "Analysis of the use of social media in Higher Education Institutions (HEIs) using the Technology Acceptance Model," International Journal of Educational Technology in Higher Education, **14**, 5, 2017.

[28] T. E. Africa, "The Effect of Covid-19 on Education in Africa and its Implication for the Use of Technology," 1-81, 2020, doi:10.5281/zenodo.4018774.

[29] E. J. Sintema, "Effect of COVID-19 on the Performance of Grade 12 Students: Implications for STEM Education," EURASIA Journal of Mathematics, Science and Technology Education, **16**, 1-6, 2020, doi: 10.29333/ejmste/7893.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 113-121 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Analyzing the Adoption of E-payment Services in Smart Cities using Demographic Analytics: The Case of Dubai

Raed Said[1], Anas Najdawi[*,2], Zakariya Chabani[1]

[1]*Canadian University Dubai, Faculty of Management, Dubai, 117781, UAE*

[2]*Amity University Dubai, Faculty of Management, Dubai, 345019, UAE*

ARTICLE INFO

Article history:
Received: 04 December, 2020
Accepted: 22 February, 2021
Online: 10 March, 2021

Keywords:
e-payment Technologies
e-payment Adoption
Demographic Analytics
Fintech
Digital Transformation

ABSTRACT

This paper is an extension of previous research that has been done on factors affecting digital payment adoption in the UAE. This study focuses on analyzing which relevant demographic factors affect new e-payment technologies, mainly in the smart city Dubai, with more complexities and dynamics of variables that affect users' behavior toward adopting new technologies. The current research included a wider range of demographic factors compared to previous studies. Quantitative methods were conducted using a survey of 270 individuals living and working in Dubai. This study revealed that e-payment adoption is very high, which could be aligned with the national digital transformation strategy of the UAE. The results of the chi-square test for independence indicate that using e-payment technologies is positively associated with the level of education and the level of income. This is confirmed by the fact that the UAE's demographic shape is identified by its high-income groups, positively influencing the residents' e-payment adoption. Surprisingly, the significant results for independence were not found between using e-payments and the gender, marital status, age group, and the current professional position in Dubai. This research's contribution adds to both academia and industry in the digital transformation and technology adoption field. Based on the results, it is recommended for decision-makers to leverage education, digital literacy, and income to accelerate moving toward a cashless economy. However, not having statistically significant differences between the rest demographic variables and adoption will encourage businesses and e-payment service providers to deliver new innovative e-payment models and technologies in a smart city context.

1. Introduction

The fourth industrial revolution (Industry 4.0), with its emerging technologies, brought new innovative technologies to reshape the whole economy and our life [1]. However, it is creating many concerns about security, privacy, and governance. It is crucial to adopt a socio-technical approach to manage digital transformation and conduct effective change management strategies to expand it from an enterprise to a society level, especially when it comes to smart cities as more complex systems in the digital economy.

The United Arab Emirates (UAE) case in incubating digital transformation using emerging technologies is considered one of the most exciting cases to be analyzed and investigated in the Middle East. Especially after the launch of the national digital transformation agenda and several futuristic acceleration initiatives toward the digital economy. All of that had created an atmosphere to accept and adapt faster to technological disruptions, mainly in the fast-growing E-Commerce sector [2]. With growth expected to reach 23% annually between 2018 and 2022 with an estimated 27 billion dollars only for E-Commerce transactions in 2022 based on a study done in collaboration between Dubai Economy Department and Visa [3].

The UAE's strategic goal to shift from an oil-dependent economy to a knowledge-based one, firstly accelerated pace of digital transformation projects across all the government sectors, followed by digital innovation in several business sectors to renovate their models and realign with the digital faster. For instance, Emaar properties initiated noon.com, and Amazon

[*]Corresponding Author: Anas Najdawi, anas.najdawi@gmail.com

acquired souq.com as evidence of the enormous potential of digital business in this region [4]. Additional factors that enabled the UAE to grow at a higher rate regarding e-commerce and digital payment could be linked with the UAE consumer demographic profile. Such profile is featured by a young internet-savvy population, with high social media usage that reaches 99% and 91% for the internet penetration. Along with more than 66% for smartphone penetration and above 7 hours per day for time spent online, which all of this considered higher than the most mature e-commerce markets worldwide [3]. Moreover, the availability of efficient and reliable payment technologies is crucial to digital business growth in any country. The momentum of digital transformation in the payment services in the UAE attracted Big Tech companies such as Apple, Google, Alibaba, and Samsung to provide their digital payment services as well as local companies such as Etisalat, Beam, and local banks to compete in this sector [5].

The fourth industrial revolution we live in with a broad spectrum of emerging technologies is expected to play an indispensable role in reshaping the whole payment industry. For example, artificial intelligence technologies, particularly machine learning algorithms and face recognition, are currently used to conduct payments as a new trend in China [6]. Also, blockchain technology enabled the creation of cryptocurrencies as a way of payment used by several countries [7], [8]. The success of new E-payment services based on such emerging technologies depends on the maturity of such technologies, environment, and end-users acceptance, which are extremely dynamic and changing fastly, hence require a continuous re-evaluation of relevant demographic factors.

This paper is an extension of work initially presented in the IEEE International Conference on Digitization (ICD) [9], which recommended the need to adopt a socio-technical design approach for digital transformation and innovation the payment services in e-commerce which is evolving exponentially in this region. That initial work recommended the need to reinvestigate to which extent demographics variables are relevant and directly affect e-payment adoption in general and especially for the smart cities context in developing countries. The lack of studies in the current literature about emerging new e-payment technologies in developing countries' smart cities is apparent. We believe this area was not addressed in the existing literature according to our best knowledge or only tackled technical perspective without considering the social factors. Thus the current research aims to provide a more in-depth analysis of how demographic variables influence the adoption of existing and new e-payment technologies in Dubai as one of the most promising smart cities in the whole region.

The paper is organized as follows. A literature review on related work about e-payment adoption is introduced in section two, followed by section three for the used research methods. Results and discussions are covered in section four; last, the conclusion and the implications of these study findings to the industry and potential new directions for research are presented in section five.

2. Literature Review

Day by day, the diffusion of new e-payment technologies is expanding globally. The World Bank reported that small retailers in 2015 received 15 trillion dollars via technological payment channels, representing 44% of global transactions [10]. The percentage of people who are using cash decreases substantially; for instance, cash users in North America dropped from 16% to 11% in 2020 compared to 2018. In emerging economies, the rate of cash usage is still comparatively higher [11].

It is noticeable that there is a growth in e-payment adoption in countries with active governmental digital transformation and ongoing smart city programs such as China, India, UAE, and Saudi Arabia. For instance, In 2019, both UAE and France headed other countries in delivering a complete set of Citizen to Government (C2G) e-payment services [12]. That reflected how the UAE government is serious about moving into a cashless economy, as revealed by the Central Bank of UAE in 2018 [13]. However, the same report reflected that the UAE ranked 27 among 72 analyzed countries regarding government e-payment adoption ranking. This indicates that more efforts are needed in this full digital transformation journey.

2.1. Demographic Variables and e-Payment Adoption

Even though many studies about e-payment adoption were found in the literature, few studies associated with the adoption of e-payment in the UAE were found, some analyzed the E-payment adoption principally through SWOT and TOWS matrix [14]. Others used the extended TAM model to analyze the factors affecting the adoption of e-payment systems through a survey of university students in the UAE. In [15], the author investigated the factors affecting the adoption of smart government services in the UAE compared to other Arab countries [16], [17].

Authors studied how people from different socio-economic classes would behave differently toward using e-payment technologies and found that people with higher-income use e-payment technologies more than lower-income people since last have less access to ICTs. [18] Additionally, authots found that the level of education and age affected the behavior of users toward e-payment systems obliquely; because the older people are, the more education they get, and the more income they receive, which is proof that it impacts the use of e-payment. However, the study revealed that there is no relationship between e-payment usage behavior and any commodities and services (i.e., food, bills, transportation, etc.)

Their work also studied the correlation between using e-payment and specific channels such as credit and debit cards for online transactions, E-Money, and Money transfer via internet banking channels. The results showed a positive relationship between Credit and Debit Cards usage on the internet and the behavior toward e-payment systems, And a negative relationship between the behavior toward e-payments and e-money. Regarding the relationship between user behavior and money transfer via internet banking is found to be partially supported.

Other studies studied the usage of digital wallets via analyzing the relationship between awareness about digital wallets and demographic variables such as age, gender, income, etc. According to the study, people in the age group of 36-45 are the most interested in learning and using e-wallets. While the people in the group age 26-36 are more interested in using e-wallets than using credit and/or debit cards due to its continent features. As for

the relationship between the awareness of using e-wallets and gender, males are more interested in using e-wallets than Females; additionally, males prefer using e-wallets over cash use [19].

Other authors surveyed Ghana to investigate the influence of demographics on user's behavior toward using e-payment systems and found no significant relationship between gender and the use of e-payment. However, there is a relationship between age and the behavior to use e-payment systems, more precisely, the satisfaction of using the systems; because young people have more expectations from e-payment services, they are not satisfied easily. The level of education is not significantly related to the behavior toward e-payment systems usage; however, according to the study, users with a higher level of education feel less secure using the e-payment services [20].

A different study Investigated how socio-demographic factors affect usage of e-payment systems in Malaysia. According to the study, older people are less motivated to use e-payments. Married people use e-payment more than unmarried people due to the fact that married people have higher living expenses; additionally, they use e-payment systems more frequently than unmarried people in order to save time. The level of education is a significant factor affecting the use of e-payment services; according to the researchers, people with high education levels tend to use more e-payment services compared to less educated people. Additionally, the income level was found to be an important factor to affect the use of e-payment services; people with lower income levels are unlikely to use e-payment services compared to people with a high level of income [21].

The primary demographic variables, such as gender, age, educational level, and income, were studied to explain their impact on e-payment adoption [22]. The different categories of the population were examined to demonstrate this, involving students, employees, professionals, and business people. The research was designed to estimate the individual opinion related to the e-payment systems and their frequency along with demographic aspects. As a result, the educational process has a direct impact on the adoption of new e-payment systems.

The impact of age and occupation on e-payment systems usage and their influence on anticipated benefits and speed was examined by [23]. The study revealed that the significance of age on e-payment usage was not that considerable compared to the occupation variable. Furthermore, the paper demonstrated the substantial difference in the perception of e-payment technologies among such categories as business, retired employees and employees, students and homemakers; the three last-mentioned categories have a clear awareness about benefits and speed. Besides, these people have different opinions about facilitating conditions.

A research determined that e-payment card acceptance is not significantly affected by such factors as age, gender, education, and marital status of respondents. However, none of the previous studies investigated how demographic variables could impact e-payment adoption in the context of smart cities with a higher level of digital transformation and maturity of e-government services such as UAE [24].

2.2. E-Payment Channels

Previous studies showed that quick, easygoing, and trustable-payment techniques played an essential role in the success and spread of e-commerce worldwide [25].

Examining e-payment adoption requires an understanding of the new emerging technologies evolving fast and disrupting the whole sector, creating both new opportunities and challenges such as face-recognition algorithms, Blockchain, and implantable/ wearable devices that definitely will gain greater acceptance in the future [8], [26].

There are main e-payment channels identified in the previous studies summarized in Table 1 [25], [27], [28] which include but are not limited to the following.

Table 1: The Main e-Payment Channels Identified in Previous Studies

Payment Channels	Description	Examples
Cards payment	This channel includes cards issued by banks such as debit cards, credit cards, internet shopping cards, or issued by shops such as prepaid cards, gift cards, and loyalty cards that can be used in certain shops at the cashier. [29]. Most of these cards are using Radio frequency identification technology (RFID).	Bank cards, e-Dirham card, transportation Nol card for transportation, and emirates ID.
Web Wide Web (WWW) Payment	This method requires the use of the internet and websites to make a payment, mainly through internet banking websites or making online transfers or pay through a third party intermediary [30]	This includes PayPal or Cashu, as well as debit accounts using bank transfers.
Mobile Devices	This channel includes all smart mobiles of devices to conduct a payment mainly through several technologies such as digital wallets and mobile applications, SMS, and NFC. [31] [32]	Examples include mobile payment apps like Samsung Pay, Ali Pay, Apple Pay, bank applications, or payment through telecommunication providers such as Etisalat or DU.
Financial Services Kiosks	This approach includes conducting payment transactions through self-service digital kiosks or stand-alone-machines. [33]	Examples of this approach include both non-financial services such as paying bills or purchase a product or a service, as well as financial ones such as money transfer or payment using ATM.
Biometric payments	These approaches include using authentication based on biometric features of the body that are hard to copy and impersonate. [34]	Using several biometrics for an authentication mechanism, including faceprint, fingerprint, voice, and iris recognition
Cryptocurrency	This includes payment that is based on Blockchain technology as the main component of the IT infrastructure. [8]	Among the most well-known cryptocurrencies can be Bitcoin, Ethereum, etc.
Wearable devices	This approach of payment is still new amongst others and includes wearing devices or objects to substitute in the near future card and mobile payment. [26]	Real examples include payment through digital rings or smartwatches.

All previous studies reflected in general the significance of individuals' demographics on technology acceptance/adoption, such as e-payment technologies [35]. However, up to our best

knowledge, no sufficient studies have been made, including new emergent e-payment technologies with regards to demographic variables in smart cities as a special unite of analysis, where the adoption tends to be more complex, dynamic, and faster than other cities [36]. The primary objective of this study is to fill such gap and investigate the role of demographic factors in adopting several e-payment technologies in the case of Dubai as a case of smart city, having in mind the following main questions:

- To which extent various e-payment technologies are adopted in Dubai nowadays?
- How is the e-payment adoption affected by demographic variables?

3. Methods

3.1. Data Collection

The study was carried out using a self-administered online questionnaire, which consisted of two major sections. The first section comprises ten questions intended to collect demographic data and some information about using different e-payments technologies. The demographic information section included eight questions designed to collect information about the characteristics of the participants. Information was gathered about residents' income, current professional position, age, marital status, gender, and educational level. The second section contains 28 statements meant to examine those factors that influence the perception of using e-payment. The items in this section essentially required the respondents to choose to what extent they agreed with each statement. A Likert scale of one to five (1 = strongly disagree, 2 = disagree, 3 = agree to some extent, 4 = agree, and 5 = strongly agree) was used so that participants could rate their degree of agreement regarding the survey statements.

3.2. Sample and Population

The participants in this study were individuals living and working in Dubai. In this research, a nonprobability sample technique is used for collecting responses. To achieve the objectives of the current research, the researchers of this work used the convenience sampling technique, which is a type of non-probability sampling where population elements are selected for inclusion in the sample based on the ease of access. In some cases, judgment sampling is used where the researchers' judgment is used for selecting participants who are considered as representative of the population. A sample size of 270 eligible respondents was collected to conduct the study from the targeted population of individuals living and working in the Emirate of Dubai.

The pilot study was done on 15 individuals in Dubai before its dissemination to satisfy the face validity. Cronbach alpha, which is a measure of reliability based on the internal consistency of the constructs, was calculated and found to be 0.913. This indicating that the data collection instrument has attained a relatively high level of reliability. Hence, all items are retained.

3.3. Data processing and analysis

The respondents' collected data are recorded and coded in the Statistical Product and Service Solutions (SPSS v.25). For achieving the purpose of the study, a quantitative approach of research is used. Additionally, we used percentages technologies for counting all the yes-category of the items listed in the research tool instrument to measure each factor that influences the perception of using e-payment methods. The researchers used a t-test to perform a comparative analysis of using e-payment technologies between the gender and marital status, where each one of them has two categories. They also used the ANOVA test to compare the different educational level categories, individual income, and age groups. To validate the stated hypotheses, the researchers used the Chi-Square test of independence to test the association between the demographic variables used in this study and the use of e-payment technologies.

3.4. Research Hypotheses

Since the researchers in this paper mainly aimed at investigating the influence of demographic variables such as income, current professional position, age, marital status, gender, and educational level on the adoption of e-payment technologies, this study attempts to test and verify the following six hypotheses.

H_1: There is a relationship between gender and adoption of e-payment

H_2: There is a relationship between profession and adoption of e-payment

H_3: There is a relationship between age and adoption of e-payment

H_4: There is a relationship between education and adoption of e-payment

H_5: There is a relationship between marital status and adoption of e-payment

H_6: There is a relationship between income and adoption of e-payment

4. Results and Findings

4.1. Demographic Characteristics of The Respondents

We conducted an online survey to analyze the adoption of e-payment technologies in Dubai based on different demographic variables. A total of 379 respondents visited the survey's link, and 270 respondents (71.2%) completed the survey.

Table 2 summarizes the demographic characteristics of the respondents to the survey. The table shows that 52.6 % of the participants were males, 10.4 % were born before 1980 (Generation X), 32.6% were born between 1980 and 1994 (Generation Y), and 57% were born in 1995 or above (Generation Z). Twenty percent of the participants were married, and in terms of education, 26.3 % of them had High School or lower, 12.6 Diploma, 41.5% Bachelor's degree, 11.5% Masters and 8.1% were Ph.D. holders. In terms of the monthly income, the study found that 55.2% of the respondents have an income below 10000 Dirhams. 14.3% from 10000 to less than 20000, 18.2% from 20000 to less than 30000, 5.9% from 30000 to less than 40000 and 6.3% having a monthly income of at least 40000 Dirhams

4.2. E-Payments Channels

The enormous development of e-payment technologies in Dubai led to an increase in adopting such technologies by its

residents. Table 3 shows that out of 270 respondents, 84.3% of the individuals in Dubai stated that they are now using e-payment technologies, only 8.6% never used it before, and 7.1% used it in the past they are not using it anymore. This indicated that the majority of individuals in Dubai having experience in e-payment technologies.

Table 2: Demographic Characteristics of the Sample

Criterion	Category	Number	Percent
Gender	Male	142	52.6
	Female	128	47.4
Age Group	Less than 1980	28	10.4
	1980-1994	88	32.6
	1995 and above	154	57
Marital Status	Single	215	79.6
	Married	55	20.4
Education	High school or lower	71	26.3
	College (2 years program)	34	12.6
	Bachelor's degree	112	41.5
	Masters	31	11.5
	Ph.D.	22	8.1
Monthly Income	Below 10000	149	55.2
	10000 up to less than 20000	38	14.1
	20000 up to less than 30000	47	17
	30000 up to less than 40000	18	6.7
	40000 or above	13	4.8
Total		270	100

Table 3: The Frequency Use of e-Payments

Category	*N*	Percent
I am now using e-payment	226	84.3
In the past, I used to deal with e-payment, but am not using it now	19	7.1
I have never used it before	23	8.6
Total	268	100.0

Despite the enormous development of payment technologies for the governmental sectors in Dubai, the individuals') acceptance rates for the different e-payment channels have been found to be satisfactory. According to the survey's results, about 76 percent of individuals gave preference to Contactless Cards. However, 24 percent of them prefer to use cash until now. Furthermore, the acceptance rate of Online Payment(s) was found to be 66 percent, while 46 percent of individuals relied on Mobile Payments. Wearable Payment Devices and Digital Currency are seen to have lesser excitement among individuals, where 5 percent of the respondents believe that they would be using each one of these techniques. The Website Domain Payments and ATM through Different Mobile Applications are also seen to have even lesser excitement among individuals where only one percent of the individuals in Dubai believe that they would be using each one of these techniques. Finally, the rate of acceptance of Artificial Intelligence Payments is only 0.6 percent. Even we expected that Artificial Intelligence Payments would not be used at all at the time of data collection of this research since that technology was not available but in a few smart cities around the world. As per the diversity of people in Dubai, we expect that those who used that technology initially comes from these cities and will make these technologies transfer faster to Dubai based on government support to attract new technologies for their residents.

Out of the 51 nationalities that participated in this study's survey, the acceptance rates of different payment channels for the 11 dominant nationalities living in Dubai are summarized in Table 4. The table revealed that most of the dominant races selected Contactless Cards as their first choice with a percentage exceeding 80% for the Canadians, Indians, Pakistanis, Lebanese and Saudi Arabians, while it was between 68 and 77% for UAE, Egyptian, Iranian, Chinese and Jordanians. The Online Payments are selected as a first choice for the Syrian (85%) and Chinese (75%), and as a second choice for the Canadian (80%) and most of the dominant nationalities (50 to 69%), but not for the Lebanese, who selected Mobile Payments as a second choice. Only the Chinese have chosen Mobile Payments as their first choice, and all the other dominant nationalities set this digital mode of payments as their third choice with percentages of users around 50%.

Table 4: The Distribution of e-Payment Channels Based on the Dominant Nationalities in Dubai

	e-Payments Channels						
	Online Payment	Contactless Cards	Mobile Payment	Payment Kiosks	Wearable Payment Devices	Digital Currency	Artificial Intelligence payment
UAE	60% (53)	77% (53)	55% (53)	2.8% (53)	4% (53)	3.5% (53)	15% (53)
Jordan	64% (25)	68% (25)	56% (25)	25% (25)	0% (25)	4% (25)	8% (25)
Egypt	69% (13)	77% (13)	53% (13)	31% (13)	0% (13)	7% (13)	7% (13)
Syria	85% (13)	61% (13)	46% (13)	54% (13)	7% (13)	15% (13)	0 (13)
Lebanon	36% (11)	82% (11)	46% (11)	18% (11)	0% (11)	0% (11)	9% (11)
Saudi Arabia	66% (6)	100% (6)	17% (6)	16% (6)	0 (6)	0 (6)	0 (6)
Iran	62% (8)	75% (8)	50% (8)	25% (8)	0% (8)	0% (8)	0% (8)
Canada	80% (15)	87% (15)	40% (15)	7% (15)	6% (15)	0 (15)	0 (15)
India	62% (25)	80% (25)	38% (24)	13% (24)	8% (24)	8% (24)	4% (24)
China	75% (4)	75% (4)	100 (3)	75% (4)	50% (4)	50% (4)	25% (4)
Pakistan	50% (12)	91% (12)	42% (12)	25% (12)	8% (12)	0 (12)	0 (12)

4.3. The Use of e-Payment technologies for Regular Services

The participants in the survey were asked to identify the most frequent uses of e-payments for regular services, such as paying utility bills, food, clothing, electronics, phone internet, insurance, travel, education, road toll, and entertainment. The analysis of data revealed that most individuals in Dubai use e-payments for food (77%), clothing (71%), entertainment (67%), phone internet (57%), utility bills (48%), and travel for leisure (41%), whereas the paying for road roll, education, and electronics were used by about 35% of the individuals in Dubai for each one of these regular services.

For investigating the most frequently used e-payment techniques -based on location, the survey respondents were asked which e-payment techniques they use most regularly and from where. As observed in Table 5, more than 47% of Dubai individuals used Online Payments most frequently from home, 45.1% at shopping malls, 26.6% at work, 24.9% at the bank, and 19.2% at school. More than 48% of individuals are using Contactless Cards most frequently at home, 58.1% at shopping malls, and around 25% at each of the work, banks, and schools. The use of Mobile Payment and Payment Kiosks ranked third and fourth, were 33.3% and 17.4%, respectively, of the individuals in Dubai, are using these two e-payment techniques from home. Almost 35% and 20% of the individuals are using them at shopping malls. Mobile Payments are used by about 20% of the individuals in Dubai at either the bank or at work. However, Payment Kiosks for almost half of this number for the same regular services. Less than 5% of individuals are using Wearable Payment Devices, Digital Currency, or Artificial Intelligence Payment at any one of the mentioned regular services

Table 5: Percentages of the Most Frequently Used e-Payment Channels based on the location

	e-Payment Channels						
	Online Payment	Contactless Cards	Mobile Payment	Payment Kiosks	Wearable payment devices	Digital Currency	Artificial Intelligence payment
Home	128/ 47.4	131/48.5	90/33.3	47/17.4	13/4.8	11/4.1	11/4.1
Malls	122/ 45.1	157/58.1	93/34.4	53/19.7	12/4.4	10/3.7	13/4.8
Work	72/26.6	69/25.5	56/20.7	33/12.2	11/ 4.1	5/1.8	5/1.8
Bank	67/24.9	70/25.9	55/20.3	28/10.3	8/2.9	8/2.9	7/2.6
School	52/19.2	66/24.4	31/11.4	22/8.1	6/2.2	5/1.8	4/1.5

4.4. Comparative analysis of using e-payment techniques between groups

This part of the research identified the difference in using e-payment technologies among the different categories of each demographic variable used in this study. The investigation is based on the differences in using e-payments values between the different categories of each variable by analyzing and examining the Mean scores (M) and Standard Deviation (SD) of each category's values supported by performing the two independent samples t-test and ANOVA test.

As predicted, findings of the study show that males are more interested in using e-payment technologies (M = 3.91, SD = 0.81834) than females (M = 3.72, SD = 0.90083). As shown in Table 6, results from an independent samples t-test indicated that the difference between the two genders was significant, t(268) = 1.810, p = 0.071 at a 10% alpha level.

Likewise, the married participants were outperformed (M = 4.0242, SD = 0.67237) the singles (M = 3.7690, SD = 0.89818) in using e-payment technologies. Moreover, the results from an independent samples t test showed that the difference between the two groups of the marital status was significant, t(108) = -2.332, p = 0.021.

Table 6: Group mean scores in relation to the gender and marital status variables

Variable	Group	Mean (M)	Standard Deviation (SD)	t-cal.	p
Gender	Male	3.9108	0.81834	1.810	0.071*
	Female	3.7213	0.90083		
Marital Status	Single	3.7690	0.89818	-2.332	0.021
	Married	4.0242	0.67237		

*The mean difference is significant at 0.10 levels.

Table 7: Descriptive Statistics for the Using of e-payment for the Educational level, income, and the Age variables

	N	Mean	Standard Deviation	Std. Error	95% Confidence Interval for Mean	
					Lower Bound	Upper Bound
High school or lower	71	3.7793	.85994	.10206	3.5758	3.9829
College	34	3.8333	.71657	.12289	3.5833	4.0834
Bachelor's degree	112	3.7143	.91573	.08653	3.5428	3.8857
Masters	31	4.0323	.85355	.15330	3.7192	4.3453
Ph.D	22	4.1818	.71067	.15152	3.8667	4.4969
Below 10000	149	3.7987	.85847	.07033	3.6597	3.9376
10000 up to less than 20000	38	3.8333	.88955	.14430	3.5409	4.1257
20000 up to less than 30000	46	3.8986	.83109	.12254	3.6517	4.1454
30000 up to less than 40000	18	3.9074	.73924	.17424	3.5398	4.2750
40000 or above	13	3.9692	1.14167	.31664	3.0793	4.4591
Less than 1979	28	3.9881	.79857	.15092	3.6784	4.2977
1980-1994	88	3.7311	.99918	.10651	3.5194	3.9428
1995 and above	154	3.8420	.78493	.06325	3.7170	3.9670

Additionally, the ANOVA test's large p-values indicate no significant differences among the different categories of the participants' income, educational levels, and age groups regarding the use of e-payment technologies. Moreover, the descriptive statistics presented in Table 7 show that participants with high education levels tend to use more e-payment technologies compared to less educated participants increased slightly and gradually from 3.77 for the High school or lower educational levels to 4.18 for the Ph.D. holders. Even though the p-value of the ANOVA test was large for the monthly income, the study found that this variable is slightly affecting the use of the e-payment technologies; people with lower income levels are unlikely to use e-payment services compared to people with a high level of income. Similarly, the study of e-payment technologies found that the participants born before 1979 are the most interested in using e-payment technologies, with an average of almost 4. While the

participants born between 1980-1994 are least interested in using e-payments technologies with an average of 3.73, and the young participants born after 1995 are in the middle with an average of 3.84.

5. Testing Hypotheses and Discussion

This study uses a sample of 270 individuals living and working in Dubai on a self-administered online questionnaire to assess how e-payment adoption is affected by demographic variables. The study tested and verified five hypotheses that assumed the adoption of e-payment technologies is influenced by gender, current professional position, age, educational level, marital status, and income level.

5.1. Hypotheses Testing

This section discusses the association between the use of e-payments and demographic variables aiming to investigate the importance of individual's demographics on the adoption of e-payment technologies, Findings of the study show that about 55% of individuals the Ph.D. holders and 52% of the Masters' holders are using Mobile Payments comparing with about 42% for the individuals with the lower education. This yields that the higher educated individuals in Dubai tend to use more Mobile Payments than lower educated individuals. On the other hands, Dubai's individuals with low income (less than 10000 dirhams) and students are using Online Payments significantly with a lower amount (about 60%) comparing with those individuals with high income or from none students current professional position categories (more than 71%) with a maximum percentage of users from the executives' category which reached to more than 82% of them.

Table 8: Chi-Square Test for Independence, H_0: Using e-Payment (Variable 1) and Variable 2 are Independent

Variable 2	*Chi-square*	*Asymp.Sig. (2 sided)*
Gender	3.291	0.193
Education	14.923353	0.060653
Marital Status	3.715201	0.156047
Monthly Income	18.292553	0.019137
Age	5.528197	0.237262
Current Professional Position	8.653823	0.565242

Table 8 shows the chi-square ($\chi 2$) test results for independence between the using e-payment technologies and each of the demographic variables investigated in this research. Statistically significant results at a 5% alpha level were found between the use of e-payments and the monthly income (Asymp.Sig. = 0.019137) and at a 10% alpha level between the use of e-payment technologies and the level of education of the individuals in Dubai (Asymp.Sig. = 0.060653). This indicates that the relatively high income of the individuals and the high levels of education increase the opportunities for using e-payment techniques. Therefore, the use of e-payments is positively associated with the level of education and the level of income; hence, the 4th and the 6th hypotheses are supported.

In contrast, statistically significant results for independence were not found (with a (Asymp.Sig. more than the 5% alpha level) between the using e-payments and the gender, marital status, age group, and the current professional position of the individuals in Dubai. Therefore, the association between the use of e-payment technologies and each one of these demographic variables (as stated in the 1st, 2nd, 3ed, and the 5th hypotheses) is not supported.

Results of the study show that the majority of individuals in Dubai having experience in e-payment technologies. Despite the enormous development of payment technologies for Dubai's governmental sectors, the individuals' acceptance rates for the different e-payment channels be satisfactory. The chi-square ($\chi 2$) test results for independence indicate that the use of e-payments is positively associated with the level of education and the level of income and supports the 4th and the 6th hypotheses (see Table 8). This is confirmed by the fact that the UAE's demographic profile is identified by its high-income levels, which positively influence the consumers' e-payment technologies adoption.

Surprisingly, the significant results for independence were not found between the using e-payments technologies and the gender, marital status, age group, and the current professional position of the individuals in Dubai (see Table 9). Therefore, the association between the use of e-payments and each one of these demographic variables (as stated in H_1, H_2, H_3, and H_5 hypotheses) is not supported.

Table 9: Results of Tested Hypotheses

No.	Hypothesis	Finding
H_1	There is a relationship between gender and the adoption of e-payment	Not supporting
H_2	There is a relationship between profession and adoption of e-payment	Not supporting
H_3	There is a relationship between age and adoption of e-payment	Not supporting
H_4	There is a relationship between education and adoption of e-payment	Supporting
H_5	There is a relationship between marital status and adoption of e-payment	Not supporting
H_6	There is a relationship between income and adoption of e-payment	Supporting

5.2. Discussion

The high growth rate of the digital transformation of payment transactions in the UAE across several sectors, especially in the government, backed with the vision to move for a cashless economy, forced many users to try new e-payments technologies that are not well prepared for it. By the time that created more familiarity with and acceptance to try new innovative e-payment channels [37].

This study revealed a more updated analysis of the demographic variable of this dynamic market on the mode of use and channels. Most of the dominant nationalities selected Contactless Cards as their first choice. Online Payments are selected as a second choice for most of them. Most of the dominant nationalities selected Mobile Payments as a digital mode of payments as their third choice, with percentages of users more than 56%

The present study reveals the fact that most users had experience in e-payment technologies. The research's remarkable

results revealed that about 70% of the individuals use more than one e-payment technology frequently. The most significant two categories of e-payments sectors included the food (77.4%) and clothing (70.7%), while payments for government services were the most versatile where 48.1% of the individuals pay for Utilities (Electricity, Water, Gas), 57.4% pay for Phone/Internet, and 35.6% for transportation and road toll gates.

Regarding the popularity of the payment channels, both Contactless cards and online payments have been dominating by 75.9% and 65.6% of Dubai individuals, respectively. Followed by Mobile payment (including Mobile Bank Apps, Apple Pay, Ali Pay and Samsung Pay) and Payment Kiosks (such as kiosks in shopping malls, airports, Etisalat/DU, DEWA, and Dubai Police) ranked the third and the forth with percentages of users reached to 46% and 25% of Dubai consumers (individuals) respectively.

Other new emerging e-payment channels were lower than expected in terms of diffusion, such the new and innovative payments through Virtual currency (including mobile games currencies, Linden Dollars of Second life, and Facebook coins) and Artificial Intelligence payment (such as Face Recognition payment), as well as wearable payment devices (such as apple watch, rings, etc.) are all considered to be futuristic e-payment techniques used in emerging markets. These new methods are still being utilized rarely in Dubai (about 5% for each method), and the majority of the people who are using these technologies are expats mainly from China. These technologies are expected to be more dominating due to the reliability of technology, convenience, and specialization based on the needs of younger generations (i.e., Generation Z).

The fact that the UAE's demographic profile is identified by its high-income levels influences individuals' e-payment technologies adoption positively. As per the results of this study, the individual's monthly income is mainly the most crucial demographic variable for the adoption decision of e-payments. Additionally, this study found statistically significant results between the adoption of e-payment and the level of education. Therefore, the adoption of e-payments is higher for more educated individuals than those who were less educated. While the study found that the other four investigated demographic variables (age, gender, marital status, and current professional position) had no statistical influence on the adoption of e-payment technologies.

6. Conclusion and Implications

6.1. Opportunities for Theoretical Contribution

The results of the study can be utilized by the industry of the digital transformation and technology adoption/innovation field in relooking or revamping their strategies for digital payment technologies.

E-governments should pay more attention to education and income as results prove that using e-payment technologies is positively associated with the level of education and the level of income. Also, the companies in this field should devise the policies and strategies to attract more people to smart cities like Dubai, regardless of their demographic category, as results prove that the significant results for independence were not found between using e-payments and the demographic variables investigated in this research, i.e., the gender, marital status, age group, and the current professional position of the individuals in Dubai.

Unlike the other studies, the researchers' contribution of this work is creating a proposed theoretical model with an opportunity for gap filling the demographic variables compared with the other studies. This will shed light on the need to design and create compelling digital payment services based on the socio-technical approach rather than the technical one in smart cities regardless of their citizens' demographic categories.

6.2. Implications and Future Work

This section highlights the future research based on the current work results, especially on analyzing the key factors affecting the adoption of new technologies and services provided by new Fintech services, which is evolving fast and making many questions marks on the future of the traditional e-payment service providers in the region. The implications of this study will update change managers and digital transformation practitioners, and consultants to identify relevant correlations among e-payment technologies and demographic variables for this region. This will help all organizations looking to consider digital transformation for payment to benefit from the result to design competitive strategies based on the innovation of e-payment services, which is more customizable, user-centered, Omni-channels, smart, and powered with AI and business analytics.

The current study is not without limitations of selecting one city leading the region in digital transformation, such as Dubai. In the future, it can be interesting to analyze additional cities and do comparative analysis among them. Another issue can be the focus on user's requirements/demographic analysis can be limited base on the socio-technical theory, which can include in higher levels of additional analysis requirements collected from a group of users, community, or the whole society. As we can go one step simultaneously, these limitations give implications to the researchers to move further in the current area of research. (The implications of this research will help both researchers and digital transformation practitioners, mainly in the government and leading businesses, to design and develop more effective digital payment services for people in the smart cities [35], [36] based on the socio-technical approach rather than merely technical one.

There are many opportunities for further research using the current variables of the study and the questionnaire on a broader scope. Further research may include individuals from other emirates in the UAE to explore the inter-relations among the current demographic variables and the factors affecting the adoption of e-payments in the UAE.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] D. Kimani, A. Kweku , A.-B. Rexford , U. Subhan , F.-H. Jane and K. Ja , "Blockchain, business and the fourth industrial revolution: Whence, whither, wherefore and how?," Technological Forecasting and Social Change, **161,** 120-254, 2020. doi:10.1016/j.techfore.2020.120254

[2] A. Najdawi and B. Gaur, "AI 2031': The UAE Initiative to Promote Artificial Intelligence," Case Center, Reference no. 320-0184-1, 2020.

[3] VISA, Dubai Economy, "The United Arab Emirates (UAE) eCommerce Landscape 2020: Accelerated growth during turbulent times," 1 November 2020. [Online]. Available: https://ae.visamiddleeast.com/dam/VCOM/regional/cemea/unitedarabemirates/media-kits/documents/visa_uae_ecommerce_landscape_2020_ppinion_paper_vf.pdf. [Accessed 18 Jan 2021].

[4] H. Laubscher, "Noon.com Uses Partnerships In Its Battle With Souq.com," 6 August 2018. [Online]. Available: https://www.forbes.com/sites/hendriklaubscher/2018/08/06/noon-com-uses-partnerships-in-its-battle-with-souq-com/#7bcf48b3695f.

[5] The National, "UAE e-commerce transactions to reach $16bn in 2019," 24 June 2019. [Online]. Available: https://www.thenational.ae/business/technology/uae-e-commerce-transactions-to-reach-16bn-in-2019-study-says-1.889835. [Accessed 13 Sep 2019].

[6] A. France-Presse, "Smile-to-pay: Chinese shoppers turn to facial payment technology," 4 September 2019. [Online]. Available: https://www.theguardian.com/world/2019/sep/04/smile-to-pay-chinese-shoppers-turn-to-facial-payment-technology.

[7] E. Saiedi, A. Broström and F. Ruiz, "Global drivers of cryptocurrency infrastructure adoption," Small Business Economics, 1-54, 2020. doi:10.1007/s11187-019-00309-8

[8] M. Vejačka, "Basic Aspects of Cryptocurrencies," Journal of Economy, Business and Financing, 75 - 83, 2014.

[9] A. R. Najdawi, Z. Chabani, R. Said and O. Starkova, "Analyzing the Adoption of E-Payment Technologies in UAE Based on Demographic Variables," in 2019 International Conference on Digitization (ICD), Sharjah , 244-248. 2019. DOI: 10.1109/ICD47981.2019.9105908

[10] World Bank, "Innovation in Electronic Payment Adoption: The case of Small Retailers," The World Bank, Washington, 2016.

[11] Worldpay, "Global Payments Report," Worldpay, 2018.

[12] The Economist, "The 2018 Government e-payments adoption ranking," The Economist, 2018.

[13] Noor Bank, "UAE's prepaid and digital payments eco-system," Noor Bank, Abu Dhabi, 2019.

[14] I. Al-Mayahi and S. P. Mansoor, "UAE E-goverment: SWOT analysis and TOWS Matrix," 2012 Tenth International Conference on ICT and Knowledge Engineering, Bangkok, Thailand, 201-204, 2012, doi: 10.1109/ICTKE.2012.6408556.

[15] S. A. Salloum and M. Al-Emran, "Factors affecting the adoption of e-payment systems by university students: extending the TAM with trust," International Journal Electronic Business , **14**(4), 371-390, 2018. DOI:10.1504/IJEB.2018.10019536

[16] S. A. A. Nasser and S. M. Jasimuddin, "Factors that Influence End-Users' Adoption of Smart Government Services in the UAE: A Conceptual Framework," The Electronic Journal of Information Systems Evaluation, **20** (1), 11-23, 2017.

[17] A. Takeoka Chatfield and O. Alhujran , "A Cross-Country Comparative Analysis of E-Government Service Delivery among Arab Countries," Information Technology for Development, **15**(3), 151-170, 2009. doi:10.1002/itdj.20124

[18] W. Chaiyasoonthorn and W. Suksa-ngiam, "The Diffusion and Adoption of Electronic Payment Systems in Bangkok," International Journal of E-Business Research, 102-115, 2019. DOI: 10.4018/IJEBR.2019040106

[19] P. Tiwari, V. Garg and A. Singhal, "A study of Consumer adoption of Digital Wallet special Reference to NCR," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 664-669, 2019, doi: 10.1109/CONFLUENCE.2019.8776939.

[20] W. Yaokumah, P. Kumah and E. S. A. Okai, "Demographic Influences on E-Payment Services," International Journal of E-Business Research, 44-65, 2017.

[21] H. S. Lee , Y. J. Loke and A. K. Tan, "The Demand for E-Payments in Malaysia: An Examination of Usage Intensity," The Journal of Applied Economic Research **7**(4), 371-389, 2013. doi:10.1177/0973801013500167

[22] V. Kumar, "Electronic Payment Systems :," SCMS Journal of Indian Management, 41-51, 2016.

[23] K. Vinitha and Dr.S.Vasantha, "INFLUENCE OF DEMOGRAPHIC VARIABLES," International Journal of Mechanical Engineering and Technology (IJMET), **8**(11), 265-276, 2017.

[24] H. Dehbini, "Factors influencing the adoption of electronic payment cards in urban micro-payments," Research Journal of Finance and Accounting, **6**(1), 39-47, 2015.

[25] Trappey, "MODELING TECHNOLOGY ROADMAPS OF E-COMMERCE PAYMENT SYSTEMS BASED ON PATENT INFORMATICS," International Journal of Electronic Business Management, **14**, 24-34, 2016.

[26] K.-S. Wong and M.-H. Kim, "Towards autonomous payments for internet of things," in International Conference on Internet of Things, Stuttgart, 2016. DOI: 10.29268/iciot.2016.0013

[27] A. &. C. A. Premchand, "Future of Payments–ePayments," International Journal of Emerging Technology and Advanced Engineering, **5**, 110-115, 2015.

[28] E.Turban, "Electronic Commerce Payment Systems," in Electronic Commerce 2018, Springer , 457–499, 2018, DOI:10.1007/978-3-319-58715-8

[30] J.-M. Sahut, "Internet Payment and Banks," International Journal of Business , 362-376, 2011.

[31] N. Kreyer, K. Pousttchi and K. Turowski, "Mobile Payment Procedures:Scope and Characteristics," e-Service Journal, 7-22, 2003. DOI: 10.1353/esj.2004.0005

[32] Z. Chabani, "Mobile money, a way to develop financial services," Revue d'économie et de statistique appliquée, 2016.

[33] J. Rowley and F. Slack, "Kiosks in retailing: the quiet revolution," International Journal of Retail &amp Distribution Management, **31**(6), 329-339. 2003. doi:10.1108/09590550310476060

[34] J. Ogbanufe, "Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment," Decision Support Systems, **106**(1), 1-14, 2018. doi:10.1016/j.dss.2017.11.003

[35] K. Baskaran, "The Impact of Digital Transformation in Singapore e-Tail Market," International Journal of Innovative Technology and Exploring Engineering (IJITEE), **8**(11), 2320-2324, 2019.

[36] A. Mehrotra, "Artificial Intelligence in Financial Services – Need to Blend Automation with Human Touch," 2019 International Conference on Automation, Computational and Technology Management (ICACTM), London, UK, 342-347, , 2019. doi: 10.1109/ICACTM.2019.8776741.

[37] C. T. Nikita and K. Baskaran, "A Study on the Effectiveness of Non-Monetary Retention Strategies in UAE," in 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 122-131 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Model-Driven Digital Twin Framework Development for Sulfur Dioxide Conversion Units Simulation

Amine Mounaam*, Ridouane Oulhiq, Ahmed Souissi, Mohamed Salouhi, Khalid Benjelloun, Ahmed Bichri

Ecole Mohammadia d'Ingénieurs, Mohammed V University of Rabat, Rabat, 10010, Morocco

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 14 February, 2021
Online: 10 March, 2021

Keywords:
Sulfuric acid
Sulfur dioxide conversion
Catalytic conversion
Modeling and simulation
Unisim Design
Digital twin

ABSTRACT

In the phosphate industry, sulfuric acid is a key compound in phosphoric acid and fertilizer production. Industrially, the sulfuric acid H_2SO_4 is made generally in a sequence of three main steps: burning liquid sulfur with air, catalytic oxidation of sulfur dioxide SO_2 to sulfur trioxide SO_3, and formation of H_2SO_4 by the reaction of H_2O with the SO_3. The catalytic conversion of the SO_2 into the SO_3 is considered as the crucial reaction that affects the gas emissions and the performance of the process. In this paper, an industrial SO_2 conversion unit of four catalytic beds reactors with vanadium pentoxide as a catalyst, and three heat exchangers were modeled. The model was based on heat transfer, energy and mass balance equations, and the kinetic reaction of the SO_2 catalytic conversion was proposed and calibrated using the experimental plant data. The simulation of the four catalytic beds was carried out in steady-state and dynamic mode using Unisim Design R451 simulator. The proposed model was tested and validated using the studied plant measurements, and the accuracy of the model has exceeded 97%. A graphical interface of the SO_2 conversion unit was integrated to make it suitable for industrial use and operator training. Finally, a digital twin (DT) of the studied conversion unit was developed based on an architecture integrating the plant, the virtual system, and the communication part in a Distributed Control System (DCS) context. The developed DT in this work makes it possible to simulate in real-time the SO_2 conversion unit, predict the process performance, and optimize the unit efficiency.

1. Introduction

In the era of digital transformation and digital factories, a new concept has been deployed known as digital twin (DT). It is considered an important component in achieving a smart and intelligent manufacturing plant. Digital twins' concept or the virtual representation of a physical product was introduced the first time in [1]. In [2], authors have defined the DT as a technology that refers to the method that can model and describe the performance of a physical entity. They add that DT is the tool that enables the interaction between the real physical system, and its twin in the cyber world. In manufacturing processes, DT can represent a reliable base that provides an accurate prediction of process parameters and performance, which can be used for different monitoring and optimization tasks [3, 4]. The implantation of a DT must include three important parts: the physical part, the digital part, and the physical-digital communication part that ensures the connection between the physical and the virtual product. Thus, physical product data are generated and collected from the physical part of the DT, and feed its virtual model to provide real-time simulation and prediction in real conditions [5].

In recent years, different researchers have focused on DT development and implementation for industrial issues. In [6], authors have worked on the reengineering of aircraft structural life using DT concept. They used a DT that was developed by integrating several models in a virtual software to study the design variation under different operating conditions. The virtual software model includes the most important models in the aircraft structural life field, such as the computational fluid dynamic (CFD) model, the structural dynamic model, and the fatigue cracking model. In [7], authors have investigated the DT-based geometric optimization of the centrifugal impeller (CI) with free form blades for five-axis flank milling. To do so, several experiments and variations were performed to adjust the virtual model of the studied CI to its real physical part. The virtual model in their case was developed based on the geometric modeling, the aerodynamic parameters evaluation, and the machining optimization. In [8], a DT model was proposed for hollow glass production lines. The developed DT makes it possible to simulate the production line behavior with real process data and optimize the design of the

*Corresponding Author: Amine Mounaam, amine.mounaam@um6p.ma

production line. Indeed, DT aims to capitalize on advances in modeling and simulation aspects, since all observations and results that will be generated by the DT are, in fact, predicted by the virtual model of the studied system.

In the literature, several studies have been carried out to develop models that can simulate sulfuric acid processes with high accuracy. The most used process in sulfuric acid manufacturing plants is the contact process [9]. This process is based on three important steps: liquid sulfur burning, sulfur dioxide SO_2 catalytic conversion, and sulfur trioxide SO_3 absorption. The SO_2 catalytic conversion stage is considered as the most critical step in the sulfuric acid contact process. Considering the importance of this step, numerous studies have been performed to describe and optimize the SO_2 conversion reactor based on process modeling and simulation aspects. In [10], authors have presented a dynamic model to simulate the SO_2 catalytic oxidation over the vanadium pentoxide V_2O_5 catalyst. The obtained model was validated using an experimental setup of a fixed bed reactor, and a new process design was proposed to ensure zero-mission. In [11], authors have presented a model to simulate the SO_2 catalytic converter using a pseudo homogeneous perfect plug flow model, and the model has been solved using the COMSOL Multiphysics software. In [12], authors have simulated the SO_2 oxidation reactor by a series of tanks, and simulation results were validated using measurement collected from the pilot and industrial reactors. Another interesting work has been published recently in The Canadian Journal of Chemical Engineering, in which the SO_2 converter model was developed based on mass and energy balance equations, and simulation results were validated using industrial measurement [13].

In addition to the SO_2 catalytic converter, and due to the exothermicity of the SO_2 conversion reactor, heat exchangers (HE) are indispensable in sulfuric acid industries. In SO_2 conversion units, HE devices are used to adjust the temperature of the gas leaving each catalytic bed of the converter, before feeding the next catalytic bed. The cooling step between the SO_2 converter beds is necessary to reach a high conversion rate [14]. Thus, the Shell and Tube Heat Exchanger (STHE) is the most used type of HE in this industrial process. In this context, many researchers have focused on STHE modeling and simulation to study and optimize the heat transfer performance within these devices. For example, in [15], authors have developed a dynamic one-dimensional model of multi-pass STHE, and model equations were solved using the finite volume method. In [16], flow and temperature fields modeling and simulation in a small HE were investigated using the CFD package of ANSYS Fluent 6.3. Also, they tested different turbulence models, knowing: Spalart-Allmaras model, two k-ε standard and realizable models. It was concluded that the k-ε realizable model was the best to simulate the studied HE, assuming a fine mesh and a first order discretization. Using Aspen-HYSYS V7.3 simulator, an optimization work of the air heating unit within the paddy drying process has been performed [17]. Simulation tasks of the studied HE were performed in Aspen HYSYS under many operating conditions, while the design task was carried out in Aspen Exchanger Design and Rating (EDR).

This paper represents an extension of the work that has been presented in the 5th International Conference on Renewable Energies for Developing Countries (REDEC), in which authors have modeled and simulated a STHE with different approaches, using ANSYS-Fluent, COMSOL Multiphysics software, Unisim Design R451 simulator, and Matlab-Simulink [18]. All the presented models were validated using an experimental setup under a wide range of operating conditions. However, the model developed in Unisim Design simulator was selected as the best model to simulate the studied system. Besides, the model was used to simulate an industrial sulfuric acid cooling unit and has shown a good accuracy with the plant measurement.

In the present study, a dynamic model of an industrial sulfuric acid conversion unit was proposed and simulated using Unisim Design simulator. The simulated unit comprises a SO_2 conversion reactor with four catalytic beds and three HE. The kinetic parameters of the SO_2 catalytic conversion reaction were determined using experimental data. The simulation was performed in steady-state and dynamic mode, and simulation results were validated using experimental measurement from the studied SO_2 conversion unit. Additionally, a DT architecture of the studied system was proposed. The physical system of the DT was represented by the industrial plant unit, while the virtual part was represented by the developed SO_2 conversion unit model. The communication part of the DT was developed based on the Communication Object Model (COM), to ensure the connection between the real unit and its virtual representation in Unisim Design simulator. Also, a graphical interface was developed to make the developed DT useful for industrial use.

2. Sulfur dioxide conversion unit description

In the sulfuric acid double contact process, the sulfuric acid is produced in three principal steps: liquid sulfur combustion, SO_2 conversion, and SO_3 absorption.

Firstly, the wet air is dried in a drying tower using the circulating sulfuric acid, to absorb moisture contained in the air. Secondly, the liquid sulfur is burned in a sulfur burner with the dry air, to produce the SO_2 necessary for conversion step. The gas mixture feeds the conversion unit in which the SO_2 is converted into SO_3, using the vanadium pentoxide catalytic V_2O_5. Finally, the SO_3 produced by the SO_2 catalytic conversion is absorbed using the diluted circulating sulfuric acid. The three exothermic reactions governing the sulfuric acid double contact process are given as follows:

$$S + O_2 \rightarrow SO_2 \quad (1)$$

$$S + \frac{1}{2}O_2 \rightleftharpoons SO_3 \quad (2)$$

$$SO_3 + H_2O \rightarrow H_2SO_4 \quad (3)$$

The sulfur dioxide conversion step is considered as the main stage of each sulfuric acid manufacturing plant, and its role is to ensure the conversion of the SO_2 generated by the liquid sulfur combustion into SO_3. Each catalytic conversion unit comprises generally a reactor of several catalytic beds and heat exchangers that are used to regulate the inlet temperature of each catalytic bed. In this study, the SO_2 catalytic converter is a four catalytic beds reactor, with three heat exchangers as shown in Figure 1.

The typical configuration of the SO_2 converter consists of a vertical cylindrical adiabatic reactor. The dimensions of the equipment, as well as the height of each catalytic bed, are variable according to the conversion unit operating parameters. The packed catalytic beds of the converter are the main components SO_2 conversion reactor. Each one of the packed beds is supported by perforated metal grids that help in retaining the catalyst and allowing the gas to flow through. The gases flow from the top to the bottom of the beds. Ceramic rollers are placed on the catalyst to avoid any movement that may be caused by the gases flow, and to facilitate the exit of the gases after passing through the catalyst. The vanadium pentoxide V_2O_5 catalyst used in the SO_2 catalytic conversion reactor is highly porous and presents a specific surface in which the active phase is deposited. It is considered as the typical and highly catalyst used in the SO_2/SO_3 conversion step, with an operating temperature between 370 °C and 630 °C [19]. The conversion reaction in each catalytic bed is identical, but the amount of the catalyst required for the reaction changes according to the SO_3/SO_2 ratio in the feed. For example, in the first three catalytic beds, the amount of catalyst increases as the SO_3/SO_2 ratio increases.

The gas mixture feeds the first catalytic bed at a temperature of 400~440 °C. Since the catalytic conversion of SO_2 is an exothermic reaction, the temperature of the gas mixture increases as the amount of the generated SO_3 increases until reaching the equilibrium point. At the equilibrium point, the forward and the backward reaction rate are equal. To increase the forward reaction rate, the gas leaving the first catalytic bed is cooled to a lower temperature value using a superheater heat exchanger. After passing through the second catalytic bed, the mixture gas is cooled again before feeding the third catalytic bed. Cooling the gas before feeding each catalytic bed is a strategy that is used to further increase the SO_2 conversion rate. The gas leaving the third catalytic bed passes through an intermediate absorption, in which the SO_3 generated by the SO_2 catalytic conversion is absorbed by the circulating sulfuric acid. This strategy is used to reduces the amount of the reaction products, which also increases the forward rate of the conversion reaction.

3. Dynamic model of the conversion unit

In this study, the Unisim Design R451 simulator was used to model and simulate the studied SO_2 conversion unit. To develop the virtual model of the studied flowsheet under the Unisim Design simulator, it is necessary to go through two environments. The first environment that must be configured is the basis environment of the simulator, in which the configuration starts by specifying the chemical components involved in the studied process. For the components that are not presented in the simulator components library, they can be defined as new hypo-component using their specific chemical properties, such as the molar weight and the critical temperature and pressure value. For the SO_2 conversion unit, all components required to perform this simulation are available in the simulator components library. Another specification that must be defined in the basis environment is the fluid-package, which represents the equation of state that will be used to calculate and predict the fluids' properties. Thus, the Peng-Robinson fluid-package was selected for the gaseous phase, and the Non-Random Two-Liquid model (NRTL) was selected for the liquid phase. Table (1) summarizes the chemical components and the fluid-packages used to perform this simulation.

Table 1: Unisim Design components list for the SO_2 conversion unit process simulation.

Component name	Component formula	Fluid-packages
Oxygen	O_2	Peng-Robinson
Nitrogen	N_2	
Sulfur dioxide	SO_2	
Sulfur trioxide	SO_3	
Water	H_2O	NRTL
Sulfuric acid	H_2SO_4	

Finally, the chemical reactions involved in the process must be defined in the simulator using the stoichiometric coefficients of the

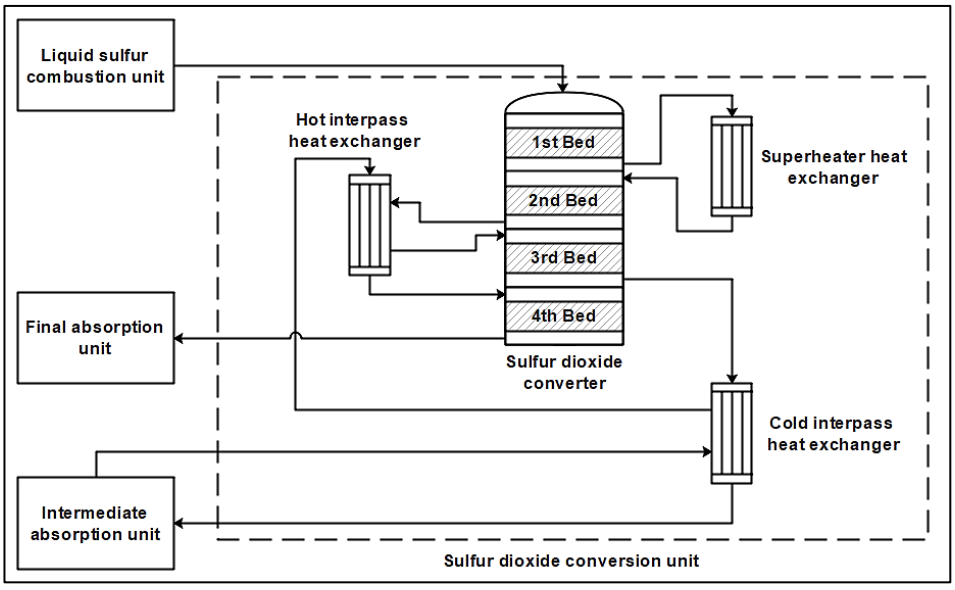


Figure 1: Sulfur dioxide conversion unit

reactants and the products. Note that the type of the defined chemical reaction must be also specified. In the Unisim Design simulator, five types of chemical reactions are provided: conversion reaction, equilibrium reaction, kinetic reaction, heterogeneous reaction, and simple rate reaction. For the SO_2 catalytic conversion reaction, the heterogenous reaction type was selected. Thus, the reaction rate of the SO_2 catalytic conversion is defined in the simulator as follows [14]:

$$r = \frac{K_1.P_{O_2}.P_{SO_2}.\left(1-\frac{P_{SO_3}}{K_2.P_{O_2}^{\frac{1}{2}}.P_{SO_2}}\right)}{\left(1+K_3.P_{SO_2}+K_4.P_{SO_3}\right)^2} \quad (4)$$

where:

- r : kinetic conversion rate (kmole/m^3.s);
- P_i : partial pressure of the component i (atm);

Constants K_1, K_2, K_3 and K_4 have been determined experimentally using industrial data to fit with the plant measurement, and are given as follows:

$$K_1 = \exp\left(15.31 - \frac{45501}{R.T}\right) \quad (5)$$

$$K_2 = \exp\left(-10.68 + \frac{93943}{R.T}\right) \quad (6)$$

$$K_3 = \exp\left(-9.95 + \frac{71655}{R.T}\right) \quad (7)$$

$$K_4 = \exp\left(-71.74 + \frac{437270}{R.T}\right) \quad (8)$$

Secondly, the simulation environment of the simulator must be configurated by developing the flowsheet of the studied process, using material and energy streams, and equipment models. Table (2) shows the equipment models that have been used in the SO_2 conversion unit simulation:

Table 2: Unisim Design equipment models for the studied process simulation

Equipment model	Description
Plug flow reactors	Catalytic conversion
Shell and tube exchangers	Heat transfer

3.1. Catalytic converter model

The SO_2 converter was modeled with a series of reactors, in which each catalytic bed was modeled with a plug flow reactor (PFR). The PFR consists of a tubular reactor with a cylindrical pipe form, in which the transport mechanism is supposed ideal plug flow type. The following assumptions were considered in the model development: (1) the internal and external transfer resistances to mass and heat transfer are neglected, (2) the gas flow in the reactor is one-dimensional in the axial direction z, (3) the gas is radially isotropic, (4) the axial mixing in each PFR is supposed negligible. Under these assumptions, the proposed dynamic model of the SO_2 conversion reactor is given by the following mass and energy conservation equations:

Component i mass balance in the catalytic bed k:

$$\frac{\partial C_{i,gas}^k}{\partial t} = -\frac{4.\dot{Q}_{gas}^{k,in}}{\pi.{D_r}^2}.\frac{\partial C_{i,gas}^k}{\partial z} + \nu_i.\rho_{bulk}^k.r \quad (9)$$

Energy balance in the catalytic bed k:

$$\left[\varepsilon^k.\rho_{gas}.C_{p,gas} + \left(1-\varepsilon^k\right).\rho_{cat}^k.C_{p,cat}\right].\frac{\partial T_{gas}^k}{\partial t} = -\frac{4.\dot{Q}_{gas}^{k,in}}{\pi.{D_r}^2}.\rho_{gas}.C_{p,gas}.\frac{\partial T_{gas}^k}{\partial z} + (-\Delta_r H).\nu_i.\rho_{bulk}^k.r \quad (10)$$

Boundary and initial conditions:

$$z = 0 \ , \ C_{i,gas}^k(t) = C_{i,gas}^{k,in}(t) \ , \ T_{gas}^k(t) = T_{gas}^{k,in}(t) \quad (11)$$

$$z = H_r \ , \ \frac{\partial C_{i,gas}(t)}{\partial z} = 0 \ , \ \frac{\partial T_{gas}(t)}{\partial z} = 0 \quad (12)$$

$$t = 0 \ , \ z > 0 \ \ C_{i,gas}^k(z) = 0 \ , \ T_{gas}^k(z) = T_{bed}^k \quad (13)$$

where:

- $C_{i,G}^k$: molar concentration of the component i in the gas (mol/m3);
- ν_i : stoichiometric coefficient of the component i in the conversion reaction;
- ρ_{bulk}^k : bulk density (kg/m3);
- r : reaction rate (kmole/kg cat.h);
- $C_{p,G}$: heat capacity of the gas (J/kg.K);
- $C_{p,cat}$: heat capacity of the catalyst (J/kg.K);
- ε^k : void fraction of the bed k;
- T_G^k : temperature of the gas in the bed k (K);
- ρ_G : density of the gas (kg/m3);
- $\Delta_r H$: conversion reaction heat (J/mole).

In equation (9), the term on the left-hand side represents the component i mass accumulation in the gas. The first term on the right-hand side represents the component i mass contribution, and the second term represents the mass consumption of the components i. In equation (10), the term on the left-hand side represents the heat accumulation in the gas, the first term on the right-hand side represents the heat contribution, and the second term represents the heat generated by the catalytic reaction.

In the SO_2 catalytic conversion, the conversion rate is defined as the amount of SO_2 that has been converted into SO_3, and it is given by:

$$\tau_{SO_2}^k = \frac{\dot{Q}_{gas}^{k,in}.C_{SO_2,gas}^{k,in} - \dot{Q}_{gas}^{k,out}.C_{SO_2,gas}^{k,out}}{\dot{Q}_{gas}^{k,in}.C_{SO_2,gas}^{k,in}} \quad (14)$$

At the equilibrium state, and using the reaction stoichiometry and the mass conservation, the molar concentration of the

components involved in the conversion reaction can be expressed according to the SO_2 conversion rate τ_{SO_2} as follows:

$$\dot{Q}_{gas}^{k,out}.C_{SO_2,gas}^{k,out} = \left(1-\tau_{SO_2}^{k}\right).\dot{Q}_{gas}^{k,in}.C_{SO_2,gas}^{k,in} \quad (15)$$

$$\dot{Q}_{gas}^{k,out}.C_{O_2,gas}^{k,out} = \dot{Q}_{gas}^{k,in}.C_{O_2,gas}^{k,in} - \frac{1}{2}.\tau_{SO_2}^{k}.\dot{Q}_{gas}^{k,in}.C_{SO_2,gas}^{k,in} \quad (16)$$

$$\dot{Q}_{gas}^{k,out}.C_{SO_3,gas}^{k,out} = \dot{Q}_{gas}^{k,in}.C_{O_2,gas}^{k,in} + \tau_{SO_2}^{k}.\dot{Q}_{gas}^{k,in}.C_{SO_2,gas}^{k,in} \quad (17)$$

Assuming that the mixture gas is ideal, the SO_2 molar concentration $C_{SO_2,gas}^{k}$ can be given according to its molar fraction $X_{SO_2,gas}^{k}$ using the equation (18):

$$C_{SO_2,gas}^{k} = X_{SO_2,gas}^{k}.\frac{R.T_{gas}^{k}}{P_{gas}^{k}} \quad (18)$$

$$\sum_{i=1}^{4} C_{i,gas}^{k} = \frac{P_{gas}^{k}}{R.T_{gas}^{k}} \quad (19)$$

By summing the equations from (15) to (17) and using the equation (19), the gas flowrate $\dot{Q}_{gas}^{k,out}$ at the outlet of the catalytic bed k can be calculated by (20):

$$\dot{Q}_{gas}^{k,out} = \left(1-\frac{R.T_{gas}^{k,in}}{2.P_{gas}^{k,in}}.\tau_{SO_2}^{k}.C_{SO_2,gas}^{k,in}\right).\frac{P_{gas}^{k,in}.T_{gas}^{k,out}}{T_{gas}^{k,in}.P_{gas}^{k,out}}.\dot{Q}_{gas}^{k,in} \quad (20)$$

3.2. Heat exchangers model

In this study, STHE used in the SO_2 conversion unit was modeled based on the heat balance between fluids in the shell and the tube side. In the STHE model development, the following assumptions were considered: (1) pressure drops are negligible, (2) fluids are single-phase, (3) fluids are incompressible, (4) fluids flows are one dimensional, (4) fluids flows are radially isotropic (5) radiations heat transfer are negligible, (5) heat capacities of fluids and solids are constant. Under these assumptions, the STHE model can be written in terms of energy balance equations in the tube side and the shell side [18]:

$$\rho_{tb}.\dot{V}_{tb}.C_{p,tb}.\frac{dT_{tb,out}}{dt} = \rho_{tb}.\dot{V}_{tb}.C_{p,tb}.\left(T_{tb,in} - T_{tb,out}\right) - \dot{Q}_{ex} \quad (21)$$

$$\rho_{sh}.\dot{V}_{sh}.C_{p,sh}.\frac{dT_{sh,out}}{dt} = \rho_{tb}.\dot{V}_{tb}.C_{p,sh}.\left(T_{sh,in} - T_{sh,out}\right) + \dot{Q}_{ex} - \dot{Q}_{loss} \quad (22)$$

where:

- ρ_{tb}, ρ_{sh} : densities of fluids in the tube side and the shell side, respectively (kg/m3);
- $\dot{Q}_{tb}, \dot{Q}_{sh}$: volume flow rate of fluids in the tube side and the shell side, respectively (m3/h);
- $C_{p,tb}, C_{p,sh}$: heat capacities of fluids in the tube side and the shell side, respectively (J/kg.K);
- $T_{tb,in}, T_{tb,out}$: inlet and outlet temperature of fluid in the tube side (K);
- $T_{sh,in}, T_{sh,out}$: inlet and outlet temperature of fluid in the shell side (K);
- $\dot{Q}_{ex}, \dot{Q}_{loss}$: heat transfer and heat loss in the heat exchanger, respectively (W).

The heat transfer between the tube side and the shell side, and the heat loss in the atmosphere are defined as bellow:

$$\dot{Q}_{ex} = U_{ex}.A_{ex}.\left(T_{tb,out} - T_{sh,out}\right) \quad (23)$$

$$\dot{Q}_{loss} = U_{loss}.A_{loss}.\left(T_{sh,out} - T_{amb}\right) \quad (24)$$

where:

- U_{ex}, U_{loss} : heat transfer coefficients (W/m².K);
- A_{ex}, A_{loss} : heat transfer areas (m²);
- T_{amb} : ambient temperature (K).

By substituting (23) and (24) in (21) and (22), the dynamic model of the STHE can be described by the following equation:

$$X' + AX = B \quad (25)$$

With:

$$A = \begin{bmatrix} 1+\frac{U_{ex}.A_{ex}}{\rho_{tb}.\dot{Q}_{tb}.C_{p,tb}} & -\frac{U_{ex}.A_{ex}}{\rho_{tb}.\dot{Q}_{tb}.C_{p,tb}} \\ -\frac{U_{ex}.A_{ex}}{\rho_{sh}.\dot{Q}_{sh}.C_{p,sh}} & 1+\frac{U_{ex}.A_{ex}}{\rho_{sh}.\dot{Q}_{sh}.C_{p,sh}}+\frac{U_{loss}.A_{loss}}{\rho_{sh}.\dot{Q}_{sh}.C_{p,sh}} \end{bmatrix}$$

$$B = \begin{bmatrix} T_{tb,in} \\ T_{sh,in}+\frac{U_{loss}.A_{loss}}{\rho_{sh}.\dot{Q}_{sh}.C_{p,sh}} \end{bmatrix}$$

$$X = \begin{bmatrix} T_{tb,out} \\ T_{sh,out} \end{bmatrix}$$

The overall heat transfer coefficients can be calculated using the exchanger sizes and material properties:

$$\frac{1}{U_{ex}.A_{ex}} = \frac{1}{h_{tb}.A_{tb,i}} + \frac{\ln\left(\frac{r_{tb,o}}{r_{tb,i}}\right)}{2\pi.L_{tb}.(r_{tb,o}-r_{tb,i}).\lambda_{tb}} + \frac{1}{h_{sh}.A_{tb,o}} \quad (26)$$

$$\frac{1}{U_{loss}.A_{loss}} = \frac{1}{h_{sh}.A_{sh,i}} + \frac{\ln\left(\frac{r_{sh,o}}{r_{sh,i}}\right)}{2\pi.L_{sh}.(r_{sh,o}-r_{sh,i}).\lambda_{sh}} + \frac{1}{h_{amb}.A_{sh,o}} \quad (27)$$

where:

- h_{tb}, h_{sh} : heat transfer coefficient of fluids in the tube side and the shell side, respectively (W/m².K);
- h_{amb} : heat transfer coefficient of the air (W/m².K);
- $r_{tb,i}, r_{tb,o}$: inner and outer radius of the tube (W/m².K);
- $r_{sh,i}, r_{sh,o}$: inner and outer radius of the shell (W/m².K);

- L_{sh}, L_{sh} : length of the shell and the tube, respectively (m);
- $A_{tb,i}, A_{tb,o}$: tube side inner and outer areas (m²);
- $A_{sh,i}, A_{sh,o}$: shell side inner and outer areas (m²).

Heat transfer coefficients in fluids circulating in the tube and shell can be estimated using the Prandtl (Pr), Reynolds (Re), and Nusselt (Nu) numbers:

$$Pr = \frac{\mu.C_p}{\lambda} \quad (28)$$

$$Re = \frac{\rho.U.D_h}{\mu} \quad (29)$$

$$Nu = \frac{h.D_h}{\lambda} = f(Re, Pr) \quad (30)$$

where:

- μ : fluid dynamic viscosity (kg/m.s);
- C_p : specific heat (J/kg.K);
- λ : fluid thermal conductivity (W/m².K);
- ρ : fluid density (kg/m³);
- U : fluid velocity (m/s);
- D_h : characteristic diameter (m).
- h : convective heat transfer coefficient (W/m².K);
- f : function of Re and Pr.

Several correlations can be used to relate these three dimensionless numbers, depending to the geometry and the flow type, knowing the free convection at vertical wall or the horizontal plates, and the forced convection on flat plate. Thus, numerous equations are available to calculate the Nu number according to Re and Pr numbers, such as Dittus-Boelter equation [20], Sieder-Tate equation [21] and Gnielinski equation [22].

4. Model Simulation

The conversion unit simulation of the studied sulfuric acid plant was performed under the Unisim Design R451 simulator, using PFR and STHE models. The chemical components and reactions involved in the SO_2 conversion unit, and the appropriate fluid-packages were defined in the basis environment. Firstly, the simulation was configurated and performed in steady state around an operating point before running the simulation in dynamic mode. Key stream properties of the gas mixture feeding the first catalytic bed are summarized in Table (3). For the PID controllers used to adjust the inlet temperature of each catalytic bed, the setpoint in the three first catalytic beds was configurated at 440 °C, while it was set to 390 °C in the last catalytic bed. Equipment sizes and material properties were also introduced in the simulator to simulate the dynamic response of the unit. The flowsheet simulation of the studied SO_2 conversion unit under the Unisim Design R451 simulator is shown in Figure 2. Note that all the simulations performed in this study were carried out in a Dell Precision 5820 desktop, using 32 Go of RAM and Intel® Xeon® W-2123 CPU @ 3.60 GH processor.

Table 3: Gas mixture properties of the stream 'G1' feeding the first bed

Stream name	G1
Temperature (°C)	440
Pressure (kPa)	150
Flowrate Q(m³/h)	572,6
% N_2 (molar fraction)	79,16
% O_2 (molar fraction)	9,7
% SO_2 (molar fraction)	10.87
% SO_3 (molar fraction)	0.27

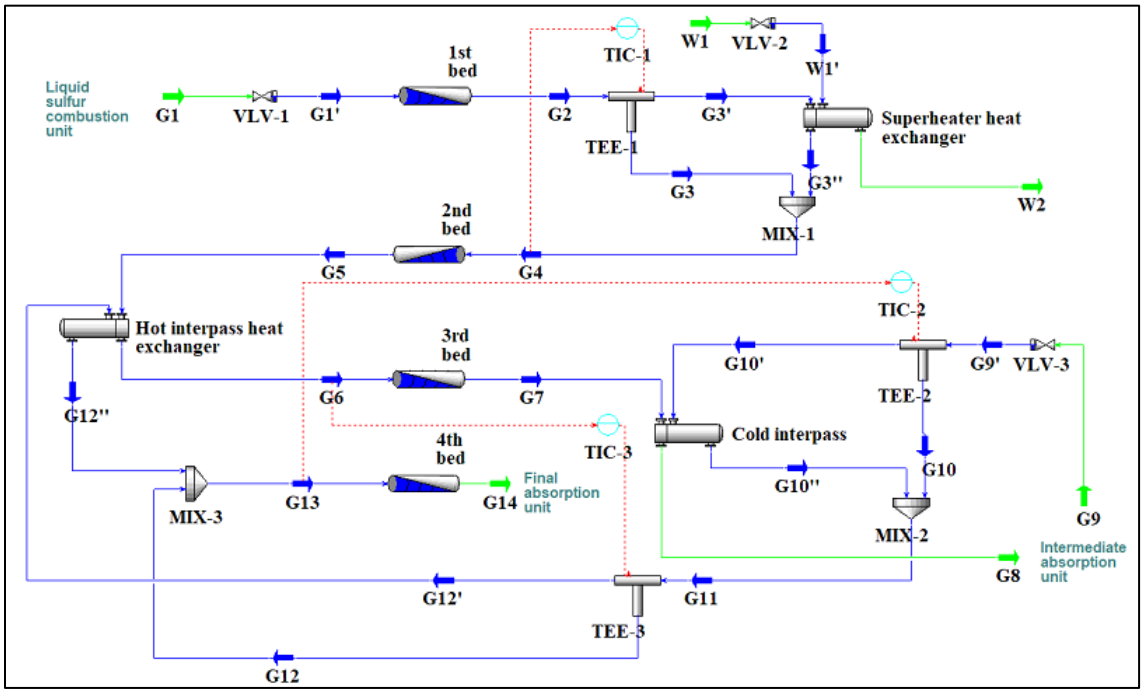


Figure 2: SO_2 conversion unit model simulation under Unisim Design

According to the steady-state simulation results, the gas mixture leaves the first catalytic bed at a temperature of 634.7 °C, since the catalytic conversion of SO_2 into SO_3 is a heat generating reaction. The conversion rate at the first catalytic bed was calculated by the model and found at the value of 63.43 %. As mentioned in the conversion unit description, the gas mixture is cooled before feeding the next catalytic bed to increase the SO_2 conversion rate. The outlet temperature value of the second and the third catalytic bed was 522 °C and 460.5 °C, and the conversion rate has reached 89.99 % and 96.59 %, respectively. At the outlet of the third catalytic bed, the gas mixture is sent to the intermediate absorption unit, in which the SO_3 produced in the first three catalytic beds is absorbed by the circulating sulfuric acid. The intermediate absorption step represents another way to well increase the SO_2 conversion rate. The absorption rate of the SO_3 in the intermediate absorption unit is about 99.98 %. The gas mixture feeds the last catalytic bed at the temperature of 390 °C and the remained SO_2 is converted into SO_3 before feeding the last absorption unit. The outlet temperature of the fourth catalytic bed was about 403.8 °C, and the cumulative conversion rate of SO_2 has reached the value of 99.97 %. Note that the gas transport along the four packed catalytic beds is accompanied by a pressure drop, and it is calculated using the Ergun equation. Table (4) summarizes the steady state simulation results of the SO_2 catalytic conversion reactor.

Table 4: Steady state simulation results of the SO_2 catalytic converter.

Reactor beds	1st bed	2nd bed	3rd bed	4th bed
Inlet temperature (°C)	440	440	440	390
Outlet temperature (°C)	634.7	522	460.5	403.8
Pressure drop (kPa)	1.707	1.735	2.084	1.662
Cumulative conversion rate (%)	63.43	89.99	96.59	99.97

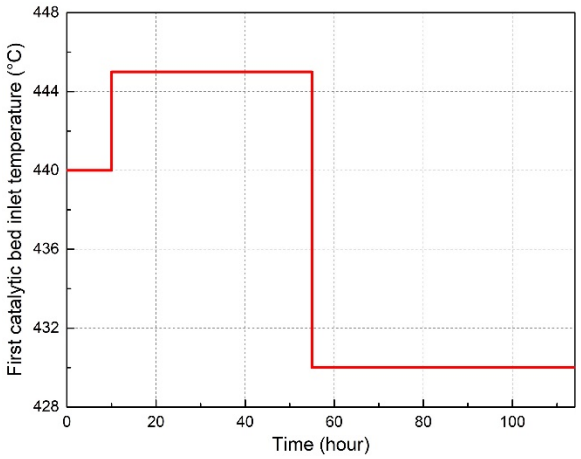


Figure 3: First catalytic bed inlet temperature variations for the dynamic simulation

To simulate the dynamic response of the studied system, slight variations were applied on the first catalytic inlet temperature as shown in Figure 3. In this study, only the dynamic response of one catalytic bed was considered since the model is the same and remains applicable for all the other catalytic beds. Afterward, the SO_2 conversion rate and the outlet temperature of the first catalytic bed have been calculated and observed using the dynamic simulation of the studied unit. Simulation results of the first catalytic bed dynamic response are presented in Figure 4 and Figure 5. Firstly, the inlet temperature of the first catalytic bed has been set to the value of 445 °C. As shown in the dynamic simulation results, the outlet temperature increased from 634.7 °C to 637 °C after a rise time 28.5 hours, and with a time constant of 10 hours. Then, a new steady state has been established at this point. When it comes to the SO_2 conversion rate, it decreased to the value of 62.68 %. After decreasing the inlet temperature of the gas mixture at the first catalytic bed from 445 °C to 430 °C, the system has evolved towards a new steady state point. The transit response of the system after this variation can be also described approximately with a time constant of 9.75 hours and a rise time of 29 hours. After that, the outlet temperature decreased from 637 °C to the value of 629.5 °C, while the SO_2 conversion rate increased from 62.68 % to the value of 64.91 %.

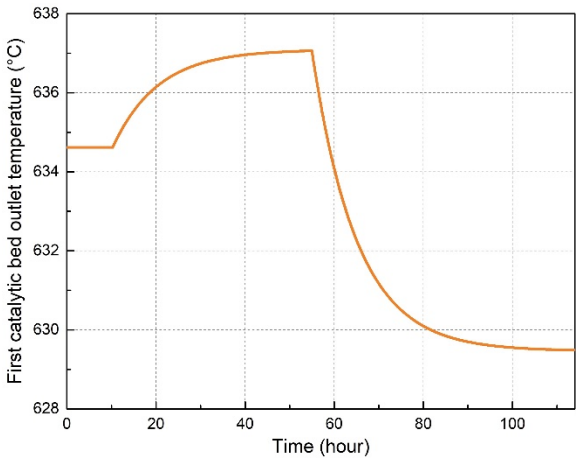


Figure 4: First catalytic bed outlet temperature dynamic response

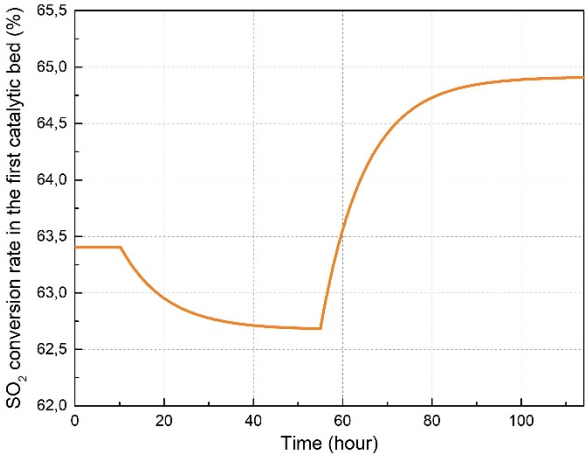


Figure 5: First catalytic bed SO_2 conversion dynamic response

5. Digital Twin architecture development and simulation

There are different understandings of Digital Twins. In some papers, the focus is on simulation [23, 24]. Others argue that DT integrates three dimensions: physical, virtual, and communication parts [25, 26]. In this paper, based on the three-dimension model, the proposed architecture involves three main components: the physical system, the virtual system, and the communication protocol. The physical system represents the real sulfuric acid

process. The virtual system is the developed and simulated model. For communication, the Open Platform Communications Unified Architecture (OPC UA), an industrial communication protocol that enables interoperability and connectivity of devices with different protocols [27], is used. In [28], a literature review discussing digital twins, it was mentioned that one of the main gaps in digital twins is the integration with the control system.

In the proposed architecture, the DT is integrated into a DCS context. The DT is installed on a server in the engineering room of the process distributed control system (DCS) (Figure 6) A copy of the graphical interface of the DT is also integrated into the control room for operators' daily uses. Using the OPC UA, an OPC server is developed to read real time data from field control stations (FCS) (Figure 7), which are connected to the real process. Then an OPC client is used to read data from the server and feed it to the core simulator developed using Unisim Design R451. A friendly user interface, inspired by the supervisory interfaces, is then integrated to show the results in both the engineering room and the control room. The communication between the modules of the DT, the OPC client, the core simulator, and the graphical interface is ensured based on the COM protocol. To communicate with the FCSs, the DT unit uses the appropriate DCS interface card. Regarding the Human Interface Station DT (HIS-DT), the communication with the DT uses the Distributed COM (DCOM) protocol.

The proposed DT of the studied SO_2 conversion unit was implemented and tested in a period of 24 hours. Measurements of the liquid sulfur and the air supply properties (flowrate, temperature, and pressure) were collected online with a sampling time of 1 min. Thus, the simulation of the studied unit was performed in real-time using the generated data and the actual operating parameters of the plan as inputs. In Figure. 8 is shown the temporal variation of the four beds outlet temperature obtained by the DT simulation and the plant measurements. To compare and validate the simulation results, some statistical measures were calculated based on data illustrated in Figure. 8, knowing the determination coefficient R^2, the maximum error ME, the mean absolute error MAE, and the root mean square error RMSE. Table (5) regroups the results of the statistical comparison between the simulation results and the measurement values. Referring to Table (5), it is shown that results obtained by the simulation are in good agreement with the plant data with 97 %, which validates the model employed in the development of the SO_2 conversion unit DT.

Table 5: Statistical measures of the simulation results and plant data.

Measure	1st bed	2nd bed	3rd bed	4th bed
R-squared	0.9708	0.9838	0.9842	0.9787
Maximum error (°C)	11.14	6.32	8.52	6.20
Mean absolute error (°C)	7.76	3.97	6.20	5.16
Root mean square error (°C)	2.07	2.40	1.89	1.25

6. Conclusion

In this work, a digital twin framework of an industrial SO_2 conversion unit of four catalytic beds reactors was developed. The proposed digital twin integrates the plant, the virtual system, and the communication part. In the virtual system, Unisim Design R451 simulator was used to model and simulate the studied SO_2 conversion unit. For a smooth industrial integration, a graphical interface, inspired by the operator's graphical displays, was used. The graphical interface makes it possible to visualize the most important parameters obtained by the unit simulation, knowing the outlet temperature, the conversion rate and the pressure drop within the four catalytic bed. Regarding the communication part,

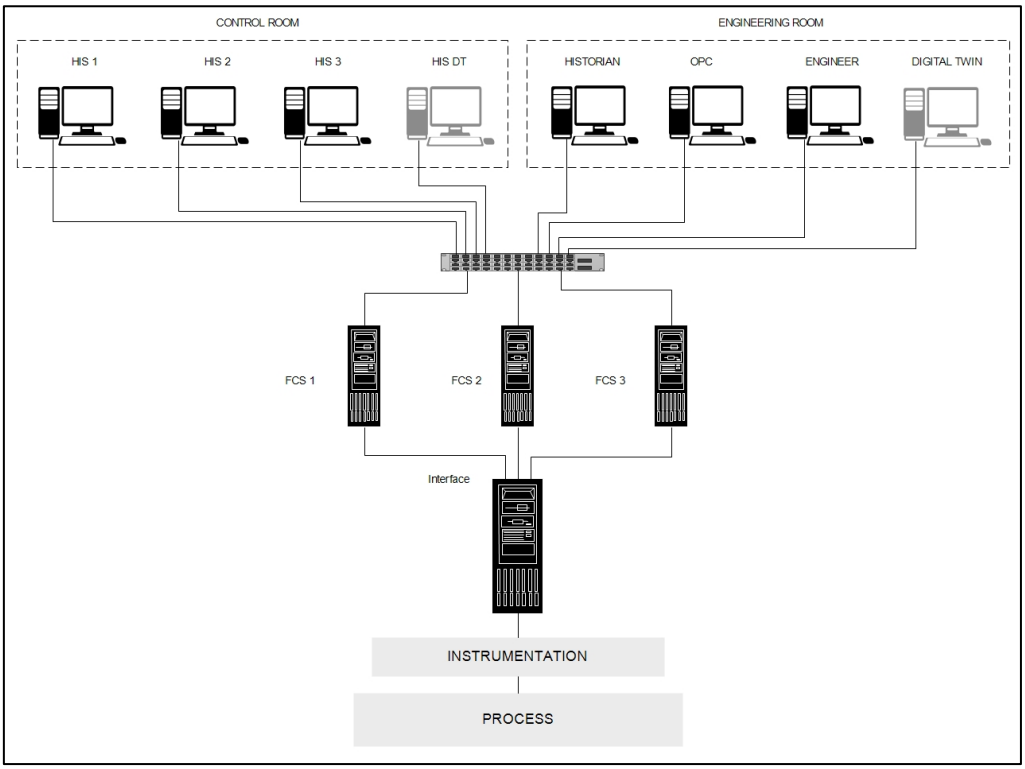


Figure 6: Digital Twin integration with DCS

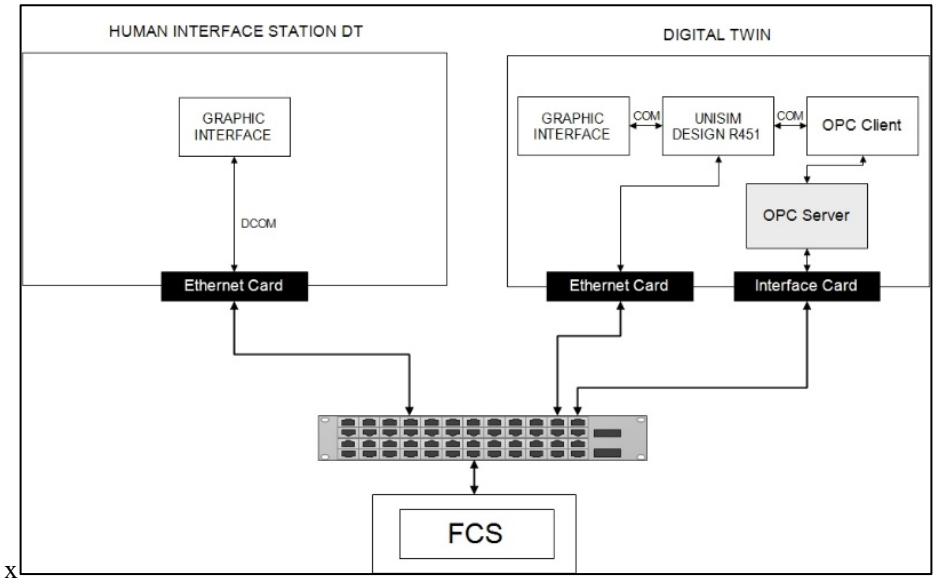


Figure 7: Digital Twin proposed architecture

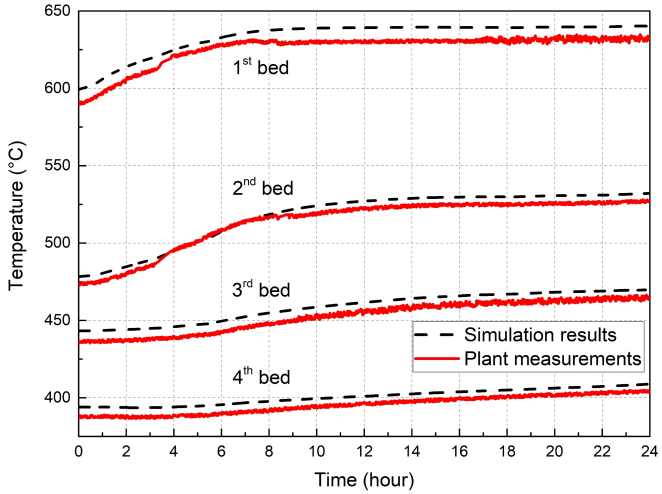


Figure 8: Real-time simulation results of the proposed SO_2 conversion unit model

an OPC server is developed to read process real time data. Then an OPC client is used to feed it to the core simulator. The simulation model was tested and validated using data from the studied sulfuric acid plant, and the comparison has shown an accuracy that exceeds 97%. The digital twin will then be used to improve operations and simulate production scenarios for real time optimization and decisions.

In future works, the proposed approach will be combined with suitable control strategies for performance improvement of the studied industrial sulfuric acid plant.

References

[1] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, F. Sui, "Digital twin-driven product design, manufacturing and service with big data," International Journal of Advanced Manufacturing Technology, **94**(9–12), 3563–3576, 2018, doi:10.1007/s00170-017-0233-1.

[2] C. Zhuang, J. Liu, H. Xiong, "Digital twin-based smart production management and control framework for the complex product assembly shop-floor," International Journal of Advanced Manufacturing Technology, **96**(1–4), 1149–1163, 2018, doi:10.1007/s00170-018-1617-6.

[3] T.H.J. Uhlemann, C. Lehmann, R. Steinhilper, "The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0," Procedia CIRP, **61**, 335–340, 2017, doi:10.1016/j.procir.2016.11.152.

[4] G.L. Knapp, T. Mukherjee, J.S. Zuback, H.L. Wei, T.A. Palmer, A. De, T. DebRoy, "Building blocks for a digital twin of additive manufacturing," Acta Materialia, **135**, 390–399, 2017, doi:10.1016/j.actamat.2017.06.039.

[5] A. Ustundag, E. Cevikcan, Managing The Digital Transformation, 2018.

[6] E.J. Tuegel, A.R. Ingraffea, T.G. Eason, S.M. Spottswood, "Reengineering aircraft structural life prediction using a digital twin," International Journal of Aerospace Engineering, **2011**, 2011, doi:10.1155/2011/154798.

[7] Y. Zhou, T. Xing, Y. Song, Y. Li, X. Zhu, G. Li, S. Ding, "Digital-twin-driven geometric optimization of centrifugal impeller with free-form blades for five-axis flank milling," Journal of Manufacturing Systems, (June), 0–1, 2020, doi:10.1016/j.jmsy.2020.06.019.

[8] H. Zhang, Q. Liu, X. Chen, D. Zhang, J. Leng, "A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line," IEEE Access, **5**, 26901–26911, 2017, doi:10.1109/ACCESS.2017.2766453.

[9] A.O. Oni, D.A. Fadare, S. Sharma, G.P. Rangaiah, "Multi-objective optimisation of a double contact double absorption sulphuric acid plant for cleaner operation," Journal of Cleaner Production, **181**, 652–662, 2018, doi:10.1016/j.jclepro.2018.01.239.

[10] R. Günther, J.C. Schöneberger, H. Arellano-Garcia, H. Thielert, G. Wozny, Design and modeling of a new periodical-steady state process for the oxidation of sulfur dioxide in the context of an emission free sulfuric acid plant, 1677–1681, 2012, doi:10.1016/B978-0-444-59506-5.50166-8.

[11] H. Nouri, A. Ouederni, "Experimental and modeling study of sulfur dioxide oxidation in packed-bed tubular reactor," International Journal of Innovation and Applied Studies, **3**(4), 1045–1052, 2013.
[12] P.A. Sørensen, M. Møllerhøj, K.A. Christensen, "New dynamic models for simulation of industrial SO2 oxidation reactors and wet gas sulfuric acid plants," Chemical Engineering Journal, **278**, 421–429, 2015, doi:10.1016/j.cej.2014.09.023.
[13] J. He, J. Zhang, H. Shang, "Dynamic modelling and simulation of the sulphur dioxide converter in an industrial smelter," Canadian Journal of Chemical Engineering, **97**(6), 1838–1847, 2019, doi:10.1002/cjce.23446.
[14] A. Mounaam, Y. Harmen, Y. Chhiti, A. Souissi, M. Salouhi, M. El Khouakhi, "UniSim-design simulation and analysis of a sulphuric acid manufacturing plant with double absorption process," SIMULTECH 2020 - Proceedings of the 10th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, (Simultech), 91–100, 2020, doi:10.5220/0009832300910100.
[15] F. Zaversky, M. Sánchez, D. Astrain, "Object-oriented modeling for the transient response simulation of multi-pass shell-and-tube heat exchangers as applied in active indirect thermal energy storage systems for concentrated solar power," Energy, **65**, 647–664, 2014, doi:10.1016/j.energy.2013.11.070.
[16] E. Ozden, I. Tari, "Shell side CFD analysis of a small shell-and-tube heat exchanger," Energy Conversion and Management, **51**(5), 1004–1014, 2010, doi:10.1016/j.enconman.2009.12.003.
[17] J. Janaun, N.H. Kamin, K.H. Wong, H.J. Tham, V. V. Kong, M. Farajpourlar, "Design and simulation of heat exchangers using Aspen HYSYS, and Aspen exchanger design and rating for paddy drying application," IOP Conference Series: Earth and Environmental Science, **36**(1), 2016, doi:10.1088/1755-1315/36/1/012056.
[18] A. Mounaam, Y. Harmen, Y. Chhiti, A. Souissi, M. Salouhi, K. Benjelloun, M. Elkhouakhi, L. Deshayes, "Dynamic behavior analysis of tubular heat exchanger: Experimental and theoretical study," 2020 5th International Conference on Renewable Energies for Developing Countries, REDEC 2020, **5**, 1–6, 2020, doi:10.1109/REDEC49234.2020.9163862.
[19] M.J. King, W.G. Davenport, M.S. Moats, Overview, Elsevier, Oxford: 1–9, 2013, doi:10.1016/b978-0-08-098220-5.00001-0.
[20] M. Yang, X. Wang, Z. Wang, Z. Li, Y. Zhang, "Correlation for turbulent convection heat transfer in elliptical tubes by numerical simulations," Frontiers in Heat and Mass Transfer, **11**, 2018, doi:10.5098/hmt.11.7.
[21] Y. Matsuda, O. Kawanami, R. Orimo, K. Uete, A. Watanabe, Y. Egami, H. Yamaguchi, T. Niimi, "Simultaneous measurement of gas-liquid interface motion and temperature distribution on heated surface using temperature-sensitive paint," International Journal of Heat and Mass Transfer, **153**, 119567, 2020, doi:10.1016/j.ijheatmasstransfer.2020.119567.
[22] S. Ebrahim Ghasemi, A. Akbar Ranjbar, "Numerical thermal study on effect of porous rings on performance of solar parabolic trough collector," Applied Thermal Engineering, **118**(April), 807–816, 2017, doi:10.1016/j.applthermaleng.2017.03.021.
[23] S. Weyer, T. Meyer, M. Ohmer, S. Dominic, T. Meyer, M. Ohmer, S. Dominic, T. Meyer, M. Ohmer, D. Gorecky, "Future Modeling and Simulation of CPS-based Factories : Future Factories : Future Factories : Example from the Automotive Industry Future an Factories : an Example from the Automotive Example from the Automotive an Example from the Automotive Industry," IFAC PapersOnLine, **49**(31), 97–102, 2016, doi:10.1016/j.ifacol.2016.12.168.
[24] T. Gabor, M. Kiermeier, M.T. Beck, A. Neitz, "A Simulation-Based Architecture for Smart Cyber-Physical Systems."
[25] Q. Qi, F.E.I. Tao, S. Member, "Digital Twin and Big Data Towards Smart Manufacturing and Industry 4 . 0 : 360 Degree Comparison," IEEE Access, **6**, 3585–3593, 2018, doi:10.1109/ACCESS.2018.2793265.
[26] V. Souza, S. Member, R. Cruz, W. Silva, S. Lins, V. Lucena, "A Digital Twin Architecture Based on the Industrial Internet of Things Technologies," 2019 IEEE International Conference on Consumer Electronics (ICCE), 1–2.
[27] L.F. C. Cimino, E. Negri, "OPC Unified Architecture," 2009.
[28] C. Cimino, E. Negri, L. Fumagalli, "Computers in Industry Review of digital twin applications in manufacturing," Computers in Industry, **113**, 103130, 2019, doi:10.1016/j.compind.2019.103130.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 959-966 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Standalone Operation of Modified Seven-Level Packed U-Cell Inverter for Solar Photovoltaic System

Kishan Bhushan Sahay[1,*], Pankaj Kumar Singh[1], Rakesh Maurya[2]

[1]*Department of Electrical Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur,273010, India*

[2]*Department of Electrical Engineering, Sardar Vallabhbhai National Institute of Technology, Surat, 395007, India*

ARTICLE INFO

Article history:
Received: 29 October, 2020
Accepted: 16 November, 2020
Online: 08 December, 2020

Keywords:
Cascaded H-Bridge (CHB)
Modified Packed U-Cell (MPUC)
Packed U-Cell (PUC)
Solar Photovoltaic System

ABSTRACT

In this paper, a modified configuration of Single-Phase Seven-Level Packed U-Cell (PUC) Multilevel Inverter for solar photovoltaic system is presented & investigated for standalone operation. The Seven-Level MPUC Inverter comprises of six semiconductor switches & two DC links which generates seven voltage levels at the inverter output. The maximum amplitude of inverter output voltage is more than that of available maximum DC link value (i.e. sum of two DC link values). The voltage levels generated in the MPUC inverter is more with reduced number of components counts as compare to conventional multilevel inverter topologies like Cascaded H-Bridge (CHB) and Neutral Point Clamped (NPC). The proposed inverter finds application in PV System where green power is derived from two PV panels with different power rating and voltage rating connected to DC links. Several design considerations viz. the RMS value of inverter output voltage, switching frequency & voltage rating of semiconductor switches are taken into consideration to prove the advantages of the MPUC inverter. The simulation results of the MPUC inverter shows the appropriate operation of multilevel inverter.

1. Introduction

Owing to growing number of consumers in addition to large scale power industries, the power grid has experienced large energy demand from last decade. Traditional bulky transformers are significantly replaced by power electronic devices because a lot of research work is happening in power electronic field these days resulting in advancement of semiconductor device technology. Power converters are utilized as a conversion medium for transferring green power either for grid or a standalone load. Economic pricing of power semiconductor devices makes it beneficial to produce and allow them in competitive market. Recently, the use of a greater number of switches in the converter causes no significant rise in pricing. This is the reason that the traditional two-level converters having high harmonics and power losses are being interchanged by the multilevel inverters having less switching frequency [1, 2].

Now, the focus of researchers has moved towards developing various topologies of multilevel inverters and their control strategies. So, the foremost attention is to reduce the component counts in these inverters. Fewer the component counts, lower the cost and power loss [3]-[5]. The two most popular commercially available multilevel inverter (MLI) topologies are CHB and NPC inverter topologies [6].

Recently, the flooding of green power into the grid has become a prime concern [7]. Most of the countries have initiated to implement the solar power panels at broad scale to cater the local energy need. With the use of power electronic converters, the DC power is converted into AC power which is further transferred to the grid or a standalone load.

Thus, the improvement in conversion efficiency and reduction in power loss of these converters are tackled in several reports [8-9]. Conversion of DC voltage of PV system into AC waveform with low harmonics to be used by the grid or standalone load with high conversion efficiency and small size filtering requirement can be fulfilled by single-phase multilevel inverters (MLI) [10].

The key problem with most of the topologies are to use numerous isolated DC sources which makes it difficult for its practical implementation. An isolated DC source can be made by an AC transformer and diode bridge rectifier which is a costly affair and become bulky for manufacturing. Consequently, topologies having single DC source i.e. NPC (Neutral Point

*Corresponding Author: Kishan Bhushan Sahay, kishansahay16@gmail.com

Clamped) inverter and five-level PUC (PUC5) inverter [11] are commercially more viable than other new topologies. Nevertheless, some topologies e.g. CHB are utilized for very high-power applications owing to necessity of large number of DC sources.

In latest survey, seven-level PUC inverter topology was introduced which has many advantages like low component counts and numerous output voltage levels [12]. The key problem with PUC inverter is its maximum output voltage level which is actually the maximum available DC link value of the topology. In recent times, a seven-level modified PUC inverter topology has been presented which provides an alternative for those applications where several Photovoltaic panels are available for connecting it to isolated DC links. In [13], the authors demonstrated that Photovoltaic panels having different voltage rating & power rating can be coupled with two DC link of modified PUC inverter topology.

In this paper, modification of Seven Level Packed U-Cell inverter is done for solar PV systems. MPUC inverter topology is similar to CHB inverter having unequal DC sources where two cells produces seven voltage levels. The introduced inverter has a smaller number of semiconductor switches than the CHB inverter with equivalent performance. In MPUC inverter, two unequal DC sources (i.e. two PV panels with different power rating) are required to generate seven output voltage levels by means of reduced harmonic and uniformly distributed voltage levels. The maximum inverter output voltage level is sum of amplitudes of two dc sources. So, authors call it as Modified PUC Inverter (MPUC).

The seven-level PUC inverter is elaborated in section II. Subsequently, seven-level MPUC topology is described in section III. The rms value of seven level inverter output voltage for standalone operation is discussed in section IV. The results obtained from simulation of MPUC inverter has been given in section V. Conclusion is specified under section VI.

2. Seven-Level Packed U-Cell Inverter

The template is used to PUC inverter topology was first introduced by Al-Haddad in 2010 & later it was advanced by Hani Vahedi in 2015 [14]. PUC inverter is among the most promising multilevel inverters which takes advantages of both Flying Capacitor Converter (FCC) & CHB inverter. This topology is known as Packed U-Cell as the shape of each inverter unit is in U form. In Each U Cell, it contains two semiconductor switches & one DC voltage source. PUC topology work only with one DC source & a smaller number of active and passive components. Six semiconductor switches with anti-parallel diodes are used to achieve seven voltage levels which is evident from figure 1.

These 6 switches are divided into two legs having three switches in each leg. In one leg, there are three main switches (Q_1, Q_2 & Q_3) and in another leg, there are three complimentary switches of main switch ($Q_{1'}$, $Q_{2'}$ & $Q_{3'}$).

Various voltage levels are achieved either through DC voltage sources available or through their series opposition provided the amplitude of second DC source (V_2) is equal to one third of amplitude of first DC voltage source (V_1) i.e. $V_1 = 3V_2$. Thus, inverter output voltage has seven-voltage levels i.e. 0, $\pm V_2$, $\pm 2V_2$, and $\pm 3V_2$. From the Table I, it can be concluded that there are 3 positive voltage levels, 3 negative voltage levels and 2 zero voltage levels (viz. state 4 and 5). All eight switching states are shown graphically in Figure 2. Zero voltage level can be achieved only when the load is either shorted through main switches or through complimentary switches. Redundant zero voltage states are utilized by appropriate means just for reduction in switching frequency at the time when there are changeover in between positive & negative half cycles of AC signal.

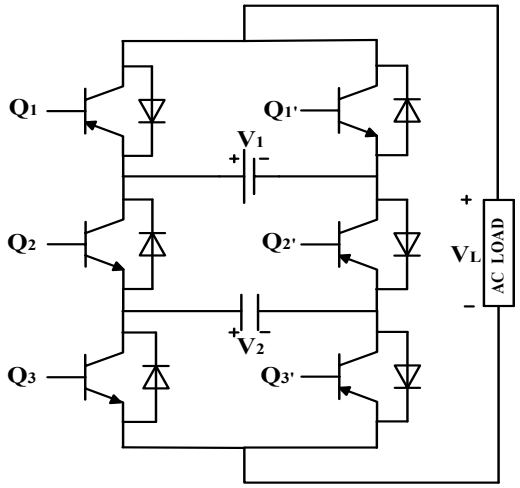


Figure 1: Packed U-Cell (PUC) Inverter Topology

Table 1: Complete Switching States & Voltage Levels

Switching States	Q_1	Q_2	Q_3	$Q_{1'}$	$Q_{2'}$	$Q_{3'}$	$V_{inv} = V_L$
1	1	0	0	0	1	1	V_1
2	1	0	1	0	1	0	$V_1 - V_2$
3	1	1	0	0	0	1	V_2
4	1	1	1	0	0	0	0
5	0	0	0	1	1	1	0
6	0	0	1	1	1	0	$-V_2$
7	0	1	0	1	0	1	$V_2 - V_1$
8	0	1	1	1	0	0	$-V_1$

From above switching states Table I, it can be concluded that irrespective of change in switching frequency and carrier frequency of adopted PWM technique, the switches Q2 & Q2' operate at line frequency (i.e. 50 Hz). This results in significant reduction of switching losses. Thus, the switches produce flow of current in the DC bus resulting in the use of a series diode for blocking the reverse voltage. To overcome this problem, an energy storage element i.e. Capacitor is used as a second DC source. The voltage across capacitor must be one third to that of V1 which can be achieved by various voltage balancing techniques. The current through the capacitor is in positive and negative parts which shows that there is charging and discharging in the capacitor [15]-[19]. Thereby achieving the desired voltage level for generating constant desired inverter output level. The biggest issue in using PUC inverter topology is its low maximum voltage level i.e. the maximum amplitude of the PUC inverter topology is maximum value of amplitude of the two voltage sources. Thus, the PUC inverter can't generate voltage more than the maximum value of the two voltage sources. So, it doesn't get application in medium or high-power applications [20]-[23].

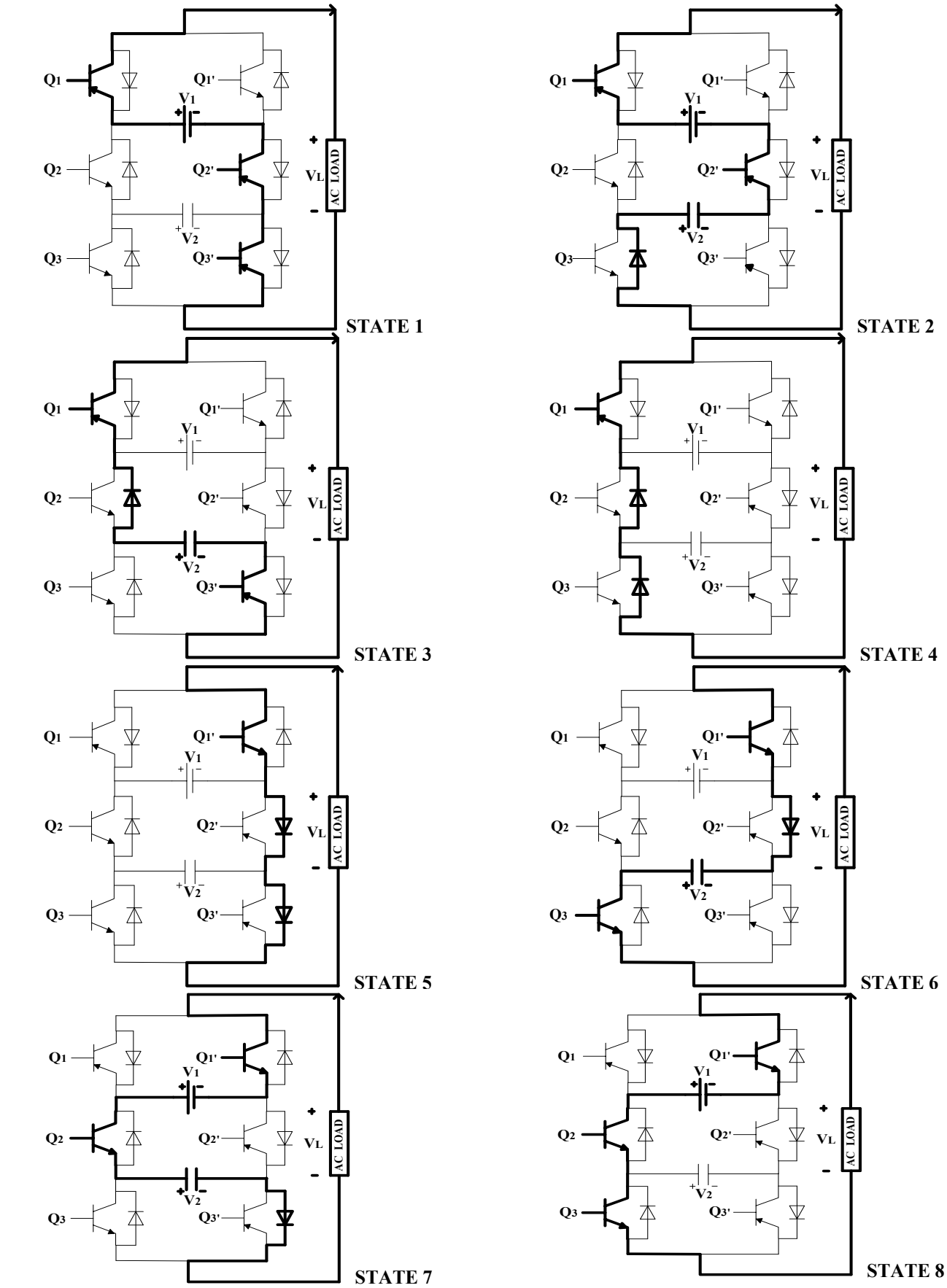


Figure 2: Switching States for PUC Inverter Topology

3. Seven-Level Modified Packed U-Cell Inverter

Due to recent advancement in solar panels and its economic cost, it becomes feasible to use number of PV panels as DC sources. Hence multilevel inverters using large number of DC sources are justified economically. The maximum output voltage in PUC inverter is not more than that of maximum DC voltage source magnitude. To solve this issue, the MPUC inverter topology is proposed as shown in figure 3. The DC source (V_2) is attached in reverse direction when compared to PUC inverter. So, the lower switches viz. Q_3 & $Q_{3'}$ have also been reversed to hinder current through diodes. Thus, with appropriate use of gate pulses, the path of current flow can be allowed or hindered. Various Switching states & output voltage levels for MPUC inverter is shown in Table II. From table II, it can be concluded that maximum inverter output voltage of MPUC inverter is nothing but total addition of two DC voltage sources i.e. V_1+V_2 which can be obtained by joining two voltage sources in series. Thus, we are getting more voltage amplitude by this proposed topology. It is evident that seven uniform voltage level is produced by MPUC inverter provided the second DC source (V_2) amplitude is two times the first DC source (V_1) amplitude i.e. $V_2 = 2V_1$. So, the output voltage levels have following levels: 0, $\pm V_1$, $\pm 2V_1$, $\pm 3V_1$. Thus, the maximum output voltage amplitude will be $3V_1$ while maximum DC source is equal to $2V_1$. The voltage level $3V_1$ is obtained by joining two DC voltage source V_1 & V_2 in series which appear on load. The voltage level $2V_1$ can be achieved by making second DC voltage source V_2 to be appear on the load. The voltage level V_1 can be achieved by making first voltage source V_1 to be appear on the load [24]-[26].

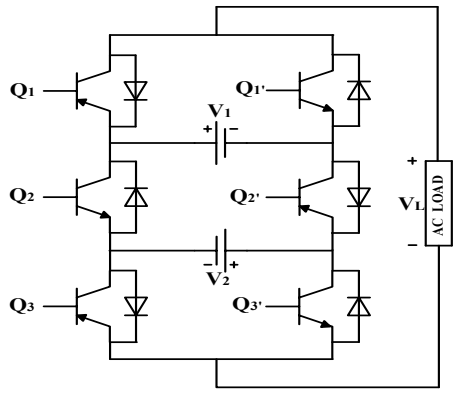


Figure 3: MPUC Inverter Topology

Table 2: Complete Switching States & Voltage Levels

Switching States	Q_1	Q_2	Q_3	$Q_{1'}$	$Q_{2'}$	$Q_{3'}$	$V_{inv} = V_L$
1	1	0	0	0	1	1	$V_1 + V_2$
2	1	0	1	0	1	0	V_2
3	1	1	0	0	0	1	V_1
4	1	1	1	0	0	0	0
5	0	0	0	1	1	1	0
6	0	0	1	1	1	0	$- V_1$
7	0	1	0	1	0	1	$- V_2$
8	0	1	1	1	0	0	$- V_1 - V_2$

Eight switching states are presented graphically in figure 4. Here are two redundant zero voltage states viz. state 4 and 5 which is used in the reduction of switching frequency. When zero voltage level is required to be generated at the output, the modulation technique selects the appropriate state in state 4 and 5 to give gate pulse to switches which gives less change in Turn-On & Turn-Off mode of operation resulting in significant reduction in switching frequency. From the switch states shown in table II, it can be concluded very clearly that irrespective of change in switching & carrier frequency, the two switches viz. Q_2 & $Q_{2'}$ operate at the line frequency i.e. 50 Hz.

4. Design Consideration of Standalone Operation

A mathematical analysis has been done for the standalone application of MPUC inverter supplying power to single-phase load. Here we aim to determine general expression for output voltage waveform of seven-level MPUC inverter based on figure 5.

The general expression for determining RMS voltage is given by:

$$V_r = \sqrt{\frac{1}{T}\int_0^T v_0^{\,2}(t)\,dt} \tag{1}$$

where Vr is RMS voltage, v_0 is time function of voltage and T is time period of the periodic function.

Thus equation (1) can be rewritten as follows:

$$V_r = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} v_0^{\,2}(t)\,dt}$$

$$= \sqrt{\frac{1}{2\pi}\left[4\int_{\theta_1}^{\theta_2} V_1^{\,2}\,d\omega t + 4\int_{\theta_2}^{\theta_3} V_2^{\,2}\,d\omega t + 2\int_{\theta_3}^{\theta_4}\left(V_1+V_2\right)^2 d\omega t\right]}$$

$$= \sqrt{\frac{1}{2\pi}\left[4V_1^{\,2}\alpha + 4V_2^{\,2}\beta + 2\left(V_1+V_2\right)^2\gamma\right]} \tag{2}$$

Since $V_2 = 2V_1$

$$V_r = \sqrt{\frac{1}{2\pi}\left[4V_1^{\,2}\alpha + 4\left(2V_1\right)^2\beta + 2\left(3V_1\right)^2\gamma\right]}$$

$$= V_1\sqrt{\frac{2\alpha + 8\beta + 9\gamma}{\pi}} \tag{3}$$

where α , β and γ are angles at which reference signal crosses voltage levels as fixed lines between the carriers. The RMS value of output voltage can also be determined in term of modulation index (m_a) and maximum output voltage level ($V_{max} = 2V_1$) which is given as follows:

$$V_r = 0.725 \times m_a \times V_{max} \tag{4}$$

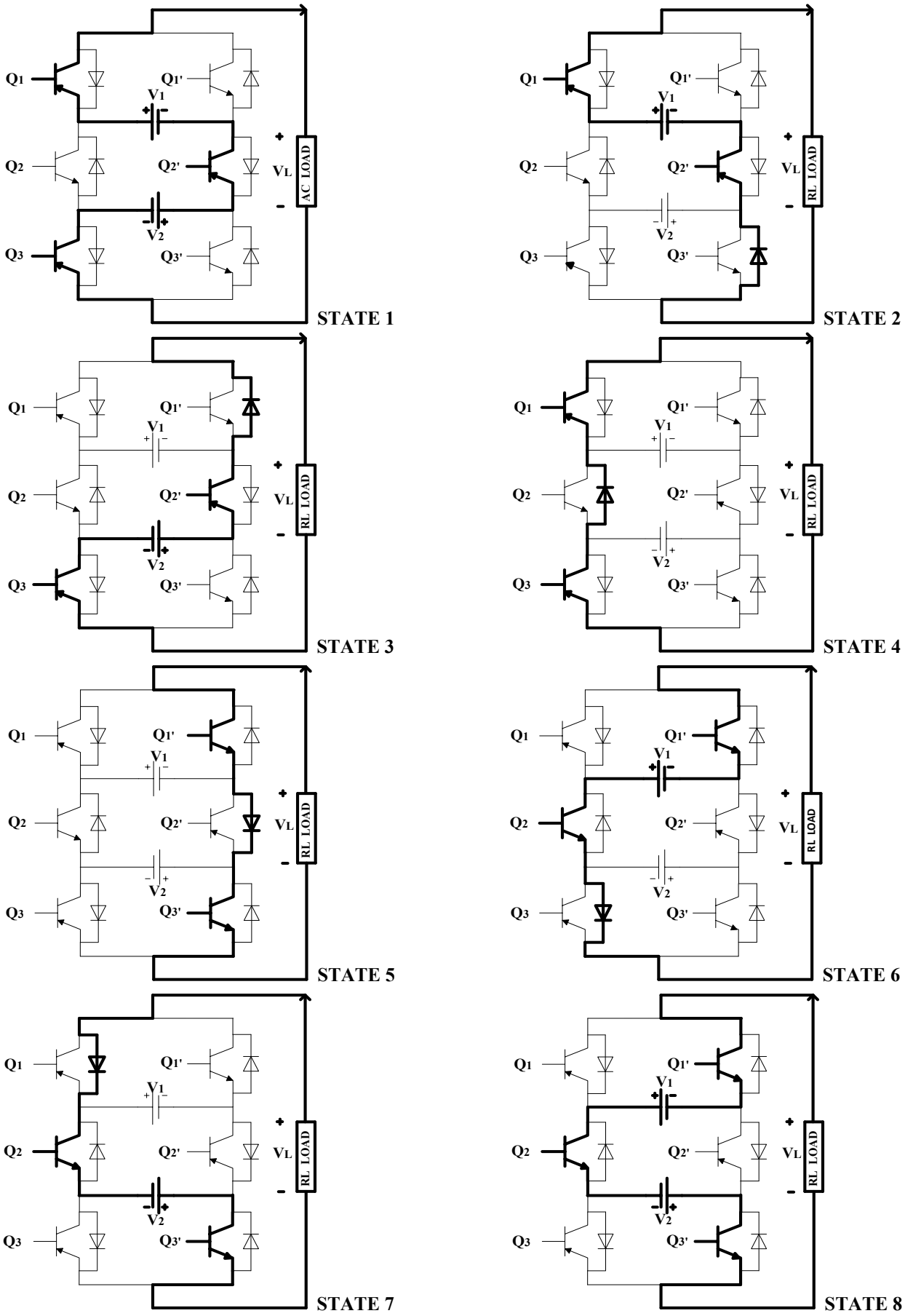


Figure 4. Switching States for MPUC Inverter

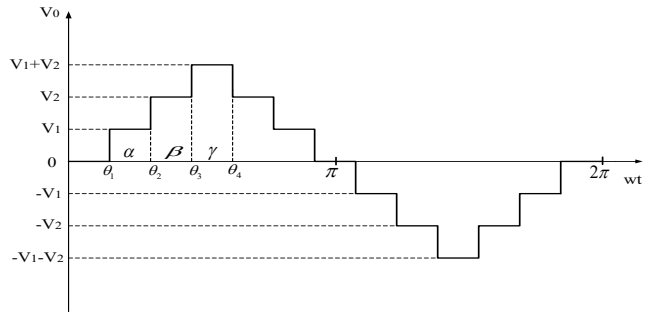


Figure 5: Seven-Level Inverter Output Voltage Waveform

5. Simulation & Results

The introduced MPUC inverter topology is simulated in standalone mode of operation to validate the above analysis. Six IGBT switches are taken in the simulation model. The switching algorithm adopted is elucidated in [16] where four level shifted carriers are used for modulating the reference signal. Thus, gate pulses are given to all the six pulses. Table III gives complete view of simulation parameters.

Here two DC sources viz. V_1 and V_2 are joined to different DC links of PV panels having different power & voltage rating and standalone load has been taken at the output. At higher switching frequency, passive components size reduces significantly.

Here the MPUC inverter is connected to RL load because most of the AC loads are inductive in nature. The parameters of RL load is specified into table III and results obtained from simulation has been given in figure 6.

The inverter output voltage waveform i.e. V_{inv} has seven uniform voltage levels having maximum of 120 Volt i.e. the total addition of two DC sources. The uniformly distributed & symmetric structure of inverter output voltage & current shows proper selection of modulation technique. THD of inverter output voltage waveform with no use of any filter is 23.96 % which is evident from figure 7. THD for output current waveform is 1.24 % which is evident from figure 8. Thus, the current THD is less than 5 % which is as per IEEE STD 519. Thus, the filtering requirement will be small and economical for MPUC inverter applications especially renewable energy.

For determining switching frequency of MPUC inverter topology, the inverter output voltage as well as gate pulses for switches Q_1, Q_2 & Q_3 has been shown for one time period of output voltage which is given in figure 9. Based on number of switching pulses for one cycle, we can determine the switching frequency of MPUC inverter switches which is given in table IV. Since each switch in one U cell operate in complimentary, so both the switches in a U cell will have same frequency. During switching time, the upper switches viz. Q_1 & Q_1, have to block voltage V_1 & lower switches viz. Q_3 & Q_3, have to block voltage V_2. So, the voltage rating of upper switches and lower switches are V_1 and V_2, respectively. The two middle switches viz. Q_2 & Q_2, have to suffer higher voltage than other switches i.e. V_1+V_2. So, the voltage rating of two middle switches are V_1+V_2.

The middle switches operate online frequency i.e. 50 Hz. So middle switches operating online frequency and high power are placed into middle U cell. Thus, a comparative study based on component counts is performed among popular Cascaded H-Bridge (CHB) inverter & MPUC inverter. The main limitation of MPUC inverter is that the unequal shading of Photovoltaic panels may cause deterioration of power balance between two DC links which may lose voltage control.

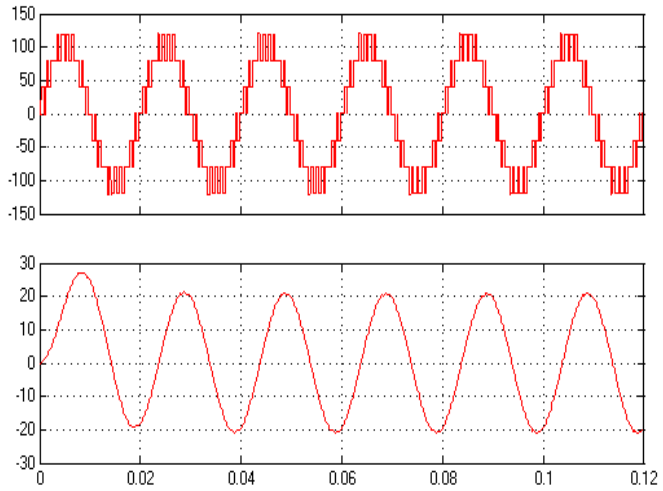


Figure 6: MPUC Inverter Output Voltage & Current Waveform

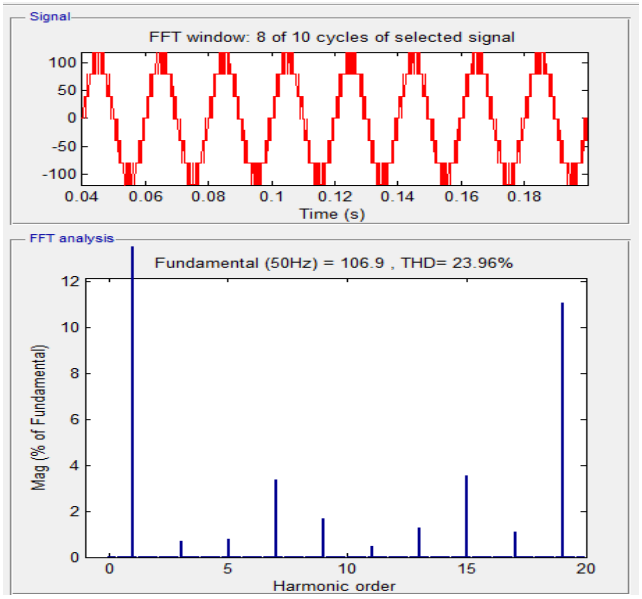


Figure 7: FFT Analysis for Inverter Output Voltage

Table 3: Parameters Used in Simulation

Parameters	Values
First DC Source (V_1)	40 Volt
Second DC Source (V_2)	80 Volt
Frequency (f)	50 Hz
Switching Frequency (f_{PWM})	1 KHz
Resistance (R)	2 ohm
Inductance (L)	15 mH

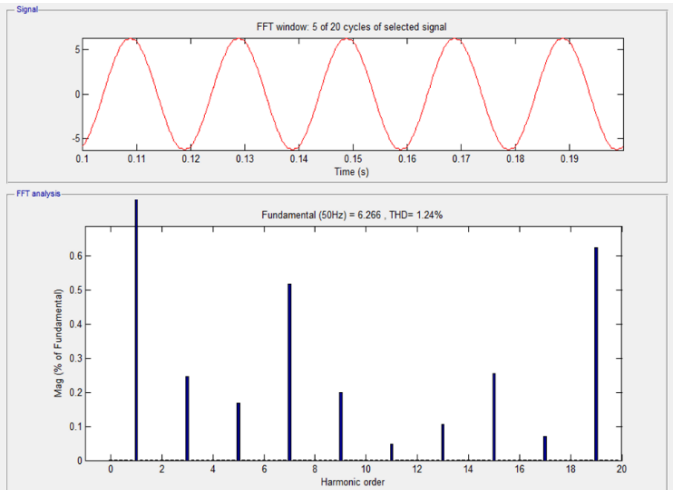


Figure 8: FFT Analysis for Inverter Output Current

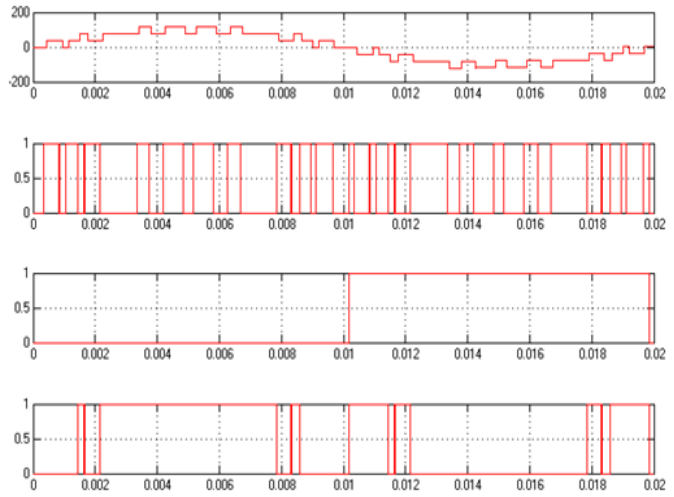


Figure 9: One Complete Cycle of (i) Inverter Output Voltage (ii) Q_1 Gate Pulse (iii) Q_2 Gate Pulse (iv) Q_3 Gate Pulse

Table 4: Switching Frequency of Inverter Switches

Q_1 & Q_1'	1250 Hz
Q_2 & Q_2'	50 Hz
Q_3 & Q_3'	500 Hz

Table 5: Comparative Study of Component Counts

Topology	Number of Switches	DC Sources
7-level CHB (Equal DC Sources)	12	3
7-level CHB (Unequal DC Sources)	8	2
7-level MPUC Inverter	6	2

6. Conclusion

In this paper, the modified form of Packed U-Cell inverter topology has been presented and validated on MATLAB Simulink. The proposed MPUC topology can generate seven levels of inverter output voltage with acceptable harmonic contents. Unlike PUC topology, the proposed topology is capable to produce maximum output voltage level which is more than the available maximum voltage source. In this topology, the DC source amplitudes gets added so that it can supply more power at the output. The switching algorithm adopted in MPUC inverter is such that it can operate at less switching frequency. Furthermore, this topology is targeted for PV applications for delivering power from different voltage rating PV panels. To show this, simulation results are given to verify THD for inverter output voltage & output current. Likewise, the comparison of MPUC inverter & CHB inverter proved the suitability of proposed inverter for advantages like fewer component counts, economical manufacturing cost and compact package owing to reduced filter size.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] D. Heydt, M.V. West, "Directions of esearch on Electric Power Quality," IEEE Trans. on Power Delivery, **8**(1), 429-436, 1993, doi: 10.1109/61.180365.

[2] S. Bakran, "Comparison of Modular Multilevel Converter based HV DC-DC-converters,"18th EPE'16 ECCE Europe, EISBN:978-9-0758-1524-5, 1-10, 2016 doi : 10.1109/EPE.2016.7695259.

[3] V. Al-Haddad, L. Rahmani, "Cascaded Multilevel Inverter with Multicarrier PWM technique and Voltage Balancing Feature," IEEE 23rd ISIE, Istanbul, 2155-2160, 2014, doi : 10.1109/ISIE.2014.6864951.

[4] Abolhasani, "Modular multipulse rectifier transformers in symmetrical cascaded H-bridge medium voltage drives," IEEE Tran. on PE, **27**(2), 698-705, 2012, doi: 10.1109/TPEL.2011.2161593.

[5] R. Bernet et al., "A Survey on Neutral-Point Clamped Inverters,"IEEE Trans. on IE, **57**(7), 2219-2230, 2010, doi: 10.1109/TIE.2009.2032430.

[6] S. Vahedi et al., "Modified Selective Harmonic Elimination Employed in Four-leg NPC Inverters," IEEE IECON, Dallas, TX, USA, Electronic ISBN:978-1-4799-4032-5, 2014, doi: 10.1109/IECON.2014.7049291.

[7] B. Singh, "Solar PV Array Fed Brushless DC Motor Driven Water Pump,"IEEE 6th ICPS, New Delhi, EISBN:978-1-5090-0128-6 2016, 1-5, doi: 10.1109/ICPES.2016.7584057.

[8] Enerdata, global energy statistical yearbook, 2015 [Online].

[9] L.B. Hiscock, "National survey report of pv power application in Canada," International Energy Agency,2013.

[10] J. Fahimi, "Multi-port power electric interface for renewable energy sources,"24th Annual IEEE Applied Power Electronics Conference and Exposition, Washington DC, 2009, doi: 10.1109/APEC.2009.4802680.

[11] V. Al-Haddad, "Multilevel electric power converter," US Patent 20110280052, 2011.

[12] V. Al-Haddad, "Real-Time Implementation of a Seven Level Packed U-Cell Inverter with a Low-Switching-Frequency Voltage Regulator," IEEE Trans. on Power Electr., **31**(8), 5967-5973, 2016, doi: 10.1109/TPEL.2015.2490221.

[13] V. Hani et al., "PUC converter review: Topology, Control and Applications," 41st Annual Conference of the IEEE Industrial Electronics Society, 4334-4339, 2015, doi: 10.1109/IECON.2015.7392774.

[14] J.I. Metri et al., " Real-Time Implementation of Model-Predictive Control on Seven-Level Packed U-Cell Inverter", IEEE Transactions on Industrial Electronics , **63**(7) ,4180 - 4186, 2016, doi: 10.1109/TIE.2016.2542133.

[15] S. Arazm et al., " Phase-shift modulation technique for 5-level packed U-cell (PUC5) inverter ", IEEE CPE-Powereng, 2018, doi: 10.1109/CPE.2018.8372577.

[16] V. Hani et al., "Modified seven level pack U-cell inverter for photovoltaic applications" IEEE Journal of Emerging and Selected Topics in Power Electronics, **6**(3), 1508-1516, 2018, doi: 10.1109/JESTPE.2018.2821663.

[17] S. Najafi, "Standalone operation of modified seven-level packed U-cell (MPUC) single-phase inverter" MDPI, **8**(3) , 2019, doi: 10.3390/electronics8030268.

[18] R. Vivekanand et al., "Improved Solar PV Cell Matlab Simulation Model and Performance Analysis," i-Manager's Journal on Electrical Engineering,

9(3), 12-19, 2016, doi: 10.26634/jee.9.3.4883

[19] S. Suneet et al., "A low cost power quality management tool for different load types,"IEEE 6th International Conference on Power Systems, Electronic 2016, doi: 10.1109/ICPES.2016.7584045.

[20] S. Deeksha et al., "Basic concepts of superconducting fault current limiter," IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems, ISBN:978-1-4673-8587-9,2016, doi: 10.1109/ICPEICES.2016.7853069.

[21] S. Bhandari et al., "Optimization Techniques In Modern Times & Their Applications", EISBN: 978-1-5386-2317-6, 6th IEEE iEECON2018 Krabi Thailand, 2018, doi: 10.1109/IEECON.2018.8712308.

[22] S.K. Bhushan et al., "Implementation of MPPT Technique in PV Array for a Varying Load by Modeling and Simulation,"International Electrical Engineering Congress (iEECON), 2018, doi: 10.1109/IEECON.2018.8712313.

[23] P. Shubham, S.K. Bhushan, "Modeling of Solar Energy Grid Integration System Using Typhoon HIL," International Electrical Engineering Congress (iEECON), 2018, doi: 10.1109/IEECON.2018.8712253 .

[24] S.K. Bhushan, M. Goel. "Grid Integrated Solar Photovoltaic Array Power Plant Modeling and Simulation," International Conference and Utility Exhibition on Green Energy for Sustainable Development, EISBN:978-9-7482-5799-0, 2018, doi: 10.23919/ICUE-GESD.2018.8635688.

[25] S.K. Bhushan et al., "Modeling and Simulation of Grid Connected 1 MW Solar PV Power Plant." 2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE). IEEE, 2018, doi: 10.23919/ICUE-GESD.2018.8635745.

[26] R. Vivekanand, K. Sahay. "Power Quality Improvement of Grid Connected PV System Using Dynamic Voltage Restorer," i-Manager's Journal on Power Systems Engineering, **4**(1), 5-13, 2016, doi: 10.26634/jps.4.1.5966.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 981-987 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

NPC five level inverter using SVPWM for Grid-Connected Hybrid Wind-Photovoltaic Generation System

Elamri Oumaymah[*,1], Oukassi Abdellah[1], Bouhali Omar[2], El Bahir Lhoussain[1]

[1]*Systems and Applications Engineering Laboratory (LISA), National School of Applied Sciences, Cadi Ayyad University, Marrakech, 40000, Morocco*

[2]*Mechatronic Laboratory, Jijel University, Jijel, 18000, Algeria*

ARTICLE INFO

Article history:
Received: 16 September, 2020
Accepted: 26 November, 2020
Online: 08 December, 2020

Keywords:
Wind power generation
Solar power
PMSG
Five-level NPC inverter
SVPWM
Grid injection

ABSTRACT

This paper covers a new topology, a synchronous wind turbine generator, and a solar photovoltaic generator. The Permanent Magnet Synchronous Generator is linked to the grid by back-to-back voltage source converters (BtB VSC), consisting of a two-stage rectifier and a five-stage Neutral Point Clamped inverter, operated by the Space Vector Pulse Width Modulation technique. A solar photovoltaic system has a direct interface with the capacitor of the DC link of the BtB VSCs, with no additional DC/DC conversion stage, while the efficiency of the system is maximized. This work features a Perturb and Observe algorithm Maximum Power Point Tracking for to extract the optimum power from both wind turbine and solar PV generator. The application of Space Vector Pulse Width Modulation control scheme is used for harmonics compensation THD. A proportional integral regulator is used to regulate the DC link voltage. A prototype is tested under various conditions, wind velocity variations as well as under several photovoltaic solar isolations. The present work has been treated using Matlab/Simulink. Simulation results proved efficiency provided by this controlling approach: reduced distortion of harmonic and increased power factor.

1 Introduction

Today, the major energy sources used for generating electricity are fossil fuel resources [1]. To reduce the dependency on fossils based energy it necessary for power generation to move towards the renewable resources including solar energy, wind power, hydroelectricity, and biomass. In literature, most of the distributed production systems are uniquely focused on one type of sustainable resource, e.g., a wind power like in [2] or solar power like in [3].

The combination and the penetration of the renewable resources, face considerable challenges due to growing power industry. The injection into the grid of the energy yielded by wind and solar photovoltaic generators, is proven to be one of the technological solution [4, 5]. Their combination offers additional benefits such as increasing the overall operational efficiency and power reliability, optimal use of lands resources, and improves the capital investments [5]. Grid-connected wind/Photovoltaic (PV) cogeneration is not largely discussed [6]–[7], whereas in literature, most of wind-PV hybrid systems are proposed for standalone off-grid applications [5]. The optimization method of generation of power using hybrid system used for battery charging is developed in [8], to increase the stability and efficiency. In [9], the author have proposed a small scale hybrid solar- wind power generation system to reduce Electricity shortage in Egypt.

Much research was conducted to track maximum power point (MPP) of the wind using methods based on Perturb and Observe (P&O) algorithm [6], neural network technique, or other approaches based on fuzzy logic. Due to the nonlinear voltage-current properties of PV solar panels, the necessity of MPP tracking has become apparent. MPP tracking schemescan be accounted for as perturb and observe (P&O), incremental conductance method, Fractional Open-Circuit Voltage, Neural Network, Fuzzy Logic Control, Current Sweep [10]. The combination of multilevel inverter and renewable energy sources are advantageous. In [11], the authors have proposed a hybrid cascaded nine level inverters for photovoltaic applications. The proposed system in [12] are based on Maximum Power Point Tracking (MPPT) combined with multilevel inverter to achieve the maximum power. In [6], the system comprises back-to-back voltage source converters with no extra dc/dc conversion stages to connect wind and photovoltaic generators to the power grid. Moreover, the control of the currents on the machine- and grid- side are carried out using hysteresis regulators, introducing higher harmonic contents

[*]Corresponding Author: Elamri Oumaymah, Email: oumaymahelamri@gmail.com

from the variable switching frequency. The objective in [13] is to realize a photovoltaic system with a MPPT linked to the 11 KV grid by employing a new control technique in order that it is possible to transfer maximum active power from the PV to the grid without injecting harmonics. The study proposes in [14, 15] a control strategy for a single-phase, three-phase photovoltaic system connected to a distribution network. The proposed technology is based on the modification of the conventional control in current mode to adjust the power factor of the photovoltaic system and control the voltage of the converter link.

This paper extends the work originally presented in ICEIT'4 [2], where the modeling and simulation of a wind turbine string, based on a vector-controlled five-level inverter, was presented. Its objective was the optimization of the energy produced, adjust the DC link and control power transferred into the grid. Motivated by advantages of hybrid production systems, presented paper covers a new topology to interface both is Permanent Magnet Synchronous Generator (PMSG)-based wind energy and PV generators with power grid. Note a blocking diode is series connected with each PV string to avoid reverse flow of current. Two P&O algorithms, namely P&O Wind and P&O PV, are used to optimize the energy extracted by turbine and solar system, respectively. Note that perturbations could impact correct operation of the receivers attached to the grid if the voltage output wave is not sinusoidal. To avoid this problem, we have implemented in this chain a two-stage rectifier followed by a five-stage Neutral Point Clamped (NPC) inverter regulated by the Space Vector Pulse Width Modulation (SVPWM) technique, without additional DC/DC conversion stages[16]. The converter control strategy is detailed in which the presented study was treated using Matlab/Simulink. Simulation results proved efficiency provided by this controlling approach: decrease distortion of harmonic and increased power factor.

2 Wind Photovoltaic Generator Side Model

2.1 Wind turbine model

Mechanical energy absorbed by a wind turbine is defined by [2]:

$$P_m = \frac{1}{2}.\rho.C_p(\lambda,\beta).V_v^3.R^2.\pi \tag{1}$$

where: C_p is the coefficient of the rotor that is a non-linear function of the tip speed ratio λ and the pitch angle of blades β, R and V_v are the radius of blades and the wind speed, respectively.

2.2 PMSG model

PMSG in the dq marker is modelised using these equations [2]:

$$V_d = -R_s.I_d - \frac{d\varphi_d}{dt} + \omega.(L_{sq}.I_q + \varphi_f) \tag{2}$$

$$Vq = -R_s.I_q - \frac{d\varphi_q}{dt} - \omega.(L_{sd}.I_d + \varphi_f) \tag{3}$$

The mechanical equation is as shown below:

$$J.\frac{d\Omega}{dt} + f_c.\Omega = C_r - C_m \tag{4}$$

Where: R_s and L_s are stator-winding resistance and inductance, respectively, I_d and I_q are the stator currents, Ω is Rotational speed of the PMSG, φ is stator phase magnetic flux, f_c and J are the friction coefficient and motor inertin, respectively. φ_f: the flux generated by the permanent magnets across stator windings, C_m is the mechanical torque .

2.3 Photovoltaic Generator

The model of the PV generator is presented in Appendix whereas it is modeled as following:

$$V_{pv} = -R_{dc}.I_{pv} - \frac{d\,I_{pv}}{dt}.L_{dc} + Vdc \tag{5}$$

The I_{op}-V_{op} characteristics of the PV array, there is an optimal operating point (I_{op}-V_{op}) corresponding to the maximum generated PV power, at any solar irradiance level.

3 Control Strategy

The suggested wind-PV system appliances two appropriate control techniques for machine-side Voltage Source Rectifier Control (VSRC) and grid-side Voltage Source Inverter Control (VSIC).The proposed system operates efficiently despite solar insolation changes and wind speed variability. The next section deals with control strategies.

3.1 Voltage Source Rectifier Control (VSRC)

Figure 1 visualizes the VSRC schema. Momentary torque is regulated through the application of the vector control schema. Vector Control technology treats the technique of control of AC machines. This control is based on eliminate the coupling between the inductor and the induced by dissociating the static current into two components (i_{sd},i_{sq}). So that, the reference direct axis current (i_{sd}) is controlled relative to desired flux, while reference quadrature axis current (i_{sq}) is adjusted according to the desired torque. Which allows for a function comparable to a dc machine, with separate excitement [4].

The Voltage Source Rectifier (VSR) equivalent circuit is presented in Figure 1. The control of the VSR is split over two parts. The first is the direct and quadrature reference current generation (i_{sdref} and i_{sqref}) and generation of switching signals for VSR.

3.1.1 Generation of i_{sdref} and i_{sqref}

In proposed system, the electromagnetic torque is controlled by controlling the current i_{sqref} while the current i_{sdref} is set to zero.

For each wind speed, the speed of the rotor has only one value for maximum wind power. Using MPPT P&O algorithm, the reference speed Ω_{ref} is generated.

Speed is regulated by the proportional integral (PI) speed controller, as illustrated in Figure 1, and its output is considered as i_{sqref}.

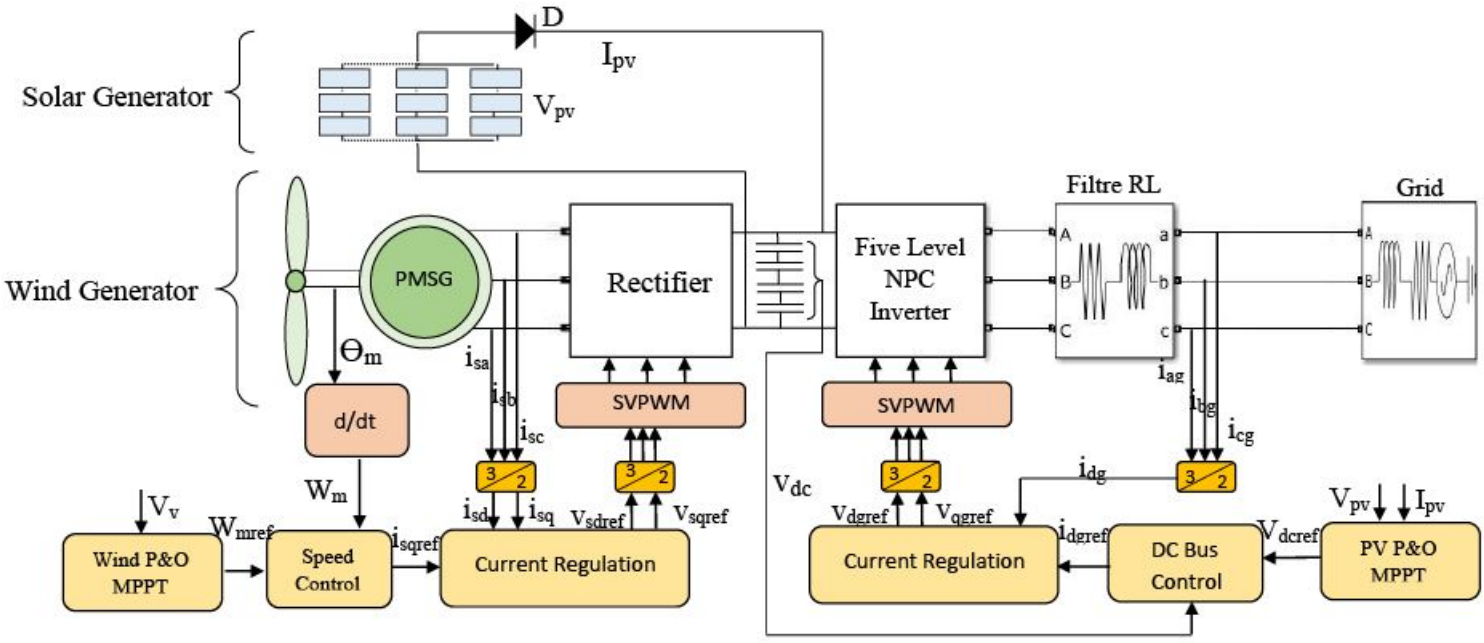


Figure 1: The proposed Wind-PV System Configuration

$$\begin{cases} i_{sdref} = 0 \\ i_{sqref} = \frac{2.Te_{ref}}{3.p.\varphi} \\ Te_{ref} = K_p.(\Omega_{ref} - \Omega) + K_i.\int(\Omega_{ref} - \Omega)\,dt \end{cases} \quad (6)$$

With: K_p and K_i are the proportion and integral coefficients of PI speed controller.

3.1.2 Generation of switching signals for VSR

The PI current controller is employed to ensure stator currents generated by PMSG correspond to references in (6), as shown in Figure 1 and (7).

$$\begin{cases} V_{sdref} = K_p.(i_{sdref} - i_{sd}) + K_i.\int(i_{sdref} - i_{sd})\,dt - \\ \Omega.L_s.i_{sq} \\ V_{sqref} = K_p.(i_{sqref} - i_{sq}) + K_i.\int(i_{sqref} - i_{sq})\,dt + \\ \Omega.L_s.i_{sd} + \Omega.\varphi \end{cases} \quad (7)$$

With: K_p and K_i are proportion and integral coefficients of PI current controller.

Three phase reference stator voltages (V_{saref}, V_{sbref}, V_{scref}) are generated from V_{sdref} and V_{sqref} using inverse Parks transform. They fed the SVPWM controller.

Voltage reference is determined using:

$$V_{ref} = \frac{2}{3}(V_{saref} + e^{-j2\frac{\pi}{3}}V_{sbref} + e^{-j4\frac{\pi}{3}}V_{scref}) \quad (8)$$

The rectifier delivers eight states, consisting of two zero vectors and six active vectors. The complex diagram is split by the active vectors in six sectors with 60 degrees forming each hexagon, with voltage reference vector corresponding to a pair of two active adjacent vectors (T_a, T_b) and a zero vector (T_0) of commutation time T_s [17].

Commutation times are determined by the equations below:

$$\begin{aligned} T_a &= T_s.m.sin(\frac{\pi}{3} - \theta) \\ T_b &= T_s.m.sin(\theta) \\ T_0 &= 1 - T_a - T_b \end{aligned} \quad (9)$$

$$\text{With: } m = \frac{(\sqrt{3})}{V_{dc}}|V_{ref}| \text{ modulation index} \quad (10)$$

3.2 Voltage Source Inverter Control (VSIC)

Figure 9 shows the Voltage Source Inverter (VSI) commutation configuration, uses the SVPWM strategy. VSI control is structured in two sections. The first section is the active power component i_{dgref} generation, the second is the VSI switching signal generation.

3.2.1 Generation of i_{dgref} and i_{qgref}

VSI system is regulated by using vector control scheme in which a phase locked loop (PLL) is applied to converter to synchronize it with power grid. Active and reactive power regulation are expressed by:

$$\begin{cases} P = 1.5(V_{dg}.i_{dg} + V_{qg}.i_{qg}) \\ Q = 1.5(V_{qg}.i_{dg} - V_{dg}.i_{qg}) \end{cases} \quad (11)$$

This control has the advantage of decoupling active and reactive control, whereas to regulate DC bus voltage V_{dc}, current i_{dg} is used and to regulate reactive power transited Q, current in quadrature i_{qg} is used. .

$$\begin{cases} P = 1.5V_{dg}.i_{dg} \\ Q = -1.5V_{dg}.i_{qg} \end{cases} \quad (12)$$

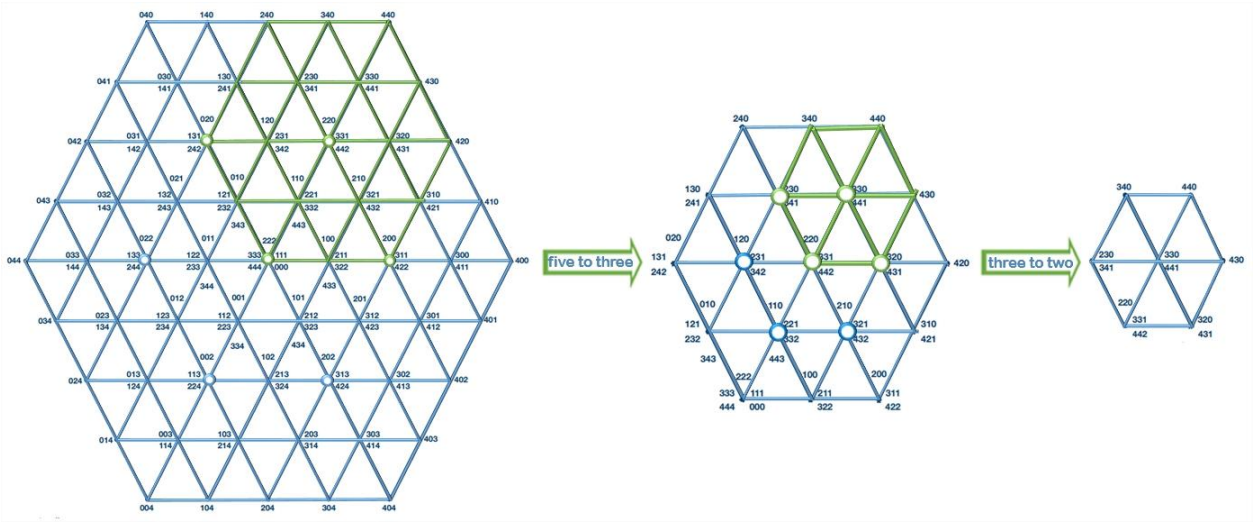


Figure 2: Decomposition of the five-level inverter vector diagram

We can therefore, through the imposition of reference currents, provide referencing for active power P_{ref} and reactive power Q_{ref} fixed in zero that maintains a unity power factor (PF), we have:

$$\begin{cases} i_{dgref} = \frac{P_{ref}.V_{dg} - Q_{ref}.V_{qg}}{V_{dg}^2 + V_{qg}^2} \\ i_{qgref} = \frac{P_{ref}.V_{qg} + Q_{ref}.V_{dg}}{V_{dg}^2 + V_{qg}^2} \end{cases} \quad (13)$$

The DC link reference (V_{dcref}) is obtained by the PV P&O MPPT algorithm as shown in figure 1, corresponding to its maximum PV power output voltage at various irradiation levels. The regulation of DC bus voltage is performed by PI regulator and its output is considered as the active power component i_{dgref}. The i_{dgref} is given as,

$$i_{dgref} = K_p.(V_{dcref} - V_{dc}) + K_i.\int (V_{dcref} - V_{dc})\,\mathrm{d}t \quad (14)$$

With: K_p and K_i are the proportion and integral coefficients of DC-bus voltage controller PI, respectively.

3.2.2 *Generation of switching signals for VSI*

The i_{dgref} and i_{qgref} are compared with i_{dg} and i_{qg}, respectivly, to obtain the currents error and are given as,

$$\begin{cases} i_{dge} = i_{dgref} - i_{dg} \\ i_{qge} = i_{qgref} - i_{qg} \end{cases} \quad (15)$$

The currents error are fed to PI controllers and the output of decoupled controller is given as:

$$\begin{cases} V_{dgref} = -K_i.\int i_{dge}\,\mathrm{d}t - K_p.i_{dge} + \\ \Omega_g.L_f.i_{qg} + v_{dg} \\ V_{qgref} = -K_i.\int i_{qge}\,\mathrm{d}t - K_p.i_{qge} + \\ \Omega_g.L_f.i_{dg} + v_{qg} \end{cases} \quad (16)$$

The three phase reference grid voltages (V_{agref}, V_{bgref}, V_{cgref}) are generated from V_{dgref} and V_{qgref} using inverse Parks transform.

This approach is founded on geometrically splitting of a vector diagram to several hexagons that represent low-level vector diagrams, as shown in figure 2 [18].

To attain this objective, two actions need to be completed:
- Depending of voltage reference vector position V_{ref}, one of six referenced hexagons is chosen;
- V_{ref} is moving to center of the hexagon selected in the first action.
Following these actions two times would reduce the five-level inverter's vector diagram to the two-level inverter's vector diagram. Next, we will employ the classical method of vector modulation of the two-level inverter, discussed in the section 3.1.2.

4 Simulation results

For the hybrid system, a simulation model shown at figure 1 has been modulated with Matlab/Simulink® in order to validate and evaluate system performance. Simulation parameters are as shown in the appendix A. Penetration of wind and PV energy is studied under different weather conditions.As shown in Figure 3 and 4, the wind speed increases from 20 to 8, and finally drops to 5 m/s, at t= 0.3 and 0.7s, respectively. Solar irradiance level increases from 400 to 1000, than drops at 500, and increases to 300 W/m, at t= 0.3, 0.5, and 0.7s, respectively.

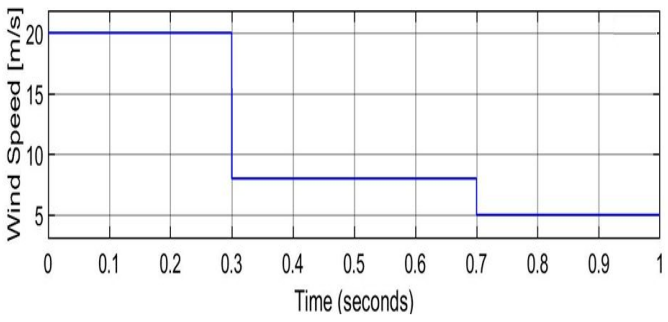


Figure 3: Cp(λ,β)curves typical for different inclination angles

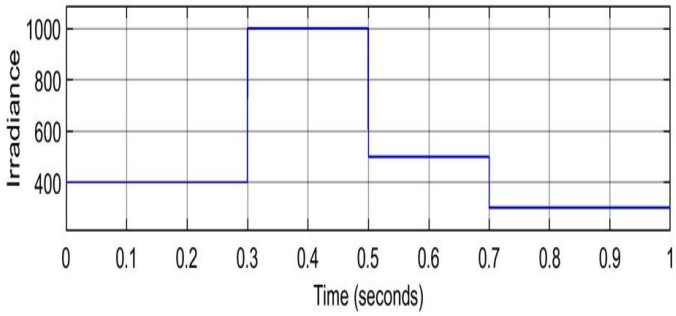


Figure 4: I_{pv} V_{pv} photovoltaic characteristics at different irradiation levels

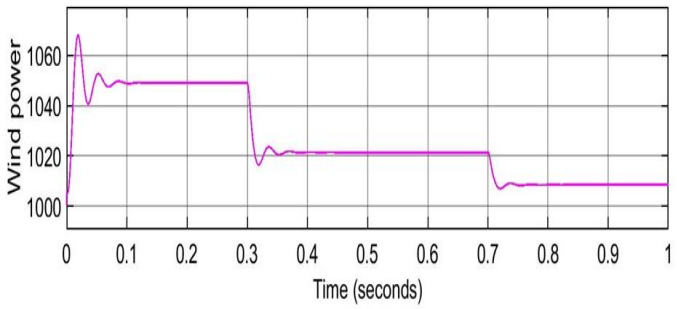


Figure 7: wind power

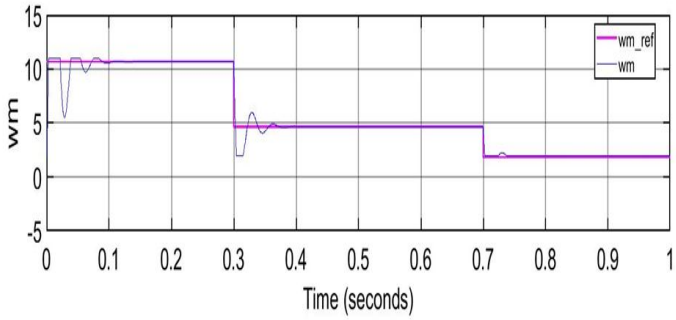


Figure 5: Generator rotation speed

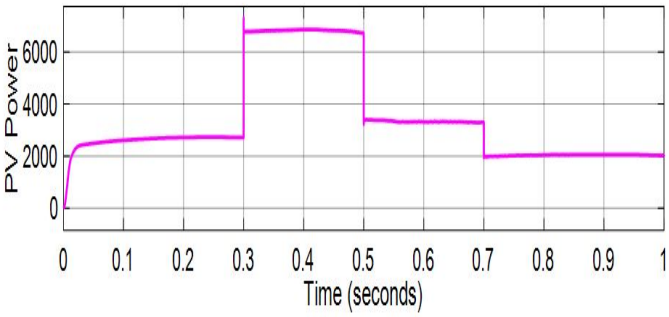


Figure 8: PV power

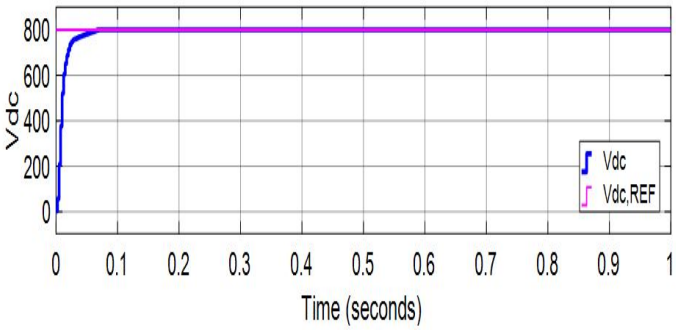


Figure 6: DC-link voltage

Optimal Ω_{ref} and V_{dcref} are generated by Wind and PV MPPT P&O, respectively. As depicted in Figure 5 and 6, the corresponding variables are then regulated.

Under these conditions, Ω_{ref} and V_{dcref} are both refracted on the corresponding PV and wind powers, as shown at figures 7 and 8.

Three-phase grid voltages before filtering are shown in figure 9, representing five-level inverter output voltage waveform obtained from capacitors. Economically, it is not interesting to use a rectifier for each capacitor, so it is necessary to ensure the balance of the continuous bus, as depicted in Figure 10 .

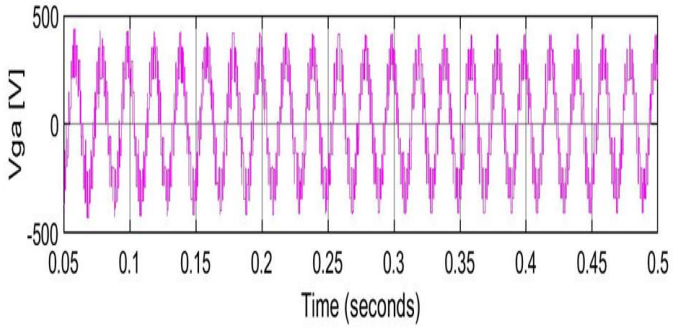


Figure 9: Output voltages of the three-phase inverter at five levels

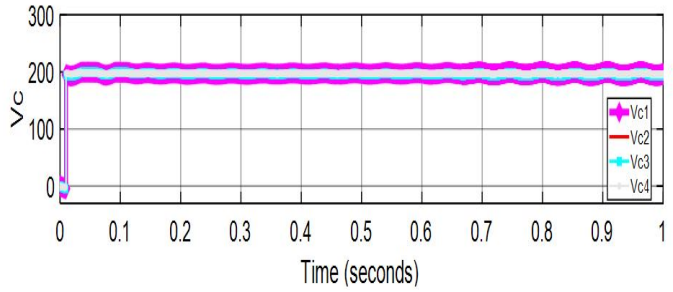


Figure 10: Capacitor voltage

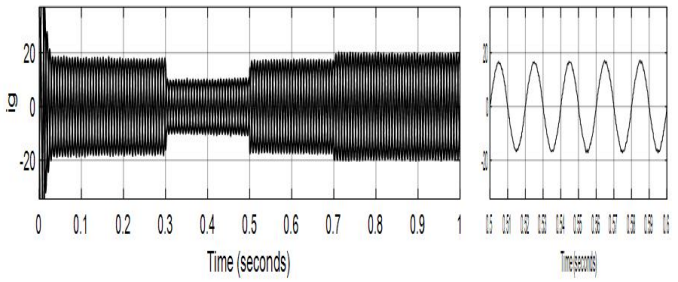


Figure 11: Grid current

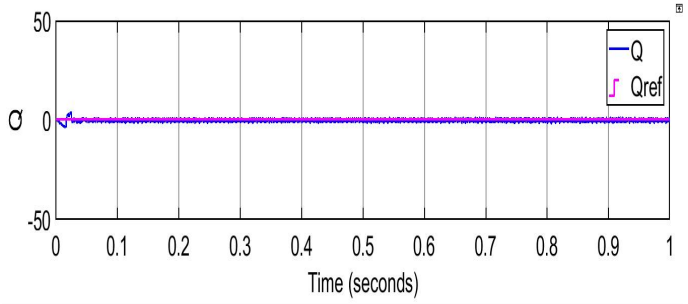


Figure 12: Reactive power

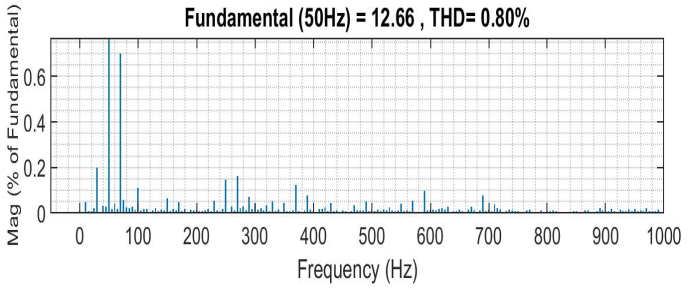


Figure 13: THD of grid currents injected

The AC current injected into grid is indicated in Figure 11. During entire operating range, the reactive power is set to zero, then a unity PF is maintained, as shown in in Figure 12.

The figure 13 illustrates the THD currents injected in the grid. According to THD analysis above, the total harmonic distortion at Fc=1KHz given THD=0.80%. Therefore, as the number of levels is increased, THDs are minimized in comparison with classical multi-level inverters [19].

5 Conclusion

In this paper, the modeling and simulation of wind-power hybrid systems connected to the grid, using vector control systems were presented.A hybrid system based on wind and photovoltaic energy was set up with a five-level inverter. The VSRC's purpose consists of extracting the maximum wind power, and the VSIC is responsible for extracting maximum PV power, and balance of dc-link capacitor in order to transfer produced power to utility-grid. This way, reactive power is fixed at zero, then a unity Power Factor PF is maintained. The proposed system has enhanced efficiency and reliability via a hybrid renewable energy system, the maximum power is extracted from independent MPPT using P&O algorithm, and reduces current THD. The results of the simulations showed the correspondence of the quantities assessed with the reference quantities and allowed the validation of the mathematical models.

References

[1] A. Demirbas, "Global Renewable Energy Projections," Energy Sources, Part B: Economics, Planning, and Policy, **4**(2), 212–224, 2009, doi:10.1080/15567240701620499.

[2] E. Oumaymah, O. Abdellah, E. A. Mustapha, "The Injection of Wind Power into a Grid Using a Multi-Level Inverter Controlled by SVPWM, booktitle = 2020 International Conference on Electrical and Information Technologies (ICEIT)," 1–6, IEEE, Rabat, Morocco, 2020, doi:10.1109/ICEIT48248.2020.9113201.

[3] L. Wang, D. Zhang, J. Duan, J. Li, "Design and Research of High Voltage Power Conversion System for Space Solar Power Station," in 2018 IEEE International Power Electronics and Application Conference and Exposition (PEAC), 1–5, IEEE, Shenzhen, 2018, doi:10.1109/PEAC.2018.8590543.

[4] F. Chishti, S. Murshid, B. Singh, "LMMN-Based Adaptive Control for Power Quality Improvement of Grid Intertie Wind–PV System," IEEE Transactions on Industrial Informatics, **15**(9), 4900–4912, 2019, doi:10.1109/TII.2019.2897165.

[5] F. Chishti, S. Murshid, B. Singh, "An Intelligent Control for Synchronous Generator Based Wind-PV-Battery Microgrid System," in 2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), 1–6, IEEE, Chennai, India, 2018, doi:10.1109/PEDES.2018.8707685.

[6] A. A. A. Radwan, Y. A.-R. I. Mohamed, "Grid-Connected Wind-Solar Cogeneration Using Back-to-Back Voltage-Source Converters," IEEE Transactions on Sustainable Energy, **11**(1), 315–325, 2020, doi:10.1109/TSTE.2019.2890828.

[7] S. Z. Hassan, H. Li, T. Kamal, M. Awais, "Stand-alone/grid-tied wind power system with battery/supercapacitor hybrid energy storage," in 2015 International Conference on Emerging Technologies (ICET), 1–6, IEEE, 2015.

[8] R. S. S. Singh, M. Abbod, W. Balachandran, "A Design Scheme of Control/Optimization System for Hybrid Solar — Wind and Battery Energy Storages System," in 2016 51st International Universities Power Engineering Conference (UPEC), 1–6, IEEE, Coimbra, 2016, doi:10.1109/UPEC.2016.8114093.

[9] I. Elsayed, I. Nassar, F. Mostafa, "Optimization and Economic Evaluation of Small Scale Hybrid Solar/Wind Power for Remote Areas in Egypt," in 2017 Nineteenth International Middle East Power Systems Conference (MEPCON), 25–30, IEEE, Cairo, 2017, doi:10.1109/MEPCON.2017.8301158.

[10] K. Saidi, M. Maamoun, M. Bounekhla, "Simulation and analysis of variable step size P&O MPPT algorithm for photovoltaic power control," in 2017 International Conference on Green Energy Conversion Systems (GECS), 1–4, IEEE, Hammamet, Tunisia, 2017, doi:10.1109/GECS.2017.8066265.

[11] N. Kalaiarasi, S. S. Dash, S. Paramasivam, R. Bayindir, "Hybrid Cascaded Nine Level Inverter Using Dspace Controller for Standalone Photovoltaic Applications," in 2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA), 745–750, IEEE, San Diego, CA, 2017, doi:10.1109/ICRERA.2017.8191159.

[12] B. S. N. Raj, R. Siddharth, S. M. Shyni, "MPPT with Bi-Directional DC-DC Converter and Multi-Level Inverter for Grid Connected Hybrid System," in 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC), 773–777, IEEE, Melmaruvathur, 2017, doi:10.1109/ICCPEIC.2017.8290467.

[13] P. K. Hota, B. Panda, B. Panda, "Fault Analysis of Grid Connected Photovoltaic System," American Journal of Electrical Power and Energy Systems, **5**(4), 35, 2016, doi:10.11648/j.epes.20160504.12.

[14] A. Yazdani, P. Dash, "A Control Methodology and Characterization of Dynamics for a Photovoltaic (PV) System Interfaced With a Distribution Network," IEEE Transactions on Power Delivery, **24**(3), 1538–1551, 2009, doi: 10.1109/TPWRD.2009.2016632.

[15] P. P. Dash, A. Yazdani, "A Mathematical Model and Performance Evaluation for a Single-Stage Grid-Connected Photovoltaic (PV) System," International Journal of Emerging Electric Power Systems, **9**(6), 2008, doi: 10.2202/1553-779X.2033.

[16] A. Ghosh, F. Shahnia, "Applications of Power Electronic Devices in Distribution Systems," in J. A. Martinez-Velasco, editor, Transient Analysis of Power Systems, 248–279, John Wiley & Sons, Ltd, Chichester, UK, 2014, doi:10.1002/9781118694190.ch7.

[17] M. M. Irfan, P. H. K. Prasad, P. V. Rao, "Simulation of five-level five-phase SVPWM voltage source inverter," in 2010 International Conference on Power, Control and Embedded Systems, 1–5, IEEE, Allahabad, 2010, doi: 10.1109/ICPCES.2010.5698659.

[18] K. Himour, K. Ghedamsi, E. M. Berkouk, "New alternative of design and control of grid connected PV-Storage systems with the five level diode clamped inverter," in 2013 Conférence Internationale des Energies Renouvelables (CIER'13), volume 2, 10, 2013.

[19] Yaosuo Xue, Liuchen Chang, "Closed-loop SPWM control for grid-connected buck-boost inverters," in 2004 IEEE 35th Annual Power Electronics Specialists Conference (IEEE Cat. No.04CH37551), 3366–3371, IEEE, Aachen, Germany, 2004, doi:10.1109/PESC.2004.1355070.

A Appendix

The system Parameters: Wind turbine: $p = 122\ Kg/\ m^3$, $R = 412\ m$, $J = 0.0002\ Kg.m$; PMSG: P= 2 kW, $R_s = 12\ m$, $L_{sq} = 0.0125\ mH$, $L_{sd} = 0.0057\ mH$,P = 4; Solar PV :$N_s = 40$, $N_p = 2$, $I_{phc} = 3.45$ A, $G_{ref} = 1000\ W/m^2$, Open circuit voltage = 20.1 V, Saturation diode current = 4.05e-7 A, Reference temperature = 24 $°C$; DC link: $U_c = 50\ V$, $C = 2200\ pF$; Grid connection: $L_f = 1\ mH$, $R_f = 0.1\ \Omega$; Grid: U= 338 V and f =50 Hz.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 363-369 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Vietnamese Text Classification with TextRank and Jaccard Similarity Coefficient

Hao Tuan Huynh[1], Nghia Duong-Trung[2], Dinh Quoc Truong[1], Hiep Xuan Huynh[*1]

[1]*College of Information & Communication Technology, Can Tho University, Can Tho city, 94000, Vietnam*

[2]*Software Engineering Department, FPT University, Can Tho city, 94000, Vietnam*

ARTICLE INFO

Article history:
Received: 03 September, 2020
Accepted: 28 October, 2020
Online: 20 November, 2020

Keywords:
TextRank
Vietnamese
Jaccard Similarity
Representative Keywords
Text Classification

ABSTRACT

Text classification is considered one of the most fundamental and essential problems that deal with automatically classifying textual resources into pre-defined categories. Numerous algorithms, datasets, and evaluation measurements have been proposed to address the task. Within the era of information redundancy, it is challenging and time-consuming to engineering a sizable amount of data in multi-languages manually. However, it is time-consuming to consider all words in a text, but rather several key tokens. In this work, the authors proposed an effective method to classify Vietnamese texts leveraging the TextRank algorithm and Jaccard similarity coefficient. TextRank ranks words and sentences according to their contribution value and extracts the most representative keywords. First, we collected textual sources from a wide range of Vietnamese news websites. We then applied data preprocessing, extracted keywords by TextRank algorithm, measured similarity score by Jaccard distance and predicted categories. The authors have conducted numerous experiments, and the proposed method has achieved an accuracy of 90.07% on real-world datasets. We have proved that it is entirely applicable in practice.

1 Introduction

This paper is an extension of work initially presented in IEEE RIVF 2020 [1] as an invited paper.

Within the era of information redundancy, it is challenging and time-consuming to manually engineer a sizable amount of multilingual data. For example, an electronic library might quickly identify documents for archiving and management. Besides, it is often less accurate and overwhelming if humans are involved in the process. Therefore, apply machine approaches to automate the procedure of text classification is mandatory. It makes the classified results more reliable and less subjective. It helps alleviate human involvement and information overload and enhances knowledge retrieval efficiency. There are many text classification studies such as, e.g. Bayes [2], decision trees [3], K-nearest neighbor [4], and neural network [5]. The literature shows that automated text classification is one mainstream research in natural language processing [6, 7]. Many research papers have been conducted to solve such problems as email-messages filtering [8], topic modeling [9], geo-localization [10], and document categorization [11]. The flowchart of standard text classification is presented in Figure 1.

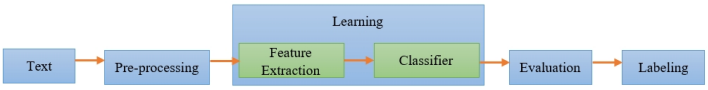


Figure 1: The flowchart of regular text classification where pre-processing and featuring perform an essential role.

The curse of dimensionality affects accuracy and computing time that arise when analyzing and organizing textual data in high-dimensional spaces [9]. However, machine learning models do not need to learn all tokens in the texts to categorize them. Instead, the text's label can be identified through its necessary tokens, which contribute the most to the text's meaning. Consequently, if we can extract the text source's main keywords, we might accurately classify it into an appropriate topic. The TextRank algorithm [12, 13] allows extracting the list of representative keywords of textual contents. On that consideration, the authors propose the method of automatic Vietnamese text classification based on the representative keyword analysis of the text. Textual datasets have been downloaded from several news websites with 15 main topics. Then, keyword sets that represent each topic are built. The system will extract that text's specific keywords when it is necessary to identify the topic

[*]Corresponding Author: Hiep Xuan Huynh, College of Information & Communication Technology, Can Tho University, Can Tho city, 94000, Vietnam, Email: hxhiep@ctu.edu.vn

of a text. The class of the text will be determined by computing the coverage of the representative keywords. It would be the right way for text classification applications in areas such as electronic library management. Although many Vietnamese documents are electronically available, no previous research has applied keywords-based text classification to them.

This work portrays our endeavor to construct a practical framework for addressing Vietnamese texts' classification task on the news. To the best of our knowledge, in this extended article, the authors have made several contributions. First, we discuss the paper [1] in a lot more detail by extending the related works and technical background. Second, we extend the experiments by investigating more scenarios.

2 Related Work

An early effort to address the task of Vietnamese text classification was conducted more than a decade ago [14]. In that paper, the authors solved automatically categorizing the problem, given textual sources into predefined categories. A comparison between statistical N-Gram language modeling and bag of words approaches has been investigated on their collected dataset. Several researchers have applied the idea of spam filtering into Vietnamese text sources [15]. Short messages such as conversational texts have also been exploited by addressing the task of suggestion intents [16]. The authors proposed a user suggestion intent definition in general from conversational texts at a functional segment unit. The task of automatic text categorization has been studied by comparing the performance of several term weighting schemes rather than analyzing the actual classification task [17]. Regarding Vietnamese sources, full-text representation has been exploited by many other research papers [18, 19, 20]. Thoughtfully learning the literature, we could claim that this is the first attempt to featuring Vietnamese texts using the idea of representative keywords.

3 Proposed System

3.1 Design Concept

The Vietnamese text classification system proposed by the authors includes two main components. One is the keyword extraction module, and the other is the comparison with the training set to identify the topic of a new document. We present the overall design of the system in Figure (**??**). The process begins with the text sources featured in the representative keyword vectors. For the training set, the vectors will be marked with the topic label. For the test data, the keyword extraction module is applied to convert the original text data into its keyword-based representation. Next, the system compares it to training data by calculating similarity scores. Finally, the prediction is assigned to the test data.

3.2 Text Pre-processing

Vietnamese is the only language in which every syllable is pronounced separately and is represented by a written word. This feature is evident in all aspects of phonetics, vocabulary and grammar. Data pre-processing is the first important step of any data mining process. It makes data in its original form easier to observe and explore. For the problem of text classification, due to specific characteristics, each language has its own challenges. The preprocessing process will help improve sorting efficiency and reduce the complexity of the training algorithm. Depending on the purpose of the classifier, we will have different preprocessing methods, such as

- Convert text to lowercase and correct spelling errors.
- Remove punctuation marks (if no sentence separation is performed).
- Remove special characters ([], [.], [,], [:], [”], [”], [;], [/], [[]], [$\tilde{]}$, [$\acute{]}$, [!], [@], [#], [$], [%], [$\hat{]}$, [&], [*], [(], [)]).
- Separate tokens by compound words (Vietnamese).
- Remove the stopwords, e.g. the words that appear most in the text that are not meaningful when participating in text classification. We utilize a list of 1942 Vietnamese stopwords[1] in our data processing.
- For the tokenization step, we utilize vnTokenizer [21] in our research. The comparison of tokenization accuracy achievable with different software is beyond the scope of this research paper.

3.3 Vietnamese Text Tokenization

Phonetic characteristics In Vietnamese, there is a special type of unit called ”tieng” or a sound of the thing. Phonetically, each ”tieng” is a syllable. For example, the word ”student” is translated into two syllables ”sinh vien” which are two separate words. As a result, these two words should come together to form a meaning token.

Vocabulary characteristics Each ”tieng”, in general, is a meaningful element. Continuing the previous example, the words ”sinh” and ”vien” have their own meaning when coming alone. But when they come together to form a single word ”sinh vien”, it has the meaning of student as in English. The vocabulary of Vietnamese is based on single words (one syllable) and the countless combination of them. Creating new words is very easy and flexible. If we pronounce a sound, we could write it down as a word.

Grammatical characteristics The Vietnamese words do not change morphology. For example, verbs in Vietnamese do not have -ed, -s, -ing forms. This feature will dominate other grammatical characteristics. When words combine with words into sentences, it is important to know the word order, the word phrase, and keywords for tenses recognition. Sorting words in a certain order is one of the main processes to express syntax relations.

[1] https://github.com/stopwords/vietnamese-stopwords/blob/master/vietnamese-stopwords.txt

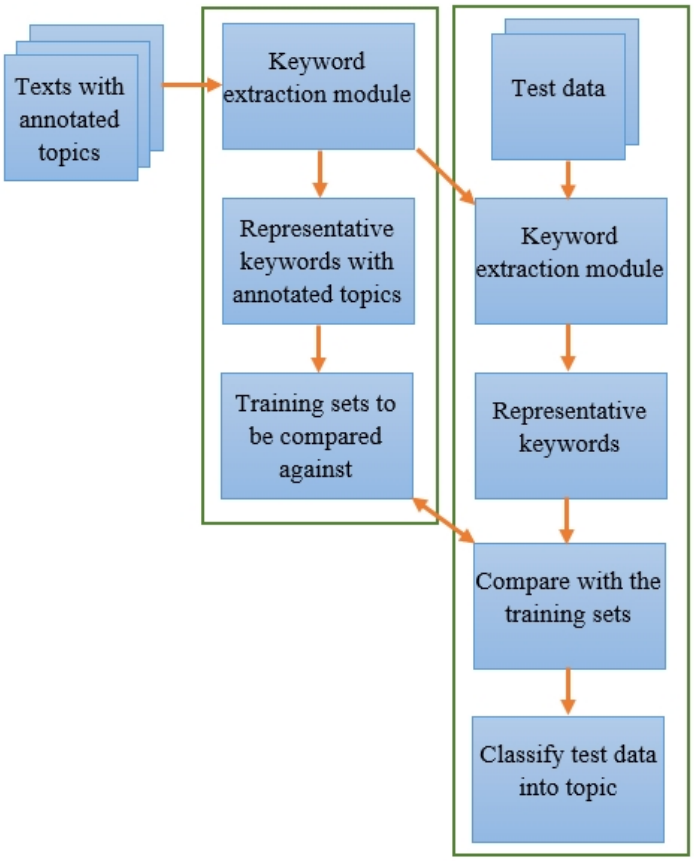


Figure 2: Our proposed Vietnamese text classification procedure.

3.4 Keyword Extraction by TextRank

Algorithm TextRank [22] was developed based on the main idea of the PageRank algorithm that Google search engine uses to rank website cite zhou2019chip, langville2008google, berkhout2016google. The bottom line of the TextRank algorithm is to use graphs to represent text and to score important information about the structure in which the text is represented by keywords. In other words, the TextRank algorithm processes a group of keywords representing the entire text. TextRank ranks words by their importance, arranges them in descending order of computed value, and extracts the most important words. The number of important words is a hyperparameter which is determined by the user prior to the TextRank algorithm execution. This algorithm is successfully applied for keyword extraction based on key value from a single text and this is also the advantage of TextRank.

The TextRank algorithm represents the textual source as a graph $G = (V, E)$ where V is the set of vertices and E is the set of edges in the graph. E is derived from a subset of V x V. Each vertex of the graph G corresponds to one word extracted from the text. An edge between any two vertices is created when their words appear in the text at any position between 2 and N. The value for the importance of the vertex V_i is calculated using the following formula:

$$S(V_i) = (1-d) + d * \sum_{j \in \text{In}(V_i)} \frac{1}{|\text{Out}(V_j)|} S(V_j)\,, \tag{1}$$

where $d = [0, 1]$. In our experiments, d is set to 0.85 by default [22]. $\text{In}(V_i)$ is the collection of vertices that point to it, and $\text{Out}(V_i)$ is the collection of vertices that vertex V_i points to it. The TextRank algorithm can be presented in the Algorithm (1).

Algorithm 1 Implementation of TextRank algorithm.

Input: Textual data
Output: Extracted keywords K

1: Build the graph $G = (V, E)$
2: Compute the edges' importance score by Equation (1)
3: Sort the edges by their scores
4: Select top K edges

3.5 Keyword Extraction for Topics

At this stage, we build a sample set of keywords for each topic from a set of keywords extracted from the subject-labeled texts in the training set. The model of building a set of keyword samples for topics can be summarized as Figure (3). The keywords of the topic are calculated by the statistical method of the number of occurrences of each word in the list of keyword sets of training text. Words that are keywords of one topic cannot be keywords of another topic.

3.6 Similarity Measurement by Jaccard Distance

Mathematically, there are many ways to calculate the similarity between any two keyword lists R_i and any R_j, provided they are of the same length. However, in the context of the similarity between two documents, we do not need to include all words in the text but only the representative keywords T. The number of keywords represented will be a lot less than the entire word in the text. Then, a weighted version of the Jaccard [23] distance is determined as follows:

$$\text{Jaccard}(R_i, R_j) = \frac{1}{T} \sum_{d=1}^{T} \frac{\left|R_{i,d} \cap R_{j,d}\right|}{\left|R_{i,d} \cup R_{j,d}\right|}, \tag{2}$$

4 Experiments

4.1 Datasets

Data are collected from highly reputable Vietnamese websites. We used Teleport pro [2] software for automated data collection. The downloaded data is converted to plain text file and saved to the corresponding folders with the folder name as predefined theme name. Specifically, data is downloaded from the website [3] [4] [5] [6] with 15 main topics. These topics are summarized in Table (1). The

[2] http://tenmax.com
[3] http://vnexpress.net
[4] http://dantri.vn
[5] http://tuoitre.vn
[6] http://yahoo.com.vn

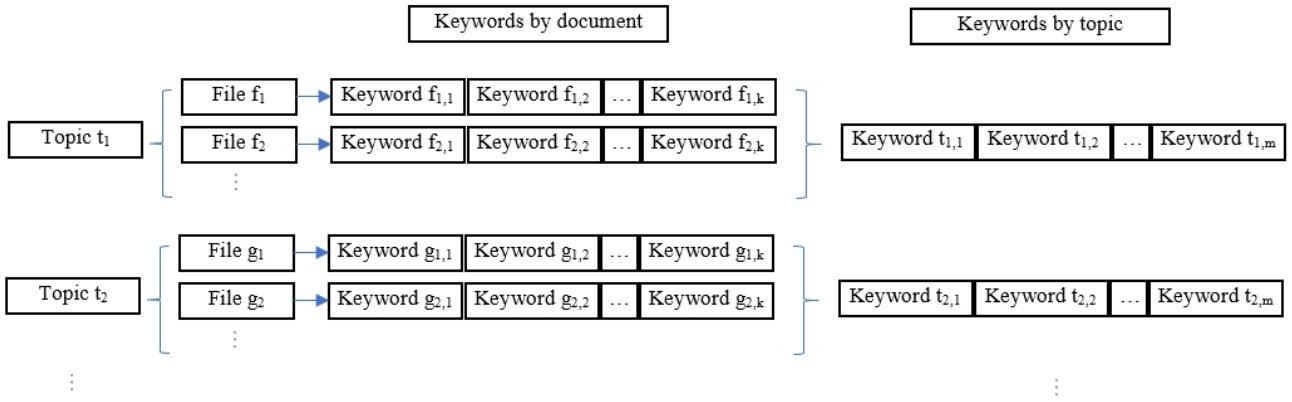


Figure 3: The keyword extraction for topics.

collected data was distributed into 80% training data group and 20% test data group.

Table 1: Summary of 15 topics

#	Topic	Description
1	Music	Singers, songs, musicians, performers, audiences.
2	Cuisine	Food, restaurants, cooks, menus.
3	Movies	Cinemas, genres, actors/actresses.
4	Tourism	tourist destinations, resorts, attractions, hotels.
5	Family	Marriage, family, grandpa, father, mother.
6	Education	Education and training, enrollment, students.
7	Science	Research, inventions, scientific discoveries.
8	Business	market, buying and selling, gold prices, currencies.
9	Beauty	beauty salon, care, therapies.
10	Motorcycle	motorbikes, cars, prices, comparison, repair.
11	Law	Legal, criminal, police events.
12	Health	health information, hospitals.
13	Sports	soccer matches, games, scores, comments.
14	Fashion	Field of fashion, costumes, designers, stylists.
15	Computers	Digital technology, computers, operating systems.

4.2 Evaluation Metric

We define the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). We also define m_+ the total of condition positives, m_- the total of condition negatives, $\hat{m}_+$ the total predicted condition positives, $\hat{m}_-$ the total predicted condition negatives, and m the total population. Then, we compute the sensitivity or recall by using:

$$\text{recall} = \frac{\text{TP}}{m_+} . \tag{3}$$

We compute precision as follows:

$$\text{precision} = \frac{\text{TP}}{\hat{m}_+} . \tag{4}$$

Then, we compute F1-score as follows:

$$\text{F1-score} = 2\frac{\text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} . \tag{5}$$

Table 2: Statistics on all experimental scenarios.

# Keywords	Accuracy (%)	Training time (s)	Test time (s)
10	75.27	2155	1111
15	82.47	2156	1116
20	85.87	2162	1114
25	87.40	2168	1115
30	88.13	2178	1116
35	88.00	2177	1118
40	89.40	2184	1119
45	89.90	2191	1119
50	90.07	2199	1120
55	90.07	2205	1112
60	90.03	2213	1123
65	90.06	2219	1127
70	90.07	2307	1125
all	**95.40**	9279	1362

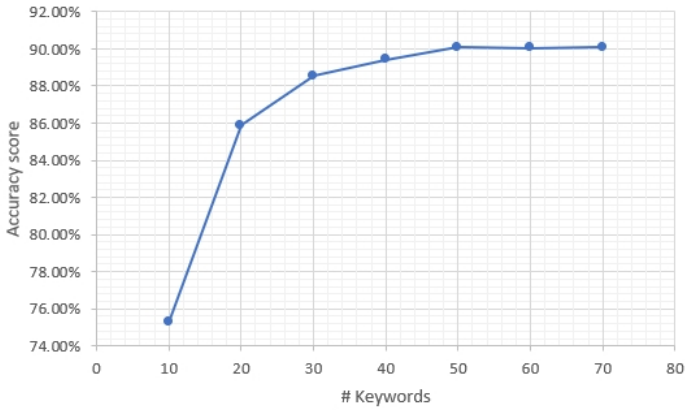


Figure 4: The correlation between the number of keywords and accuracy score.

4.3 Experimental Results

In the experiments, the authors have designed several scenarios where the number of keywords is varied to test the model's performance. The number of keywords $T = \{10, 20, 30, 40, 50, 60, 70\}$ and

in an extreme scenario where all the keywords selected by the TextRank algorithm are used. The experimental results are presented in Table (3), Table (4), Table (5), Table (6), Table (7), Table (8), Table (9), and Table (10) respectively.

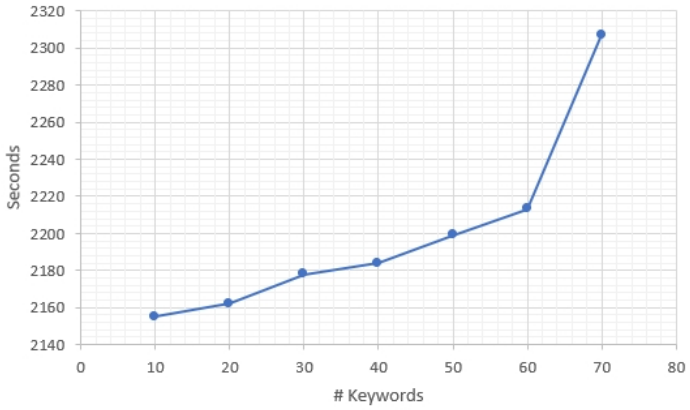


Figure 5: The correlation between the number of keywords and training time.

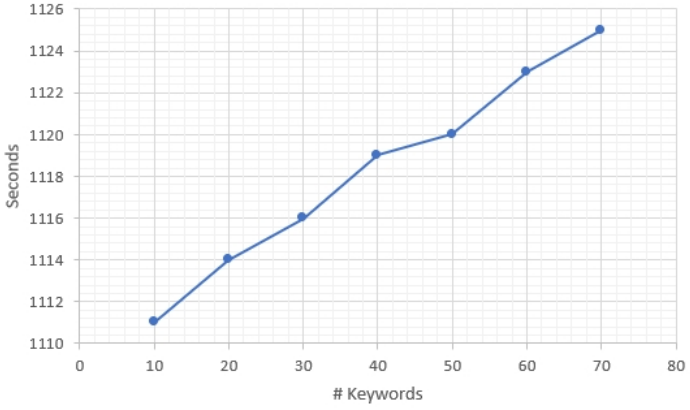


Figure 6: The correlation between the number of keywords and test time.

Table 3: Test set's confusion matrix with 10 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**87**	0	4	0	7	0	0	0	0	0	1	0	0	1	0	87
2	0	**73**	0	2	6	0	11	2	4	0	0	0	0	2	0	73
3	6	0	**93**	0	1	0	0	0	0	0	0	0	0	0	0	93
4	5	2	3	**60**	14	0	0	6	2	0	2	1	2	3	0	60
5	4	2	2	0	**78**	2	0	2	0	0	6	2	0	0	2	78
6	4	2	0	0	2	**80**	0	0	0	0	10	0	2	0	0	80
7	4	2	3	5	9	5	**49**	0	7	2	0	1	3	1	9	49
8	4	1	0	2	7	0	2	**77**	0	2	0	1	2	0	2	77
9	0	1	1	0	2	0	4	1	**83**	0	0	6	2	0	0	83
10	0	0	0	1	0	0	0	4	1	**91**	1	0	1	1	0	91
11	0	1	1	1	7	0	1	6	0	0	**82**	1	0	0	0	82
12	4	0	0	3	9	1	5	0	6	1	5	**66**	0	0	0	66
13	2	1	0	1	18	1	0	0	1	0	0	0	**74**	1	1	74
14	6	0	10	2	2	0	2	2	3	4	1	0	1	**67**	0	67
15	2	0	2	0	1	0	1	10	3	10	0	0	0	2	**69**	69
Average accuracy																**75**

The authors also make statistics on all experimental scenarios in Table (2). Note that *all* in the Table means all the tokens in a particular text. The texts' length is varied, and the average number of tokens is 500. We plotted the correlation between the number of keywords and accuracy scores in Figure (4). The accuracy of 90% is stable at 50 keywords. In Figure (5), we observe a considerable increase from 60 to 70 keywords. While in Figure (6), the test time grows eventually.

Table 4: Test set's confusion matrix with 20 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**96**	0	1	0	3	0	0	0	0	0	0	0	0	0	0	96
2	1	**85**	0	1	5	0	3	2	0	1	1	0	0	1	0	85
3	0	0	**95**	1	3	0	0	0	1	0	0	0	0	0	0	95
4	0	0	1	**83**	5	0	0	10	0	0	0	0	0	1	0	83
5	0	0	0	2	**92**	0	0	0	0	0	6	0	0	0	0	92
6	0	0	2	0	7	**77**	0	2	0	0	10	0	0	0	2	77
7	0	0	0	3	5	2	**78**	3	3	2	0	2	1	0	1	78
8	0	0	2	1	2	0	2	**89**	0	1	2	0	0	0	1	89
9	0	0	0	0	4	0	2	1	**81**	3	0	7	0	2	0	81
10	1	0	0	0	0	0	1	1	0	**95**	0	0	1	1	0	95
11	1	0	0	0	8	0	0	5	0	0	**85**	1	0	0	0	85
12	0	1	0	2	14	2	5	2	6	1	3	**63**	0	0	1	63
13	0	0	0	0	4	0	0	0	0	0	0	0	**96**	0	0	96
14	5	0	3	1	3	0	0	0	1	1	0	0	0	**86**	0	86
15	1	0	2	0	0	0	2	3	0	5	0	0	0	0	**87**	87
Average accuracy																**86**

Table 5: Test set's confusion matrix with 30 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**95**	0	2	0	3	0	0	0	0	0	0	0	0	0	0	95
2	0	**91**	0	2	7	0	0	0	0	0	0	0	0	0	0	91
3	0	0	**95**	0	3	0	0	0	0	0	0	0	0	2	0	95
4	0	0	2	**84**	5	0	0	8	0	0	1	0	0	0	0	84
5	0	0	0	0	**94**	0	0	0	0	0	4	2	0	0	0	94
6	0	0	0	0	12	**74**	2	6	0	2	4	0	0	0	0	74
7	0	0	0	0	2	0	**88**	5	3	1	0	1	0	0	0	88
8	0	0	2	1	1	1	3	**91**	0	0	1	0	0	0	0	91
9	0	0	0	0	4	0	0	0	**88**	2	0	4	0	2	0	88
10	0	0	0	0	0	0	0	3	0	**96**	0	0	1	0	0	96
11	0	0	0	0	5	0	0	7	0	0	**88**	0	0	0	0	88
12	0	3	0	2	15	1	5	5	5	0	2	**62**	0	0	0	62
13	0	0	0	0	4	0	0	0	0	0	0	1	**95**	0	0	95
14	4	0	2	1	0	0	0	2	1	0	0	0	0	**90**	0	90
15	0	0	0	0	1	0	1	5	0	2	0	0	0	0	**91**	91
Average accuracy																**88**

Table 6: Test set's confusion matrix with 40 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**91**	0	1	0	7	0	0	0	0	0	0	0	0	1	0	91
2	0	**95**	0	2	3	0	0	0	0	0	0	0	0	0	0	95
3	0	0	**94**	1	3	0	0	0	0	2	0	0	0	0	0	94
4	0	0	1	**87**	1	0	0	11	0	0	0	0	0	0	0	87
5	0	0	0	0	**92**	2	0	0	0	0	4	2	0	0	0	92
6	0	0	0	0	12	**78**	2	2	0	0	6	0	0	0	0	78
7	0	0	1	1	0	0	**90**	4	2	1	0	1	0	0	0	90
8	0	0	1	0	2	0	0	**94**	0	1	2	0	0	0	0	94
9	0	0	0	1	3	0	0	0	**90**	1	0	2	0	3	0	90
10	0	0	0	0	0	0	0	4	1	**95**	0	0	0	0	0	95
11	0	0	1	1	5	0	0	3	0	0	**89**	0	0	0	1	89
12	0	3	0	3	14	0	3	11	5	0	0	**61**	0	0	0	61
13	0	0	0	0	4	0	0	0	0	0	0	0	**96**	0	0	96
14	2	0	1	0	0	0	0	2	0	0	0	0	1	**94**	0	94
15	0	0	0	0	1	1	0	2	0	1	0	0	0	0	**95**	95
Average accuracy																**89**

Table 7: Test set's confusion matrix with 50 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**94**	0	1	0	4	0	0	0	0	0	0	0	0	1	0	94
2	1	**96**	0	1	1	0	0	1	0	0	0	0	0	0	0	96
3	6	0	**87**	0	2	0	0	1	0	2	0	0	0	2	0	87
4	0	0	0	**89**	3	0	0	8	0	0	0	0	0	0	0	89
5	0	0	0	0	**96**	0	0	0	0	0	2	2	0	0	0	96
6	0	0	0	0	14	**82**	2	0	0	0	2	0	0	0	0	82
7	0	0	1	3	0	0	**87**	3	1	4	0	1	0	0	0	87
8	0	0	0	2	1	0	0	**95**	0	0	2	0	0	0	0	95
9	0	0	0	0	1	0	0	0	**90**	1	0	4	1	3	0	90
10	0	0	0	1	0	0	1	4	1	**93**	0	0	0	0	0	93
11	0	0	1	0	6	0	0	2	0	0	**91**	0	0	0	0	91
12	0	1	0	1	13	0	3	6	3	0	1	**72**	0	0	0	72
13	0	0	0	0	4	0	0	0	0	0	0	0	**96**	0	0	96
14	3	0	2	0	1	0	0	0	0	0	0	0	0	**94**	0	94
15	2	0	0	0	1	1	0	5	1	1	0	0	0	0	**89**	89
Average accuracy																**90**

Table 8: Test set's confusion matrix with 60 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**93**	0	1	0	4	0	0	1	0	0	0	0	0	1	0	93
2	0	**94**	0	2	4	0	0	0	0	0	0	0	0	0	0	94
3	2	0	**89**	2	4	0	0	1	0	0	0	0	0	2	0	89
4	0	0	0	**87**	3	0	0	10	0	0	0	0	0	0	0	87
5	0	0	0	0	**94**	2	0	0	0	0	4	0	0	0	0	94
6	0	0	1	1	12	**74**	2	2	0	0	8	0	0	0	0	74
7	0	0	1	0	0	0	**87**	5	0	6	0	1	0	0	0	87
8	0	0	2	2	1	0	0	**93**	0	0	2	0	0	0	0	93
9	0	1	0	1	0	0	0	0	**93**	1	0	2	0	2	0	93
10	0	0	0	1	0	0	1	3	0	**95**	0	0	0	0	0	95
11	0	0	0	1	1	0	0	3	0	0	**95**	0	0	0	0	95
12	0	1	0	2	7	0	5	6	4	0	3	**72**	0	0	0	72
13	0	0	0	0	4	0	0	0	0	0	0	1	**95**	0	0	95
14	3	0	1	0	0	0	0	2	0	0	0	0	0	**94**	0	94
15	0	0	0	0	2	0	0	2	0	1	0	0	0	0	**95**	95
Average accuracy																**90**

Table 9: Test set's confusion matrix with 70 keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**94**	0	1	0	2	0	1	1	0	0	0	0	0	1	0	94
2	0	**93**	0	1	3	0		0	0	0	0	2	0	1	0	93
3	0	0	**87**	2	6	0	0	1	0	0	0	0	0	4	0	87
4	0	0	0	**94**	0	0	0	6	0	0	0	0	0	0	0	94
5	0	0	0	0	**94**	0	0	0	0	0	4	2	0	0	0	94
6	0	0	0	3	10	**75**	2	4	0	0	6	0	0	0	0	75
7	0	0	0	0	0	0	**91**	5	0	4	0	0	0	0	0	91
8	0	0	1	1	1	0	0	**95**	0	0	2	0	0	0	0	95
9	0	0	0	1	3	0	0	0	**90**	1	0	2	0	3	0	90
10	0	0	0	1	0	0	1	4	0	**94**	0	0	0	0	0	94
11	0	1	0	0	1	0	0	2	0	0	**95**	1	0	0	0	95
12	0	1	0	0	10	1	3	6	8	0	4	**67**	0	0	0	67
13	0	0	0	0	4	0	0	0	0	0	0	1	**95**	0	0	95
14	2	0	2	0	0	0	1	1	0	0	0	0	0	**94**	0	94
15	0	0	0	0	1	0	0	4	1	0	0	0	0	1	**93**	92
Average accuracy																**90**

5 Conclusion

In this article, the authors have described a proposed approach that allows text classification based on the solution of extracting specific representative keywords of the text. We discussed the proposed system in detail from the abstract design, text pre-processing, and Vietnamese characteristics. Then we described the TextRank algorithm based on graphs to score important information about the text's structure. Intensive experiments have been conducted to prove the stability and robustness of the proposed system. High accuracy of 90.07% has been achieved. Although many Vietnamese documents are electronically available, this is the first to conduct text classification based on keywords. This research portrays our endeavor to construct a practical framework for addressing Vietnamese texts' classification tasks on news websites.

Table 10: Test set's confusion matrix with all keywords selected by the TextRank algorithm.

Topic ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Accuracy (%)
1	**92**	0	0	0	8	0	0	0	0	0	0	0	0	0	0	92
2	0	**92**	0	2	6	0	0	0	0	0	0	0	0	0	0	92
3	0	0	**86**	2	12	0	0	0	0	0	0	0	0	0	0	86
4	0	0	0	**96**	2	0	0	2	0	0	0	0	0	0	0	96
5	0	0	0	0	**99**	0	0	0	0	0	1	0	0	0	0	99
6	0	0	0	0	16	**84**	0	0	0	0	0	0	0	0	0	84
7	0	0	0	0	0	0	**99**	1	0	0	0	0	0	0	0	99
8	0	0	1	0	0	0	0	**99**	0	0	0	0	0	0	0	99
9	0	0	0	0	1	0	0	0	**99**	0	0	0	0	0	0	99
10	0	0	0	1	1	0	1	0	0	**97**	0	0	0	0	0	97
11	0	0	0	0	1	0	0	0	0	0	**99**	0	0	0	0	99
12	0	0	0	0	2	0	0	0	1	0	0	**97**	0	0	0	97
13	0	0	0	0	4	0	0	0	0	0	0	0	**96**	0	0	96
14	0	0	1	0	0	0	0	0	0	0	0	0	0	**99**	0	99
15	0	0	0	0	1	0	1	0	0	0	0	0	0	0	**98**	98
Average accuracy																**95**

References

[1] H. T. Huynh, N. Duong-Trung, X. S. Ha, N. Q. T. Tang, H. X. Huynh, D. Q. Truong, "Automatic Keywords-based Classification of Vietnamese Texts," in 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), 1–3, IEEE, 2020, doi:10.1109/RIVF48685.2020.9140761.

[2] L. Zhang, L. Jiang, C. Li, G. Kong, "Two feature weighting approaches for naive Bayes text classifiers," Knowledge-Based Systems, **100**, 137–144, 2016, doi:10.1016/j.knosys.2016.02.017.

[3] C. C. Aggarwal, Data classification: algorithms and applications, CRC press, 2014.

[4] S. Jiang, G. Pang, M. Wu, L. Kuang, "An improved K-nearest-neighbor algorithm for text categorization," Expert Systems with Applications, **39**(1), 1503–1509, 2012, doi:10.1016/j.eswa.2011.08.040.

[5] J. Nam, J. Kim, E. L. Mencía, I. Gurevych, J. Fürnkranz, "Large-scale multi-label text classification—revisiting neural networks," in Joint european conference on machine learning and knowledge discovery in databases, 437–452, Springer, 2014, doi:10.1007/978-3-662-44851-9_28.

[6] V. B. Kobayashi, S. T. Mol, H. A. Berkers, G. Kismihók, D. N. Den Hartog, "Text classification for organizational researchers: A tutorial," Organizational research methods, **21**(3), 766–799, 2018, doi:10.1177/1094428117719322.

[7] N. Duong-Trung, Social Media Learning: Novel Text Analytics for Geolocation and Topic Modeling, Cuvillier Verlag, 2017.

[8] G. Xu, J. Qi, D. Huang, M. Daneshmand, "Detecting spammers on social networks based on a hybrid model," in 2016 IEEE International Conference on Big Data (Big Data), 3062–3068, IEEE, 2016, doi:10.1016/j.neucom.2015.02.047.

[9] N. Duong-Trung, L. Schmidt-Thieme, "On Discovering the Number of Document Topics via Conceptual Latent Space," in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM '17, 2051–2054, Association for Computing Machinery, New York, NY, USA, 2017, doi:10.1145/3132847.3133086.

[10] N. Duong-Trung, N. Schilling, L. Schmidt-Thieme, "Near real-time geolocation prediction in twitter streams via matrix factorization based regression," in Proceedings of the 25th ACM international on conference on information and knowledge management, 1973–1976, ACM, 2016, doi:10.1145/2983323.2983887.

[11] D. Kim, D. Seo, S. Cho, P. Kang, "Multi-co-training for document classification using various document representations: TF–IDF, LDA, and Doc2Vec," Information Sciences, **477**, 15–29, 2019, doi:10.1016/j.ins.2018.10.006.

[12] R. Niu, B. Shen, "Microblog User Interest Mining Based on Improved TextRank Model," Journal of Computers, **30**(1), 42–51, 2019, doi:10.3966/199115992019023001005.

[13] N. Akhtar, M. S. Beg, H. Javed, "TextRank enhanced Topic Model for Query focussed Text Summarization," in 2019 Twelfth International Conference on Contemporary Computing (IC3), 1–6, IEEE, 2019, doi:10.1109/IC3.2019.8844939.

[14] V. C. D. Hoang, D. Dinh, N. Le Nguyen, H. Q. Ngo, "A comparative study on vietnamese text classification methods," in 2007 IEEE International Conference on Research, Innovation and Vision for the Future, 267–273, IEEE, 2007, doi:10.1109/RIVF.2007.369167.

[15] T.-H. Pham, P. Le-Hong, "Content-based approach for Vietnamese spam SMS filtering," in 2016 International Conference on Asian Language Processing (IALP), 41–44, IEEE, 2016, doi:10.1109/IALP.2016.7875930.

[16] T.-L. Ngo, K. L. Pham, H. Takeda, S. B. Pham, X. H. Phan, "On the identification of suggestion intents from vietnamese conversational texts," in Proceedings of the Eighth International Symposium on Information and Communication Technology, 417–424, 2017, doi:10.1145/3155133.3155201.

[17] V. T. Nguyen, N. T. Hai, N. H. Nghia, T. D. Le, "A Term Weighting Scheme Approach for Vietnamese Text Classification," in International Conference on Future Data and Security Engineering, 46–53, Springer, 2015, doi:10.1007/978-3-319-26135-5_4.

[18] N. H. D. Tri, N. T. Quan, N. V. Tien, N. T. Hung, "Xay dung mo hinh phan tan cho phan lop khoi luong lon van ban theo chu de (in English: building distributed model for classification massive text data by topic)," PROCEEDING of Publishing House for Science and Technology, 2017, doi:10.15625/vap.2016.000104.

[19] B. K. Linh, N. T. T. Ha, N. T. N. Tu, D. T. Tinh, "Phan loai van ban tieng Viet dua tren mo hinh chu de (in English: vietnamese text classification based on topic modeling)," PROCEEDING of Publishing House for Science and Technology, 2017, doi:10.15625/vap.2016.00065.

[20] T. Ngoc Phuc, P. Tran Vu, P. Cong Xuyen, N. Vu Duy Quang, "Phan loai noi dung tai lieu Web tieng viet (in English: classification of vietnamese texts on the web)," Vietnam Journal of Science and Technology, **51**(6), 669–680, doi:10.15625/2525-2518/51/6/11629.

[21] N. T. M. Le Hong Phuong, A. R. Huyen, A. Ho Tuong Vinh, "Hybrid Approach to Word Segmentation of Vietnamese Texts," 2008, doi:10.1007/978-3-540-88282-4_23.

[22] R. Mihalcea, P. Tarau, "Textrank: Bringing order into text," in Proceedings of the 2004 conference on empirical methods in natural language processing, 404–411, 2004.

[23] G. Rebala, A. Ravi, S. Churiwala, An introduction to machine learning, Springer, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1334-1342 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

SEA: An UML Profile for Software Evolution Analysis in Design Phase

Akram Ajouli[1,2,*]

[1]*Department of Computer Science, College of Computer and Information sciences, Jouf University, Sakaka 2014, KSA*

[2]*High Institute of Applied Sciences and Technology of Gafsa, Gafsa University, Gafsa, 2100, Tunisia*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 14 February, 2021
Online: 28 February, 2021

Keywords:
Software Evolution
Tyranny of decomposition
Modular Maintenance
Software Concerns
UML Profile

ABSTRACT

Software evolution is one of the software process activities that occupies a major percentage of software development cost. Since requirements change continually and new technologies emerge, software should be adapted to satisfy these new changes to continue to survive. Despite software evolution being performed after software validation and deployment, software developers should predict at earlier stages how software would evolve in the future to avoid surprises. Although many works focus on how to enhance the program structure to facilitate maintenance tasks, only few works treat software evolution in earlier phases of software development process. In this direction, we propose an UML profile that permits to tackle software maintenance issues at the early phases of software development process. The proposed approach helps software developers to predict in design phase the kind of maintenance tasks that could occur in the future.

1 Introduction

According to [1], software evolves and changes continually in order to support new emerging technologies and satisfy customers and business needs. Some studies estimate that software evolution could reach more that 60% of software development total cost [2]. This issue pushes software engineering community to focus on how optimizing software evolution cost.

In fact, software maintenance cost varies depend on the kind of maintenance tasks that will be performed. For example, if the source code to be maintained is crosscutting among software modules, maintenance tasks performed on that source code are as well, the thing that could generate an extra cost. But, if the source code to be maintained is encapsulated in a single module, software maintenance applied to that source code is modular (less time and less cost).

Duality among crosscutting maintenance and modular maintenance motivates the community to provide solutions to keep software maintenance being modular. For example, some design patterns [3] and some program structuring/decomposition [4]–[8] facilitate software reuse and maintenance. But, even though some solutions could assure modular software maintenance, the challenge still existing since software maintenance could be modular for the same program with respect to a given architecture/structure but not with respect to another. A such problem is known as tyranny of decomposition or expression problem [9] (see Sect. 3 for more details). To resolve this problem, many works ([10]–[14]) tried to attenuate the effect of tyranny of decomposition on software maintenance but results still relative and they treat this issue only in source code level.

In this paper, we illustrate how software developers could tackle software maintenance issues at the design phase of software development process. The early treatment and analysis of software maintenance tasks that could occur later permits to get a clear overview about how the program will evolve in the future. The main goal is to specify from the beginning the program structure that is more adaptable for future maintenance tasks. This could be realized by: first, identifying what kind of maintenance tasks is more probable to occur and then recommending the program architecture that is more adaptable to this kind of maintenance tasks. To concertize our approach, we define an UML profile that helps software developers to analyze and predict the kind of future software maintenance tasks that could occur.

The rest of this paper is organized as follows: first, we explore some related works (Sect. 2). Then, we give some explanations about tyranny of program decomposition problem. We show also how this problem affects modular maintenance (Sect. 3). Next, we present how we have defined *SEA* UML profile (Sect. 4). After that,

*Corresponding Author: Akram Ajouli, Department of Computer Science, College of Computer and Information sciences, Jouf University, Sakaka 2014, KSA , Contact No: +966547205946 & Email: asajouni@ju.edu.sa

we show how we validated *SEA* by applying some stereotypes in a real case study (Sect. 5). Finally, we resume our contributions and we propose some future works (Sect. 6).

2 Related work

In fact, this paper is an extension of the paper [15] (presented in the International Conference on Computer and Information Sciences (ICCIS) in 2019). In [15], the approach consists of proposing an UML profile called *MODEM* used to model maintenance tasks in design phase. This profile provides some UML stereotypes that permits to software developers to model and analyse maintenance tasks before occurring. The extension we propose in this work consists of adding more details and semantics to the previous work. We have modified some stereotypes, added new stereotypes and added tagged values in order to better analyse and predict in design phase future maintenance tasks that could occur. Further more, comparing to the previous work, we validate the extended UML profile by apply it to a real case study.

In [16], the author proposed a process of eight steps helping to supervise maintenance tasks that could occur later. Their approach simulates communications between stakeholders in order to design maintenance tasks, apply them and test software again before delivering the final version. In [17], the author proposed also an approach that aims to treat maintenance tasks in early phases of software development. They propose a software maintenance life cycle model that relies on a process of four steps. During this process, stakeholders interact between each other to study and analyze maintenance tasks to do in order to guarantee a consistent planning for performing maintenance tasks. The process is incremental and iterative and it stops when maintenance tasks analysis proves that software maintenance requests percentage becomes under a given threshold.

Although these works show a relevant contribution in planning maintenance tasks at earlier stages of software development, both of them involve different stakeholders to express needed maintenance tasks which could disturb software development process. In addition, maintenance issue is not expressed explicitly in models as we propose in our approach.

In [18], the author proposed an UML profile that permits to support requirements evolution. Their method is tool supported and consists of extending requirements engineering process by making it supporting requirements evolution. Although this approach facilitates requirements changes management, it still limited to specification phase. In our work, we focus more on design phase because we treat a problem which is related to software architecture rather than software requirements.

Some other works like [19]–[22] focus on analyzing and expressing design patterns explicitly at early phases of software development by defining the appropriate UML profiles. These works aim to enhance design patterns implementation comprehension for programmers since design pattern does not appear explicitly on source code. These works focus more on improving software comprehension rather than treating software evolution issue.

Some other works such as [13] and [23] treat maintenance issues in implementation phase by providing reversible transformations among dual design patterns (composite and visitor) in order to keep maintenance tasks modular. In [24], the author proposed also a relevant approach that treats software maintenance issue in source code level. Their approach consists first of exploring the best machine learning techniques used for predicting software maintenance and then improving their performance. Although these works show important contributions in treating software maintenance issues, they still limited to implementation phase while we focus more on tackling this issue in earlier software process activities.

3 Background

We present in this section an overview about potential challenges that software evolution could face. One of the issues that could affect software evolution is tyranny of decomposition or expression problem [9]. In fact, although software modularity enhances program structure and facilitates software maintenance, a modular program structure could be biased to one kind of modular maintenance and not to another. A maintenance task is called modular when changes are mostly encapsulated in one module. Contrary to this, if changes touches many modules at the same time, the maintenance task is called crosscutting.

For example, Figure 1 illustrates a program composed of two classes *Circle* and *Rectangle* that extend an abstract class *Shape*. The program contains also two methods *print()* and *show()* which represent the business code of the program. The structure of the program is a base form of Composite design pattern in which each data type is encapsulated in a module. A such program structure decomposes the program with respect to data-types axis.

The decomposition of the program (Figure 1) permits to perform modular maintenance with respect to data-types. For instance, if we add a new shape (a Triangle for example) to the hierarchy of shapes shown in Figure 1, we first create a new class, then, we redefine the code of methods *print()* and *show()* only in this new class. In this case, the maintenance task is modular (low cost). But, if we add a new feature to the program (for example a method *color()*), we have to implement it in all classes which causes a crosscutting maintenance (extra cost). We deduce that the architecture of the program shown in Figure 1 allows a modular maintenance with respect to data-types and not with respect to functions (methods). Hence, the program structure is dominated by the data-types decomposition axis (a tyranny of decomposition according to data-types).

Another axis of decomposition of the program described above is illustrated in Figure 2. This program represents the Visitor design pattern implementation of the program shown in Figure 1. The Visitor pattern encapsulates each business code in a single module: the business code *print* is encapsulated in module/class *Print* and the business code *show* is encapsulated in module/class *Show*. This program structure decomposes the program with respect to functions which permits to add a new function without touching other modules: for example if we add a feature *color* to the program, we add a new class in which we implement this new feature for all shapes. Thus, Visitor design pattern permits to perform modular maintenance with respect to functions since it provides a functions decomposition axis.

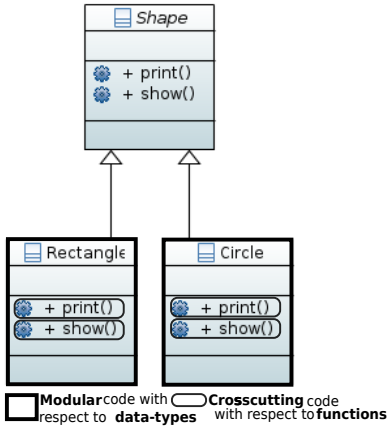


Figure 1: Decomposition axis of a simple variety of Composite design pattern.

But, if we add a new shape to the program shown in Figure 2, many modules would be changed since the new shape should be visited by both *print* and *show* business codes. In this case, the program maintenance is crosscutting. We say that the program structure is dominated by functions decomposition axis (a tyranny of decomposition according to functions).

The above two examples explain well the problem of tyranny of decomposition. A such issue still always a main problem that affects software evolution. In this work, we aim to contribute in attenuating tyranny of program decomposition effect on software maintenance. Our target is to help software developers in analyzing and predicting in earlier phases of software development the kind of maintenance tasks that could occur later and in recommending the right architecture (decomposition axis) of the program that could keep maintenance being modular.

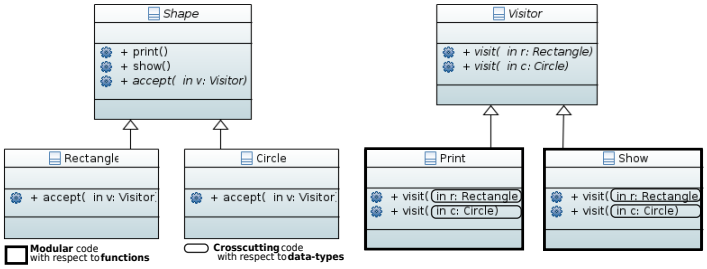


Figure 2: Visitor structure corresponding to the program shown in Figure1.

4 SEA specification

4.1 Overview

Many works have treated software evolution issue at the implementation phase. In this work, we tackle this issue in design phase of software development process. Since UML (Unified Modeling Language) [25] does not provide customized semantic constructs related to software maintenance, we propose an UML profile [26] that supports analyzing and predicting maintenance tasks that could occur in the future.

The UML profile we propose is called *SEA* (in short *SEA* for *S*oftware *E*volution *A*nalysis) (see Fig. 3) . *SEA* is defined in Papyrus [27] EMF-based editor and it provides the following features:

- tracing software evolution history by adding some information to each software version. This feature could help to get an overview about maintenance tasks characteristics, the thing that could provide some recommendations about choosing the right program structure.
- studying and deciding which parts of a given software are more stable than others. This feature could help to predict at earlier phases the program structure evolution.
- expressing explicitly some semantic terms related to software maintenance in class diagrams such as data-type module, function module, composite design pattern...etc. This feature enhances software comprehension and helps to identify the decomposition axis of a given software.

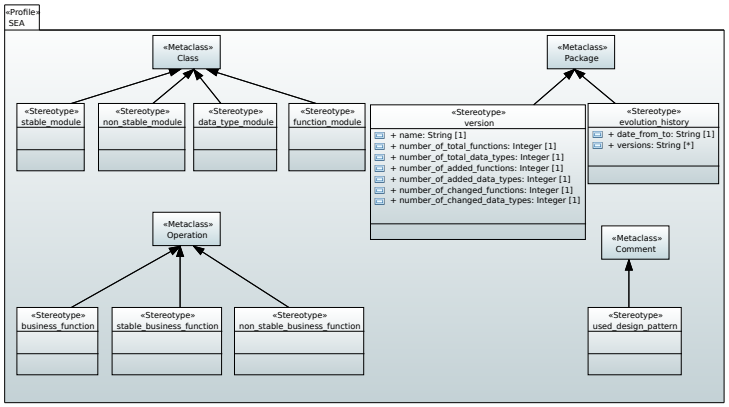


Figure 3: *SEA*: an UML profile for analyzing software evolution.

4.2 Stereotypes for analyzing software modularity

Marking modular and crosscutting software parts. Modularity in software architecture helps to optimize software maintenance cost. It permits also to encapsulate concerns and minimize crosscutting concerns effect. Thus, expressing modularity concepts explicitly in models and diagrams helps software developers to explore at early stages of software development the decomposition axis that is dominant. This helps to recommend the right kind of future maintenance tasks that are modular and to avoid those which are crosscutting. In this direction, *SEA* permits to mark explicitly in class diagram which elements represent modular concerns and which elements represent crosscutting concerns.

Until now, *SEA* supports two concerns (or two decomposition axes): data-types concern and functions concern. In this context, we propose the following stereotypes:

- <<*data_type_module*>>: it extends the meta class *Class* and it is used to precise classes that belong to data-types concern when software architecture is built according to Composite design pattern.
- <<*function_module*>>: it extends the meta class *Class* and it is used to precise classes that belong to functions concern

and that encapsulate program business code when software architecture is built according to Visitor design pattern.

- <<*business_function*>>: it extends the meta class *Operation* and it is used to mark methods that contain program business code. It could be used especially in Composite pattern structure to show that business code is crosscutting among different modules. This feature helps software developers to be aware of the cost of maintenance tasks performed with respect to functions.

Fig 4 shows an example of how to precise explicitly modules/classes that belong to the concern data-types by applying <<data_type_module>> to class diagram. In the example, we see explicitly that classes *Rectangle* and *Circle* are marked as data-type modules. In the same example, methods *print()* and *show()* are marked explicitly as business code. This shows visually that business code is crosscutting among program modules which gives an indication that maintenance tasks performed with respect to functions are costly.

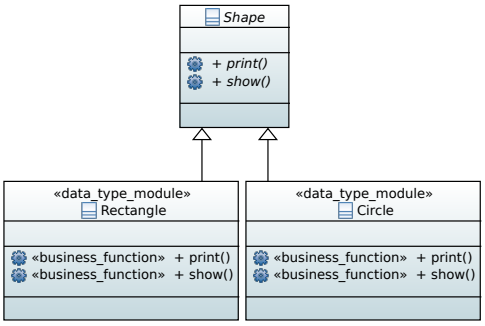


Figure 4: Use of *SEA* in marking modular data-types concern.

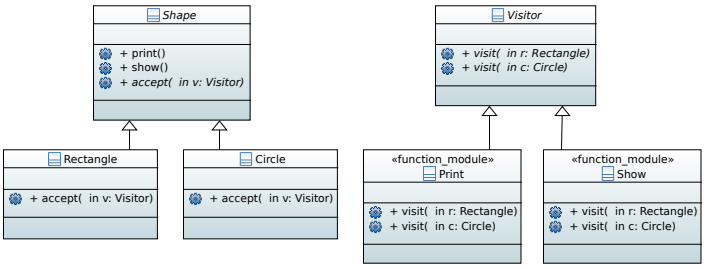


Figure 5: Use of *SEA* in marking modular functions concern.

Fig 5 shows how to use <<function_module>> stereotype to indicate that business code is encapsulated. This is used when the program is structured with respect to Visitor design pattern. We remark that classes *Show* and *Print* encapsulate respectively *show* and *print* business codes. The fact that these two classes are marked explicitly as functions indicates that software maintenance performed with respect to functions is not costly.

Precising modular program decomposition visually on diagrams helps to guide software developers to perform the right kind of maintenance that respects the existing concern. If they have to perform a crosscutting maintenance, they could be at least ready early to apply some tools or techniques that transform automatically the actual architecture to another one that supports modular maintenance (for example using reversible transformations among Composite and Visitor [13]).

Marking stable and changeable software parts. In order to predict future software evolution at early stages, one should predict which software parts tend to change sharply. For this purpose, we propose the following stereotypes:

- **stable_business_function vs non_stable_business_function**: these two stereotypes extend the meta class *Operation*. The former is used to mark functions/methods that would not be sharply changed. The latter marks probable functions that could change drastically in the future.

- **stable_module vs non_stable_module**: these two stereotypes extend the meta class *Class*. The former is used to mark classes/modules that would not be sharply changed, or will have small changes that do not influence the software architecture. The latter marks classes/modules that could radically change.

Fig. 6 shows how to apply stereotypes mentioned above to a class diagram. For example, after marking software parts as stable or non stable, one could identify which parts seems to be more stable than others. In case of finding that stable data-types modules number is greater than stable functions number, then, the decision could be biased to choose Visitor pattern as a future software architecture (since business code tend to change more than data-types).

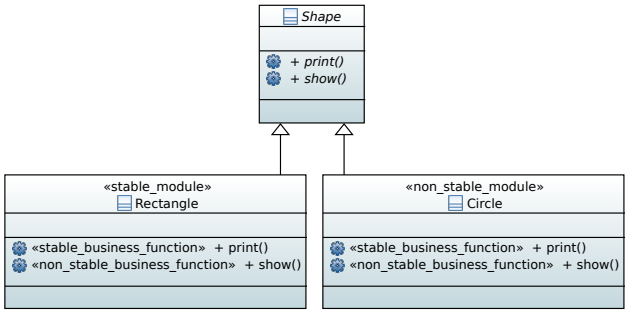


Figure 6: Use of *SEA* in marking changeable and stable software parts.

4.3 *Stereotypes for software evolution history tracing*

With time, software evolution could cause software architecture degeneration [28]. So, having a backup for each version is necessary and helps to get back to the non degenerated version. This is already existing by saving old versions. In our approach, we propose the stereotype <<evolution_history>> (extends meta-class *Package*) which permits to precise explicitly on a package including all software versions that this package contains software evolution history. The proposed stereotype provides two tagged values: <<date_from_to>> and <<versions>>. The former is used to indicate visually the period of time during which versions included in a package stereotyped <<evolution_history>> are built. The latter is used to list all versions that are built in the period of time indicated by <<date_from_to>>. These tagged values permit to add valuable

visual data that could be used to assess the stability of software during a period of time.

Fig. 7 shows how to apply <<evolution_history>> stereotype: the package *software_X* (stereotyped <<evolution_history>>) includes three sub-packages that include respectively three versions of *software_X*.

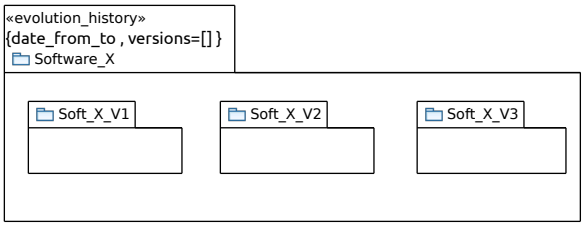


Figure 7: Tracing software evolution with <<evolution_history>> stereotype.

The second stereotype that is related to software evolution history is the stereotype <<version>> which extends meta-class *Package*. It provides information about each version of a given software. It contains seven tagged values:

- *name*: version name
- *number_of_total_data_types*: total number of classes of current version
- *number_of_total_functions*: total number of methods of current version
- *number_of_added_data_types*: number of classes that have been added to the previous version to get the current version
- *number_of_added_functions*: number of methods that have been added to the previous version to get the current version
- *number_of_changed_data_types*: number of classes that have been changed to get the current version
- *number_of_changed_functions*: number of methods that have been changed to get the current version

Tagged values mentioned above permit to provide visual information about the behavior of software evolution. Getting a visual access to added classes and methods number during switching among software versions helps to get an overview about the recommended software architecture that could optimize maintenance cost. For example, if the percentage of *number_of_added_data_types* comparing to *number_of_total_data_types* is always greater than the percentage of *number_of_added_functions* comparing to *number_of_total_functions* and the current architecture is implemented with respect to Visitor pattern, then it will be better to switch the program structure into Composite pattern in order to keep maintenance modular with respect to data-types (see more details in Sect. 5).

Fig. 8 shows how to apply <<version>> stereotype: the package *soft_X_V1* (stereotyped <<version>>) shows the seven defined tagged values that permits to note helpful information about software evolution behavior related to *soft_X_V1*.

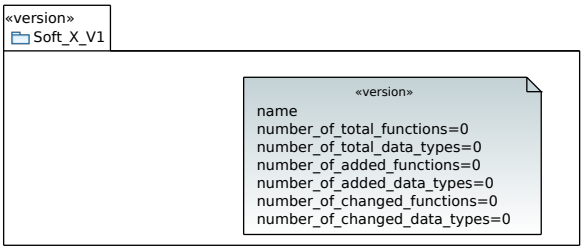


Figure 8: Adding helpful information to a software version by applying <<version>> stereotype.

4.4 Visualizing design patterns

Many programmers face some difficulties when implementing design patterns. Thus, visualizing in design phase some information about the design pattern that will be implemented later orients programmers early to cover any lack of knowledge about the indicated design pattern. In addition, marking design patterns explicitly on diagrams enhances software comprehension.

Marking information about chosen design patterns on diagrams could also assist used techniques in design pattern detection [29] and provide relevant data for data analysis and classification.

To reify this idea, we propose the stereotype <<used_design_pattern>> that extends the meta-class *Comment*. It could be associated to the package that includes the hole software project and indicates information about the chosen design patterns to implement. Provided information includes pattern name, pattern uses, modular maintenance tasks permitted, crosscutting maintenance tasks...etc. Fig. 9 shows how to apply <<used_design_pattern>> on comments.

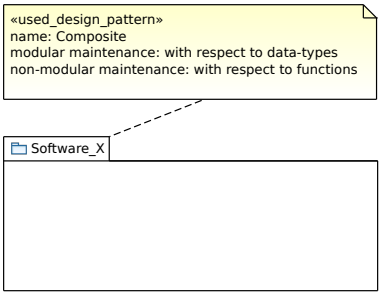


Figure 9: Use of *SEA* in visualizing information about chosen design patterns.

5 Application of *SEA* on a real software

5.1 An overview about JHotDraw evolutions

We study the utility of *SEA* in analyzing *JHotDraw* [30] evolutions. *JHotDraw* is a framework for two-dimensional graphics used for structured drawing editors. It is implemented in Java and it is based on Erich Gamma's *JHotDraw*, which is copyright 1996, 1997 by IFA Informatik and Erich Gamma [31].

We define the following metrics to use them in the rest of this section:

- V_i: a version i of *JHotDraw* where 1 <= i <= 13 (we study here 13 versions of *JHotDraw*)

- CNV_i: classes number in version V_i
- MNV_i: methods number in version Vi
- ACV_1: added classes number to version $V1$ to obtain version V_i

  $ACV_1 = CNV_i - CNV_1$
- AMV_1: added methods number to version $V1$ to obtain version V_i

  $AMV_1 = MNV_i - MNV_1$
- ACV_{i-1}: added classes number to version $Vi-1$ to obtain version V_i

  $ACV_{i-1} = CNV_i - CNV_{i-1}$
- AMV_{i-1}: added methods number to version $Vi-1$ to obtain version V_i

  $AMV_{i-1} = MNVi - MNV_{i-1}$
- $RACV_1$: rate of added classes number to version $V1$ to obtain version V_i

  $RACV_1 = \frac{ACV_1}{CNV_1}$
- $RAMV_1$: rate of added methods number to version $V1$ to obtain version V_i

  $RAMV_1 = \frac{AMV_1}{MNV_1}$
- $RACV_{i-1}$: rate of added classes number to version $Vi-1$ to obtain version V_i

  $RACV_{i-1} = \frac{ACV_{i-1}}{CNV_{i-1}}$
- $RAMV_{i-1}$: rate of added methods number to version $Vi-1$ to obtain version V_i

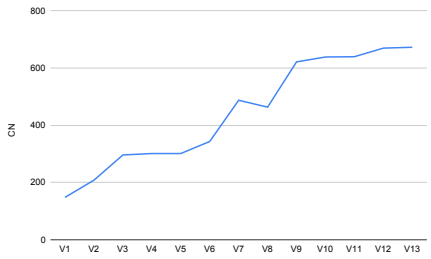
  $RAMV_{i-1} = \frac{AMV_{i-1}}{MNV_{i-1}}$

The table shown in Fig. 10(a) illustrates *JHotDraw* evolution during 11 years. During this period of time, there were 13 versions of *JHotDraw* (5.2, 5.3, 5.4 b1, 5.4 b2, 6.0 b1, 7.0.8, 7.0.9, 7.1, 7.2, 7.3.1, 7.4.1, 7.5.1 and 7.6). We call here each version by V_i where $1 <= i <= 13$. The first column of the table mentions the version name (V_1, V_2,...,V_{13}), the second column mentions classes number in V_i (CNV_i) and the third column mentions methods number in V_i (MNV_i).

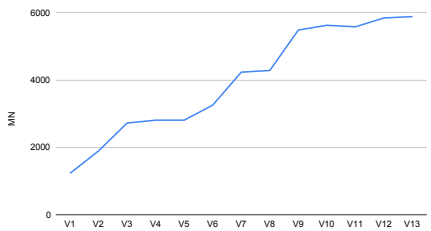
Like any software, *JHotDraw* is influenced by Lehman law [1] as shown by the shape of both curves of CNV_i progress (Fig. 10(b)) and MNV_i progress (Fig. 10(c)). We remark that classes number, which was 148 in V_1, increases to reach 672 in V_{13}, and methods number jumps from 1229 in V_1 to 5885 in V_{13}. Considering these evolutions, we apply some stereotypes of *SEA* profile to show its utility in analyzing software evolution during earlier stages of software process.

V_i	CNV_i	MNV_i
V_1	148	1229
V_2	208	1896
V_3	296	2723
V_4	301	2809
V_5	301	2809
V_6	343	2809
V_7	487	4234
V_8	463	4285
V_9	621	5486
V_{10}	638	5627
V_{11}	639	5582
V_{12}	669	5845
V_{13}	672	5885

(a) Classes number and Methods number of each version (data extracted from [32]).



(b) Evolution of *JHotDraw* classes number (CNV_i) during 11 years.



(c) Evolution of *JHotDraw* methods number (MNV_i) during 11 years.

Figure 10: Variations of classes number and method numbers during *JHotDraw* evolution.

5.2 *Using* SEA *to trace* JHotDraw *evolutions*

To make *JHotDraw* evolutions easy to analyze in the design activity of software process, we apply stereotypes <<evolution_history>> and <<version>>. The former is applied to a package that includes all versions of JHotDRaw during 11 years and the latter is applied to each package that includes a version V_i of *JHotDraw*.

Application of <<evolution_history>> stereotype The target of applying <<evolution_history>> stereotype is to add information about the period of time in which *JHotDraw* versions are built and also about number and names of versions. This information is represented by defined tagged values that characterize the mentioned

stereotype.

Fig. 11 shows how <<evolution_history>> stereotype and its tagged values are applied. By using this stereotype, software developers could create a package that includes all versions of *JHotDraw* and marks this package as a special package to trace *JHotDraw* evolutions history. In addition, they could note information such as "date_from_to=2000_to_2011" and list different versions that has been elaborated during this period.

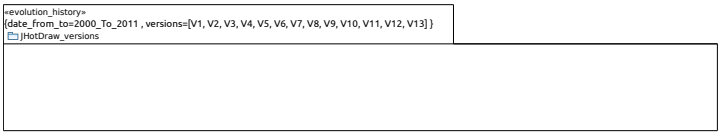


Figure 11: Application of <<evolution_history>> stereotype to JHotDRaw.

In fact, the expressiveness provided by the stereotype <<evolution_history>> permits to assess the stability of *JHotDraw*. The stereotype <<evolution_history>> provides information about the rate of versions number of *JHotDraw* by a given period of time (here 13 versions by 11 years) which could evaluate the impact of *JHotDraw* evolution on total development cost. It motivates also software developers to focus on studying the switch from V_i to V_{i+1} and deduce lessons for future evolutions.

Application of <<version>> stereotype As detailed in Sect. 4, the stereotype <<version>> has 7 tagged values. We consider here only *name*, *number_of_total_functions*, *number_of_total_data_types*, *number_of_added_functions* and *number_of_added_data_types*. For each version V_i of *JHotDraw*, software developers could apply the stereotype <<version>> to the package that contains V_i and they could assign a value to each tagged value. For example, Fig. 12 shows the main package that is stereotyped as <<evolution_history>> and inside of it the set of packages stereotyped as <<version>> that represent the 13 versions of *JHotDraw*.

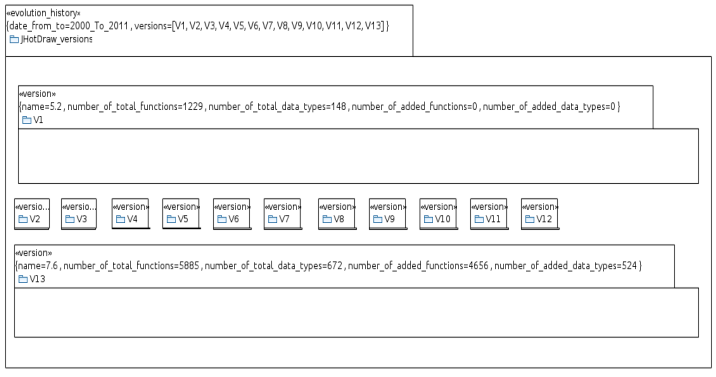


Figure 12: Application of <<version>> stereotype to JHotDRaw.

For each version V_i, we assume that software developers mark:

- name of the version (example: for version V_1, name is *5.2*)
- *number_of_total_data_types* (example: for version V_{13}, *number_of_total_data_types*= 672)
- *number_of_total_functions* (example: for version V_{13}, *number_of_total_functions*= 5885)
- *number_of_added_data_types* (example: for version V_{13}, *number_of_added_data_types*= 524)
- *number_of_added_functions* (example: for version V_{13}, *number_of_added_functions*= 4656)

5.3 *Utility of* SEA *in* JHotDraw *evolutions analysis*

As shown in the two last paragraphs, the application of stereotypes <<evolution_history>> and <<version>> adds explicit information about each version of *JHotDraw* (the thing that UML does not provide). This information makes packages that includes different versions being more expressive and informative about how *JHotDraw* (or any other software) evolves. This helps software developers to analyze easily the growth of *JHotDraw* at design phase and to take the right design decisions.

Let's assume a scenario in which a software developer analyzes information marked explicitly on packages that represents *JHotDraw* versions (each package is stereotyped as <<version>>):

***JHotDraw* evolution analysis scenario based on *SEA* application** We associate to each tagged value a metric from those presented above:

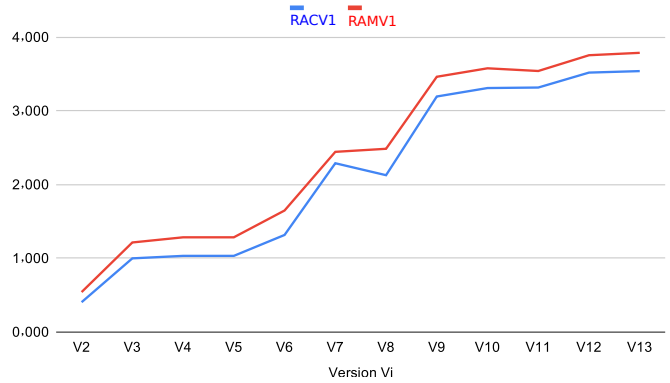
- *number_of_total_data_types* is represented by CNV_i
- *number_of_total_functions* is respresented by MNV_i
- *number_of_added_data_types* is represented by ACV_1
- *number_of_added_functions* is respresented by AMV_1

The table shown in Fig. 13(a) resumes values that are calculated basing on data collected from packages stereotyped as <<version>>. It shows variations of $RACV_1$ and $RAMV_1$ for each version V_i. The curve shown in Fig. 13(b) visualizes the progress of $RACV_1$ and $RAMV_1$ during 11 years. We remark that in all versions $RAMV_1$ is greater than $RACV_1$ which means that the percentage of added methods in each version comparing to initial number of methods in V_1 is greater than the percentage of added classes comparing to initial number of classes in V_1.

The fact that $RAMV_1$ is always greater than $RACV_1$ during 11 years of *JHotDraw* evolution indicates that maintenance tasks are applied more on methods. We could deduce in this case that software developers should recommend to use Visitor design pattern if it is not already used in *JHotDraw*.

V_i	$RACV_1$	$RAMV_1$
V_2	0.4	0.54
V_3	1	1.21
V_4	1.03	1.28
V_5	1.03	1.28
V_6	1.31	1.64
V_7	2.29	2.44
V_8	2.12	2.48
V_9	3.19	3.46
V_{10}	3.31	3.57
V_{11}	3.31	3.54
V_{12}	3.52	3.75
V_{13}	3.54	3.78

(a) Rates of added classes and methods number in a version Vi comparing to first version of JHoDraw $V1$.



(b) Rates evolution of added classes and methods number in a version V_i comparing to first version of JHoDraw V_1.

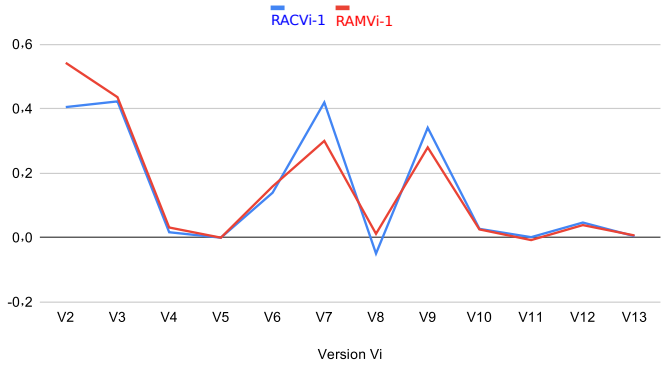
Figure 13: Study of Rates evolution of added classes and methods number in a version Vi comparing to first version of JHoDraw $V1$.

Let's now interpret evolutions from one version to its successor version. The table illustrated in Fig. 14(a) shows variations of $RACV_{i-1}$ and $RAMV_{i-1}$. We remark that $RACV$ of version V_2 is 0.4 and its $RAMV$ is about 0.54 which means that version V_2 of *JHotDraw* grows by 40% of V_1 classes number and 54% of V_1 methods number. Hence, maintenance tasks that switches V_1 to V_2 are applied more on methods (functions decomposition axis). This case occurs 5 times against 7 times in which maintenance tasks are applied more on classes (data-types decomposition axis). So, the switch from version to version shows that maintenance tasks change their behavior: they are applied 5 times with respect to functions decomposition axis against 7 times with respect to data-types decomposition axis. In this case, Composite design pattern seems to be the best choice for the future versions.

Lessons By analyzing *JHotDraw* evolutions as done above, we deduce that long-term evolution shows that Visitor design pattern is the best choice to optimize maintenance cost. But, variation in targeted decomposition axis due to the switch from version to version indicates that *JHotDraw* should change its architecture in order to keep maintenance being modular (software developers could use automatic reversible transformations among Composite and Visitor each time they remark that maintenance tasks would change behavior when switching from version to version).

V_i	$RACV_{i-1}$	$RAMV_{i-1}$
V_2	0.4	0.54
V_3	0.42	0.436
V_4	0.017	0.031
V_5	0.0	0.0
V_6	0.14	0.15
V_7	0.42	0.3
V_8	-0.049	0.012
V_9	0.34	0.28
V_{10}	0.027	0.025
V_{11}	0.002	-0.007
V_{12}	0.047	0.039
V_{13}	0.004	0.006

(a) Rates of added classes and methods number in a version V_i comparing to version V_{i-1}.



(b) Rates evolution of added classes and methods number in a version Vi comparing to version V_{i-1}.

Figure 14: Study of Rates evolution of added classes and methods number in a version Vi comparing to version V_{i-1}.

6 Conclusion and future work

We have defined an UML profile called *SEA* which could help software developers to analyze and predict maintenance tasks in early activities of software process. To reify our approach we defined some stereotypes that support treating maintenance issue in design phase of software development. *SEA* treats three main issues related to software evolution:

- Analyzing and predicting maintenance tasks by defining stereotypes that allow to trace software evolution history. They permit also to precise software parts that are more probable to change in the future and those which could stay stable. This feature helps to recommend the right software architecture that supports changes without extra cost.
- Marking explicitly on class diagrams software parts that are

encapsulated in modules and those that are crosscutting. For example, in Composite pattern, components classes represent data-types modules and methods/functions represent crosscutting concern. This feature permits to precise visually and easily modular concerns and crosscutting concerns, the thing that permits to identify quickly which kind of maintenance tasks is preferred.

- Visualizing in design phase design patterns that will be implemented later. This feature enhances software comprehension and helps to assimilate the chosen design pattern use and properties before implement it (sometimes programmers face difficulties to understand some design patterns). This could also give an overview about maintenance tasks that could be modular with respect to a given design pattern and those which could not. In the last case, software developers could use appropriate tools to switch program structure into another design pattern that is more convenient (such as automatic reversible transformation among Composite and Visitor [13]).

SEA is partially validated by apply it on *JHotDraw* evolution. This partial validation shows that it is possible to analyze software evolution history and predict the kind of maintenance tasks that could occur in the future. Our approach validation still partial until applying all *SEA* stereotypes to *JHotDraw* and also to other case studies. As a future work, we look for developing an algorithm based on supervised data analysis and classification to exploit data provided by *SEA* in models and diagrams to predict automatically future maintenance tasks kind.

Conflict of Interest The authors declare no conflict of interest.

References

[1] M. M. Lehman, "Laws of software evolution revisited," in 5th European Workshop on Software Process Technology (EWSPT'96), volume 1149/1996 of *LNCS*, 108–124, Springer, 1996.

[2] L. Erlikh, "Leveraging legacy system dollars for e-business," in IT Professional,, 17–23, 2000.

[3] E. Gamma, R. Helm, R. Johnson, J. Vlissides, Design patterns: elements of reusable object-oriented software, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1995.

[4] D. L. Parnas, "On the criteria to be used in decomposing systems into modules," Commun. ACM, **15**, 1053–1058, 1972.

[5] G. M. Rama, N. Patel, "Software modularization operators," in IEEE International Conference on Software Maintenance (ICSM), 2010.

[6] C. Szyperski, Component Software: Beyond Object-Oriented Programming, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 2002.

[7] G. Kiczales, E. Hilsdale, "Aspect-oriented Programming," SIGSOFT Softw. Eng. Notes, **26**(5), 313–, 2001, doi:10.1145/503271.503260.

[8] A. Ajouli, "A Shadow Structure for Modularity of Java Program Evolution," in 2015 41st Euromicro Conference on Software Engineering and Advanced Applications, 39–42, 2015, doi:10.1109/SEAA.2015.28.

[9] P. Wadler, "The Expression Problem," 1998, note to Java Genericity mailing list.

[10] K. B. Bruce., "Some challenging typing issues in object-oriented languages." Electronic Notes in Theoretical Computer Science,, **82**, 2003.

[11] C. Clifton, T. Millstein, G. T. Leavens, C. Chambers, "MultiJava: Design rationale, compiler implementation, and applications," ACM Trans. Program. Lang. Syst., **28**, 517–575, 2006.

[12] J. Garrigue., "Programming with polymorphic variants." in ML Workshop,, 1998.

[13] A. Ajouli, J. Cohen, J.-C. Royer, "Transformations between Composite and Visitor Implementations in Java," in Software Engineering and Advanced Applications (SEAA), 2013 39th EUROMICRO Conference on, 25–32, 2013.

[14] J. Cohen, R. Douence, A. Ajouli, "Invertible Program Restructurings for Continuing Modular Maintenance," in Software Maintenance and Reengineering (CSMR), 2012 16th European Conference on, 347 –352, 2012, doi:10.1109/CSMR.2012.42.

[15] A. Ajouli, K. Henchiri, "MODEM: an UML profile for MODEling and Predicting software Maintenance before implementation," in 2019 International Conference on Computer and Information Sciences (ICCIS), 1–5, 2019, doi:10.1109/ICCISci.2019.8716421.

[16] R. Yongchang, X. Tao, L. Zhongjing, C. Xiaoji, "Software Maintenance Process Model and Contrastive Analysis," 2011, doi:10.1109/ICIII.2011.324.

[17] H.-J. Kung, C. Hsu, "Software Maintenance Life Cycle Model," 113 – 121, 1998, doi:10.1109/ICSM.1998.738499.

[18] I. Cote, M. Heisel, "A UML Profile and Tool Support for Evolutionary Requirements Engineering," 161 – 170, 2011, doi:10.1109/CSMR.2011.22.

[19] H. Marouane, C. Duvallet, A. Makni, R. Bouaziz, B. Sadeg, "An UML profile for representing real-time design patterns," Journal of King Saud University - Computer and Information Sciences, **30**(4), 478 – 497, 2018.

[20] J. Dong, S. Yang, "Visualizing design patterns with a UML profile," in HCC, 2003.

[21] G. M. Ana Garis, Daniel Riesco, "Defining the Proxy Design Pattern using UML Profile," in Software Engineering Project of National University of San Luis, 2006.

[22] K. Ngee Loo, S. Peck Lee, "Representing design pattern interaction roles and variants," **6**, 2010, doi:10.1109/ICCET.2010.5486125.

[23] J. Cohen, A. Ajouli, "Practical use of static composition of refactoring operations," in ACM Symposium On Applied Computing, 6 pages, Portugal, 2013.

[24] R. Malhotra, K. Lata, "An empirical study on predictability of software maintainability using imbalanced data," Software Quality Journal, **28**, 2020, doi:10.1007/s11219-020-09525-y.

[25] M. Gogolla, Unified Modeling Language, 3232–3239, Springer US, Boston, MA, 2009, doi:10.1007/978-0-387-39940-9_440.

[26] L. Fuentes, A. Vallecillo, "An introduction to UML profiles," UPGRADE, The European Journal for the Informatics Professional, **5**, 2004.

[27] "Eclipse Papyrus," https://projects.eclipse.org/projects/modeling.mdt.papyrus.

[28] L. Hochstein, M. Lindvall, "Combating architectural degeneration: a survey," Inf. Softw. Technol., **47**, 643–656, 2005.

[29] N. Tsantalis, A. Chatzigeorgiou, G. Stephanides, S. T. Halkidis, "Design Pattern Detection Using Similarity Scoring," IEEE Trans. Softw. Eng., **32**, 896–909, 2006, doi:10.1109/TSE.2006.112.

[30] E. Gamma, I. Informatik, "JHotDraw as Open-Source Project," http://www.jhotdraw.org/.

[31] "Source Forge: JHotDraw project," https://sourceforge.net/projects/jhotdraw/.

[32] K. Johari, A. Kaur, "Effect of Software Evolution on Software Metrics: An Open Source Case Study," SIGSOFT Softw. Eng. Notes, **36**(5), 1–8, 2011, doi:10.1145/2020976.2020987.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 552-558 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Using Supervised Classification Methods for the Analysis of Multi-spectral Signatures of Rice Varieties in Panama

Javier E. Sánchez-Galán[1,2], Fatima Rangel Barranco[3], Jorge Serrano Reyes[3,4], Evelyn I. Quirós-McIntire[5,2*CorrespondingAuthor:*], José Ulises Jiménez[6], José R. Fábrega[6,2,*]

[1]*Facultad de Ingeniería de Sistemas Computacionales (FISC), Universidad Tecnológica de Panama, P.O. Box 0819-07289 El Dorado, Panamá*

[2]*Sistema Nacional de Investigación (SNI), SENACYT, Panamá*

[3]*Facultad de Ingeniería Eléxctrica (FIE), Universidad Tecnológica de Panama, P.O. Box 0819-07289 El Dorado, Panamá*

[4]*Centro de Producción e Investigaciones Agroindustriales (CEPIA), Universidad Tecnológica de Panama, P.O. Box 0819-07289 El Dorado, Panamá*

[5]*Instituto de Investigación Agropecuaria de Panamá (IDIAP), Coclé, Panamá*

[6]*Centro de Investigaciones Hidráulicas e Hidrotécnicas (CIHH), Universidad Tecnológica de Panama, P.O. Box 0819-07289 El Dorado, Panamá*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 08 March, 2021
Online: 17 March, 2021

Keywords:
Chemometrics
Hyper-spectral Imaging
Machine Learning
Spectroscopy
Rice (Oryza sativa)

ABSTRACT

In this article supervised classification methods for the analysis of local Panamanian rice crops using Near-Infrared (NIR) spectral signatures are assessed. Neural network (Multilayer Perceptron-MLP) and Tree based (Decision Trees-DT and Random Forest-RF) algorithms are used as regression and supervised classification of the spectral signatures by rice varieties, against other crops and by plant phenology (days after planting). Also, satellite derived spectral signature is validated with a field collected spectral model. Results suggest that MLP networks, either for regression or classification, were more efficient (RMSE of 8.78 and 0.068, respectively) than either tree based methods to regress/classify the rice spectral signature (RMSE of 19.37, 19.09 and 0.979, respectively). The validation made using satellite derived spectral signatures resulted in MLP models with RMSE of 0.216 and 7.318, respectively, leaving room for further improvement of the models. This work aimed to present a practical example of the employment of recent supervised classification algorithms for the determination of regression and classification models from reflectance spectral signatures in local rice varieties.

1 Introduction

The electromagnetic spectrum interacts with matter in a balanced relationship that indicates that the light reaching a body will be proportional to the light that is transmitted (measured through transmittance), absorbed (measured through absorbance) and reflected (measured through reflectance) in the body per in a unit of time. These measurements provide a signal that can be plotted, thus visually showing the relationship between the incident and reflected radiation flux.

These measurements are generally called spectral fingerprints or spectral signatures. Their measurement in agricultural crops is of special interest in precision agriculture, since they can provide a look into the crop health and phenological state. This information allows to determine how acceptable are the growth and development of the crop at a given planting stage.

This paper is an extension of the work originally presented in the 2019 XLV Latin American Computing Conference (CLEI) [1]. It will address other ML methods to achieve supervised classification of rice spectral signatures in Panama.

Supervised classification Machine Learning (ML) methods, such as Artificial Neural Networks (ANN), are commonly used for the analysis of Near-Infrared (NIR) spectral signatures [2]–[3]. ANN allow a way to identify relevant characteristics from a set of inputs (in this case spectral signal) and map them to corresponding output targets (physical or logical values associated with the object of under study). The model is learned through weight adjustments between connected units, called neurons, which are arranged in different

*Jose R. Fabrega; jose.fabrega@utp.ac.pa

layers. ANN are well regarded method. Especially in agriculture to determine quality parameters [4]–[5]. Deep neural networks (DNNs), on the other hand a widely used in NIR spectral analysis [6]–[7]. DNNs are basically ANN with a vast number of hidden layers, which allow them to learn complex mapping relationships between inputs and outputs [8, 9].

Other ML methods rely on tree structures that partitions sample groups and the data set of features in order understand the relationship between labelled samples. Among these methods the most notable are: Decision Trees (DT) [10, 11] and Classification and Regression Trees (CART) and Random Forest (RF) [12]. The latter being of special interest since it is suited to understand massive data sets, by aggregating/averaging decision trees [13]. As ANNs and DNNs, tree based methods have been used to understand the relationships between sets of NIR spectra [14]–[15].

The objectives of this study are twofold: 1) Explore methods for the classification of spectral signatures of rice varieties, using networks and tree based algorithms; 2) Be able to train/test/validate a model that relates to field acquired spectral signatures with satellite signatures. The rest of this article is structured as follows: Section 2, presents the spectral database and the algorithms used for the classification. Section 3, shows the results obtained and discusses the implications to the model characterization. Section 4, provides overall conclusions of the article and focuses on suggesting future work that could be used as starting point for upcoming analysis.

2 Materials and Methods

2.1 Spectral Database

As described in [1], the spectral signatures were collected on-site from 3 different plots locations using a portable Spectroradiometer. The raw measurements were organized into a database consisting of signatures in the 350 nm to 1050 nm wavelength range (447 wavelengths points per signature). For each one of these 1453 signatures the database also contained information about the plot management, including crop variety, and more importantly *days after planting* (DAP), at which the spectral signature was collected.

The most prominent crop in the database is rice with 1348 signatures, with onion and pimento having less than 60 signatures, and other crops such as tomato, and maize having less than 15 signatures in total. Three local varieties of rice were more prominent in the plots: IDIAP-38, IDIAP 52-05 and FL137-11, the first two being experimental varieties from the Instituto de Investigaciones Agroindustriales de Panama (IDIAP), while the latter being a commercial variety.

Spectral signatures were normalized using unit variance scaling. Later, the total number of wavelengths (spectral features), was reduced from the original 447 dimensions by using Principal Components Analysis (PCA). Only the first 100 components were kept for each signature.

2.2 Satellite Derived Spectral Signature Database

This study made use of satellite images from *PlanetScope*, as a source of external signatures and among other reasons for models validation. The exact polygons, where field spectral signatures were acquired, were treated as a source for spectral signatures.

A successful comparison between the spectral signatures and their satellite derived counterpart was made by normalizing matching dates (or closest days available) in which there were also field measurements and by integrating (summarize) the reflectance values obtained in the field into the satellite 4 band format, as follows: blue (455 - 515 nm), green (500 - 590 nm), red (590 - 670 nm), and infrared (780 - 860 nm). Finally, all the satellite derived spectral signatures were organized into a database for later use.

2.3 Regression and Supervised Classification of Spectral Signatures

Since one of the goals of this study was to find a model capable of predicting the days after planting (DAP), just by using the spectral signatures two families of models were selected. The first being a (continuous) regression models family in which the targets were numerical values representing the DAP. The second being a (discrete) binary classification model family, which makes possible to make rice predictions from non-rice related signatures, coded as 0 and 1,

Three regression models were tested to predict the DAP of the signatures:

1. Multilayer Perceptron Regressor (MLPR)
2. Decision Tree Regression (DTR)
3. Random Forest Regression (RFR)

Two classification models were test to predict the rice and non-rice:

1. Multilayer Perceptron Classification (MLPC)
2. Random Forest Classification (RFC).

In all models, spectral signatures were divided as 70% and 30% for training and testing purposes, respectively.

2.4 Cost functions and Evaluation Metrics

Two cost functions were compared for all models, the Mean Square Error - MSE (1), and the Mean Absolute Error - MAE (2). Two measures were used to evaluate the performance of all models, the Root Mean Square Error - RMSE (3), and the coefficient of determination, also known as: R-squared - R^2 (4).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{1}$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2} \tag{4}$$

For binary classification models, evaluation of their performance was made via a different set of metrics. Generally, the performance

of these models are tied to the model capacity to provide true predictions: true positive (TP) and true negative (TN), but also, prediction errors or false prediction is accounted via: false positive (FP) and false negative (FN). These counts are organized into a confusion matrix, shown in Table 1.

Table 1: Theoretical confusion matrix between actual and predicted classes.

		Actual	
		Positive	**Negative**
Prediction	Positive	TP	FP
	Negative	FN	TN

Using the confusion matrix, metrics that use the true predictions and prediction errors can be determined, among them:

- Accuracy: a measure of the number of true predictions made by the model. That can be calculated with the following formula: $\frac{\#TP+\#TN}{\#TP+\#TN+\#FP+\#FN}$.
- Precision: a measure that can be used to evaluate the proportion of positive predictions that were correctly classified. It can be calculated using the formula: $\frac{\#TP}{\#TP+\#FP}$.
- Sensitivity (Recall): a measure that can be used to evaluate the proportion of the actual positive predictions (observations) that were correctly classified. It can be calculated with the formula: $\frac{\#TP}{\#TP+\#FN}$.
- F1 Score: a measure of the model accuracy, but considering both the precision and recall. In general, a model can have a bad F1 Score (closer to 0) or a good one (closer to 1). It can be calculated employing the formula: $2 * \frac{Precision*Recall}{Precision+Recall}$.

2.5 *Software Implementation*

All plots were made using matplotlib. Database normalization, dimensionality reduction, regressions and classification methods and their errors were implemented using the Scikit-Learn library in Python [16]. The trees structures generated by Decision Tree and Random Forest algorithms were plotted using *graphviz* (https://graphviz.readthedocs.io/).

3 Results and Discussion

3.1 *Characteristics of the Spectral Signatures*

Figure 1 shows the complete set of reflectance spectral signatures present in the database. Figure 2 shows the reflectance spectral response of the local varieties for the Juan Hombron plot.

Figure 2A, shows the spectral signal variation of the FL137-11 material, from 48 to 116 DAP. Figure 2B, shows the variation of the IDIAP-38 material in 13 and 19 DAP, and also the variation of the IDIAP 52-05 material from 13 to 116 DAP. In all cases, the spectral signature has less variation in the visible range (400-700 nm) and more activity in the near-infrared part of the spectrum (700-100 nm) as the DAP increases.

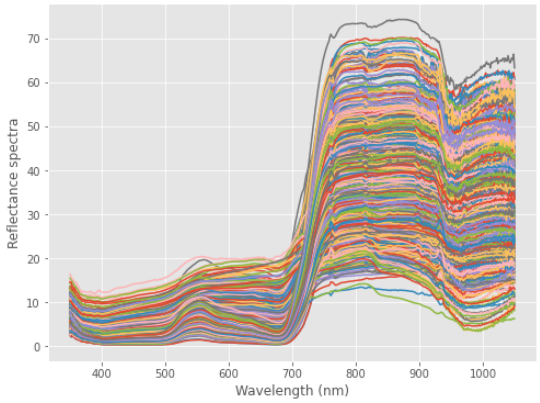


Figure 1: Rice Spectral Signatures in the Database

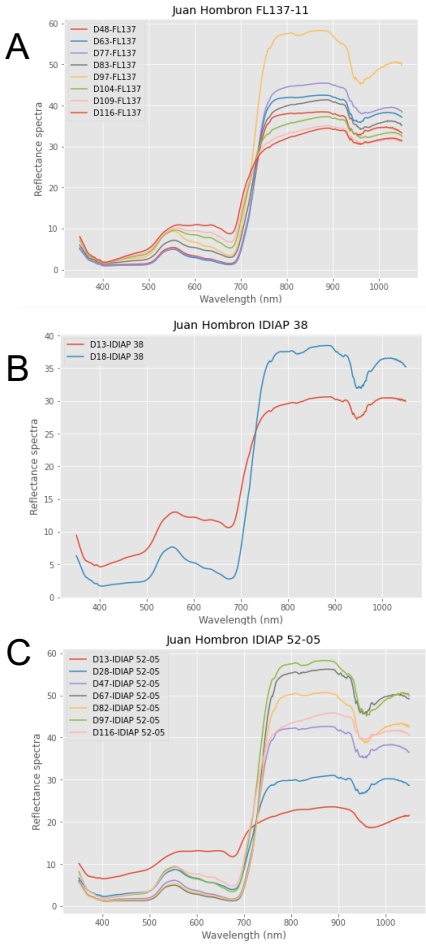


Figure 2: Spectral signatures of Rice varieties in Juan Hombron. A) FL137-11, B) IDIAP 38 and C) IDIAP 52-05.

Figure 3 shows the average spectra of rice (red) as compared against other crops in the database. It is notable that rice and onion have similar spectral signatures, with lower reflectance values in the NIR region. While tomato, maize and onion have a similar signature with higher reflectance values in the same range.

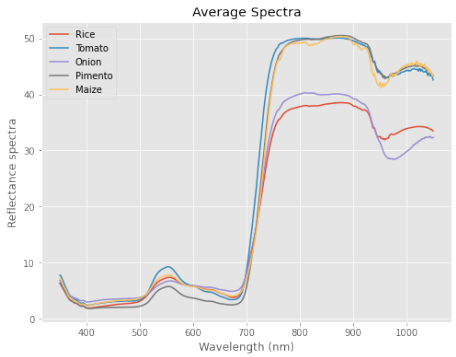


Figure 3: Comparison between Spectral Signatures of Crops in the Database

3.2 *Regression Models applied to Rice Spectral Signatures*

The Multilayer Perceptron Regressor network was set up to have 500 neurons in the hidden layer, tolerance of $1e^{-4}$ and using as ADAM as optimizer. For the Decision Tree Regressor, MSE was set to be the criterion, and the default settings were followed. The Random Forest Regressor was set up to build 50 trees (estimators), as in in the Decision tree, the default settings were followed. Error metrics for the three algorithms are shown in Table 2.

Table 2: Resulting Errors for Regression Models

Metrics	MLPR	DTR	RFR
R^2	**0.92**	0.56	0.57
MAE	6.148	10.807	13.783
MSE	77.095	375.400	364.779
RMSE	**8.780**	19.375	19.099

The MLPR model achieved the higher R^2 with 0.92 and the lower RMSE values with 8.78. The two tree based regression models showed to behave similarly with R^2 of 0.56 and 0.57, respectively. Also, having RMSE values of around 19, in both cases. Figure 4 shows the loss curve per iterations for the MLPR model, with a final loss of 0.08. The network training was stopped after 1492 iterations and resulted in an overall MLP training score of 0.99.

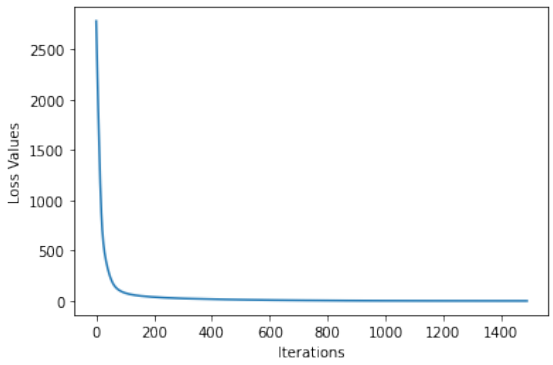


Figure 4: Loss Curves for MLPR Model

3.3 *Classification Models applied to Rice Spectral Signatures*

The Multilayer Perceptron Classifier network was set up to have 200 neurons in the hidden layer, tolerance of $1e^{-4}$ and using ReLU as an activation function. As for the Random Forest Classifier was set up to build 50 trees (estimators), here also, the default settings were followed. Error metrics for the two algorithms are shown in Table 3.

Table 3: Resulting Errors for Classification Models

Metrics	MLP Classifier	RFC
R^2	**0.94**	0.47
MAE	0.005	0.039
MSE	0.005	0.039
RMSE	**0.068**	0.197
Acc.	**0.995**	0.961
F1 Score	**0.998**	0.979

The MLPC model achieved a better R^2 than the RFC, with a value of 0.94 and the lower RMSE of the two with a value of 0.006. The two algorithms perform in the 90% having accuracy values of 0.99 and 0.96, respectively. Also having high F1 scores, with values of 0.99 and 0.97, respectively. Figure 5 shows the loss curve per iterations for the MLPC model, with a final loss of 0.006, after 166 iterations.

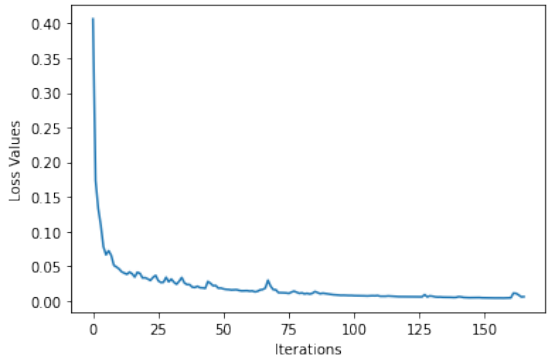


Figure 5: Loss Curve for MLPC Model

Figure 6 shows the final tree structure of the RFC. In warmer colors it shows the statistically significant decision nodes. In general, the RFC models were not as accurate as MLPC models, but they help us determine the most relevant wavelengths for classification, Table 4 provides a list of the Top-5 wavelengths that could be used as to classify the signatures. Interestingly, all these wavelengths are in the visible range between 400 and 600 nm.

A partial conclusions that can be addressed is that when predicting the DAP from the spectral signatures, it is very important to reduce the parameters to increase the precision of the model. It is important to notice that also, that when classifying crops (rice or other), there was no need to apply PCA since the model is quite accurate and has fewer losses. It even starts at a loss value of 0.4, close enough to 0, in the first iterations, as shown in Figure 5.

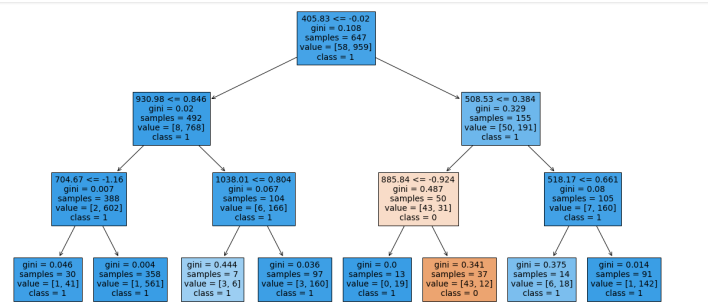


Figure 6: Resulting Tree and Rules learned by the Random Forest Classifier

Table 4: Summary of Top-5 Predictors from RFC model

ID	Wavelength	Importance (P-value)
111	534	0.031
95	508	0.028
118	545	0.023
34	409	0.023
139	579	0.028

3.4 Regression and Classification of Satellite Derived Signatures

The database for this experiment was comprised of the satellite spectral signatures and also the field (on site) spectral signals. The satellite spectral signatures are the ones compared using the Spectral Angle Mapper (SAM) method, as described in [1]. As for the (on site) spectral signals they had to be integrated to four wavelength bands that are used to represent satellite images. Table 5, shows the distribution of the spectral signature in the expanded database, with both spectral signatures collected on site and from the satellite image.

Table 5: Distribution of Spectral Signatures

Signature Type	Quantity
Rice spectral signatures (on site)	459
Rice spectral signatures (from satellite)	1303
Forest & road spectral signatures (from satellite)	337

A neural network model capable of distinguishing between the rice, and not-rice (namely, forest and road satellite signatures) was trained using a Multilayer Perceptron Classifier. The network was set up to have 600 neurons in the hidden layer, tolerance of $1e^{-4}$ and using ReLU as an activation function. Figure 7A shows the loss curve for the MLPC model, having a final loss value of 0.11 after 733 iterations.

Subsequently, a model was trained with the field spectral signatures reduced to 4 bands and the satellite spectral signatures. The field collected spectral signatures were used for training and testing (calibration), while the satellite spectral signatures were used for validation of the model. A Multilayer Perceptron Regressor was used. The network was set up to have 1000 neurons in the hidden layer, tolerance of $1e^{-4}$ and using as ADAM as optimizer.

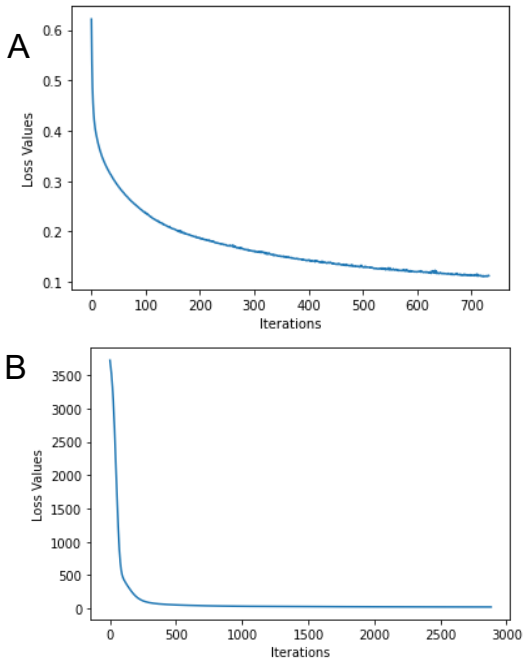


Figure 7: Model Loss Curve fror the MLPC (A) and MLPR (B) methods

Figure 7A shows the loss curve for the MLPC model, having a final loss value of 0.11 after 733 iterations. Figure 7B shows the loss curve of the MLPR model, it had final loss value of 23.74 after 2884 iterations.

As it is shown in Table 6, the MLPR model achieved the highest R^2 with 0.88 and an overall score of 0.90. However, the MLPC model had a lower RMSE value with 0.216.

Table 6: Resulting Errors for Satellite Derived Models

Metrics	MLPC	MLPR
R^2	0.711	**0.886**
MAE	0.046	53.564
MSE	0.046	5.101
RMSE	**0.216**	7.318

The validation of the model, using satellite spectral signatures resulted in having a R^2 of 0.54, a MAE of 14.998, MSE of 413.237 and RMSE of 20.3282.

A result from this subsection is that when validating the model to predict days after planting , the error is still quite large. It might be possible to make it lower by having a larger amount of field spectral signatures, since this would give more information to the model longer learning time .

4 Conclusions

The aim of this work was to be able to train supervised classification Machine Learning models in order to further understand rice reflectance signatures. A database of spectral signatures collected using a portable spectroradiometer and a set of satellite image derived signatures were employed as input for these models. As output or targets for the training of these models, few variables were

used, among them: days after planting (DAP), and the possibility of distinguishing rice from other crops or from other geographical features.

In order to develop these regression and classification models, neural network based algorithms (Multilayer Perceptrons) and tree based algorithms (decision trees and random forest), were used.

For both cases, regression and classification of rice spectral signatures, the use of Multilayer Perceptron models are recommended. Our results suggest that MLPR is the best method to make a model able to regress on the DAP just from the spectral signatures. For the classification of rice and non-rice signatures MLPC was one again the best resulting method. The tree based models, although providing less useful models, were able t suggest a list of most relevant predictors and a visualization of their results. It seems that the most useful wavelengths to build this models are found mostly in the 500-600 nm wavelength range, which corresponds to the green and yellow colors of the visible spectrum. These wavelengths are in line to changes in the panicle, as described by [17].

For the work on satellite derived signatures both MLPC and MLPR show interesting results over 0.7 and 0.88, respectively. However, the error metrics on the MLPR model are quite large. A large error that can be attributed to normalization effect between satellite and on site measurements, and total reflectance correction that exist between both spectral signatures. Suggesting that more samples should be added to the model to produce a more robust model.

One of the limitations of this study was the relatively small size of the database of spectral signatures employed. It is known that Machine Learning algorithms need more samples in order to have robust results. The need of more spectral signatures is evident by the poor generalization obtained, especially for the validation of satellite derived spectral signals.

For future work, few changes to the approach employed should be done. For instance, having different percentages in the training and testing splits. This work shows results only with 70%/30% split. It could be beneficial to test the limit of prediction by changing the percentages and/or using cross-validation techniques. Second, a class imbalance is evident in the spectral database, due to the focus on rice signatures by the original objective , leaving other crops or other spectral signatures astray. In order to make this database more balanced pseudo replicates need to be made from the non-rice signatures. In [18], the author presents an interesting approach for feature selection and replication using Variational Auto-Encoder (VAE) type of network, that could be pursued. Also in this line, the author presents a complete framework for Data Augmentation (DA) approach for convolutional neural network (CNN) based deep learning chemometrics [19], that seems worth exploring for this application. Third, recent literature shows a shift from deriving spectral signals from satellite images, to working directly over the Hyperspectral images (HSI) [20]. Even using deep learning frameworks for identifying soil characteristics [21], rice varieties [22], rice phenology [23], and even focusing on data imbalance in the image of domain [24].

The value of this work rest about on the fact that it presents a practical example of using supervised classification algorithms for the determination of regression and classification models from reflectance spectral signatures of a local crop. This study provides a basis for future works in deep learning based on chemometrics, and in particular the possibility of predicting crop conditions and characteristics from spectral signatures or satellite images.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The authors wish to thank the Secretaría Nacional de Ciencia, Tecnología e Innovación de Panamá (SENACYT) for its support through the project IDDS 15-184. The Sistema Nacional de Investigación (SNI) of SENACYT supports research activities by E.Q-M., J.R.F. and J.E.S.-G.. The authors acknowledge administrative support provided by Universidad Tecnológica de Panamá (UTP).

References

[1] J. E. Sánchez-Galán, J. S. Reyes, J. U. Jiménez, E. I. Quirós-McIntire, J. R. Fábrega, "Supervised Classification of Spectral Signatures from Agricultural Land-Cover in Panama Using the Spectral Angle Mapper Algorithm," in 2019 XLV Latin American Computing Conference (CLEI), 1–7, IEEE, 2019, doi: 10.1109/CLEI47609.2019.235101.

[2] W. Wang, J. Paliwal, "Generalisation performance of artificial neural networks for near infrared spectral analysis," Biosystems Engineering, **94**(1), 7–18, 2006, doi:10.1016/j.biosystemseng.2006.02.001.

[3] X. Meng, X. Meng, "The BP Neural Network Design Applied on the Classification of the Apples," in 7th International Conference on Education, Management, Information and Mechanical Engineering (EMIM 2017), Atlantis Press, 2017, doi:10.2991/emim-17.2017.37.

[4] A. C. Mutlu, I. H. Boyaci, H. E. Genis, R. Ozturk, N. Basaran-Akgul, T. Sanal, A. K. Evlice, "Prediction of wheat quality parameters using near-infrared spectroscopy and artificial neural networks," European food research and technology, **233**(2), 267–274, 2011, doi:10.1007/s00217-011-1515-8.

[5] M. Abdipour, M. Younessi-Hmazekhanlu, S. H. R. Ramazani, et al., "Artificial neural networks and multiple linear regression as potential methods for modeling seed yield of safflower (Carthamus tinctorius L.)," Industrial crops and products, **127**, 185–194, 2019, doi:10.1016/j.indcrop.2018.10.050.

[6] B. T. Le, "Application of deep learning and near infrared spectroscopy in cereal analysis," Vibrational Spectroscopy, **106**, 103009, 2020, doi:10.1016/j.vibspec.2019.103009.

[7] X. Zhang, T. Lin, J. Xu, X. Luo, Y. Ying, "DeepSpectra: An end-to-end deep learning approach for quantitative spectral analysis," Analytica chimica acta, **1058**, 48–57, 2019, doi:10.1016/j.aca.2019.01.002.

[8] Q. V. Le, "A Tutorial on Deep Learning Part 1: Nonlinear Classifiers and The Backpropagation Algorithm," Google Brain, 2015.

[9] Q. V. Le, "A tutorial on deep learning part 2: Autoencoders, convolutional neural networks and recurrent neural networks," Google Brain, 1–20, 2015.

[10] S. L. Salzberg, "C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993," Machine Learning, **16**(3), 235–240, 1994, doi:10.1007/bf00993309.

[11] C. Kingsford, S. L. Salzberg, "What are decision trees?" Nature biotechnology, **26**(9), 1011–1013, 2008, doi:10.1038/nbt0908-1011.

[12] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, "Introduction To Tree Classification," in Classification And Regression Trees, 18–58, Routledge, 2017, doi:10.1201/9781315139470-2.

[13] G. Biau, E. Scornet, "A random forest guided tour," Test, **25**(2), 197–227, 2016, doi:10.1007/s11749-016-0481-7.

[14] M. Clavaud, Y. Roggo, K. Dégardin, P.-Y. Sacré, P. Hubert, E. Ziemons, "Global regression model for moisture content determination using near-infrared spectroscopy," European journal of pharmaceutics and biopharmaceutics, **119**, 343–352, 2017, doi:10.1016/j.ejpb.2017.07.007.

[15] F. B. de Santana, W. B. Neto, R. J. Poppi, "Random forest as one-class classifier and infrared spectroscopy for food adulteration detection," Food chemistry, **293**, 323–332, 2019, doi:10.1016/j.foodchem.2019.04.073.

[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, **12**, 2825–2830, 2011.

[17] J. Serrano, J. Fábrega, E. Quirós, J. Sánchez-Galán, J. U. Jiménez, "Análisis prospectivo de la detección hiperespectral de cultivos de arroz (Oryza sativa L.)," KnE Engineering, 69–79, 2018, doi:10.18502/keg.v3i1.1414.

[18] Q. Wang, L. Li, X. Pan, H. Yang, "Classification of Imbalanced Near-infrared Spectroscopy Data," in 2020 12th International Conference on Advanced Computational Intelligence (ICACI), 577–584, IEEE, 2020, doi: 10.1109/icaci49185.2020.9177516.

[19] E. J. Bjerrum, M. Glahder, T. Skov, "Data augmentation of spectral data for convolutional neural network (CNN) based deep chemometrics," arXiv preprint arXiv:1710.01927, 2017.

[20] F. Arias, , M. Zambrano, K. Broce, C. Medina, H. Pacheco, Y. N. and, "Hyperspectral imaging for rice cultivation: Applications, methods and challenges," AIMS Agriculture and Food, **6**(1), 273–307, 2021, doi:10.3934/agrfood.2021018.

[21] J. Padarian, B. Minasny, A. McBratney, "Using deep learning to predict soil properties from regional spectral data," Geoderma Regional, **16**, e00198, 2019, doi:10.1016/j.geodrs.2018.e00198.

[22] S. Weng, P. Tang, H. Yuan, B. Guo, S. Yu, L. Huang, C. Xu, "Hyperspectral imaging for accurate determination of rice variety using a deep learning network with multi-feature fusion," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 118237, 2020, doi:10.1016/j.saa.2020.118237.

[23] M. Liu, X. Liu, L. Wu, X. Zou, T. Jiang, B. Zhao, "A modified spatiotemporal fusion algorithm using phenological information for predicting reflectance of paddy rice in southern China," Remote Sensing, **10**(5), 772, 2018, doi: 10.3390/rs10050772.

[24] J. F. R. Rochac, N. Zhang, L. Thompson, T. Oladunni, "A Data Augmentation-assisted Deep Learning Model for High Dimensional and Highly Imbalanced Hyperspectral Imaging Data," in 2019 9th International Conference on Information Science and Technology (ICIST), 362–367, IEEE, 2019, doi: 10.1109/icist.2019.8836913.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 488-498 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Novel Approach for Evaluating Eddy Current Loss in Wind Turbine Generator Step-Up Transformers

Bonginkosi Allen Thango*, Jacobus Andries Jordaan, Agha Francis Nnachi

Department of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 03 February, 2021
Online: 17 March, 2021

Keywords:
Wind energy
Transformer
Eddy currents
Harmonics
Temperature rise

ABSTRACT

South Africa is aiming to achieve a generation capacity of about 11.4GW through wind energy systems, which will contribute nearly 15.1% of the country's energy mix by 2030. Wind energy is one of the principal renewable energy determinations by the South African government, owing to affluent heavy winds in vast and remote coastal areas. In the design of newfangled Wind Turbine Generator Step-Up (WTGSU) transformers, all feasible measures are now being made to drive the optimal use of active components with the purpose to raise frugality and to lighten the weight of these transformers. This undertaking is allied with numerous challenges and one of them, which is particularly theoretical, is delineated by the Eddy currents. Many times the transformer manufacturer and also the buyer will be inclined to come to terms with some shortcomings triggered by Eddy currents. Still and all, it is critical to understand where Eddy currents emanate and the amount of losses and wherefore the temperature rise that may be produced in various active part components of WTGSU transformers. This is the most ideal choice to inhibit potential failure of WTGSU transformers arising from excessive heating especially under distorted harmonic load conditions. In the current work, an extension of the author's previous work, new analytical formulae for the Eddy loss computation in WTGSU transformer winding conductors have been explicitly derived, with appropriate contemplation of the fundamental and harmonic load current. These formulae allow the distribution of the skin effect and computation of the winding Eddy losses as a result of individual harmonics in the winding conductors. These results can be utilized to enhance the design of WTGSU transformers and consequently minimize the generation of hotspots in metallic structures.

1. Introduction

South Africa remains engaged in diversifying the power mix that will make it possible to reduce the subjection to coal power generation. The ongoing decommissioning of various coal power plants resulting from end of service life, can potentially make way for a radically distinct power mix as opposed to the current monopolization by coal power plants which have an installed capacity of 40GW [1]. At the present day, South Africa's Department of Energy through the Renewable Energy Independent Power Producer's Procurement Programme (REI4P) has deployed a total of 64 renewable energy facilities, of which 25 are wind energy constituting of 961 wind turbines with a total generation capacity of 2,1 GW [2]. Wind energy currently contributes 52% of the country's renewable energy capacity [3]. The wind energy tariff is currently at about R0.62/kWh, in which is about 45% less than the tariff for a coal plant. Each turbine generator in a wind energy facility is furnished with a step-up transformer at the bottom of the tower as shown in Figure 1 (a), which transforms its output voltage to a desired medium voltage (MV) collector level. In South Africa, these Wind Turbine Generator Step-Up (WTGSU) transformers are failing at an unnerving rate and utility owners are confound to ascertain the potential causes for these prevailing deficiencies. A typical WTGSU transformer failure arising from an insulation failure which is attributable to a disc-to-disc short circuit failure to generate a spark that ignites the transformer oil as shown in Figure 1 (b).

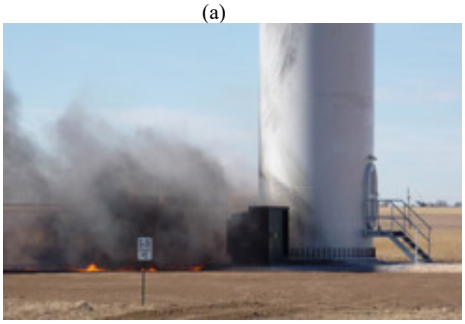
Field experience is indicating that these failures are triggered by the use of regular distribution (RD) transformers that cannot meet the required operational requirements of the wind energy environment. With the objective to pledge future dependability of wind energy in South Africa, WTGSU transformers must integrate these requirements particularly with regard to a deformed loading

*Corresponding Author: Bonginkosi Allen Thango, thangotech@gmail.com

cycle, harmonics and distortion, de-rating, switching over-voltages and stray gassing.



(a)



(b)

Figure 1: (a) WTGSU transformer during service; (b) Ignition of a WTGSU transformer.

Nonlinear loads and sporadic wind turbine generators are infamous for producing harmonic load currents and call for WTGSU transformers to connect them safely to the electrical network [4-5]. Transformers are commonly designed under the presumption of sinusoidal loading conditions [6-7]. It is widely known that distorted harmonic load currents accentuates the service losses in transformer and ultimately the temperature rise and generation of hotspots in metallic structures, resulting in degradation of insulation materials and untimely failures as shown in Figure 1 (b) [8-10].

Transformer windings are conceivably the most crucial part of WTGSU transformers, electrically and mechanically, and they are one of the major contribution of untimely failures in WTGSU transformers [11]. The contribution of winding conductors to WTGSU transformer failures is about 80%, which is a significant portion among the failures caused by other transformer parts [11]. In view of the increasing proclivity of nonlinear loads in the power system, the WTGSU transformers are more notably susceptible to fail. As a result, to minimize power losses and excessive heating in winding conductors, magnetic shielding [12-13], and laminating the core material [14-15] has been proposed in the literature.

At tender stage, it is essential to evaluate the transformer winding stray losses triggered by the EMFs. In this way, the inception of new formulae is recommendable and endeavor to enhance the computation and evaluation methods. A number of studies have been undertaken recently which take these objectives into account [16-20]. Thus, analytical techniques have demonstrated to be efficacious in computing the electromagnetic service losses. In a broader sense, there are two commonly used techniques in the transformer manufacturing industry to evaluate the winding Eddy current losses: (i) direct computation of the EMF using Maxwell's equations in [21-22] and (ii) use of Poynting's theorem in [21-22]. Early attempts to compute the EMF using both these methods are presented by authors in [23]. In the book Transformer Engineering: Design and Practice [24], the authors described these methods and how they can be embedded into manufacturers internal design programs to compute the Eddy currents. Design programs based on the Finite element Method to compute the Eddy currents are witnessed in the publications [25-28].

Owing to the increasing renewable energy market and applications of nonlinear loads, the evaluation of the winding stray losses in WTGSU transformer necessitates the consideration of the harmonic load current spectrum (HLCS), with appropriate computation of the electromagnetic fields (EMFs). Consequently, it is essential to study the impact of the HLCS on the winding stray losses and associated temperature rise of active part components.

In the current work, an extension of the author's previous work [29], comprehensive analytical formulae to calculate the winding Eddy losses in the presence of harmonics by taking into account of the winding conductor dimensions, EMF and skin effect are derived. The acquired formulae ensure that the examination of the contribution of individual harmonics to the power losses is done. Therefore, these are effective formulae that can produce rapid results for electrical designers without the requirement of costly and high-class computational resources.

The remainder of this work is structured as follows. Section 2 outlines the challenges and electrical design considerations of WTGSU transformers. In Section 3, a winding Eddy loss formula is derived under normal operating conditions by taking into account the fundamental frequency, EMF, local flux density and winding conductor material properties. In section 4, a winding Eddy loss formula is derived under harmonic load conditions and takes into account the winding conductors, EMFs, skin effect, additionally, a new and simplified harmonic loss factor (HLF) is derived to account for the additional losses that will be seen by the WTGSU transformer during service. In Section 5, a case scenario of a WTGSU transformer supplying a distorted harmonic load is presented and corresponding losses are computed using proposed formulae. These results are then compared with the method for calculating transformer losses recommended by the IEC standard and simulation results obtained by Finite Element Method. Finally, the conclusion is presented in Section 6.

2. Electrical design concerns

A large portion of wind turbine transformers currently in service are afflicted with numerous electrical concerns. These are prevalently on account of the unpredictable wind speed round the clock during their service. These inefficacies are the main drivers of the premature failures of these transformers.

2.1. Distorted load cycle

Wind turbines depend to a large extent upon regional wind and other weather patterns, and their annual average capacity factor is close to 40% [30]. A majority of the power producers in the past projected the operational loading would be more than 50%. The comparatively light loading of WTGSU transformer present two

distinct and operationally momentous challenges that must be embedded into the design philosophy of WTGSU transformer.

The foremost challenge is that the WTGSU transformer's relatively low average capacity factor falsify purchasing adjudication and invalidate classical economic models for regular distribution transformers. On account of the ideational nature of wind energy facility developments in South Africa, the Engineering, Procurement and Construction (EPC) developer is most frequently concerned with a cheaper initial transformer price. The utility owner who takes over responsibility of the facility from the EPC take a greater interest in the Total ownership cost (TOC) of the WTGSU transformer over its designed service lifetime. These costs include the initial purchase cost of the transformer, the cost of service no-load and load losses and routine maintenance costs [31-32]. The TOC models developed by authors in [31-34] do not take into consideration the relatively low average capacity factor of WTGSU transformers. Bearing in mind that the loading cycle of WTGSU transformers is so distinct from regular distribution transformers, EPCs must be mindful of the obsolete loss capitalization formulae when evaluating the TOC for WTGSU transformers.

Another challenge is that the WTGSU transformers are susceptible to regular thermal load cycling on account of the sporadic characteristic of the wind turbines. The latter give rise to the generation of hotspots in the transformer active part components including core, windings, clamping structure, flitch plates, tank walls et cetera. Recurrent thermal stresses drives the immersion of nitrogen gas into the hot insulating oil and it is emancipated as the insulating oil cools down. This phenomena results in the production of combustible gaseous bubbles that can potentially generate partial discharges and consequently damage the cellulose insulation. Additionally, sporadic thermal cycling can expedite the aging of electrical connections within the transformer tank.

2.2. Harmonics and distortion

The transformer design principle is formed on the rationale of generating an alternating EMF from a steady sinusoidal power supply voltage to instigate the flow of current and voltage potential in the winding conductors across that EMF.

During service, a steady sinusoidal wave shape is not practical. In the power system, the voltage and current wave shapes are distorted from the theoretical sinusoidal wave shapes. In practical terms, the total harmonic distortion (THD) ranging between 1 and 2 is prevalent at the point of common coupling (PCC). The application of non-linear loads in the facility can further contribute to distortion of the voltage and current wave shapes. These accumulative distortions reiterate very cycle, embellishing peaks that mount the voltage and current wave shapes and arise at distinct frequencies from the fundamental frequency of 50 Hertz (Hz) as shown in Figure 2.

The adverse effects of harmonics is that they trigger an upsurge in the copper loss, winding Eddy loss and stray loss in other metallic components. Eddy current and circulating current generate overheating in the transformer active part components, which must be treated by a sufficient cooling system design to thermal aging and untimely failures of WTGSU transformers during service. The wind turbine generator output, much like other renewable energy sources is intermittent and will produce distorted harmonic wave shapes. In this regard, the WTGSU transformer is operated with solid state controls which furnish additional harmonics and distortion.

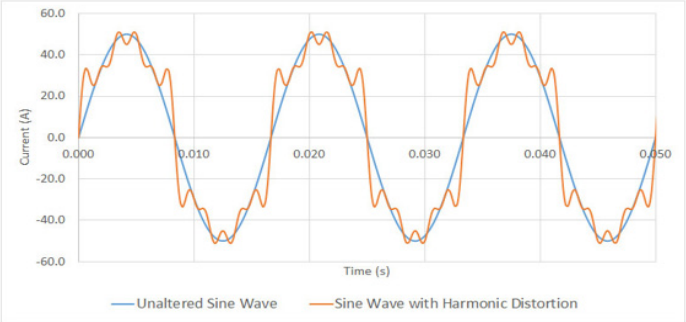


Figure 2: Current wave shapes: With and without distortion

Electrical generator systems that employ electronic switching devices and circuits also present specific problems for WTGSU transformers. From a manufacturer perspective, design philosophies to minimize the Eddy current loss to redress for harmonic load currents are imperative. While harmonic filtering is particularly not a function of a WTGSU transformer, magnetic shields strategically positioned between the windings may act as filters to mitigate the distorted harmonic currents into the collector bus. As a result, magnetic shielding should be deemed compulsory in the design of WTGSU transformers.

2.3. Switching Over-Voltages

A major concern with vast, wind turbine arrays is the necessity for connecting individual WTGSU transformers to the utility collector bus, leading to lengthy runs of cables. In large-scale wind facilities with a cluster of wind turbine towers, the WTGSU transformers may be positioned near the tower base as shown in Figure 1 in order to minimize the costs of large and lengthy runs of copper cables. Problems including excessive voltage drop, cable losses and risk of ground faults may arise from lengthy cable runs.

In view of the fact that wind energy facilities are in remote areas, they are more susceptible to lightning surges. As a result, surge arrestors must be mandatory for all WTGSU transformers.

Other things being equal, the biggest concern about switching over-voltages may be that generated by the wind turbines. On a 24-hours period, depending on the intermittent nature of wind; the wind turbines may be connected and disconnected online and offline when the wind speed increase and decline respectively. The latter may also occur when circuit breakers opens and closes the WTGSU transformer from the collector bus circuit. The switching phenomena by circuit breaker operation trigger transient over-voltages into the WTGSU transformers. This occurrence is aggravated by the application of vacuum circuit breakers, which have high-speed switching times. The fusion of the transient over-voltages and capacitance of the copper cables either HV or LV WTGSU transformer side may result in stationary waves and ringing that exceed the sinusoidal voltage amplitude as shown in Figure 3.

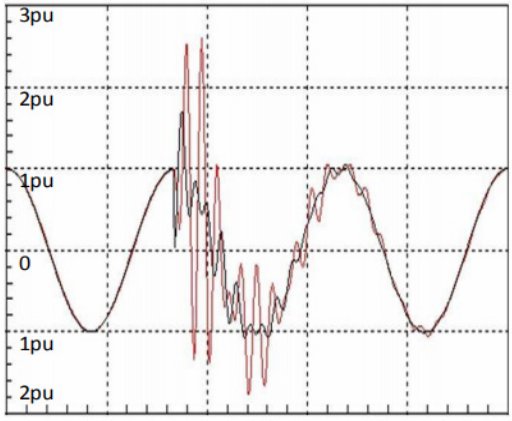


Figure 3: Typical switching surge [35]

Excessive transient over-voltages can trigger WTGSU transformer insulation failures. The transient over-voltages as shown in Figure 3 contains fast rise-time and high voltage amplitudes which clash with the resonant frequencies on the windings and can introduce electrical stresses exceeding the dielectric capability of the windings.

3. Winding Eddy loss under normal conditions

During service the winding conductor is immersed in an alternating magnetic field and subsequently the Eddy current start to flow. The energy loss dissipated as thermal energy on account of the Eddy current bring forth the Winding Eddy losses. The phenomena in rectangular winding conductors is demonstrated in Figure 4.

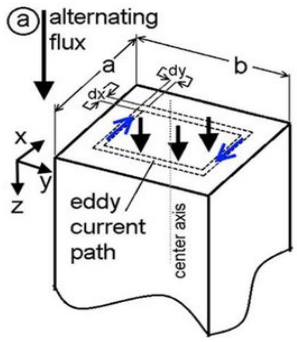


Figure 4: Harmonic current profiles [36]

The derivation of the winding Eddy loss formulae is reliant on factors suchlike frequency, local flux density and the winding conductor material properties. Provided that a winding conductor has the parameters as indicated in Figure 4 namely the, length, height and thickness are denoted by L, h and τ respectively. Then the alternating magnetic flux characteristic that is perpendicular to the height and thickness can be described by the eq. (1).

$$B = B_{max} \sin \omega t \tag{1}$$

Taking into consideration a closed-path in a clock-wise direction as shown in Figure 4, of thickness dx and distance x, a presentation of a single conductor with an induced voltage in closed-path is established. The area of this closed-path is expressed as follows in Eq. (2).

$$A = 2 \times h \tag{2}$$

The amount of magnetic flux passing through the unit area in eq. (2) is expressed as follows in Eq. (3).

$$\emptyset = B \times A \tag{3}$$

It follows then, that the amount of magnetic flux characteristic that is perpendicular to the height and thickness in the closed-path can be described by the eq. (4).

$$\emptyset = B_{max} \times \sin \omega t \times 2hx \tag{4}$$

The induced voltage in the closed-path can be expressed in accordance with the Faradays law of induction as expressed in Eq. (5).

$$E = 4.44 \times f \times \emptyset_{max} \tag{5}$$

Equating the constant 4.44=√2 π yields the formula in Eq. (6)

$$E = \sqrt{2}\pi \times f \times \emptyset_{max} \tag{6}$$

Taking into account the amount of magnetic flux passing through the closed-path unit area the eq. (6) yields the formula expressed in Eq.(7).

$$E = \sqrt{2}\pi \times f \times B_{max} \times A \tag{7}$$

Substituting the area of the closed-path in the induced voltage formula in Eq. (7) yields eq. (8).

$$E = \sqrt{2}\pi \times f \times B_{max} \times 2hx \tag{8}$$

The resistance of the closed-path can then be expressed as follows in Eq. (9).

$$R = \rho \times \frac{(2h + 4x)}{L} dx \tag{9}$$

The winding Eddy loss in the closed-path dx is then expressed as shown in Eq. (10). Considering that the thickness of the closed-path is extremely small in relation to the height, then the distance x I negligible.

$$dp = \frac{E^2 L}{\rho \times 2h} dx \tag{10}$$

Substituting eq. (7) into eq. (10) yields the expression in Eq. (11).

$$dp = \frac{\left[\sqrt{2}\pi \times f \times B_{max} \times 2hx\right]^2 L}{\rho \times 2h} dx \tag{11}$$

Equating the constant √2 π=4.44 in Eq. (11) yields the formula in Eq. (12).

$$dp = \frac{[4\pi \times f \times B_{max} \times 2hx]^2 L}{\rho \times 2h} dx \tag{12}$$

Removing the square and expressing each terms to the exponent 2 yields eq. (13).

$$dp = \frac{4\pi^2 \times f^2 \times {B_{max}}^2 \times hL}{\rho h} \quad (13)$$

Integrating both sides of Eq. (13) and considering the closed-path then the Eddy loss can be derived using Eq. (14).

$$P_{Eddy} = 4\pi^2 \times f^2 \times {B_{max}}^2 \times hL \int_{x=0}^{\frac{\tau}{2}} x^2\, dx \quad (14)$$

Subsequently, the solution can be expressed as follows in Eq. (15).

$$P_{Eddy} = \frac{\pi^2 \times f^2 \times {B_{max}}^2 \times \tau^2}{6\rho}(hL\tau) \quad (15)$$

Dividing the Eddy loss by hLτ equates to the area per unit volume as expressed in eq. (16).

$$P_{Eddy} = \frac{\pi^2 \times f^2 \times {B_{max}}^2 \times \tau^2}{6\rho} \quad (16)$$

Finally, the Winding Eddy loss at fundamental frequency is expressed as follows in Eq. (17).

$$P_{Eddy} = K_e f^2 {B_{max}}^2 t^2 \quad (17)$$

Here,

$$K_e = \frac{\pi^2}{6\rho}$$

The skin depth of penetration at fundamental frequency 50Hz may be calculated as follows in Eq. (18).

$$\delta_R = \sqrt{\frac{\rho}{\mu \times \pi \times f}} \quad (18)$$

The depth of penetration under HLC is then computed as follows in Eq. (19).

$$\delta = \frac{\delta_R}{\sqrt{h}} \quad (19)$$

In principle, the following conclusions can be drawn with regards to the Eddy current loss in relation to the area per unit volume of winding conductor based on this derivation:

- There is a direct proportion to the fundamental frequency, conductor dimension and local flux density.
- There is an inversely proportional relationship to the resistivity of the copper conductor.

4. Transformer winding Eddy loss weighting factor

4.1. Winding Eddy loss under harmonic current loading

The winding Eddy loss along a winding height (H) can be expressed as follows in Eq. (20) [37].

$$P_{winding} = P_{DC} + P_{WEC} \quad (20)$$

The standardized winding Eddy loss is then expressed by dividing eq. (18) with the rated copper loss to yield eq. (21).

$$\emptyset = B \times A \quad (21)$$

Taking into consideration that the copper loss has a directly proportional relationship with the root mean square (RMS) load current under harmonic conditions [38-40] then first part of eq. (21) yields Eq. (22).

$$\frac{P_{DC}}{P_{DC_R}} = \left(\frac{I_h}{I_R}\right)^2 = \frac{\sum_{h=1}^{h=max} h^2 {I_h}^2}{{I_R}^2} \quad (22)$$

The contribution of the second part of Eq. (20) is as expressed in Eq. (17). The magnetic flux leakage under harmonic conditions has a directly proportional relationship to the RMS load current as expressed in Eq. (23).

$$B = K_B I \quad (23)$$

Under harmonic current conditions, the current I is expressed in Eq. (24).

$$I = \sqrt{\sum_{h=1}^{h=max} {I_h}^2} \quad (24)$$

The harmonic current in Eq. (24) is expressed as I_h. As such, the Eq. (24) yields Eq. (25).

$$P_{WEC=}\, K \sum_{h=1}^{h=max} h^2 {I_h}^2 \quad (25)$$

Here, the constant K is expressed in Eq. (17).

During service, when the transformer is operating at fundamental frequency, the equivalent rated winding Eddy loss is expressed as shown in Eq. (26).

$$P_{WEC_R} = K {I_R}^2 \quad (26)$$

Combining Eq. (25) and Eq. (26), the winding Eddy loss for a transformer operating under harmonic current can be expressed as shown in Eq. (27).

$$P_{WEC=} P_{WEC_R} \sum_{h=1}^{h=max} h^2 \left(\frac{I_h}{I_R}\right)^2 \quad (27)$$

Subsequently, combining Eq. (22) and Eq. (27) yields Eq. (28). This equation is formulated on the premise that induced current is at fundamental load current I_R of the transformer during service.

$$\frac{P_{winding}}{P_{DC_R}} = \frac{\sum_{h=1}^{h=max} h^2 {I_h}^2}{{I_R}^2} + \frac{P_{WE_R}}{P_{DC_R}} \sum_{h=1}^{h=max} h^2 \left(\frac{I_h}{I_R}\right)^2 \quad (28)$$

Further simplification of Eq. (28) yields Eq. (29).

$$\frac{P_{winding}}{P_{DC_R}} = \frac{\sum_{h=1}^{h=max} {I_h}^2}{{I_R}^2}\left[1 + \frac{P_{WE_R}}{P_{DC_R}} \frac{\sum_{h=1}^{h=max} h^2 \left(\frac{I_h}{I_R}\right)^2}{\sum_{h=1}^{h=max} {I_h}^2}\right] \quad (29)$$

The allowance made in order to take into account of the harmonic load current in relation to their effects on transformer thermal performance is then expressed as follows in Eq. (30).

$$F_{HL} = \frac{\sum_{h=1}^{h=max} h^2 {I_h}^2}{\sum_{h=1}^{h=max} {I_h}^2} = \frac{\sum_{h=1}^{h=max} h^2 \left(\frac{I_h}{I_1}\right)^2}{\sum_{h=1}^{h=max} \left(\frac{I_h}{I_1}\right)^2} \quad (30)$$

In principle, the following conclusions can be drawn with regards to the Eddy current loss under harmonic current loading:

- The winding Eddy loss is directly proportional to the RMS load current and the harmonic order. This relationship is similar in the computation of the stray loss in tank walls, core clamps et cetera except that the harmonic order is expressed to an exponent of 0.8 [35].
- The skin effect is not taken into consideration at high harmonic order. The magnetic flux under these conditions does not completely penetrate the surface of the winding conductors.

In the next section, a weighted loss factor that considers the skin effect is derived.

4.2. Winding Eddy loss under harmonic current loading

The Eddy current problem form part to the area of quasi-stationary electromagnetic effects of conductors. Insomuch that, the displacement current enclosed by winding conductors may incessantly be ignored in relation to the conductive current. This is indeed the case even at high frequencies given that in practice only winding conductors comprising high electric conductivity are used. Eddy currents give rise to uneven dissemination of current density in a studied cross sectional area of a conducting conductor. This inherently leads to rise in joule heating as opposed to the state produced by direct current (DC). The Eddy currents and related uneven dissemination of the magnetic flux are known as the skin effect. The rise in current density give rise to resistive heating as opposed to the DC resistance as well as a reduction in the inductance. In addressing the skin effect problem, this study adopts the Maxwell equations in [38] and remodel these equations to treat the quasi-stationary electromagnetic effects of conductors. These equations are expressed as follows in Eq. (31) and Eq. (32).

$$curl\ \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{H}}{\partial t} \quad (31)$$

$$curl\ \boldsymbol{H} = \sigma \quad (32)$$

On the above equations an add-on of the Ohm's law as expressed in Eq. (33).

$$\sigma = \gamma E \quad (33)$$

The electric field intensity must fulfil Eq. (34).

$$div\boldsymbol{E} = 0 \quad (34)$$

Similarly, the magnetic field must fulfil Eq. (35).

$$div\boldsymbol{H} = 0 \quad (35)$$

In investigating the dissemination of current, small but sufficiently long conductors are considered. The location of the conductors in a coordinate system as demonstrated in Figure 5. The height and ratio of these conductors fulfil the condition $h \gg 2b$ and the spatial dependence on the y-coordinate may be ignored with insignificant error. When the current flows in the z-coordinate, the magnetic field has one component as described in Eq. (36).

$$\boldsymbol{H} = jH \quad (36)$$

In the same order, the electric field intensity has a single component as expressed in Eq. (37).

$$\boldsymbol{E} = kE \quad (37)$$

In the event that a winding conductor has the conductivity $\gamma\left[{}^{1}/_{\Omega m}\right]$, the permeability $\mu\left[{}^{1}/_{Hm}\right]$ and time-variant harmoniously at angular frequency $\omega\left[{}^{rad}/_{s}\right]$, by elimination the ordinary second-order differential Maxwell equations for magnetic field intensity vector is expressed as follows in Eq. (38).

$$\frac{d^2H}{dx^2} = k^2H \quad (38)$$

Here, the constant k is expressed as shown in Eq. (39).

$$k^2 = j\omega\gamma\mu \quad (39)$$

In investigating the skin effect in winding conductors with a spasmodic magnetic flux, the rectangular copper conductors in Figure 5 are considered.

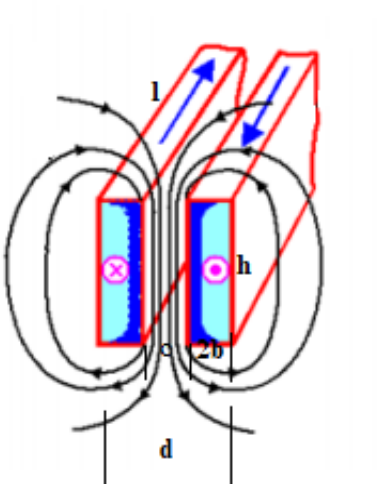


Figure 5: Two winding conductors immersed in alternating EMFs

The spasmodic magnetic flux streams along in the z-coordinate and the interaction of winding conductor 1 and 3 with respect to

winding conductor 2 is investigated in terms of the winding Eddy loss under harmonic conditions. The effect of the remainder of winding conductors will be disregarded in order to presume that the impact of winding conductor 1 and 3 will be completely offset in the center winding conductor. Presupposing a harmonic time response of the 2nd order differential equation in Eq. (38), with the initial conditions necessitated for symmetry along the edges of the winding conductors shall be valid the equation $H_{(x=b)} = H_{(x=-b)}$ and will yield the solutions in Eq. (40) and Eq. (41) as expressed below. It is evident that the solution in this case scenario is of $h \gg 2b$ nature.

$$H_0 = H_1 e^{kb} + H_2 e^{-kb} \tag{40}$$

$$H_0 = H_1 e^{-kb} + H_2 e^{kb} \tag{41}$$

Here, H_0 is the root mean square (RMS) magnetic field strength. Additionally, the condition of $x = \pm b$ is valid since the impact along the edges of the winding conductors $= \pm {}^{h}/_{2}$ is negligible if $x = \pm b$ is satisfied. The constants of integration from Eq. (40) and Eq. (41) can then be expressed as follows in Eq. (42).

$$H_1 = H_2 = \frac{H_0}{\cosh kb} \tag{42}$$

The solution of the magnetic field strength presume the format expressed in Eq. (43).

$$H = H_0 = \frac{\cosh kx}{\cosh kb} \tag{43}$$

By association, the magnetic flux density by Eq. (42) is expressed as shown in Eq. (44).

$$B = \mu H_0 = \frac{\cosh kx}{\cosh kb} \tag{44}$$

Maxwell's 1st equation in Eq. (33), can then be expressed as follows in Eq. (45).

$$\sigma = (curlH)_y = -kH_0 \frac{\sinh kx}{\cosh kb} \tag{45}$$

The constant k is introduced into the rationalization of the equilateral hyperbola as shown in Eq. (46).

$$k = \frac{1+j}{a} \tag{46}$$

The skin depth of penetration a under harmonic conditions may be calculated as follows in Eq. (47).

$$a = \sqrt{\frac{2}{\omega\gamma\mu}} \tag{47}$$

The skin depth at various harmonic orders is demonstrated in Figure 6.

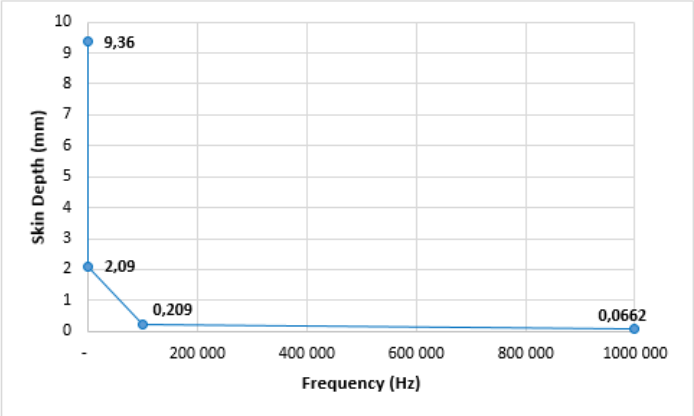


Figure 6: Depth of skin effect

The formulation of the current density is acquired and expressed as shown in Eq. (48).

$$\sigma = -\frac{1+j}{a} H_0 \frac{\sinh(1+j)\,{}^{x}/_{a}}{\cosh(1+j)\,{}^{b}/_{a}} \tag{48}$$

With the expansion of the equilateral hyperbola the modulus of the current density is expressed shown in Eq. (49).

$$|\sigma| = \frac{\sqrt{2}}{a} H_0 \sqrt{\frac{\cosh\left({}^{2x}/_{a}\right) - \cos\left({}^{2x}/_{a}\right)}{\cosh\left({}^{2b}/_{a}\right) + \cos\left({}^{2b}/_{a}\right)}} \tag{49}$$

The dissemination of the magnetic flux density and current density along the surface of the winding conductors is described by Figure 7.

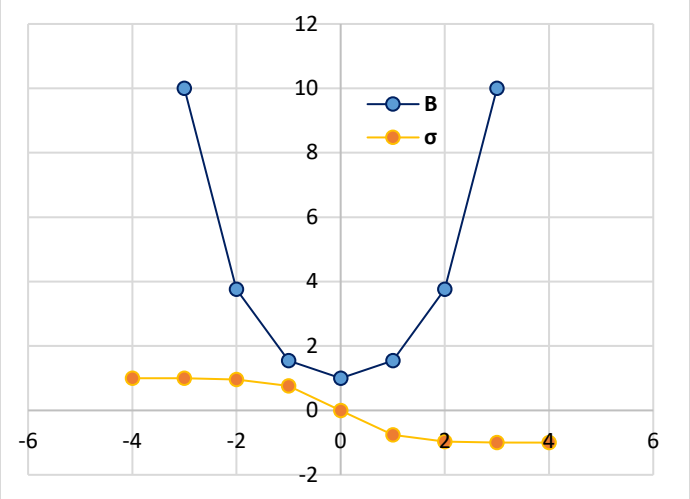


Figure 7: Distribution of the magnetic flux in the cross-section of a winding conductor

The magnetic flux leakage impinging upon one conductor is expressed by Eq. (50).

$$\Phi = \int_{-b}^{b} hBdx = \frac{2ah\mu}{1+j} H_0 \tanh(1+j)\frac{b}{a} \tag{50}$$

The peak value of the magnetic flux leakage is given by Eq. (51).

$$\sqrt{2}|\Phi| = 2ah\mu H_0 \sqrt{\frac{\cosh\left(2x/a\right) - \cos\left(2x/a\right)}{\cosh\left(2b/a\right) + \cos\left(2b/a\right)}} \quad (51)$$

Similarly, the magnetic flux density is expressed by Eq. (52).

$$B_{max} = 2\mu H_0 \frac{a}{2b} \sqrt{\frac{\cosh\left(2x/a\right) - \cos\left(2x/a\right)}{\cosh\left(2b/a\right) + \cos\left(2b/a\right)}} \quad (52)$$

By employing Eq. (49) and introducing Eq. (52), the formulation of the current density is expressed as shown in Eq. (53).

$$\frac{1}{2\gamma b}\int_{-b}^{b}|\sigma|^2 dx = \frac{{B_{max}}^2\gamma\omega^2 b^2}{6}\frac{3}{2b/a}\frac{\sinh\left(2b/a\right) - \sin\left(2b/a\right)}{\cosh\left(2b/a\right) - \cos\left(2b/a\right)} \quad (53)$$

The variable ב is introduced here to account for expression of $2b/a$ above as shown in Eq. (54).

$$ב = \frac{2b}{a} \quad (54)$$

Forthwith, the Eddy loss per unit of volume of a winding conductor can be calculated using Eq. (17), where the weighted factor will be employed as expressed in Eq. (55).

$$P = \frac{{B_{max}}^2\gamma\omega^2 b^2}{6} \times F(ב) \quad (55)$$

Here, the weighting function that takes into consideration the skin effect in winding conductors under harmonic conditions is expressed as follows in Eq. (56).

$$F_H = \frac{3}{ב} \times \frac{\sinh ב - \sin ב}{\cosh ב - \cos ב} \quad (56)$$

This function is presented graphically at different conductor dimensions as shown in Figure 8.

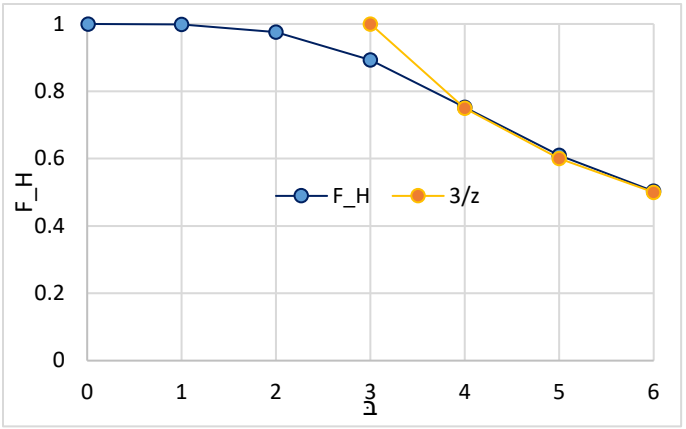


Figure 8: The function F_H and $\frac{3}{ב}$

It is evident from Figure 8, that at higher harmonic order, the magnetic flux leakage does not completely penetrate upon the surface of the conductors. The latter addresses the short coming of Eq. (30).

5. Materials and Methods

In this section, a WTGSU transformer rated at 5000 KVA, 11000/33000 volts, 3-phase, 50 cycles, oil-immersed, naturally cooled (ONAN), core type, double wound with copper conductor and fitted with on load tap changer is studied. The temperature rise in oil and winding conductors is 50 ºC and 55 ºC respectively. The rated losses under consideration are at 75ºC on normal Tap position (in Watts) as shown in Table 1.

Table 1: Rated transformer losses

Type of loss	Rated Losses (Watts)
No load	4 500
Copper	25 000
Winding Eddy	938
Other Stray	1 253
Total	31 691

The studied transformer has a full load low voltage (LV) winding current of 262,43A with a mean winding Eddy loss at the highest loss ratio under rated conditions at 15% of the Copper losses.

The load that will been seen by the studied WTGSU transformer is described as shown in Figure 9. It is preferential that this harmonic load spectrum (HLS) to which the WTGSU will be susceptible be indicated to the transformer manufacturer during bid stage. A precise evaluation for the appropriate sizing of WTGSU transformers can only be designed by estimating the particular HLS.

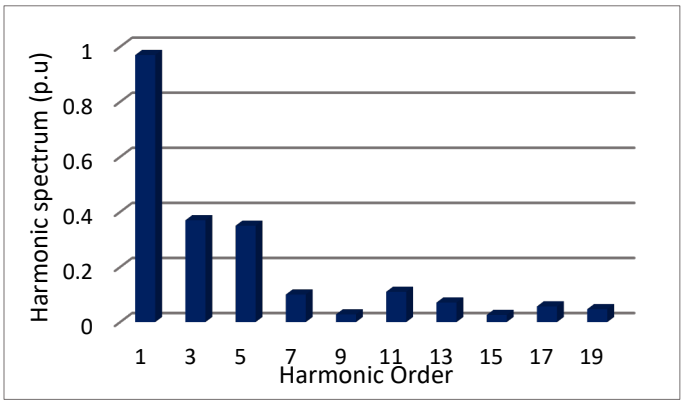


Figure 9: Harmonic Load Spectrum

The resultant HLFs and service losses are computed and discussed in the next section.

6. Results

6.1. Computation of Harmonic Loss Factor

It is handy to specify a single digit that may have recourse to evaluate the capability of the studied WTGSU transformer in facilitating power to the supplied harmonic spectrum above. This

value viz. the harmonic loss factor (HLF) illustrates the effective RMS heating on account of the harmonic spectrum. The calculation of the HLF is tabulated in Table 2. The HLF by FEM is computed using the procedure described in [38].

Table 2: Computation of HLFs.

h	I_h/I_1	$(I_h/I_1)^2$	IEC	Proposed	FEM
1	0,97	0,941	0,941	0,941	0,941
3	0,37	0,137	1,232	1,090	1,213
5	0,35	0,123	3,063	2,569	2,924
7	0,1	0,010	0,490	0,100	0,448
9	0,028	0,001	0,064	0,063	0,055
11	0,11	0,012	1,464	1,455	1,184
13	0,071	0,005	0,852	0,807	0,640
15	0,026	0,001	0,152	0,135	0,106
17	0,057	0,003	0,939	0,925	0,600
19	0,047	0,002	0,797	0,774	0,468
Total		1,234	9,994	8,860	8,576
$I_{rms\ (p.u)}$		1.111			
HLF			8,10	7,177	6,948

The ratio of the harmonic load current for each respective method and the RMS current (p.u) yields the HLFs as indicated previously.

6.2. Computation of Winding Eddy loss under harmonic current loading

The evaluation of the total service losses as a result of the HLC using the IEEE Std. C57.110-2018 is tabulated as shown in Table 3.

Table 3: Computation of total service losses: IEEE Std. C57.110-2018.

Type of loss	Rated Losses (W)	Load Losses (W)	HLF	Service Losses (W)
No load	4 500	4 500	-	4 500
Copper	25 000	30 859	-	30 859
Winding Eddy	938	1158	8,10	9 372
Other Stray	1 253	1547	1,58	2 442
Total	31 691	38 064		47 173

Under the supplied HLS, the service losses are witnessed to have a surge of about 33% and 19% from the rated and load losses respectively. The computation of the HLF for other stray loss in this work is similar to that of the winding Eddy loss as shown in Eq. (27), however these losses generated by metallic components such as tank walls, core clamp, windings, core, flitch plate et cetera are proportional to the square of the harmonic load current and the harmonic order to the exponent 0.8 as stated in [37]. The peak local load loss ratio for the HLC is computed as follows:

$$P_{LL}(p.u) = I^2{}_{rms(p.u)}\left(1 + F_{HL} \times P_{EC_R}\right) \quad (57)$$

$$P_{LL}(p.u) = 1.234 \times [1 + 0.15 \times 8.1] = 1.249 p.u$$

Given that the mean winding Eddy loss at the highest loss ratio under rated conditions is 15% of the Copper losses, then the load loss will be 1.15 p.u as shown below. The maximum permissible harmonic load current with the supplied harmonic spectrum is as follows:

$$I_{max}(p.u) = \sqrt{\frac{P_{LL_R(pu)}}{1+F_{HL}\times P_{EC_R(pu)}}} \quad (58)$$

$$I_{max}(p.u) = \sqrt{\frac{1.15}{1 + 0.15 \times 8.1}} \times 262{,}43 = 261{,}048A$$

Consequently, with the supplied harmonic spectrum the studied WTGSU transformer capability is about 261.048A.

Table 4: Computation of total service losses: Proposed method

Type of loss	Rated Losses (W)	Load Losses (W)	HLF	Service Losses (W)
No load	4 500	4 500	-	4 500
Copper	25 000	30 859	-	30 859
Winding Eddy	938	1158	7,177	8 308
Other Stray	1 253	1547	1,58	2 442
Total	31 691	38 064		46 110

Under the supplied HLS, the service losses are witnessed to have a surge of about 31% and 17% from the rated and load losses respectively. The peak local load loss ratio for the HLC is computed as follows:

$$P_{LL}(p.u) = I^2{}_{rms(p.u)}\left(1 + F_{HL} \times P_{EC_R}\right)$$
$$= 1.234 \times [1 + 0.15 \times 7.177] = 1.248 p.u$$

By applying proposed method, the maximum permissible harmonic load current with the supplied harmonic spectrum is as follows:

$$I_{max}(p.u) = \sqrt{\frac{1.15}{1 + 0.15 \times 7.177}} \times 262{,}43 = 259{,}85\ A$$

Consequently, with the supplied harmonic spectrum the studied WTGSU transformer capability is about 259,85A. The evaluation of the total service losses as a result of the HLC using FEM [38] is tabulated as shown in Table 5.

Table 5: Computation of total service losses: FEM

Type of loss	Rated Losses (W)	Load Losses (W)	HLF	Service Losses (W)
No load	4 500	4 500	-	4 500
Copper	25 000	30 859	-	30 859
Winding Eddy	938	1158	6,948	8 043

Other Stray	1 253	1547	1,58	2 442
Total	31 691	38 064		45 844

Under the supplied HLS, the service losses are witnessed to have a surge of about 31% and 17% from the rated and load losses respectively. The peak local load loss ratio for the HLC is computed as follows:

$$P_{LL}(p.u) = I^2{}_{rms(p.u)}(1 + F_{HL} \times P_{EC_R}) = 1.234 \times [1 + 0.15 \times 6{,}948] = 1.247 p.u$$

Based upon the FEM procedure [38], the maximum permissible harmonic load current with the supplied harmonic spectrum is as follows:

$$I_{max}(p.u) = \sqrt{\frac{1.15}{1 + 0.15 \times 6{,}948}} \times 262{,}43 = 258{,}70\ A$$

Consequently, with the supplied harmonic spectrum the studied WTGSU transformer capability is about 258,70A.

7. Conclusions

In the current work, an extension of the author's previous work [29], new analytical formulae for the Eddy loss computation in transformer winding conductors have been explicitly derived, with appropriate contemplation of the fundamental and harmonic load current. These formulae allow the distribution of the skin effect and computation of the winding Eddy losses as a result of individual harmonics in the winding conductors. The rectangular winding conductors considered are based on real transformer geometries and represents the configuration of the conductors carrying harmonic load current.

The new formulae were triumphantly corroborated by comparing their performance with computationally intensive FEM simulations, authenticating their adequacy and efficacy. On that account, our findings are handy for the transformer manufacturing industry, where the transformer anatomy and design necessitate specific performance and inexpensive computational resources.

Additionally, it has been demonstrated that the harmonic load currents minimizes the depth of skin effect, lead to an increase in the copper losses and winding stray losses and, thereby, consideration must be given for proper evaluation and design of fir for purpose transformers capable of facilitating wind power. These formulae can be applied in the transformer design systems for calculating the effects of harmonics on the transformer active part components. These results can be utilized to enhance the design of transformer s and consequently minimize the generation of hotspots in metallic structures. In this way, the developed formulae are a critical furtherance to the existing methodologies employed by manufactures.

The method herein provides a fundamental basis for subsequent development such as analysis of thermal performance of transformers in wind power technologies.

Future work involves the development of TOC model for WTGSU transformers to evaluate their economic life by considering the intermittent nature of wind turbine generators.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Department of Energy, "Integrated Resource Plan (IRP2019)", in 2019 Government Notice, 10-11, 2019, doi: IRP/2019/IRP-2019.pdf

[2] SAWEA," South Africa's Wind Industry. Stats and Facts SAWEA", in 2020 Stats and Figures, 2020, doi: stats-and-facts-sawea.

[3] ESI Africa,"Tracking wind energy in South Africa", in 2019 ESI Africa, 2019, doi: industry-sectors/renewable-energy/tracking-wind-energy-in-south-africa.

[4] M.A. Kamarposhti, S.B. Mozafari, S. Soleymani, S.M. Hosseini, "Improving the wind penetration level of the power systems connected to doubly fed induction generator wind farms considering voltage", Journal of Renewable and Sustainable Energy, **7**(4), 2015, doi.org/10.1063/1.4927008.

[5] J.M. Carrasco, L. Garcia Franquelo, "Power-Electronic Systems for the Grid Integration of Renewable Energy Sources: A Survey", IEEE Transactions on Industrial Electronics, **53**(4), 1002 – 1016, 2006, doi: 10.1109/TIE.2006.878356.

[6] A.A. Adly, S.K. Abd-El-Hafizb, "A performance-oriented power transformer design methodology using multi-objective evolutionary optimization", Journal of Advanced Research, **6** (3), 2015, doi: org/10.1016/j.jare.2014.08.003.

[7] M.B. Mallett, "High-Voltage Power-Transformer Design", in Transactions of the American Institute of Electrical Engineers, **62** (8), 1943, doi: 10.1109/T-AIEE.1943.5058737.

[8] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Step-Up Transformers for PV Plants: Load Loss Estimation under Harmonic Conditions," in 2020 19th International Conference on Harmonics and Quality of Power (ICHQP), Dubai, United Arab Emirates, 1-5, 2020, doi: 10.1109/ICHQP46026.2020.9177938.

[9] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Effects of Current Harmonics on Maximum Loading Capability for Solar Power Plant Transformers," in 2020 International SAUPEC/RobMech/PRASA Conference, Cape Town, South Africa, 1-5, 2020, doi: 10.1109/SAUPEC/RobMech/PRASA48453.2020.9041101.

[10] B.A. Thango, J.A. Jordaan, A.F. Nnachi, "Contemplation of Harmonic Currents Loading on Large-Scale Photovoltaic Transformers," in 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 479-483, 2020, doi: 10.1109/ENERGYCon48941.2020.9236514.

[11] J.N. Jagers, J. Khosa, P.J. De Klerk, C.T. Gaunt, "in 2007 Transformer Reliability and Condition Assessment in South African Utilities", in 2007 XVth International Symposium on High Voltage Engineering. University of Ljubljana, Elektroinstitut Milan Vidmar, Ljubljana, Slovakia, 27-31, 2007, doi:10.1.1.533.6250.

[12] V.M. Motalleb, M. Vakilian, A. Abbaspour, "Tank shielding contribution on reduction of Eddy current losses in power transformers", in 2009 IEEE Pulsed Power Conference. Washington, DC, 641-645, 2009, doi: 10.1109/PPC.2009.5386382.

[13] A.M. Milagre, M.V. Ferreira da Luz, G.M. Cangane, "3D calculation and modeling of eddy current losses in a large power transformer", in 2012 XXth International Conference on Electrical Machines. Marseille, 2012, doi: 10.1109/ICElMach.2012.6350200.

[14] S. Saha, N. Fernando, L. Meegahapola, "Multi magnetic material laminated cores: Concept and modelling", 2017 20th International Conference on Electrical Machines and Systems (ICEMS), Sydney, NSW, 2017, doi: 10.1109/ICEMS.2017.8056247.

[15] O. Bottauscio, M. Chiampi, D. Chiarabaglio, "Advanced model of laminated magnetic cores for two-dimensional field analysis", IEEE Transactions on Magnetics, **36**, (3), 561-573, 2000, doi: 10.1109/20.846219.

[16] A.I. Korolev, "On electromagnetic induction in electric conductors", arXiv: Cornell University, 2013, doi: 1303.0785.

[17] V. Sarac, G.Stefanov, " Calculation of Electromagnetic Fields in Electrical Machines using Finite Elements Method", 2011 International Journal of Engineering and Industries **2**(1), 2011, doi: 10.4156/ijei.vol2.issue1.3.

[18] A. Srikanta Murthy, N. Azis, J.Jasni, M.L. Othman, M.F. Mohd Yousof, M.A. Talib, "Extraction of winding parameters for 33/11 kV, 30 MVA

transformer based on finite element method for frequency response modelling," 2020 PLoS One, **15**(8), e0236409, doi.org/10.1371/journal.pone.0236409.
[19] J. Íñiguez, V. Raposo, M. Zazo, A. García-Flores, " The electromagnetic field in conductive slabs and cylinders submitted to a harmonic longitudinal magnetic field", American Journal of Physics, **77**, 1074, 2009, doi.org/10.1119/1.3160663.
[20] N. Amemiya, Y. Sogabe, S. Yamano, H. Sakamoto, " Shielding current in a copper-plated multifilament coated conductor wound into a single pancake coil and exposed to a normal magnetic field", Superconductor Science and Technology, **32** (11), 2019, 10.1088/1361-6668/ab3f1c.
[21] B.A. Thango, J.A. Jordaan, A.F. Nnachi, Analysis OF Distributed solar Photovoltaic System Harmonics on Transformers, Eliva Press, 2020.
[22] S.V. Kulkarni, S.A. Khaparde, Transformer Engineering: Design and Practice, Taylor & Francis Group, 2004.
[23] A. Mor, S. Gavril, "Eddy Current Effects on the Asynchronous Performance of the Hysteresis Machine", Journal of the Franklin Institute. **314**(2), 1982, doi.org/10.1016/0016-0032(82)90062-X.
[24] S.V Kulkarni, J.C Olivares, R. Escarela-Perez, V.K. Lakhiani, J. Turowski, "Evaluation of eddy losses in cover plates of distribution transformers", IEE Proceedings - Science, Measurement and Technology, **151**(5), 313-318, 2004, doi: 10.1049/ip-smt:20040632 .
[25] A. Najafi, I. Iskender, " Reducing losses in distribution transformer using 2605SA1 amorphous core based on time stepping finite element method", in 2015 International Siberian Conference on Control and Communications (SIBCON). Omsk, 2015, doi: 10.1109/SIBCON.2015.7146963.
[26] L. Betancourt, G. Martinez, D. Álvarez, J. Rosero, "Losses characterization on distribution transformer windings in frequency domain by means of finite element method (FEM): Part I", in 4th International Conference on Power Engineering, Energy and Electrical Drives. Istanbul, 2013, doi: 10.1109/PowerEng.2013.6635589.
[27] G.U. Nnachi, A.O. Akumu, C.G. Richards, D.V. Nicolae, " Estimation of no-Load Losses in Distribution Transformer Design Finite Element Analysis Techniques in Transformer Design", in 2018 IEEE PES/IAS PowerAfrica, Cape Town, 2018, doi: 10.1109/PowerAfrica.2018.8521142.
[28] L. Susnjic, Z. Haznadar, Z. Valkovic,"Stray Losses Computation in Power Transformer", in 2006 12th Biennial IEEE Conference on Electromagnetic Field Computation. Miami, FL, 2006, doi: 10.1109/CEFC-06.2006.1633280.
[29] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "A Weighting Factor for Estimating the Winding Eddy Loss in Transformers for High Frequencies," in 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 2020, doi: 10.1109/ENERGYCon48941.2020.9236472.
[30] H. Kekana, G. Landwehr, "Wind capacity factor calculator", Journal of energy South. Africa, **30** (2), Cape Town, 2019, doi: https://dx.doi.org/10.17159/2413-3051/2019/v30i2a6451.
[31] B.A. Thango, J.A. Jordaan, A.F. Nnachi, " Total Ownership Cost Evaluation for Transformers within Solar Power Plants", in 2020 6th IEEE International Energy Conference (ENERGYCon).Gammarth, Tunis, Tunisia, 2020, doi: 10.1109/ENERGYCon48941.2020.9236613.
[32] C.A. Charalambous, A. Milidonis, A. Lazari, A.I. Nikolaidis, "Loss Evaluation and Total Ownership Cost of Power Transformers—Part I: A Comprehensive Method", IEEE Transactions on Power Delivery, **28**(3), 2013, doi: 10.1109/TPWRD.2013.2262506.
[33] J.F. Baranowski, P.J. Hopkinson, "An alternative evaluation of distribution transformers to achieve the lowest TOC", IEEE Transaction Power Delivery, **7**, 1992, doi: 10.1109/61.127057.
[34] A.L. Lazari, C.A. Charalambous, "Probabilistic Total Ownership Cost of Power Transformers Serving Large-Scale Wind Plants in Liberalized Electricity Markets", IEEE Transactions on Power Delivery, **30**(4), August. 2015, 10.1109/TPWRD.2014.2365832.
[35] F. Lee and T. Wilson, "Voltage-spike analysis for a free-running parallel inverter," in IEEE Transactions on Magnetics, **10** (3), 969-972, 1974, doi: 10.1109/TMAG.1974.1058474.
[36] D. Kusiak, "The Magnetic Field and Impedances in Three-Phase Rectangular Busbars with a Finite Length", Energies 2019, **12**(8), 1419, doi.org/10.3390/en12081419.
[37] IEEE," Recommended Practice for Establishing Liquid Immersed and Dry-Type Power and Distribution Transformer Capability when Supplying Nonsinusoidal Load Currents", in IEEE Std C57.110-2018, 2018, doi: 10.1109/IEEESTD.2018.8511103.
[38] B.A. Thango, Jordaan, A.F. Nnachi, "A Novel De-rating Procedure for Distributed Photovoltaic Power (DPVP) Generation Transformers", in 2021 International SAUPEC/RobMech/PRASA Conference, 2021, doi: 10.1109/TPWRD.2013.2262506.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 126-130 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Novel Approach for Estimating the Service Lifetime of Transformers within Distributed Solar Photovoltaic (DSPV) Systems

Bonginkosi Allen Thango*, Jacobus Andries Jordaan, Agha Francis Nnachi

Department of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 24 March, 2021
Online: 20 May, 2021

Keywords:
Renewable energy
Transformer
Degree of polymerization
Furan
Service lifetime

ABSTRACT

The rapid transformation of the energy sector in South Africa towards renewable energy (RE) production calls for the management of assets to keep pace with the ongoing developments in a reliable manner. Aging assets, increasing energy needs and reliable supply of energy without load shedding are some of the challenges utilities are facing in South Africa. In resolving these challenge, imaginative solutions are required to maintain the installed assets and determining the viability of refurbishment, replacement or upgrading. In the current work, an extension of the author's previous work, a novel approach for estimating the service lifetime of transformers within Distributed Solar Photovoltaic (DSPV) Systems in South Africa is introduced. This experiential form has been derived by extensive experimental trials. The proposed approach is initially employed to evaluate the Degree of polymerization (DP) of cellulose insulation based on measured furan (2FAL) contents of 9 case scenarios. The calculated DP is then used to evaluate the service lifetime of these units. In efforts to authenticate the proposed approach, a comparative study is conducted against 5 other models. Finally, the proposed approach is compared with the results of the measured DP. It is observed that the proposed approach produce accurate estimates with an approximation not exceeding 1% and 2.2% from the measured DP and service lifetime respectively.

1. Introduction

Oil-immersed Distributed Solar Photovoltaic (DSPV) system transformers are commonly designed for a service lifetime of about 30 years [1], [2]. During service, the cellulose insulation degrades under the interactive effects emanating from electrical, thermal and chemical components. The mechanism of degradation of cellulose insulation (DCI) is reliant on the conditions such as harmonics, distorted load cycle and excessive heating that the transformer may be subjected to during their designed service lifetime. There are four underlying factors that elevate the DCI viz. moisture, oxygen, acid and excessive temperature. The oil-cellulose insulation system in DSPV transformers is prevalent on account of its cost-effectiveness and capacity to withstand severe electrical and thermal stresses [3], [4]. Due to the costly exercise of replacing field transformers, the utility owners will keep the units in service while performing periodical online monitoring and routine testing of the insulation system. Therefore, diagnostic techniques suchlike Dissolved Gas Analysis (DGA) and Furan Analysis (FA) are increasingly being recognized as efficacious and irrefutable procedures to evaluate the reliability of oil-cellulose insulation system.

The chemical structure of cellulose insulation is composed of strings of anhydroglucoses linked via glycosidic bonds [5] – [7]. These bonds have relatively low temperature resistance in relation to hydrocarbon bonds prevailing in DSPV transformer oil [8]. This degradation phenomenon triggers depolymerisation which leads to degraded forms of glucose as shown in Figure 1 [9]. The formation of glucose will further degrade to forge furans, water and combustible gases [9 – 11].

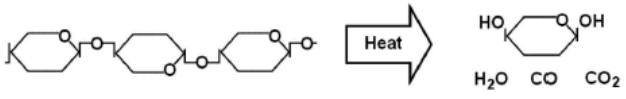


Figure 1: Thermal degradation of cellulose insulation

The generation of Furan compounds (FCs) is distinctly on the cellulose insulation and seemed to be kind of stable at elevated

*Corresponding Author: Bonginkosi Allen Thango, thangotech@gmail.com

temperatures reaching 140°C within the transformer oil [12]. The existence of FCs remains in oil-immersed transformers even at elevated temperatures geared towards providing a divulgence to the condition of the cellulose insulation.

The FCs 2-FAL, 5-hydroxymethyl-2-furaldehyde, 2-acetylfuran, 5-methyl-2-furaldehyde, and 2-furfurolare are the most widely known in the transformer manufacturing industry. The FC 2-FAL is most widely-used in the industry relating its content to the cellulose degree of polymerization (DP) which is the number of monomeric units in the cellulose polymer [13] – [14]. Furans are a constituent of the products of degradation of the transformer cellulose insulation. During service, they are moderately dissolvable in the transformer oil. The relationship between the transformer state and the concentration of the furan content in the oil is as shown in Table below.

Table 1: 2FAL concentration criterion

Status	Concentration [in ppm]
Healthy	2FAL < 0.5
Satisfactory	0.5 ≤ 2FAL ≤ 1
Warning	1.1 ≤ 2FAL ≤ 1.5
Poor	1.6 ≤ 2FAL ≤ 2
Extremely poor	2FAL ≥ 2.1

In the current work, an extension of the author's previous work [14], a novel approach for estimating the service lifetime of transformers within DSPV Systems in South Africa is introduced. This experiential form has been derived by extensive experimental trials. The proposed approach is initially employed to evaluate the DP of cellulose insulation based on measured 2FAL contents on a fleet of selected case scenarios. The calculated DP is then used to evaluate the service lifetime of these units. In efforts to authenticate the proposed approach, a comparative study is conducted against 5 other models. Finally, the proposed approach is compared with the results of the measured DP. It is observed that the proposed approach produce accurate estimates with an approximation not exceeding 1% and 2.2% from the measured DP and service lifetime respectively.

2. DP Models

In this day and age, the aging condition of transformers can be established by mathematical modelling which launch the relationship between furan concentrations, DP, and transformer service lifetime.

2.1. Existing DP models

Several authors in [15] - [19] have established mathematical formulae which relate the furan concentrations and the measure of the DP of the insulation material. An illustration of this relationship based upon this methods is presented in Figure 2. A Further in [20], mathematical formulae has been developed to describe the relationship between DP and the service lifetime. In order to attain a thread that binds the furan concentrations and the DP of the cellulose insulation (CI), the DP methods are compared at different furan contents. The comparison indicate that the methods can evaluate the DP ranging from 1 to 14.

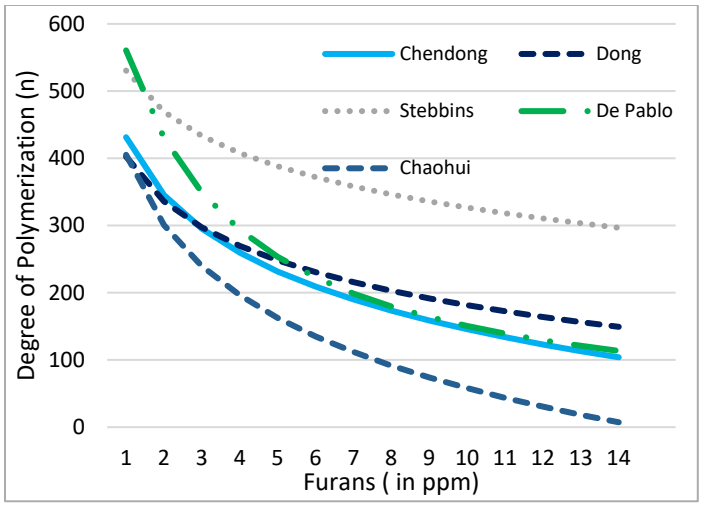


Figure 2: Correlation between the DP and Furans (2FAL)

The development, identification and behaviour of furan compounds has been described by [21], [22] and [23]. In the next sub-section a comprehensive description of the model proposed in this work is presented.

2.2. Proposed model

Based upon observational data, a fleet of mineral oil samples were analysed to develop a novel DP model for transformers serving exclusively in distributed solar photovoltaic systems. The proposed model can be applied to ascertain the DP of cellulose insulation with accuracy following 2-FAL measurement in the oil. The proposed model is as shown in Eq. (1) below.

$$DP = 763.67 - 584.23x + 398x^2 - 135.52x^3 + 17.14x^4 \quad (1)$$

Here,

x − 2-FAL

The concentration (in ppm) of the furan content 2-FAL can be easily substituted in the formulae to obtain the DP of a transformer.

2.3. Transformer Service Lifetime

The transformer age can be obtained by applying Eq. (2) and Eq. (3) below [20]. With the knowledge of the DP of the transformer, the service lifetime can be ascertained.

$$Service\ lifetime(in\ years) = 20.5\ ln\left(\frac{1100}{DP}\right) \quad (2)$$

$$\%Life\ used = \frac{log_{10}(DP) - 2.903}{-0.006021} \quad (3)$$

3. Case Scenario

In this section, the 2FAL results of 9 mineral-oil immersed DSPV transformers are considered. The gas concentrations are obtained from oil samples collected from various DSPV facilities. The studied transformers range from 100kVA to 20 000 kVA as shown in Table 2.

Table 2: Case Studies

Case	Rating (kVA)	Voltage (kV)	2FAL
A	1500	6.6/0.5	1.778
B	1500	6.6/0.5	2.377
C	100	0.525/0.4	3.00
D	150	0.525/0.4	1.601
E	20 000	6.6/0.5	0.676
F	20 000	6.6/0.5	0.231
G	20 000	88/6.6	0.143
H	4000	6.6/0.5	1.513
I	2000	6.6/0.5	1.563

The 2-FAL contents were measured during the routine maintenance of the transformers.

3.1. Degree of Polymerization

In order to determine the accuracy of the formulae presented by the authors in [15] – [19] and the proposed formula in the current work, the DP for the studied cases are evaluated and tabulated as a shown in Table 3 below. Stebbins [18] results are observed to yield a significantly protruding DP values compared to all the other methods.

Table 3: DP of studied transformers using furan compounds (2FAL)

Case	2FAL	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
A	1,778	360	454	347	480	318	393
B	2,377	324	396	319	454	275	351
C	3	295	349	297	433	240	322
D	1,601	373	474	357	489	334	405
E	0,676	480	622	440	565	464	512
F	0,231	613	731	543	660	626	648
G	0,143	673	757	589	703	699	688
H	1,513	380	485	363	494	343	411
I	1,563	376	479	360	491	338	408

These DP values are a measure of the realistic aging of the cellulose insulation, which governs the transformer's service lifetime. And so, it is possible to evaluate the service lifetime of a transformer and make comparison with real life by means of furan concentration within the oil. The latter improves the assessment of the reliability of service transformers.

3.2. Transformer Service Lifetime

Further to previous sub-section, the remaining service lifetime of the studied cases is determined using the DP results as presented in Table 4 for all models.

Table 4: Remaining service lifetime estimation

Case	2FAL	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
A	1,778	22,90	18,16	23,64	17,02	25,41	21,12
B	2,377	25,06	20,97	25,35	18,15	28,44	23,43
C	3	26,97	23,53	26,84	19,10	31,25	25,18
D	1,601	22,17	17,25	23,05	16,62	24,42	20,49
E	0,676	17,00	11,70	18,78	13,65	17,68	15,66
F	0,231	11,98	8,38	14,47	10,46	11,55	10,84
G	0,143	10,08	7,65	12,80	9,18	9,31	9,62
H	1,513	21,79	16,78	22,74	16,41	23,90	20,17
I	1,563	22,01	17,05	22,92	16,53	24,20	20,35

The relationship between the DP and the service lifetime is illustrated by Figure 3 below.

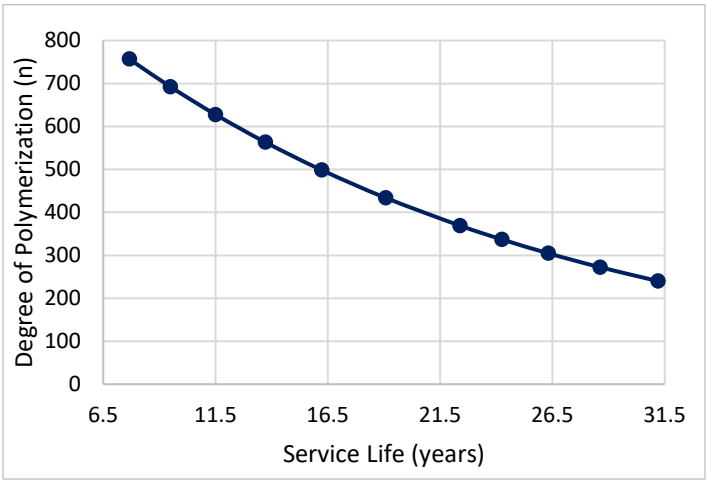


Figure 3: Correlation between the DCI and Furans (2FAL)

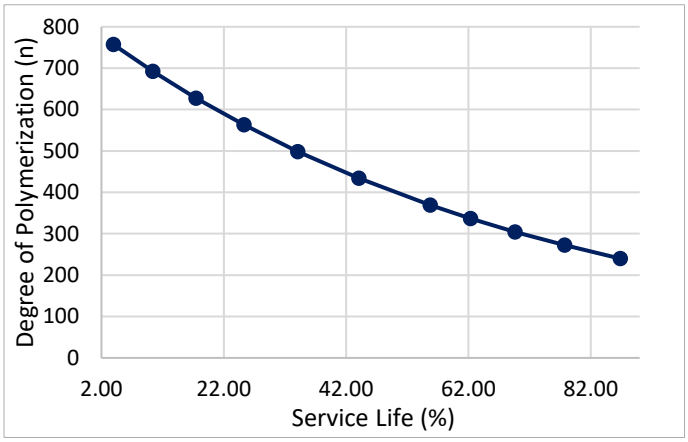


Figure 4: Correlation between the DP and %Service lifetime

It can be observed that as the transformer life is service is longer, the DP decreases. In Table 5, the percentage of the service lifetime for all cases is presented.

Table 5: Percentage of transformer service lifetime

Case	2FAL	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
A	1,778	57,58	40,90	60,18	36,89	66,42	51,32
B	2,377	65,18	50,79	66,21	40,87	77,09	59,46
C	3	71,92	59,81	71,44	44,22	86,95	65,61

Case	2FAL	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
D	1,601	55,02	37,70	58,12	35,51	62,93	49,10
E	0,676	36,83	18,18	43,10	25,03	39,23	32,13
F	0,231	19,16	6,50	27,93	13,81	17,65	15,15
G	0,143	12,48	3,94	22,07	9,31	9,77	10,88
H	1,513	53,67	36,06	57,03	34,77	61,11	47,98
I	1,563	54,44	37,00	57,66	35,19	62,15	48,62

The percentage of service lifetime is observed to correlate with the data presented in Table 4. The correlation between the DP and the percentage service lifetime is illustrated by Figure 4.

It is observed that the DP decreases with respect to the aging of the transformer.

3.3. Practical Results

In this subsection, the values of the measured DP and 2FAL of the nine studied transformers that operate in various DSPV systems are presented in Table 6.

Table 6: Percentage of service lifetime

Case	2FAL	Measured DP	Service Life	% Service Lifetime
A	1,778	389	21,30	51,96
B	2,377	352	23,35	59,16
C	3	322	25,20	65,67
D	1,601	404	20,56	49,35
E	0,676	515	15,57	31,78
F	0,231	641	11,08	15,99
G	0,143	694	9,46	10,29
H	1,513	412	20,15	47,92
I	1,563	411	20,20	48,10

3.4. Results comparison

The error of estimation for DP, service lifetime and percentage service lifetime against the measured results are presented in this subsection for the studied DSPV transformers. In Table 7, it can be observed that the proposed formula gives an accurate estimation of the DP in comparison to the other models across all nine cases.

Table 7: Error of estimation: DP

Case	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
A	7,5%	16,6%	10,8%	23,2%	18,2%	0,88%
B	8,0%	12,3%	9,3%	28,9%	22,0%	0,42%
C	8,3%	8,5%	7,7%	34,6%	25,5%	0,09%
D	7,6%	17,5%	11,4%	21,2%	17,2%	0,36%
E	6,8%	20,8%	14,5%	9,8%	9,8%	0,48%
F	4,3%	14,1%	15,3%	3,1%	2,3%	1,17%
G	3,0%	9,2%	15,1%	1,4%	0,7%	0,81%
H	7,7%	17,9%	11,9%	20,0%	16,7%	0,09%
I	8,4%	16,6%	12,4%	19,6%	17,7%	0,72%

It is observed that the proposed formula produces accurate estimates with an approximation not exceeding 1% from the measured DP. As a result, it is evident that the proposed formula achieves a higher accuracy. In Table 8, the error of estimation of the calculated and measured results of the service lifetime is presented.

Table 8: Error of estimation: Service Life

Case	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
A	7,5%	14,7%	11,0%	20,1%	19,3%	0,85%
B	7,3%	10,2%	8,6%	22,3%	21,8%	0,37%
C	7,0%	6,6%	6,5%	24,2%	24,0%	0,07%
D	7,8%	16,1%	12,1%	19,1%	18,8%	0,35%
E	9,2%	24,8%	20,7%	12,3%	13,6%	0,64%
F	8,1%	24,4%	30,6%	5,6%	4,3%	2,16%
G	6,6%	19,1%	35,4%	2,9%	1,6%	1,77%
H	8,1%	16,7%	12,9%	18,5%	18,6%	0,09%
I	8,9%	15,6%	13,5%	18,2%	19,8%	0,74%

It is observed that the proposed formula produce accurate estimates with an approximation not exceeding 2.2% from the measured service lifetime.

Table 9: Error of estimation: % of transformer service life

Case	Chendong [15]	De Pablo [16]	Dong [17]	Stebbins [18]	Chaohui [19]	Proposed
A	10,8%	21,3%	15,8%	29,0%	27,8%	1,22%
B	10,2%	14,2%	11,9%	30,9%	30,3%	0,51%
C	9,5%	8,9%	8,8%	32,7%	32,4%	0,10%
D	11,5%	23,6%	17,8%	28,1%	27,5%	0,52%
E	15,9%	42,8%	35,6%	21,2%	23,4%	1,10%
F	19,8%	59,4%	74,7%	13,7%	10,4%	5,27%
G	21,3%	61,7%	114,5%	9,5%	5,1%	5,71%
H	12,0%	24,8%	19,0%	27,4%	27,5%	0,13%
I	13,2%	23,1%	19,9%	26,8%	29,2%	1,09%

It is observed that the proposed formula produce accurate estimates with an approximation not exceeding 5.8% from the measured percentage service lifetime.

At large, the overall assessment of the proposed formula presented in this work presented the highest precision in the estimation of the DP of the cellulose insulation in comparison to the other 5 models.

4. Conclusion

The knowledge of the transformer service lifetime and the specific point in time that is suitable to replace transformers is critical for utility owners. These determinations are dependent upon the condition of the cellulose and liquid insulation. In practical terms, knowledge on the condition of insulation is crucial to undertake adequate life cycle and reliability management of transformers in the facility. In the transformer industry, the physical chemistry of oil is prevalent in the evaluation of the transformer life cycle considering that oil samples can easily be extracted at site to the laboratory for analysis.

In the current work, an extension of the author's previous work [14], a novel approach for estimating the service lifetime of transformers within DSPV Systems in South Africa was introduced. This experiential form has been derived by extensive experimental trials. The proposed approach was initially employed to evaluate the DP of cellulose insulation based on measured 2-FAL contents (in ppm) of 9 case scenarios. The calculated DP was then used to evaluate the service lifetime of these units. In efforts to authenticate the proposed approach, a comparative study was conducted against 5 other models. Finally, the proposed approach was compared with the results of the measured DP. It was observed that the proposed approach produce accurate estimates with an approximation not exceeding 1% and 2.2% from the measured DP and service lifetime respectively.

In future work, the authors will test the proposed formula with other different transformers. Further, the authors will carry out work related to the reliability analysis of DPSV transformers based upon their operating hours in service by apply the Weibull distribution technique and support vector machine to classify their reliability.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] P. Mann, "Determination of transformer life expectancy," in Electrical Engineering, **82** (8), 512-514, 1963, doi: 10.1109/EE.1963.6540988.

[2] D. Martin, Y. Cui, T. Saha, N. Lelekakis and J. Wijaya, "Life estimation techniques for transformer insulation," 2013 Australasian Universities Power Engineering Conference (AUPEC), Hobart, TAS, 1-6, 2013, doi: 10.1109/AUPEC.2013.6725457.

[3] T. A. Prevost and T. V. Oommen, "Cellulose insulation in oil-filled power transformers: Part I - history and development," in IEEE Electrical Insulation Magazine, **22** (1), 28-35, 2006, doi: 10.1109/MEI.2006.1618969.

[4] B. A. Thango, J. A. Jordaan and A. F. nnachi, "Stray Gassing of Transformer Oil in Distributed Solar Photovoltaic (DSPV) Systems," 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 484-488, 2020, doi: 10.1109/ENERGYCon48941.2020.9236522.

[5] L. Cheim, D. Platts, T. Prevost and S. Xu, "Furan analysis for liquid power transformers," in IEEE Electrical Insulation Magazine, **28** (2), no. 2, 8-21, 2012, doi: 10.1109/MEI.2012.6159177.

[6] J. R. Sans, K. M. Bilgin and J. J. Kelly, "Large-scale survey of furanic compounds in operating transformers and implications for estimating service life," Conference Record of the 1998 IEEE International Symposium on Electrical Insulation (Cat. No.98CH36239), Arlington, VA, USA, **2**, 543-553, 1998 doi: 10.1109/ELINSL.1998.694852.

[7] M. T. Imani, K. Homeier, P. Werle and G. Dräger, "How Far are Furan Compounds Reliable Indicators for Thermal Aging of Oil Impregnated Cellulose Insulation?," 2018 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), Cancun, 438-441, 2018 doi: 10.1109/CEIDP.2018.8544859.

[8] C. Zhang and J. M. K. Macalpine, "Furfural Concentration in Transformer Oil as an Indicator of Paper Ageing, Part 1: A Review," 2006 IEEE PES Power Systems Conference and Exposition, Atlanta, GA, 1088-1091, 2006, doi: 10.1109/PSCE.2006.296461.

[9] G. Xia and G. Wu, "Study on the impact of initial moisture contents on ageing characteristics of transformer oil-paper insulation," 2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE), Chengdu, 1-4, 2016, doi: 10.1109/ICHVE.2016.7800632.

[10] D. Suksawat, K. Atthaphotpong, K. Takboontam, K. Satirapattanakiat, C. Raphephat and N. Pattanadech, "Furan Analysis of Oil Impregnated Paper Aged by Chemical Stress," 2020 8th International Conference on Condition Monitoring and Diagnosis (CMD), Phuket, 422-425, 2020, doi: 10.1109/CMD48350.2020.9287264.

[11] A. Akshatha, R. Kumar A, S. Rajan J and H. Ramachandra, "Studies on chemical and dielectric phenomena in paper - oil insulation due to sulphur compounds in mineral oil," 2014 IEEE Electrical Insulation Conference (EIC), Philadelphia, PA, 281-285, 2014, doi: 10.1109/EIC.2014.6869392.

[12] "IEEE Guide for Evaluation and Reconditioning of Liquid Immersed Power Transformers," in IEEE Std C57.140-2017 (Revision of IEEE Std. C57.140-2006), 1-88, 2017, doi: 10.1109/IEEESTD.2017.8106924.

[13] "IEEE Standard Test Procedure for Thermal Evaluation of Insulation Systems for Liquid-Immersed Distribution and Power Transformers," in IEEE Std C57.100-2011 (Revision of IEEE Std C57.100-1999) , 1-37, 2012, doi: 10.1109/IEEESTD.2012.6143968.

[14] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Service Life Estimation of Photovoltaic Plant Transformers Under Non-Linear Loads," 2020 IEEE PES/IAS PowerAfrica, Nairobi, Kenya, 1-5, 2020, doi: 10.1109/PowerAfrica49420.2020.9219912.

[15] X. Chendong, Monitoring Paper Insulation Ageing by measuring Furfural Contents in Oil, 7th International Symposium on High Voltage Engineering, 26 - 30, 1991.

[16] A. De Pablo, Interpretation of Furanic Compounds Analysis - Degradation Models, CIGRE WG D1.01.03, former WG 15-01, Task Force 03, 1997.

[17] M. Dong, Z. Yan and G. J. Zhang, "Comprehensive diagnostic and aging assessment method of solid insulation in transformer," 2003 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, Albuquerque, NM, USA, 137-140, 2003, doi: 10.1109/CEIDP.2003.1254813.

[18] R. D. Stebbins, D. S. Myers and A. B. Shkolnik, "Furanic compounds in dielectric liquid samples: review and update of diagnostic interpretation and estimation of insulation ageing," Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials (Cat. No.03CH37417), Nagoya, Japan, **3**, 921-926, 2003, doi: 10.1109/ICPADM.2003.1218572.

[19] C. Lin et al., "The Aging Diagnosis of Solid Insulation for Oil-Immersed Power Transformers and Its Remaining Life Prediction", 2010 Asia-Pacific Power and Energy Engineering Conference, DOI:10.1109/APPEEC.2010.5449486.

[20] M. K. Pradhan and T. S. Ramu, "On the estimation of elapsed life of oil-immersed power transformers," in IEEE Transactions on Power Delivery, **20** (3), 1962-1969, 2005, doi: 10.1109/TPWRD.2005.848663.

[21] WG. D1.01, "Furanic Compounds for diagnosis", cigré, Electra, 4, 2012, doi: ELT_261_06-furanic-compounds-for-diagnosis.

[22] R. A. Abd El-Aal, K. Helal, A. M. M. Hassan and S. S. Dessouky, "Prediction of Transformers Conditions and Lifetime Using Furan Compounds Analysis," in IEEE Access, **7**, 102264-102273, 2019, doi: 10.1109/ACCESS.2019.2931422.

[23] H. Shertukde, Distributed Photovoltaic (DPV)-Grid Transformer Applications, CRC Press, 2017, doi: Distributed-Photovoltaic-Grid-Transformers/Shertukde/p/book/9781138073845.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 131-136 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Peculiar Stray Gassing Occurrences in Solar Photovoltaic Transformers during Service

Bonginkosi Thango*, Jacobus Jordaan, Agha Nnachi

Deptartment of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 27 March, 2021
Online: 20 May, 2021

Keywords:
Distributed Solar Photovoltaic
Transformers
Harmonics
Hotspot temperature
Stray gassing
Dissolved Gas Analysis

ABSTRACT

Distributed Solar Photovoltaic (DSP) Plants are one of the fastest growing renewable energy systems in South Africa. The primary components forming an integral part of the point of common coupling (PCC) are, the inverter used to convert dc voltage to ac voltage and step-up transformers which step-up low voltage input to the desired output level. However, DSP plant step-up transformers are considered to be one of the most sensitive equipment on the plant. These transformers are challenged with various electrical problems including abnormal levels of harmonics. The presence of harmonics in these transformers results in higher service losses thereby raising the hotspot (HS) temperature, in which, consequently introduce the stray gassing phenomena of the insulating oils. This calls for understanding of the nature of the problem and possible remediation to ensure enhanced power quality. Present work, an extension of previous work, investigate a reported case of peculiar stray gassing of transformer insulation oils during service. Initially, the harmonic spectrum of the DSP plant is presented and the related service losses at fundamental and under harmonic conditions are computed. Furthermore, the thermal performance of the transformer under these conditions is investigated. Lastly, the Dissolved Gas Analysis (DGA) results of the oil samples are presented. Novelty, findings of this work indicate that the generation of hydrogen arising from stray gassing may stem from severely hydro-treated mineral oil, but is also strenuously affected by transformer thermal aging of polymers, choice of core steel grade, zinc tank walls and vanishes. The production surplus of methane and ethane are also witnessed in the first years of service and reaches substantial concentration levels. Potentially, these occurrences also arises from the thermal aging of polymers. The authors make some recommendations to utility owners to make a distinction of stray gassing from transformer fault by means of routine inspection aside from DGA value basis withal to the increase in gas diffusion rate. Further, the authors make some significant contribution by further recommending procedures that can be employed as remedies during the design phase and manufacturing processes. Lastly, the authors highlight the need to establish standards that will provide support for transformers intended to operate in DSP applications.

1. Introduction

In line with the objectives of solving South Africa's power generation constraints, the Government's Integrated Resource Plan enacted in 2010 (IRP2010) [1] revealed the determination to produce 17.8 GW generation capacity from renewable energy resources by the year 2030. The additional electrical energy fed by the Independent Power Producers (IPPs) into the national grid has provided sustainable and supplemental solution to the heavy load shedding suffered by many South Africans. Despite this unquestionable success, the substantial growth in the South African renewable energy market including DSP plants is associated with many technical challenges. One of the foremost challenges is the irregular penetration of high voltage and current harmonics on the step-up transformers connected to the PCC. This phenomenon is a result of the switching action caused by inverters, nonlinear loads and resonances generated when the DSP plant is connected to the medium voltage (MV) network.

As a result, the inclined operational cost of the DSP plant due to high energy cost per kilowatt is incurred by the increase of the harmonic no-load and load losses in the step-up transformer. During service, the no-load losses are steady losses and are independent of the loading of the transformer [2]-[7]. The

*Corresponding Author: Bonginkosi Allen Thango, thangotech@gmail.com

magnetizing current necessary to energize the transformer core generate the no-load losses. In harmonic rich environments, no-load losses are impelled by voltage harmonics which increase hysteresis and eddy current losses in the core laminations. The last-mentioned problem initiate core vibrations, core saturation and electrical stresses in the insulation materials. The current harmonics predominantly affect the load losses of the DSP plant transformer. The load losses are dependent on the transformer loading profile and are widespread upon the copper conductors which then affect the alternating current (AC) resistance of the windings and other metallic parts [2], [3]. As a result, the winding temperature is increased and spread into the tank walls, flitch plates, core clamps and other metallic components of the active part. The hottest spot on the winding, which is generally located in the area nearby the top of the DSP plant transformer, restricts the loading capacity of the transformer. Furthermore, it determines the insulation service life and the probable risk of emancipating gas bubbles during extreme loading conditions [8]-[10]. Over time, excessive HS temperature brings upon thermal aging and the degradation of the cellulose insulation which determines the end of transformer service life.

Authors in [11]-[13] expressed the view that emancipation of gas bubbles or stray gassing in the service lifetime of the transformer is dependent on the insulating oils. For several years, DGA has been regarded as the most crucial procedure for assessing health condition of oil-immersed transformers. Drawn up on the basis regular inspection and characteristics of dissolved gases mined from transformer oil samples, an evaluation can be made on the transformer internal anomalies and severity thereof. Many authors have investigated the insulating oils status based on the DGA [14], [15] and [16]. In the past ten years, with the rise in the deployment of DSP plants in South Africa, empirical experience has indicated that dissolved gases can also materialize at relatively lower temperature (90–200°C) in the absence of anomalies or transformer fault. It should be noted that this gassing should not be misconstrued with gassing patterns of insulating oils by means of Partial discharges (PD) or catalysed reactions for case in a point of core steel grade zinc tank wall surface and coating. The DGA-evaluation tools also take into account this thermal stray gassing occurrences to permit clearer distinction between permissible stray gassing and peculiar gassing on account of PD and generation of HS temperatures in various active part components [14]-[16].

This work, an extension of previous work [16] presents the results of stray gassing results that were carried out on a 2500kV, mineral-oil filled transformer. Particularly, solution oriented on the impact of harmonic currents on the service losses and thermal performance of the studied DSP the plant transformer. Moreover, the observed stray gassing occurrences of the unit as a result of (1) the generation of HS temperatures in the active part components; and (2) the core steel grade, zinc tank wall surface and coating. The authors make a significant contribution by further recommending procedures that can be employed as remedies during the design phase and manufacturing processes.

2. Case Scenario

In this section, a harmonic study is carried out on a 2500kVA, mineral-oil filled DSP plant transformer to evaluate the no-load losses and load losses. The detailed classification of these losses are detailed by the authors in [17] and [18]. The advancement of computational power has enabled the development of tools capable of simulating complex geometries, mapping of electromagnetic fields and accurate evaluation of these transformer losses. A software tool embedding analytical methods with Finite Element Method (FEM) is used to estimate the transformer losses. The tool is automated to simulate complete transformer model geometries, magnetic field quantities and computation of eddy current losses in tank walls, flitch plate, core clamps, bus-bar's and windings. Figure 1 shows the complete model of the transformer 2-D geometry with the main parts that are susceptible to the leakage flux.

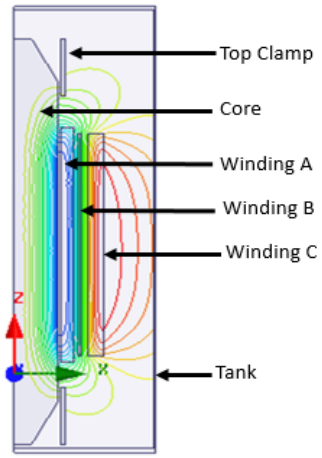


Figure 1: 2D geometry of the DSP plant transformer

2.1. Design based on the International Electrotechnical Commission (IEC) harmonic spectrum

The no-load losses and stray losses in windings and structural parts evaluated at nominal tap position are tabulated in Table 1. In order to compute these losses to account for the presence of harmonics it is necessary to obtain the values of the harmonic loss factors using the IEC 519-2014 [19] standard. Generally, from a manufactures perspective, the case of considering the harmonic spectrum proposed by this standard for a transformer intended to operate in a DSP plant is when there is no full harmonic spectrum supplied by the IPP. The harmonic spectrum recommended by IEC is shown in Figure 2.

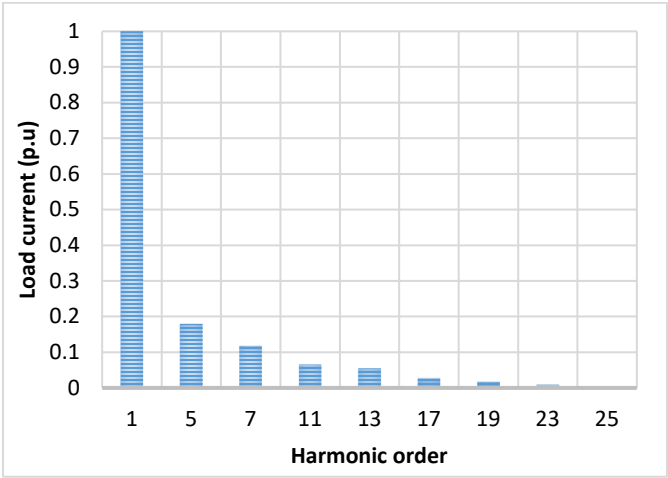


Figure 2: IEC 519-2014 Harmonic spectrum [2]

To estimate the service losses due to this harmonic spectrum, the transformer is initially designed at fundamental frequency to produce the rated losses as shown in Table 1. Thereafter, a harmonic loss factor is computed based on the supplied harmonic spectrum. The authors have presented a detailed computation of the harmonic loss factor in [17] and [18]. The Harmonic factor is in essence intended to predict the increase in the transformer losses under harmonics condition. The resultant service losses that will occur in transformer is as tabulated below.

Table 1: Tabulated loss computation under a harmonic spectrum

Type of loss	Rated losses (W)	Harmonic factor	Service losses (W)
No-load	0.180	1	0.180
I^2R	11.041	1.055	11.654
Winding Eddy	0.213	3.986	0.921
Other stray	1.194	1.252	1.495
Tank-HV side	0.249	1.252	0.312
Tank-LV side	0.494	1.252	0.619
Bus-Bar	0.013	1.252	0.129
Top yoke Clamp	0.029	1.252	0.036
Bottom yoke Clamp	0.046	1.252	0.058
Flitch plate	0.02	1.252	0.003

From both the fundamental and harmonic losses above, it can be noticed that the loss percentage error is higher for winding stray losses. Additionally, the core losses are observed to be constant, indicating that the no-load losses are independent of the harmonic currents. The minimum winding stray losses as a result of axial and radial leakage fields are computed by optimizing the selection of conductor dimensions while maintaining the permissible flux and current density of the winding [20] and [21]. The transformer winding subject to harmonic distortion is designed using continuously transposed cable epoxy (CTCE) with minimum axial and radial dimensions of each conductor to reduce the eddy losses [20] and [21]. The tank is designed using mild steel which hold a nonlinear permeability. The quantity of the magnetic field in the tank walls is evaluated by eliminating the impact of nonlinearity. Thereafter, the losses in the LV and HV tank walls are computed taking into account the effect of nonlinearity through repetitive estimation. The flitch plates are also mild steel type and contain slots at the upper and lower positions. While their loss quantity is low compared to winding eddies and tank losses under harmonic effects, the magnetic field quantity is high which results in the flitch plate being more susceptible to HS temperature.

2.2. Design based on the harmonic studies

The harmonic current level at the PCC is presented in Figure 3. The results are obtained by modelling the DSP plant in DigSILENT Power Factory as a single line diagram which is composed of the DSP plant passive components. The highest harmonic current order corresponds with the measured harmonic data. Notably, this high harmonic order exceeds the recommended IEEE spectrum presented in the last section.

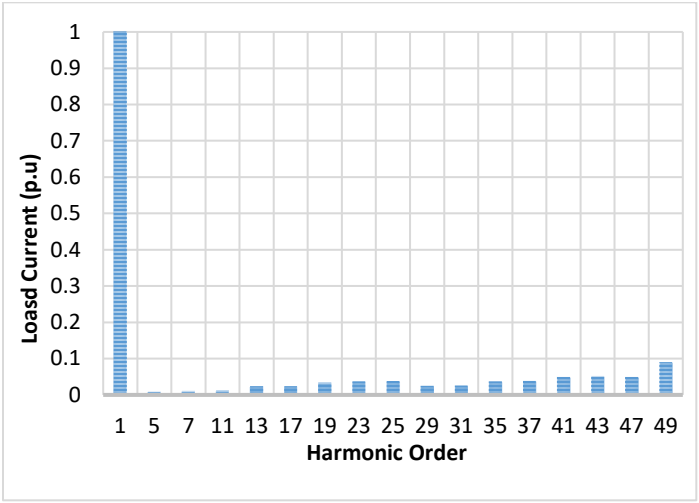


Figure 3: Simulated transformer harmonic spectrum at the PCC

The corresponding results when the design philosophy is tailored with the simulated harmonic spectrum are shown in Table 2. There are extreme differences in the percentage error between fundamental losses and the services losses. The transformer losses are observed to be even more extremely sensitive towards the simulated harmonic spectrum compared to the previous case. At large, these results outline the obligation for IPP's to conduct harmonic studies in order to provide technical information to manufactures which reflect the actual harmonic distortions that will be realised by the transformer and to design transformers suitable for such extreme conditions.

Table 2: Transformer losses based on actual DSP plant harmonic spectrum

Type of loss	Rated losses (W)	Harmonic factor	Service losses (W)
No-load	0.180	1	0.180
I^2R	11.041	1.055	11.654
Winding Eddy	0.213	3.986	0.921
Other stray	1.194	1.252	1.495
Tank-HV side	0.249	1.252	0.312
Tank-LV side	0.494	1.252	0.619
Bus-Bar	0.013	1.252	0.129
Top yoke Clamp	0.029	1.252	0.036
Bottom yoke Clamp	0.046	1.252	0.058
Flitch plate	0.02	1.252	0.003

To analyse the complete transformer harmonic response (THR), Figure 4 shows the results of the predicted transformer losses by IEC harmonic spectrum and compares them with the simulated harmonic spectrum the transformer is subjected to while in service. The transformer under inspection was designed without

the provision of a full harmonic study from the IPP, and judging from the results, the indication is that it was under-designed for the environment it was intended to operate, which would explain the reported high temperature rise and oil gassing patterns. Subsequently, this knowledge gap led to insufficient design of the winding conductor sizes and cooling system for this application.

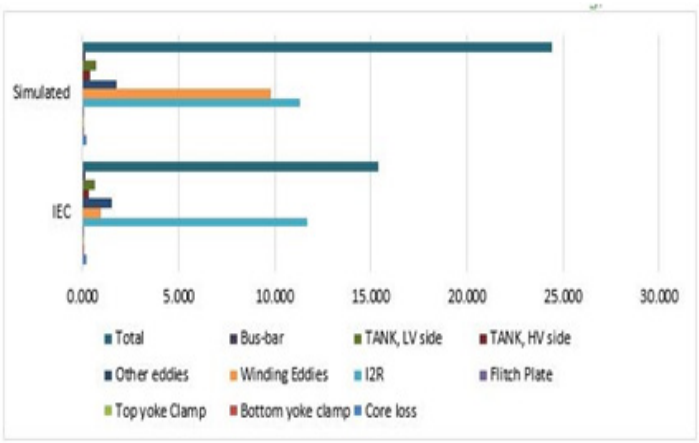


Figure 4: Loss comparison (IEC vs. Simulated Harmonic Spectrum)

In the next section, the temperature rise due to the increase in service losses is presented.

3. Thermal design

At tender stage, the transformer technical specifications supplied by IPP's to the manufacture generally include information about the permissible top oil, mean winding and HS temperatures. Electrical designers are then challenged to optimize transformer designs such that they adhere to the IEC 60076-2 [22] and IEEE C57.91-1995 [23] standards and to design the thermal and cooling requirements at an ambient temperature. During the manufacturing stage, as part of the Factory Acceptance Tests (FAT's) for newly manufactured transformers, the temperature rise test is conducted with the intent to demonstrate these temperatures, at full load and specified ambient temperature. The loading capability of the transformer is predominantly regulated by the HS temperature. In accordance with the IEC, the HS temperature is described as the mean winding temperature rise over the mean oil temperature multiply by a HS factor [22] and [23].

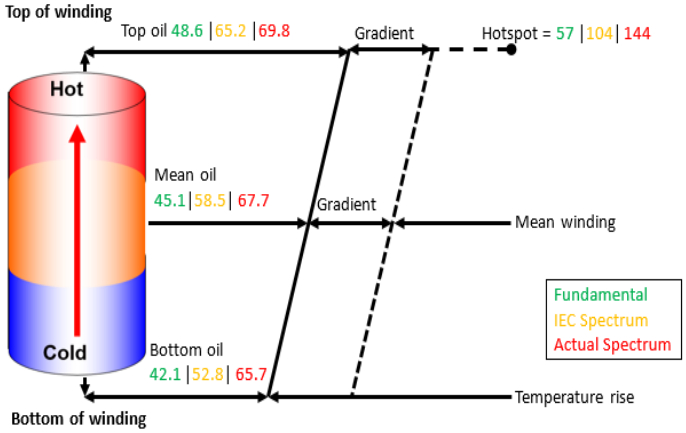


Figure 5: Simulated transformer harmonic spectrum at the PCC

A simplified thermal model of the transformer at fundamental frequency, IEC spectrum and simulated spectrum are shown in Figure 5. In this case, an IEC factor of 1.1 for DSP plant transformers is used in the calculation of the HS temperature. In Figure 5, the temperature rise within the windings is observed to increase linearly from the winding bottom to the top. A constant temperature difference between the oil duct and winding conductor along the winding is also observed. Moreover, the HS temperature is greater than the winding temperature rise as a result of the transformer losses.

The maximum error for the hot spot temperature between the fundamental values against the actual spectrum is 153%. That is 70% more than the HS temperature it was designed for under the IEC spectrum. This outcome clearly indicate that the transformer oil pumped through the winding conductors to execute the required cooling is not sufficient for transformer loading of actual harmonic spectrum. For the cooling system, the transformer oil viscosity and density vary with temperature rise, as such heat affects the fluid-flow sequence. Moreover, the quantity of the thermal fields resulting in such high HS temperature is derived from joule losses as a result of current flow in the windings and increased electromagnetic fields under harmonic conditions.

4. Stray gassing patterns

This section present findings of oil samples taken from the oil-filled hermetic DSP plant transformer under study for laboratory tests. The transformer was suspected to have oil gassing as indicated by the installed Oil Temperature Indicator (OTI). The results of Dissolved Gas Analysis (DGA) are given in Table 3. These results were taken after the transformer was in operation for about 4% of the predicted service life based on the design philosophy considering the IEC 519-2014 [19] to compute the losses, thermal and cooling requirements.

Table 3: DGA oil sample results (ppm)

Type of gas	Limits	Gas concentration
H_2	<100	1478
O_2	-	28899
N_2	-	72484
CO	<1000	28
CO_2	<15 000	860
CH_4	<80	438
C_2H_4	<150	36
C_2H_6	<35	561
C_2H_2	<7000	18
TCG	<720	2541

The adverse effect for under-designing this transformer for requirements such as suitable conductor choice and size, and cooling system can be observed from the gas concentration exceeding the permissible IEC limits as shown in red in Table 3. Due to the excessive nature and production rates of hydrogen, a possible solution would be to replace the oil in all the affected units in case there were contaminants in the oil that resulted in the production of the abnormal gassing patterns. If the observed gassing patterns persist after the oil replacement, then a redesign of the transformer units might be needed. A proper harmonic study of the DSP plant as well as the loading conditions should be conducted in order to design for the DSP conditions. For the above

case, the oil in the transformer was replaced to counter the abnormal gassing phenomena. As can be seen from Table 4, replacing the oil had a positive impact on the gases in that new DGA results showed that all of the gasses were within the specified limits. This then proved the suspected oil degeneration.

Table 4: DGA oil sample results after oil replacement (ppm)

Gasses	Limits [23]	Gas concentration
H_2	<100	150
O_2	-	19455
N_2	-	90800
CO	<1000	89
CO_2	<15 000	2291
CH_4	<80	84
C_2H_4	<150	0
C_2H_6	<35	222
C_2H_2	<7000	0
TCG	<720	545

Air and moisture are generally the main determinants of gas pressure in the transformer oil. The progression of these components are determined by the transformer tank design and loading profile. The above irregular gas bubbling patterns are observed at hot spot temperature [22] far beyond the maximum permissible temperature of 78°C by the IEC. These patterns instigate rapid ageing of oil and cellulose insulation.

5. Findings and possible remedies for stray gassing

Novelty, remedial measures proposed on the current work from a transformer manufacture perspective in addressing stray gassing problems are summarised. These measures are based on collated data from many case studies conducted in the factory. During transformer service, these studies have been found to be linked up to the generation of hydrogen arising from stray gassing, in which stems from severely hydro-treated mineral oil, but is also strenuously affected by transformer thermal aging of polymers, choice of core steel grade, zinc tank walls and vanishes.

5.1. Harmonic spectrum

It was observed that an underestimation of the harmonic spectrum may affect the application of adequate winding design to account for the actual harmonics. Moreover, an improper cooling system is designed for the reduction of the HS temperatures in the windings and structural parts. In addressing these shortcoming, during the design stage, a complete harmonic spectrum of a DSP plant must be provided as it is critical for electrical designers in the computation of transformer load losses. The authors, recommends collaboration between the IPP and the manufactures to carry out in depth harmonic studies. The latter will enable adequate winding design and sufficient cooling system for the generated hotspot temperatures.

The main problems linked up to the tank walls include apertures, bubbling, orange peel and fissures. During the transformer manufacturing assembly line, these problems arises during processes linking up the tanking and painting phase.

5.2. Gummy tracks on the transformer tank walls: After drying out

Gummy tracks on the transformer tank walls arises from the addition of a sluggish drying dissolvent resulting to poor dissipation of the dissolvent and an insufficient hardening agent to the tank walls coating. The gaps among layers have been observed to be deficient when the coating is extremely dense and the coating layer is still wet. Additionally, the gummy tracks on the tank walls has been found to be due to contaminated tanks walls prior to coating.

5.3. Apertures in transformer tank walls

Apertures in tank walls during the assembly line may arise when the tank steel sheets are not cautiously handled. This phenomena has been observed as voids of diameter of about 0.1mm. Additionally, if the coating has not been settled for a certain period, the coating viscosity will be high and potentially the air bubbles have not been cleared. Additionally, prior to the tank welding, shot blasting must not be carried for the spare parts.

An adequate drying dissolvent and hardening agent must be designated and moderately added. The manufacture must also maintain the coated tank walls pollution-free. A best practice might be coating of the top-coat after the undercoating has completely dried out. When the climate is inimical during the assembly line, then the manufacture must ascertain excellent airing. The coating must not be carried out at high temperature and moisture.

6. Conclusion

This work, presents the results of stray gassing results that were carried out on a 2500kV, mineral-oil filled transformer. Particularly, solution oriented on the impact of harmonic currents on the service losses and thermal performance of the studied DSP the plant transformer. Moreover, the observed stray gassing occurrences of the unit as a result of (1) the generation of HS temperatures in the active part components; and (2) the core steel grade, zinc tank wall surface and coating.

In the case of an IPP who is not well informed about the actual harmonic levels occurring in the DSP plant, the authors recommend that manufactures make provision of the harmonic spectrum simulation prior to the electrical design process. In any case if the IEC harmonic spectrum meet the needs of the IPP, there should be a clear written agreement between the IPP and the manufacturer regarding the maximum permissible harmonics being in accordance with the IEC spectrum.

Novelty, findings of this work indicate that the generation of hydrogen arising from stray gassing may stem from severely hydro-treated mineral oil, but is also strenuously affected by transformer thermal aging of polymers, choice of core steel material and tank walls vanishes. The production surplus of methane and ethane are also witnessed in the first years of service and reaches substantial concentration levels. Potentially, these occurrences also arises from the thermal aging of polymers. The authors make some recommendations to utility owners to make a distinction of stray gassing from transformer fault by means of routine inspection aside from DGA value basis withal to the increase in gas diffusion rate. Finally, the authors highlight the need to establish standards that will provide support for transformers intended to operate in DSP applications. At large this work identifies the shortcomings of the design considerations for

transformers intended to operate in DSP applications and recommend corrective measures to address these challenges.

The authors make a significant contribution by further recommending procedures that can be employed as remedies during the design phase and manufacturing processes.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Government Gazette," Integrated Energy Plan (IEP)", The National Energy Regulator Act, **34**, 2008, doi: wp-content/uploads/2017/04/IEP-IRP-gW-ELA-Comment-310317.pdf.

[2] B. A Thango, D.B Nyandeni, P.M Molepo, "Solar Power Plant Transformer Loss Computation under Harmonic Currents using Finite Element Method", in 9th CIGRE Southern Africa Regional Conference, 1 – 4, 2019.

[3] E. Arslan, S. Sakar, M. E. Balci," On the no-load loss of power transformers under voltages with sub-harmonics", in IEEE International Energy Conference (ENERGYCON), 228-233, 2014, doi: doi: 10.1109/ENERGYCON.2014.6850433.

[4] S.V. Kulkarni, S.A. Khaparde," Stray loss evaluation in power transformers-a review", IEEE Power Engineering Society Winter Meeting. Conference Proceedings, **3**, 2269-2274, 2000, doi: 10.1109/PESW.2000.847708.

[5] T. D. Kefalas and A. G. Kladas, "Harmonic Impact on Distribution Transformer No-Load Loss," in IEEE Transactions on Industrial Electronics, **57** (1), 193-200, 2010, doi: 10.1109/TIE.2009.2030207.

[6] E. So, R. Arseneau and E. Hanique, "No-load loss measurements of power transformers under distorted supply voltage waveform conditions," in IEEE Transactions on Instrumentation and Measurement, **52**(2), 429-432, 2003, doi: 10.1109/TIM.2003.809910.

[7] B.A. Thango, J. A. Jordaan and A.F. Nnachi, "Service Life Estimation of Photovoltaic Plant Transformers under Non-Linear Loads", 2020 IEEE PES/IAS PowerAfrica Conference, 1-5, 2020 doi: doi: 10.1109/PowerAfrica49420.2020.9219912.

[8] A. Elmoudi, M. Lehtonen and H. Nordman, "Effect of harmonics on transformers loss of life," Conference Record of the 2006 IEEE International Symposium on Electrical Insulation, Toronto, Ont., 408-411, 2006, doi: 10.1109/ELINSL.2006.1665344.

[9] K. Dursun and N. Rahmanov, "Harmonic load losses in power transformer windings using Finite Element methods," Eurocon 2013, 1526-1530, 2013, doi: 10.1109/EUROCON.2013.6625180.

[10] D. M. Said and K. M. Nor, "Effects of harmonics on distribution transformers," in 2008 Australasian Universities Power Engineering Conference, 1-5, 2008.

[11] J. Weesmaa, M. Sterner, B. Pahlavanpour, L. Bergeld, J. Nunes and K. Sundkvist, "Study of stray gassing measurements by different methods," 2013 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, Shenzhen, 184-189, 2013, doi: 10.1109/CEIDP.2013.6748192.

[12] I. Hohlein, "Unusual cases of gassing in transformers in service," in IEEE Electrical Insulation Magazine, **22** (1), 24-27, 2006, doi: 10.1109/MEI.2006.1618968.

[13] D. Martin, N. Lelekakis, J. Wijaya, M. Duval and T. Saha, "Investigations Into the Stray Gassing of Oils in the Fault Diagnosis of Transformers," in IEEE Transactions on Power Delivery, **29** (5), 2369-2374, 2014, doi: 10.1109/TPWRD.2014.2316501.

[14] F. Jakob, P. Noble and J. J. Dukarm, "A Thermodynamic Approach to Evaluation of the Severity of Transformer Faults," in IEEE Transactions on Power Delivery, **27** (2), 554-559, 2012, doi: 10.1109/TPWRD.2011.2175950.

[15] M. Phoshoko, "Predictive maintenance of transformers through stray gassing studies", 66th AMEU Convention, 2017, doi: wp-content/uploads/2017/11/Pages-from-AMEU-2017_-46-51.pdf.

[16] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Stray Gassing of Transformer Oil in Distributed Solar Photovoltaic (DSPV) Systems," 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 484-488, 2020,doi: 10.1109/ENERGYCon48941.2020.9236522.

[17] B.A. Thango, J. A. Jordaan, A.F. Nnachi, "Effects of Current Harmonics on Maximum Loading Capability for Solar Power Plant Transformers", 2020 International SAUPEC/RobMech/PRASA Conference, 1-5, 2020, doi: 10.1109/SAUPEC/RobMech/PRASA48453.2020.9041101.

[18] B.A. Thango, J.A. Jordaan, A.F. Nnachi, "Step-Up Transformers for PV Plants: Load Loss Estimation under Harmonic Conditions", 19th International Conference on Harmonics and Quality of Power (ICHQP), 1-5, 2020, doi: 10.1109/ICHQP46026.2020.9177938.

[19] IEEE 519: 2014, "Recommended practices and requirements for harmonic control in electrical power systems", 1-29, 2014, doi: 10.1109/IEEESTD.2014.6826459.

[20] T.V. Oommen, S.R Lindgren," Bubble evolution from transformer overload", IEEE/PES Transmission and Distribution Conference and Exposition. Developing New Perspectives (Cat. No.01CH37294), **1**, 137-142, 2001, doi: 10.1109/TDC.2001.971223.

[21] M.C. Hlatshwayo, The computation of winding eddy losses in power transformers using analytical and numerical methods, Masters Thsis, 2013.

[22] IEC 60076-7, "Power transformers – Part 7: Loading guide for mineral-oil-immersed power transformers", 2005.

[23] "IEEE Guide for Loading Mineral-Oil-Immersed Transformers," in IEEE Std C57.91-1995 , 1-112, 1995, doi: 10.1109/IEEESTD.1995.8684643.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 132-138 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Application of a Reusability Approach in Simulation of Heritage Buildings Performance, Taif- Saudi Arabia

Ali Alzaed*

Architectural Engineering Department, Engineering College, Taif University, Taif, 26311, Saudi Arabia

ARTICLE INFO

Article history:
Received: 26 December, 2020
Accepted: 23 February, 2021
Online: 10 March, 2021

Keywords:
Heritage building
Reusability process and attributes
Thermal comfort
Green restoration

ABSTRACT

The main purpose of this paper is to present a reusability approach that helps the designer to assess the best practice to restore a heritage building. Based on the literature review, the reusability process and attributes was used as a method to restore the heritage building. Considering these approaches helps the designer to achieve useful results in terms of the built environment and building performance; moreover, it helps the designer to identify the suitable new usage of the building. Also, the designer validated the building performance through using the TAS to assess the thermal comfort of the building after using passive techniques and design restorations. The obvious finding was the successful achievement through considering this approach and decreasing the interior temperature two degrees. This study can be assessed as one of the optimistic practices that considered the sustainability dimensions during the restoration process, as well as improving the thermal comfort of the building for the end user. The research paper provides a useful understanding of designers, restorers and researchers.

1. Introduction

This section is concerned with reviewing and describing the reusability process and its attributes. One of the conservation practice methods that has been used to preserve heritage buildings is to adaptively reuse buildings. Several researchers have stated that reusability in heritage buildings is part and parcel of restoring them. It is considered one of the solutions to revive abandoned or neglected houses, especially in old or historical neighborhoods. In addition, this method can help the designer to document the building. For this reason, understanding reusability and its attributes is an essential demand. It is stated that the change of a historical building's function and renewing it can lead to reuse by the public, as well as making it more sustainable [1]. This type of building has a mixture of cultural and historical values. It is said that there are several heritage buildings worldwide that cannot be used for the same purpose that they were built for [2]. This paper will study reusability and attributes as an approach, on the one hand. After that, the researcher will apply the results of studying this concept to one of the historical buildings in Taif in order to reuse it. That will be by determining the most prominent functional, structural and aesthetic changes, as well as taking into account the environmental, economic and social considerations of the building. In addition to that, the researcher will assess the impact of applying the reusability approach in terms of thermal comfort.

2. View of the Heritage Housing Buildings in Makkah

Despite the multiplicity of tourist attractions enjoyed by the Kingdom of Saudi Arabia, archaeological and heritage sites occupy the most prominent place. Due to the kingdom's historical heritage, there are archaeological and heritage sites in different regions of the kingdom, representing a living record of successive civilisations in the Arab region. One of the most prominent regions in Saudi Arabia is Western Region, because it is one of the richest areas for heritage places. Sharbatly House, located in Jeddah at the entrance of Haret Al-Sham on the northern side, is a good example. It consists of four floors built with excavated stone in the year 1350 AH [3]. This shows the consideration of the environment through selecting local material. There are several heritage buildings, such as Nassif house. It is located in the heart of Jeddah city and it consists of four floors, with beautiful façades adorned with doors, windows and wooden pedals. These worked in a coordinated manner and were decorated with various decorations. This house was built in 1289 AH and was completed in 1298 AH, and was built with excavated stone castings in a square shape that was cut from the coast of the Red Sea. It is classified as a high architectural landmark and a witness of the architectural thought of the old style in Jeddah [3]. This house is distinguished by using Red Sea stone,

*Corresponding Author: Ali Alzaed, Dralzaed@gmail.com

which is a local material, and is also unique because of the material of the doors and windows, which has a clear impact in terms of the built environment. Another building, Noor Waly house, is located in the old city of Jeddah, and was built in the year 1280 AH. The Islamic Museum is a landmark place located in Makkah, and is a museum specialising in Islamic monuments and Meccan heritage [3]. Its style gives the impression of the prestige of this place. The decoration of this palace was selected to be ecological. Jeddah Municipality Museum is located in Al-Balad district, Jeddah. It has been restored and furnished with many heritage collections that were from people of Jeddah until recently, as well as artefacts dating back to before the Islamic era and pieces from the Islamic era [3]. There is a wide variety of pieces for this place. The Meccan Heritage Museum is located on the ground floor of Tariq Sindi's house. It contains various types of utensils; wooden, metal, and marble tools; Meccan textiles; guns, swords, and daggers, all of which are related to weapons; a number of artefacts, inscriptions, postage stamps, records, and notebooks [3]. Based on the reviews, some of these buildings have changed to become museums or some other use. This gives us a hint about the changeability of the function, which can lead to another successful story.

Alkateb House was later known as Al-Niyabah Palace. This house is located in the Al-Salam neighbourhood in Taif, and it consists of three floors built by Hamad Abdul Wahid, the private writer of the sheriff Aoun Al-Rafeeq in 1315 AH. It became deserted in 1388 AH. This building was marked by Roman artistic and architectural influences. It consists of a main foyer, at the end of which are staircase steps with marble stone handrails leading to the upper floors, and at the back, a large irregular orchard [3] and [4]. This building is distinguished by its view and styles, which include several local materials. This link with the local environment is considered as one of the successful opportunities for reuse for a different future activity, such as a café or hotel. Alkaaki House is located in the Al-Salamah neighborhood in Taif, in front of the defence building, and it consists of three floors. It was designed in the Roman style and built in 1358 AH. This house is distinguished by its various decorations and ceilings, and there is a garden in the rear belonging to it. There is also a section for men for women, known as the Harmalik and Selimk [4]. There are many experiences that have been applied to reusing heritage buildings in historical cities. The historical city of Erbil is an example, where the author mentioned the importance of having balance between conservation and reuse of these kinds of building [5]. The location of the selected case study and its design concept shows the possibility of adaptive reuse in the future as a commercial activity. This will be verified by applying the methodologies as described in the next section.

3. Methodology

The research paper uses several methods, as follows:

- Reviewing the heritage building background in the Western Region of KSA, and the reusability approach and its attributes.
- Developing and adopting reusability as a process and as attributes for green restoration of traditional buildings in Taif City, in keeping with its traditional character. This demands documenting the building using several methods such as AutoCAD, Survey and Photo Snap to assess the current situation.
- Assessing the impact of using a reusability approach on thermal comfort through using one of the most powerful energy building tools (TAS EDSL) [6].

4. Reusability Process and Attributes

Reusability as a process has been developed by several researchers. One of them developed a proposed iterative process model [7]. The process of this model for reuse-oriented software is reusability. This model consists of five stages as follows: Stage (1) Planning Phase: to study the expected profits in terms of technical aspects, reducing the cost of management and development [7]. Stage (2) Risk Analysis: this stage has two approaches which are reactive and proactive [7]. Each one of them has several features. Proactive features are the domain stable which means it is easy to invest directly for reuse and easy to predict the product features. However, reactive features mean it can be developed based on what is required. This approach is how to reengineer and adjust the current assets. Stage (3) Core Asset Development and Evolution: this includes determination and definition of features while understanding the similarities and differences between them. Furthermore, extension of the architecture if there is a clear demand and addressing the conflicts between the quality and product functionality. Stage (4) Change Requirement and Adaptation: a new requirement that will lead to changing the technology and environment. This usually will be during development work which is not expected. For this reason, the software has to be adapted to that and capable of managing it. Stage (5) Customer Evaluation: observation of the feedback of the customer [7]. To apply these terminologies and approach, there is a clear demand to understand what the main attributes of the reusability concept are, to be the basis for more understanding of this philosophy. This will give the research a clear vision about the possibility of its application on restoration of buildings. The main attributes are illustrated in Table 1:

Table 1: Reusability Attributes

Attribute	Description	References
Flexibility	Easily modify a system or component for new uses, different from what it was originally designed for.	[8] and [9]
Maintainability	Easy-maintenance software system or component to add capabilities or change or correct faults.	[9]
Portability	No external support is required to perform the functions, and has the ability to use alternative solutions at any time.	[8] and [9]
Variability	The ability of a component to be configured which helps to decrease understandability.	[8]
Understandability	Can also be defined as the level of understanding of the system's way of working.	[9]

Size	The lines of code and length of the software.	[8] and [9]
Complexity	The difficulty in understanding the design and implementation of software, which means reusing complex software.	[9]
Scope coverage	The number of features that the component provides through a particular feature.	[8]
Availability	How easy it is to retrieve a software component.	[10]

5. Reusability Process and Attributes for Green Heritage Building

In this research, selecting reusability as process and attributes may help designers to study and evaluate the feasibility of any building reuse. The current status of a green heritage building will be evident in the following case of a heritage building in the historic center of Taif. However, the process that has been adopted is shown in Figure 1.

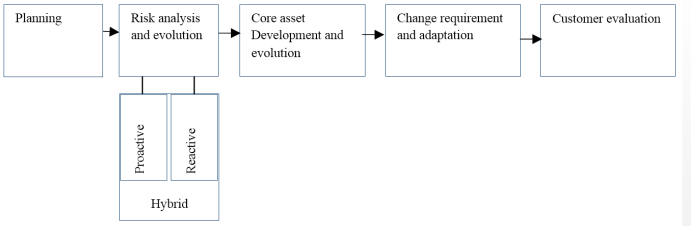


Figure 1: Reusability process

5.1. In terms of planning stage

In this stage the building was abandoned. For this reason, to assess the benefit of restoring this building, the designer has to understand its main previous function and its retirement. In addition, it has to minimize the cost of development as well as its future management. As a result, there was a clear demand to document this heritage building as a study case for this research. For this reason, the building has been visited and surveyed to discover its boundary and investigate the environmental conditions as shown in Table 2.

Table 2: Survey points

Points	E	N
1	40.4081340	21.2749090
2	40.4080430	21.2748770
3	40.4080660	21.2748220
4	40.4081550	21.2748530

In addition, it is important to understand the previous function of the building, which was residential. For this reason, there is a clear demand to know the building components, elements and features. This was done through a manual documentation device which was used to draw building floors using AutoCAD software. All building floors were accessed by natural ventilation and daylighting, except one zone in the ground floor, as shown in Figure2.

5.2. Risk Analysis Stage

Rethinking the new use of the building was done through meeting with the building investors. They were confused about the optimal use that would bring them economic returns because housing investment was not encouraging for them, compared to similar experiences in the same location. It was considered that there were no museums in the city center, which could be an ideal option. However, one of the main landmarks of Taif City is Shubra Palace Museum, not far from their location. For this reason, the decision was to make the museum with a special and distinctive style. This is in terms of identification of the proposed activity, on the one hand. On the other hand, in terms of the building it has to be identified if it is proactive or reactive. In terms of being proactive, the building is stable and reliable and only needs a low budget. At the same time, the building can be reactive as there is a demand for some adjustment and reconfiguration in the ground floor, as shown in Figure2 and Figure3.

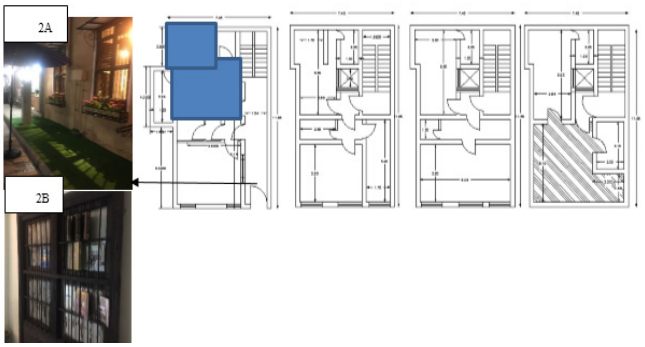


Figure 2 : Building floors.

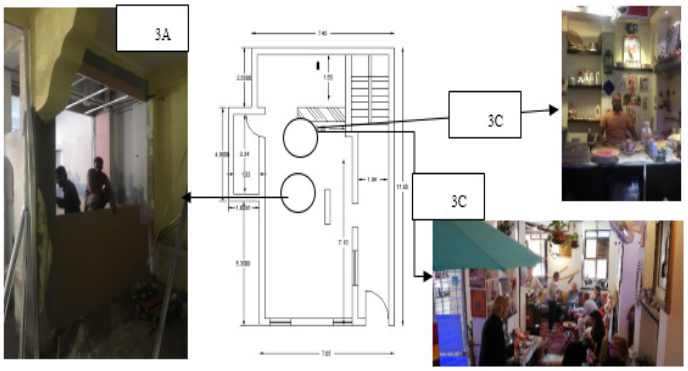


Figure 3: Ground floor before and after restoration

5.3. Core asset development and evolution

The decision was to make the museum with a special and distinctive style. This is in terms of identification of the proposed activity, on the one hand. In addition to that, the building has several attributes that support the decision. In the first place, the building style is heritage. This is obvious on the building facade which includes roshanah, which is the window made from wood. In terms of the interior, the building is capable of being a museum, as there were cavities on the wall as well as shelves. Most of the attributes considered during restoration are mentioned with more explanation, as indicated in Table 3. In addition to that, the ground floor zones have been mixed to be one zone as shown in Figure3.

The natural ventilation and daylighting did not reach the back section, which means it was not suitable for the future user.



Figure 4: Second floor restoration

5.4. Change requirement and adaptation

This happened during the building use and activation. Some of the zones have had their functions changed to be galleries, following which signs and paths on the floors were created to assist visitors. In addition, the ground floor has adapted a café to enhance activity for the museum which provides a service to visitors. The roof use has been changed to become a painting studio. All of these additions were after the building was set up with its new activity as a museum.

5.5. Customer evaluation

The overall impression of the visitors was very positive. This was evident by their enjoyment of the experience. In addition to the diversity found in the building is a café for hospitality, an art gallery, a drawing studio and a museum. All these factors enhance the visitor experience. A limitation that helps measure user satisfaction is the thermal comfort of buildings. Therefore the interior temperature of the building was evaluated before and after the restoration using the simulation software (TAS) as mentioned in the research methodology. The results of that will be explained in the following sections, and the building material and location description is shown in Table 4.

Table 3: Reusability green heritage attributes

Attributes	Descriptions
Flexibility	The designer removed the wall in the ground floor to have a large space and to let the air and light access the whole area, as highlighted in Figure 2 and illustrated in Figure 3. In addition, the skylight is adopted as the main hole as shown in Figure 4 (4E-4F), as well as the main wall being used for shelves as illustrated in Figure 4 (4C) and the staircase used as a store.
Maintainability	External electrical installations as shown in Figure 4 (4A). Plant pots easy to install, remove and maintain.
Portability	Fans and industrial lighting as shown in Figure 2 (2A).
Variability	The designer created the paths for the gallery as a guide, as in Figure 4 (4A). Indicative maps and configuration of the building were based on user needs, such as modifying the gallery to be a hospitality salon as shown in Figs. 4C and 4D, and using the heritage window (oriel window) to be bookshelves as shown in Figure 2B.
Understandability	Features of the rooms, ways of furnishing them and the building map in the distribution of its elements.
Size	Removing the wall of the ground floor and adding shade shelters as shown in Figure 2A. Some decorations on building façades may also fulfill an ecological function [11]. Agricultural ponds, wooden pergolas, shelves and gypsum decorations as shown in Figure 4G.
Complexity	Narrow space and lack of ventilation and lighting for some spaces; floors on the ground floor have more than one level; the electricity network was irregular and random; and the bathrooms were not suitable.
Scope coverage	• Outside coverage; turn it into an attraction. • Choose the right function. • Configure building activity from house to museum. • Enhance the returns through the configuration based on a coffee place, as shown in Figure 3C, and gallery (Figure 4A). Rent the rooms as hospitality places and ticketed areas, and configure the main roof as a painting studio.
Availability	All restorations and additions, such as removable wall, skylight as illustrated (4E and 4F), and shelters on the main door and windows as shown in 2A, are easy to remove and restore the building to its previous use.

6. Numerical Simulation

Distinctive physical models are utilized to depict the structure warm conduct contingent upon their particular needs, counting space warming, ventilation, cooling frameworks, inhabitants' practices, and budgetary viewpoints. The physical structure execution depends on the unraveling of the warmth move

differential conditions, which can be written regarding the energy preservation law:

$$Q_{\text{int}} + Q_{source} = Q_{out} + Q_{stock}$$

Where Qint is the heat flux entering the system, Qsource the heat flux of an eventual heat source, Qout the heat flux leaving the system and Qstock the heat flux stored. The principal in and out-coming fluxes taking place in the heat transfer are the conduction through walls, the convection (both sensible and latent heat), the radiation and the ventilation.

Countless mathematical virtual products are these days accessible to take care of such physical issues, both in fixed and dynamic conditions. The consistent state approach is ordinarily used to survey the structure energy execution and prompts long haul examinations of various situations thank to their quick computations. In any case, a significant constraint happens since the dormancy of the structure envelope is totally disregarded. Consequently, dynamic warm models ought to be favored while breaking down energy sparing arrangements.

Building demonstrating and reproduction programming TAS was utilized to foresee energy execution, improve warm and subsequently inhabitant comfort (EDSL TAS 2017). The documents used to finish the model on TAS are the arrangement sees appeared in Figure 2. At first, the model made on TAS was an imitation of the current condition of the building. Hence, the underlying produced energy model was the reference point for enhancements.

In the present study, the following modelling assumptions are adopted:

1. It was expected that the residence is involved from 6pm-8am during non-weekend days and for an entire 24 hours during the end of the week.
2. The programmed recreation of characteristic ventilation (due to windows, entryways, ventilators, and different gaps comparative with their elevation and direction) will be thought to be the reasonable portrayal of the real wind stream (EDSL TAS 2017).
3. The National Calculation Method (NCM) information base will be utilized to apply to all zones (EDSL TAS 2017).
4. The development of the TAS model is dependent on compositional drawings (Figure2). In any case, to improve the model the impact of the encompassing trees was not considered in this reproduction. At first, utilizing the accessible design drawings of the genuine structure, the floor plans were attracted the 3D-TAS. The following stage in the model improvement was to characterize zones inside each floor for warm reproduction purposes. This zone definition was critical as it impact the manner in which the model would be dissected. The indoor warm exhibition of the building, especially identified with more elevated level warm definition, was essentially estimated and investigated by contrasting the hourly normal indoor air temperature and mean brilliant temperature of zones.

These days, there are two principle ways to deal with gauge human warm solace Predicted Mean Vote (PMV) furthermore, the Percentage of Dissatisfied (PPD). These methodologies which consider a few ecological viewpoints, for example, dry bulb temperature, moistness, air speed and mean brilliant temperature (Figure 5) just as human factors, for example, warm obstruction and metabolic rate. Warm balance is inferred when the interior warmth creation in the body is equivalent to the warmth misfortune to the encompassing climate [12].

In the present study, the thermal conductivity (Watt/m2.K) is assumed to be constant for any given material, and is independent of temperature; i.e.

$$\lambda = \frac{q}{A(T_1 - T_2)}$$

The U-value for a certain structure is expressed as the ratio of the density of heat flow rate through the structure q, W/m2, and the difference between the internal and external temperature values, as follows:

$$U = \frac{q}{(T_i - T_e)}$$

It will be accepted that the Location properties of the designed building are the real current states of the building (Table 4).

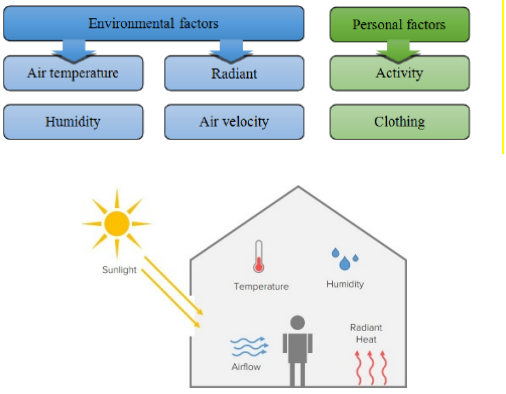


Figure 5: The important factors affecting thermal comfort in the two approaches (PMV and PPD) [13]

Table 4: Location properties of the designed building (Taif City) and green heritage building material

Location properties of the designed building (Taif City)		
No	Property	Description
1	Location	Downtown Taif, Saudi Arabia
2	Latitude	21.44 °N
3	Longitude	40.51 °E
4	Elevation	1700.0 m
5	Daily lower average air temperature	10.9 °C
6	Daily upper average air temperature	32.0 °C

7	Average air flow speed		3.8 m/s	
8	Relative humidity		46%	
Building materials				
	Layer	Width (mm)	Conductivity (W/m².°C)	Total U value (W/m².°C)
External wall	Block	300	0.25	0.73
Ground	Sand dry	1000	0.329	0.286
	Crushed aggregate	75	0.55	
	Concrete	150	0.87	
Roof	Polystyrene expanded sheet	125	0.04	0.297
	Concrete	100	1.45	
Shading	Wooden shading	10	0.14	2.848
Glazing	Type of glazing	Width (mm)	Solar reflectance	Solar absorption
	Single	10	0.070	0.115
	Solar transmittance	Emissivity	Total U value (W/m².°C)	
	0.70	0.845	5.681	

7. Results and Discussion

With respect to the impact of the reusability approach on the built environment, the main changes that have been made are to reconfigure the ground floor to be one zone, as well as to install a skylight to be a wind catcher. These simple treatments were used due to the moderate climate of Taif. In contrast to hot, dry climates that require some different treatments. An example is the evaporative cooling that was tested on an apartment building in Baghdad [14]. All these measures can lead to enhancing the thermal comfort of the building. Before the reconfiguration of the building design and through the summer season, the outdoor air temperature was 35.6 degrees at 2.00 pm and the indoor air temperature for hall one was 34.54 degrees, almost one degree different, while hall 2 was 42.87 degrees at the same time, as shown in Figure6. However, after the reconfiguration, the ground floor became one zone. The indoor air temperature becomes 33.33 degrees, which means that it decreased by approximately one degree, as shown in Figure7. This gives a clear indication of the impact of the reusability approach through mixing the zones on improving the thermal comfort for the end user. In addition, the wind catcher can add value to that and enhance the thermal comfort performance. During the wintertime, the outdoor temperature was 13.8 degrees and the indoor air temperature for hall 1 and hall 2 were 14.3 and 20.29 degrees, respectively. However, after reconfiguration of the ground floor, the indoor air temperature become 14.04 degrees, which is still higher than the outdoor temperature, as illustrated in Figure8.

The discussion above illustrates that the reusability building approach and reconfiguration can have a clear impact on the interior thermal comfort. In addition, the restorations of the design and application of some techniques such as a wind catcher can lead to improving the interior energy without any mechanical demands. Using these approaches and passive techniques is more beneficial than a huge restorations of the building, especially in heritage buildings.

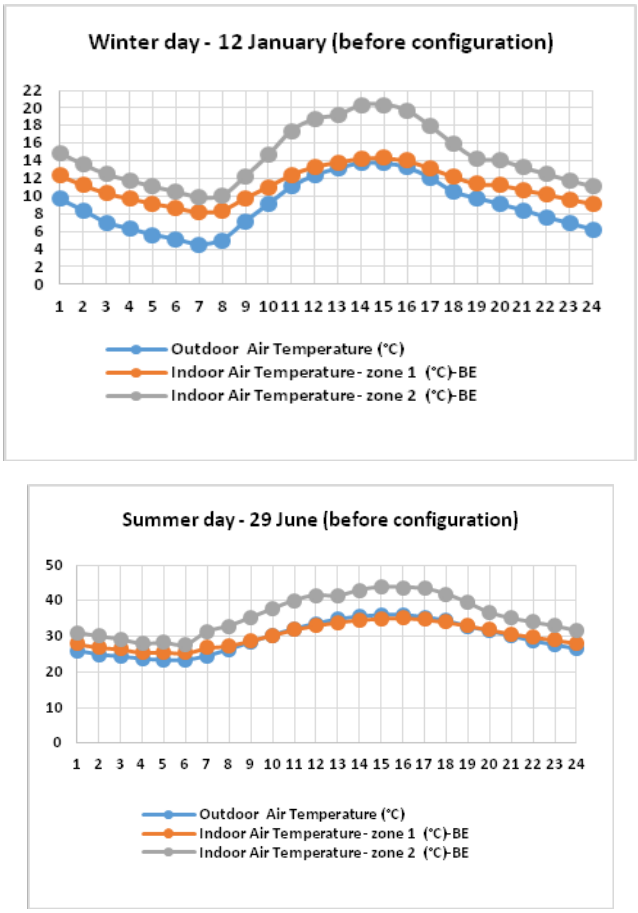
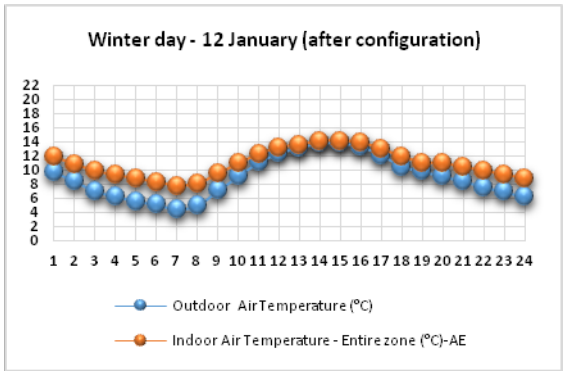


Figure 6: The result of outdoor and indoor temperature before configuration



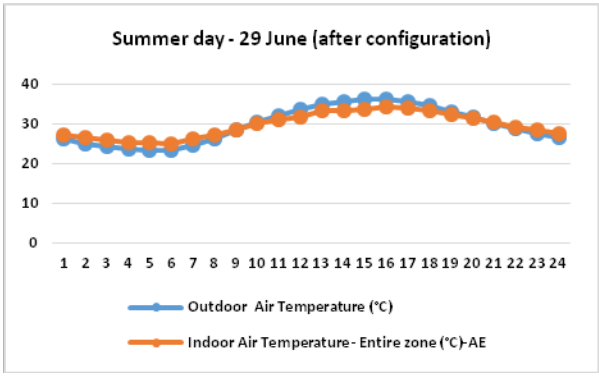
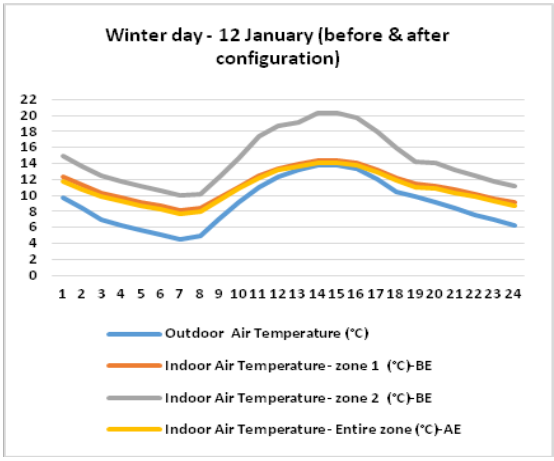


Figure 7: The result of outdoor and indoor temperature after configuration



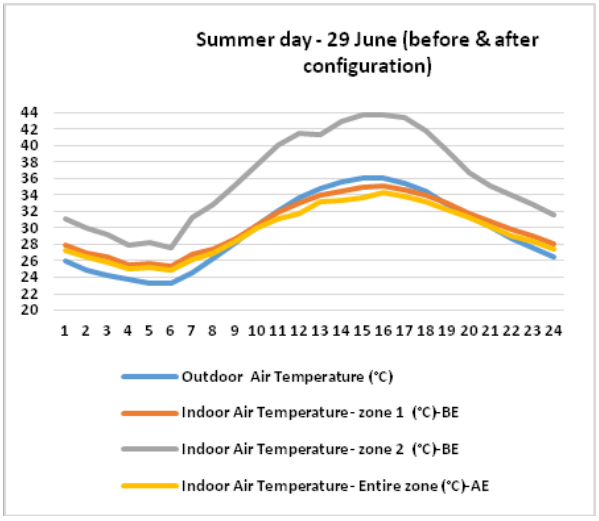


Figure 8: The result of outdoor and indoor temperature before and after configuration

8. Conclusions

The work of this research investigated the impact of using the reusability approach on the building restoration, usage and performance. This was clear in a hot and cold climate in Taif City. Usage of local material, a wind catcher, and reconfiguration of the interior design were investigated. One of the obvious results is the effect whereby the reusability approach can lead to a decrease in the interior temperature by two degrees in hot climates. In addition, using the reusability approach with some passive techniques is more beneficial than a huge design reconfiguration. The numerical simulation results indicated that due to the restoration process adopted, there is a decrease in the inside temperature of the building. However, the small value of temperature reduction can be referred to the nice weather of Taif city, where the heritage building found. The most important conclusion is such restoration can lead to temperature reduction. Any way, we think that such restoration process might be important in hot weather cities.

References

[1] M. Ulusoy, E. Erdogan, H. Erdoğan, M. Oral, "RE-USING OF THE HISTORICAL BUILDINGS IN THE CONTEXT OF SUSTAINABLITY: AN ARCHITECTURAL DESIGN STUDIO STUDY ON OLD GIRLS TEACHER TRAINING SCHOOL," in ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2013, doi:10.5194/isprsarchives-XL-5-W2-653-2013.

[2] R. Ghada, R. Amany, A.R. Riham, "Adaptive Re-Use and Sustainable Development for Existing Historic Buildings – Case Study: Buildings of Racetrack Horses in Sporting Club, Alexandria, Egypt," International Journal of Current Engineering and Technology, **7**(4), 1523–1530, 2017.

[3] A. N., Archaeological lexicon of the province of Makkah. Makkah, K.S.A, Commission publications in the revitalization of tourism Taif: 176–190, 2003.

[4] A. N., Archaeological lexicon of the province of Taif. Taif, K.S.A, Commission publications in the revitalization of tourism Taif: 109–129, 2002.

[5] A. Abbas, B. Khaznadar, "Using Adaptive Reuse as a strategy for the renovation of traditional vernacular architecture in Erbil city," 2019.

[6] EDSL. Tas, Available: Https://Www.Edsltas.Com,.

[7] J. K.s, D. Vasantha, "A New Process Model For Reuse Based Software Development Approach," Lecture Notes in Engineering and Computer Science, **2170**, 2008.

[8] A.K. Mahmood, "Reusability Assessment of Open Source Components for Software Product Lines," International Journal of New Computer Architectures and Their Applications (IJNCAA), **3**, 519–533, 2011.

[9] C. Monga, A. Jatain, D. Gaur, "Impact of quality attributes on software reusability and metrics to assess these attributes," in 2014 IEEE International Advance Computing Conference (IACC), 1430–1434, 2014, doi:10.1109/IAdCC.2014.6779536.

[10] D. Hristov, O. Hummel, M. Huq, W. Janjic, "Structuring Software Reusability Metrics for Component-Based Software Development," International Conference Software Engineering Advances, (c), 421–429, 2012.

[11] F. Alasadi, G. Al Slik, "Ornamentation and Modern Architecture in Iraq," Journal of Engineering, **25**, 117–124, 2019, doi:10.31026/j.eng.2019.06.09.

[12] ISO 7730:2005(en) Ergonomics of the thermal environment — Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria, 2005.

[13] Quick Guide: Overheating and Thermal Comfort, Available: Https://Ggbec.Co.Uk/Quick-Guide-Overheating-Thermal-Comfort/,.

[14] A. Alhosainy, I. Aljubury, "Two Stage Evaporative Cooling of Residential Building Using Geothermal Energy," Journal of Engineering, **25**(4), 29–44, 2019, doi:10.31026/j.eng.2019.04.03.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 137-145 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Review on TAM and TOE Framework Progression and How These Models Integrate

Julies David Bryan*, Tranos Zuva

School of Applied and Computer Science, Department of ICT, Vaal University of Technology, Vanderbijlpark, 1911, South Africa

ARTICLE INFO

Article history:
Received: 08 December, 2020
Accepted: 27 March, 2021
Online: 20 May, 2021

Keywords:
TAM
T-O-E
Limitations
Criticism
Trust
Reliability
Pricing
Technology context
Organisation context
Environment context

ABSTRACT

The TAM is a model that is widely used to understand IT adoption and usage process accordingly and the reason for its popularity is that the model clarifies variances like behavioral intention (BI) relevant to IT appropriation and use over a broad range of settings. The model's main factors for system use is perceived usefulness (PU) and perceived ease of use (PEOU). Likewise, the T-O-E framework is a popular framework for three stimuli that influence organizational adoption namely technology, organization, and environment. Much literature has dealt with the use of TAM and T-O-E frameworks together with their derivatives without looking at the shortfall of these models. This paper reviewed one hundred and seventeen papers that used or reviewed TAM or TOE models. The contribution of this paper is the address of the usefulness, limitations, and criticism of the two models and also how the TAM and the T-O-E frameworks can be integrated into a hybrid model using a generic framework. In conclusion these models can be used separately or as a hybrid depending on the situation at hand. In future it important to harmonize the so many factors of the models that have been suggested and used in literature.

1. Introduction

Research studies on IT adoption have rapidly evolved over the last three decades. A diverse number of theoretical models have been created such as Innovation Diffusion Theory (IDT), Theory of Planned Behavior (TBP), Diffusion of innovations (DOI), Electronic Data Interchange model (EDI), etc. Two of the most valued research studies that contribute to IT adoption are Technology Acceptance Model (TAM) and Technology-Organization-Environment (T-O-E).

In the 1970s, anticipating system growth became a niche of attraction to many researchers due to the growing technology demand and system adoption failure that goes with it [1]. According to [2] the vast majority of studies neglected to deliver solid estimates that could clarify framework acknowledgment or dismissal and recommended Technology Acceptance Model in 1985 [1]. Davis proposed that the user interaction with any system is a reaction that can be interpreted and foreseen by user inspirations, shaped by foreign catalyst [1]. As per [1] user inspirations are shaped through a foreign catalyst that abides system's features and capabilities. Davis, who counted on preceding work done by [3] continue amend his model and propose the Technology Acceptance Model in which he advises that user motivation can be driven by three factors:

- Perceived Ease of Use (PEOU)
- Perceived Usefulness (PU)
- Attitude towards using the system. (A)

According to [1] Davis's thesis on user's attitude towards a system is a major factor that determents if a user will use or reject a system. A user's attitude is leverage by perceived usefulness and perceived ease of use.

Broad literature on Information Technology (IT) appropriation demonstrates that there are a few investigations at the individual level [4]. Numerous theories and models uses IT adoption at the individual level such as Technology Acceptance Model (TAM) [1,2,4], Theory of Planned Behaviour (TPB) [3], TAM 2 [5] Unified Theory of Acceptance and Use of Technology (UTAUT) [5]. However, studies show that there is less investigation at organisational level.

According to [6] the whole process is an innovation from development by engineers and business people to the adoption and execution of those advancements by clients inside the context of

*Corresponding Author: Julies David Bryan, bryanjul@gmail.com

an organization. As per [7] the T-O-E framework performs the process of how the organisation impacts the appropriation and execution of innovations. As per [7] the T-O-E framework is an association hypothesis which clarifies three unique components association's setting impact reception choices. The components are technology, organisational and environmental context, and each of the three is set to impact technological innovation. [6] built up a structure for hierarchical reception dependent on the Contingency Theory of Organisations. According to [3] the strength of a organisation is determined by both internal and external elements which can be described as environment, organisation size, and organisation strategy. Upon decision making three key elements environmental, organisational and technological needs to be factored in, and organizational adoption is based on technology, organisation, and environment [4].

The contribution of this paper is to determine and understand usage and modifications, progression, limitations, and criticisms of the Technology Acceptance Model (TAM) and Technology-Organisation-Environment (T-O-E). The paper also looks at how these two models can be integrated to complement one another.

2. Technology Acceptance Model (TAM) Overview

As per [8], [2] established the Technology Acceptance Model (TAM) by assuming that users will perceive usefulness to technology through ease to use as they become willing to use the technology. When employees realize that the new technology will make their life easier and more productive the greater the odds are for them to use and accept it [8].

As per [1], with the rapid technological growth in the 1970s, many organizations struggled to adapt to system failure. Due to the increasing system failures, numerous researchers started to study the area of predicting systems [9]. The TAM was then proposed by [10] in 1985 during his studies at MIT Sloan School of Management [1]. According to [1], Davis proposed that system usage is a human behavior that is user motivation driven and directly influenced by external stimulation like system's components and efficiency. This can be seen in the Figure1 below.

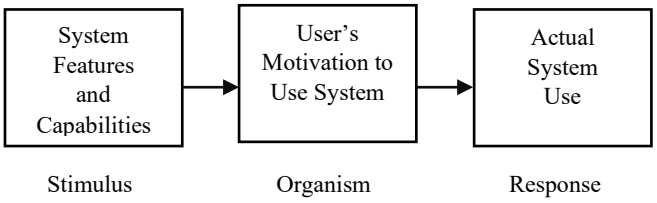


Figure 1: Model for technology acceptance [22].

Davis then continued with his research by relying on prior work done by [11] on the Theory of Reasoned Action. The refined model of Davis can be seen in Figure 2.

In Figure 2, [11] propose that user motivation can be described by three factors: Perceived Ease of Use, Perceived Usefulness, and Attitude Towards Using the systems [1]. As per the [9] Davis model, TAM hypothesized that attitude towards the system is a considerable factor whether used or rejected by the user. User's attitude is influenced by two primary convictions namely perceived usefulness and perceived ease of use.

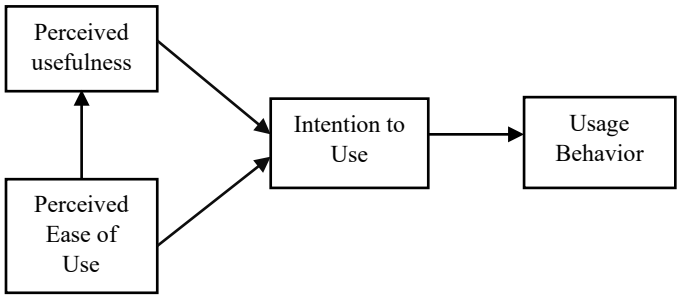


Figure 2: Original TAM [11].

According to [12] on the contrary, states that the TAM theory is based on social behavior by someone's attitude which is predicted to use an information system. [3] argue that when it comes to personal use of technology, social influences like friends and colleague is a major factor. However, when it comes to the organisations, the work climate cannot be controlled by follow workers but the organisation regulate direct the conduct of the workers as there is rule-administered conduct at organisations for utilizing a system. [8] states that TAM is simple and easy to understand and is mention by [13] to be the most popular means of measuring the scale of acceptance of technology by users.

TAM studies the individual relationship of technology acceptance, adoption, and BI to use [6,14]. As per [13] and [14], the main factors for system use with TAM is PU and PEOU. PU can be described as "the prospective user's subjective probability that using a specific application system that will increase his or her job performance within an organisational context," and PEOU is defined as "the degree to which the prospective user expects the target system to be free of effort" [2]. As per [15], PEOU influences PU because of its ease of use making TAM a popular model. It has been proven by numerous research studies that these two variables are frequently used for testing and [14] stated that these PU and PEOU variables are mentioned 40 percent for people's goal to utilize (admission) and resulting execution (transformation) of technology. TAM hasn't only been developed to predict user adoption, but it was also designed to predict user behavior after the individual has interacted with the system [16]. The TAM model is supported as a goal-based model which specifies that the goal to embrace innovation is a decent indicator of its genuine use [16]. Henceforth, it very well may be reasoned that the TAM model got solid ramifications for technology endorsements from conceptual and theoretical perspectives. [6]. [17] declares the primary reason why TAM was developed was to model user consent for IT with the desire to explain the BI system usage that was a footing for IT dealing with BIs and usage of IT [17]. [6] states that to have a better overview on IT adoption and acceptance, the primitive framework (i.e. TAM2 and TAM3) had to be extended and includes PU and PEOU categories. There were four modifications that assisted with the TAM evolution, these can be named as adjusting outer precursors; changing prescient factors; controlling mediator factors, and shifting outcome measures [13] and [18]. The underlying center segment basic to all alterations of TAM comprises of three builds: PU, PEOU, and BI [13,18] and in the centre part, BI is influenced by PU and PEOU, PU is affected by PEOU. It is concluded by [18] and [3] PU and PEOU are two

centre develops that shape client perspectives and aims in adopting a technology system.

In summary, the TAM model is widely used to understand IT adoption and usage process according to [6]. The reason for its popularity as per [16] is that the model clarifies variance like behavioural intention (BI) applicable to IT adoption of use overbroad contexts. Another reason why TAM is so famous is that it foresees an owner's IT consent by the user and the job usage [14]. [1] on the contrary, states that the reason why it is widely used is that TAM explains the determinants of user acceptance of a wide range of end-user computing technologies.

3. Limitations and Criticism of TAM

Even though TAM is popularly used, it does come with limitations and these are found in moderating and external variables [13]. TAM looks at future behaviour and not on actual behaviour [15, 18] as this also aids to its limitations. [13] confesses that TAM is known for its restricted chance of clarification and expectation, detail and absence of common-sense worth.

A crucial extension to TAM by [19] was to introduce TAM2 as TAM had some limitations in illustrating reason for a person to perceive system usefulness. For this reason, additional variables were added to the perceive usefulness variable in TAM [9]. A second important alteration to TAM by [19] was identifying the decedent to the perceived ease of use variable in the TAM model. For this reason, two new decedents namely anchors and adjustments were added. As per [9] anchors were thought-about personal computers and the usage of it, well adjustment was thought-about as direct involvement with the target system. Other limitations stated by [9] has been flagged by numerous researcher and has grouped the criticism into three categories:

3.1. Limitations in the methodology used for testing the TAM model

The biggest criticism of TAM is the model makes use of self-report use data instead of measuring against real system data. According to [20] self-report data is abstract and inaccurate when measuring system usage. Even so, many studies still make use of self-report use data. As per [21] the TAM model makes use of students as participants in a controlled environment to obtain test results and for this reason, these test results cannot be generalized to the outside world as studies have distinctive intentions like obtaining grades, rewards, etc. [20, 22].

3.2. TAM model variable and relationship limitations.

According to [23] attitude is an important factor for system usage and TAM needs to review it. The TAM model was clone in 1998, but the attitude variables were not removed as proposed by [21], instead, two additional attitude variables, affective and cognitive were added. [21]. A survey by [23] requested participants to rate their usage of spreadsheet applications. The outcome of the survey shows that the effective attitude variable did not portray the statistical connotation to anticipate system use however the results on the cognitive attitude were very significant.

3.3. TAM model theoretical foundation limitations.

According to [24], the poor hypothetical relationship was defined among the various builds detailed in TAM. He scrutinized the hypothetical quality of the goal real use connect, and saw that conduct couldn't be considered as a terminal objective. [24] claims that conduct should be treated as a way to a more terminal intention and behavior should be a fundamental target. Besides, he clarified that the aim may not hold delegate enough of real use, because the time-frame among goal and appropriation could be full vulnerabilities and different elements, that may impact a person's choice to embrace innovation. In [24], likewise scrutinized the chance of deciding conduct by including measures for perceived usefulness and perceived ease of use.

"There is a scope of investigating the role of certain other variables such as technological influences, the role of firm size in the technology make/buy decisions, the innovativeness of the firm, a firm's level of technology readiness, security, trust, marketing effort and also on evaluating the consequences of technology usage on performance such as responsiveness and financial performance" [15, 18, 25]. It must be noted by [26] that the TAM has been used in almost every IT research adoption theory and because of its frequent use, the quality of TAM could be at risk, as this could lead to degreed in the research field. Another limitation mentioned by [20] shows that TAM experimental investigations don't deliver predictable or clear outcomes and proves that demonstrates the TAM model needs to be incorporated with other IT reception models and hypotheses.

4. Technology Acceptance Model (TAM) progression

As per [13, 27, 28] the TAM is a principal adoption theory in the IT sector. TAM discloses how external variables can have an impact on adoption decision making on fundamental financial, functional, and beliefs and therefore recommends PU and PEOU as the central reason for adoption in IT [9]. An individual's goal to exploit an application is interpreted and forecast by his view of the innovation's handiness and its straightforwardness. According to [19] perceived usefulness can be determined by both perceived ease of use and predict attitudes. Over the years TAM has shown great acceptance, utilization, and reproduction but the model failed to give essential data on the user's assumption for only PU and PEOU. Based on this limitation it was necessary for the model by expanding or integrating with other IT acceptance models [9]. The integration of TAM with different models enhanced TAM by including utilization and putting premiums in explicit settings and outside factors that impact an innovation's appropriation procedure [9].

According to [7] many studies have been made to the original TAM model the causes the model to evolve. TAM-TBP was a new model introduced by [16] to integrate TAM and the theory of planned behaviour (TBP). In 2000, TAM2 was developed by [19] that extended new variables to the existing model. In 2003, the Unified Theory of Acceptance and the Use of Technology (UTAUT) model was proposed. There have been numerous studies made by scholars to alter the TAM by adding new variables. The Construct of Compatibility variable was added in 1998 by [13].

New playful factors were added by [29] to examine the acceptance of the Internet. Variables like "experience", "self-efficacy", "perceived risk", and "social influence" were added to the TAM model by [30]. An additional study made by [17] helped to extend cognitive absorption, playfulness, and self-efficacy. In 1996 two types of perceived usefulness namely near-term and

long-term were added [2]. In 2000, two new constructs were added to TAM namely perceived entertainment value and perceived presentation attractiveness, [11] was the scholar for this addition. Peer influence was then combined with TAM in 2002 by [3, 7].

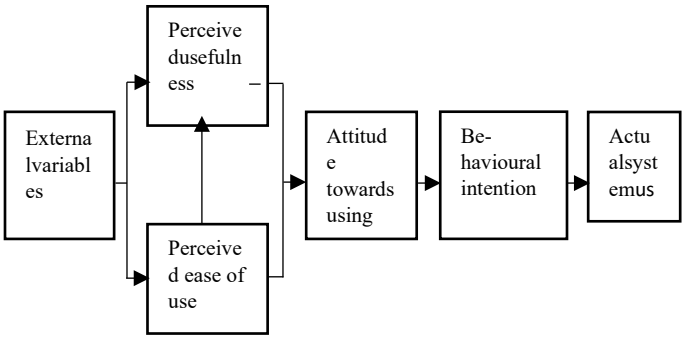


Figure 3: Technology Acceptance Model [30].

TAM has been used by scholars around the globe to perceive the acceptance of completely different sorts of information systems. A newly developed model by [6] based TAM named the shopping acceptance model (OSAM) was created to study online shopping behaviour. In 2003, in [31], the author developed an e-commerce model based on TAM with new variables trust and perceived risk.

5. Technology-Organization- Environment (T-O-E) Overview

As per [4], T-O-E Framework was developed by [6] for organisational adoption based on the Contingency Theory of Organisations. Well, it is claimed by [9] that [6] expect a nonexclusive arrangement of components to anticipate the probability of adoption. The framework proposes that an organisation should be consistent with its surroundings and environmental needs and its strength is determined by both internal and external factors like environment, organisation size, and organisation strategy [3]. Three key determinants were distinguished that influence organisational adoption: technology, organisation, and environment. It is imperative when one is making a decision, three factors of influence need to be looked into namely technology development [80] organisational conditions, business and organisational re-configuration [2], and industry environment [27].

Within the T-O-E framework, technological development presents the technologies accessible to an organisation. The organisation context outline the organisation characteristics well the environment context outlines the business field that consists of industry, competitors, regulations, and relationships with the government. As per [6], these are external factors that can have restraints and opportunities for technological innovations. The drawback of T-O-E is the assumption that the model will apply to large organisations, where customers make certain of congruity and fewer grievances, than to SMEs [9].

According to [6] there are three contexts that leverage technology innovation adoption and implementation process and these contexts of the T-O-E framework can be listed as follows.

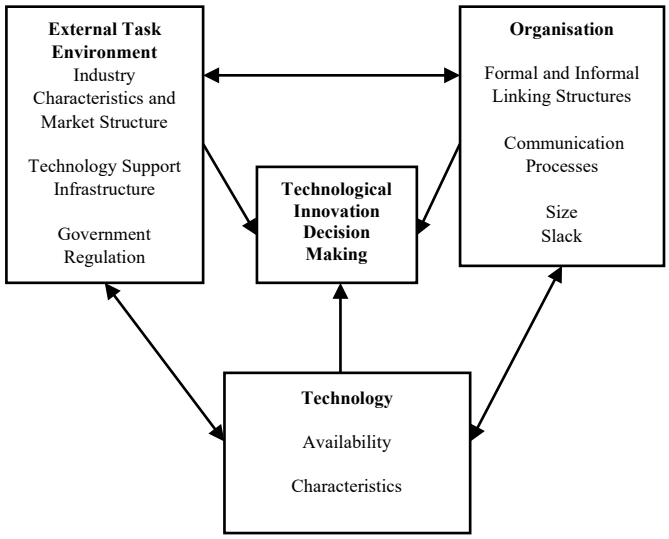


Figure 4: The technology–organization–environment framework [7].

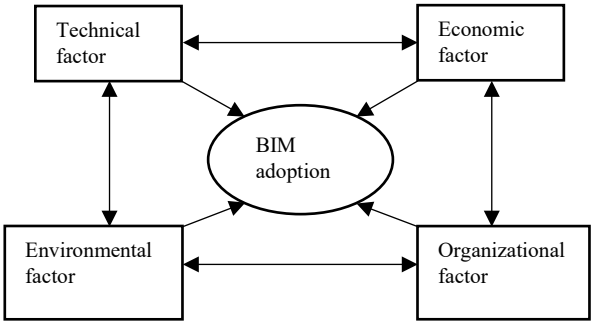


Figure 5: The derivative of T-O-E framework of BIM technology [30].

5.1. Technology context

In [2] and [16], the author defines that the technology context consisting of variables that has an impact on an individual, an organisation, and an industry's adoption of innovations and comprises of five innovation attributes as per [32] and [16] and also other attributes. Aside from innovation variables, other significant variables like system absorption, digestion, trail ability, intricacy, seen direct advantages, seen backhanded advantages and normalization has been included by several research studies while observability is found insignificant [1, 2, 9, 12, 18, 22, 32–34].

5.2. Organisation context

Adoption aptitude is impacted by explicit and spontaneous intra-hierarchical components for correspondence and supervision; along with resources and creativity of the organisation [32]. In [4], the author states that organisational context comprises of organisation scope, organisation size, and legislative belief. As per [1–3, 9, 12, 18, 22, 32, 33, 35] the most important variables of organisational context includes "financial resources, firm structure, organisational slack, innovation capacity, knowledge capability, operational capability, strategic use of technology, trust, technological resources, top management support, support for innovation, quality of human capital, organisational knowledge accumulation, expertise and infrastructure, and organisational

readiness while financial capacity and technology competence are identified as insignificant". The clarification explained by the makers is, relationship of all sizes have perceived the fundamental meaning of development for the achievement of their associations, consequently are glad to contribute enthusiastically to progressions to improve their high grounds. [12, 32, 34]; states that the role of top management is not the same and differs in context. As per [2] some inter-organisational variables context are on EDI that identifies the power of partners, an expectation in the partner and tie responsibility with a partner and dependency of the partner.

5.3. Environmental context

Environmental context is an area where the company targets business operations like government incentives and regulations [4]. The environmental context consists of variables like rivalry, relations with buyers and suppliers, and areas of the industry life cycle [2, 4, 6, 9,12, 18, 22, 24, 32-34] says the environmental context variables covers client order, serious weight, outer weight, inside weight, exchanging accomplice pressure, merchant uphold, business reliance, ecological vulnerability, data power and organisation force.

In summary, T-O-E has materialized to be a widespread framework used for theoretical perspective on IT adoption [24]. Consideration of technological, organisational, and environmental variables has made the T-O-E frame invaluable over other selection models in contemplating innovation appropriation, innovation use, and worth creation from innovation development [2, 8, 9, 32, 36].

6. Limitations and Criticism of T-O-E framework

As per [32] the T-O-E framework limitation as glossary of variables that is not well integrated or well developed and needs more research study on organisational adoption. [1] likewise featured that T-O-E system has no significant develops in the model and the factors in every unique circumstance [33] states that there lack in power of technology and adoption variance is unexplained. On the contrary [12] declares that the major constructs and the variables in the T-O-E framework are not concise and differ from context to context. For this reason as by [9,35] and [3, 12, 24, 32, 36] other variables like sociological variables, cognitive variables, technology readiness, ability to leverage IT investment through different channels professionals' experience and skills, managerial capabilities of change management, security concerns, government promotion and factors salient to the country context such as government policy/regulation, technology infrastructure, and culture is needed to refine the T-O-E framework.

7. Technology-Organisation-Environment (T-O-E) progression

According to [37], the T-O-E framework lack change for different reasons from original development. This lack of evolution can be described as follows. Firstly, as per [24], T-O-E characterizes as a universal theory as T-O-E is used as a framework in which a large group of different elements can be set. For this reason, researchers and scholars saw little need to change or refine the T-O-E framework. Second, the T-O-E structure may have seen generally little advancement because it has been seen as lined up with other adoption model theories. For this reason, pressure has been seen by the T-O-E framework and other theories for the T-O-E framework to assimilate in contesting innovations. As per [38], it can be seen that the T-O-E framework has been persistent with the theory of the diffusion on innovation (DOI). For example, DOI adoption on individual leader characteristics and internal characteristics of organisational structure is equally yok to T-O-E's organisation context element. Another consistency can be seen with DOI's external characteristics of the organisation with T-O-E's environmental context. Lastly, consistency can be seen in the technological characteristics of the innovation with T-O-E's technological context [39, 40]. Since these theories are portrayed as notably comparable, the T-O-E system has not been adjusted because of DOI [37].

As per [9] several research studies explains the essence of the T-O-E framework which can be explained as e-commerce Enterprise Resource Planning, e-business, , open systems, Knowledge Management Systems, Electronic Data Interchange etc. in Table 1.

Table 1: Organisational Adoption of Information Technologies

Author(s)	Domain
[22]	“Internet/E-Business”
[21]	“HRIS”
[2]	“Green IT Initialization”
[67]	“E-Business”
[3]	“KMS”
[32]	“Enterprise Systems”
[37]	“Mobile Commerce”
[18]	“E-Business”
[16]	“ERP”
[2]	“Web Site”
[6]	“E-Commerce”
[41]	“IT”
[32]	“E-Signature”
[3]	“E-Commerce”
[37]	“E-Business”
[2,25]	“E-Business”
[6]	“Internet”
[6]	“E-Business”
[32]	“E-Commerce”
[27]	“IT”
[2]	“Open System”

Over the past 30 years, the T-O-E framework has exhibit influence across many technological, industrial, and national/cultural contexts [7]. The T-O-E framework has been used across various industries to explain the adoption of inter-organisational systems [2, 42]. E-business [2, 6, 7, 15], electronic data interchange (EDI) [42], open systems[2], enterprise systems [32] and a broad spectrum of general IS applications [2]. The T-O-E model has been used to clarify the selection of advancements in a large group of ventures, including manufacturing [2, 15] health care [34] retail, wholesale, and financial services [15]. According to [2, 7, 15] and [6] the T-O-E model has been approved in European, American, and Asian contexts, as well as in both

developed as well as developing countries. In every one of the experimental investigations that test the T-O-E system, analysts have utilized somewhat various elements for technological, organisational, and environmental contexts. [7] claims that researchers agree with [6] that the three T-O-E contexts affect adoption, however, they have speculated that for every particular innovation or setting that is being contemplated, there is a special arrangement of components or measures. [6] argues that one relevant factor in technological context that influences the adoption of e-business is " technology readiness." [6] also states that "firm size", global scope" and "financial resources are the appropriate elements that ought to be concentrated to see how organisational context influences the adoption of e-business. When looking at environmental context influences, "regulatory environment" and "competition intensity" are suitable in the adoption of e-business.

8. Integration of TAM and T-O-E framework

As per [30] the creation of TAM by [2] was in light of the Theory of Reasoned Action (TRA) in 1989. [2] perceived that a person's behavioural intention to use a newly created system is associated with their perceived usefulness and ease of use. Meanwhile, the perceived ease of use influences perceived usefulness, and the perceived usefulness and ease of use are influenced by foreign variables.

As per [43] the T-O-E framework persuades models of technological, organisational, and environmental backgrounds to adopt and achieve technological innovations. he goes on to say the technological context indicates internal and external technologies by firms, organisational context refers to company size, organisational structure, and human resource, well environmental context looks at components like competition, partners, and industry environment that out-side the control of a firm [30].

The Technology Acceptance Model and Technology Organisation Environment System have been generally applied in I.T. and data innovation across various research studies [30]. The two models complement each other as TAM is flexible with external variables and can abduct a person's acceptance behaviours while on the other side of the coin T-O-E acknowledges the technical, environmental and organisational elements that have an effect on technology acceptance and adoption at organizational level. Many research studies where TAM and T-O-E are integrated can be found. [24] has integrated TAM and T-O-E to illustrate cloud adoption at organisational level. Another study by [34] evaluates system factors in the adoption of ERP in manufacturing companies. Lastly, [28] set up the UTAUT-T-O-E technology acceptance model framework and investigated the components influencing IT selection [30]. Figure 6 illustrates the integration of TOE External variables to TAM.

Joining the TAM in Figure 3 with the TOE structure in Figure 5, an incorporated TAM-TOE model for BIM innovation was created as seen in Figure 7.

The impact scale of each factor on different variables are determined through their logical relationship in the system to shape the immediate impact network, at that point the reason and focus level of each factor is determined to decide the circumstances and end-result between the elements. This strategy can visually demonstrate the logical relationships between factors through determinant analytics and can disentangle complex issues.

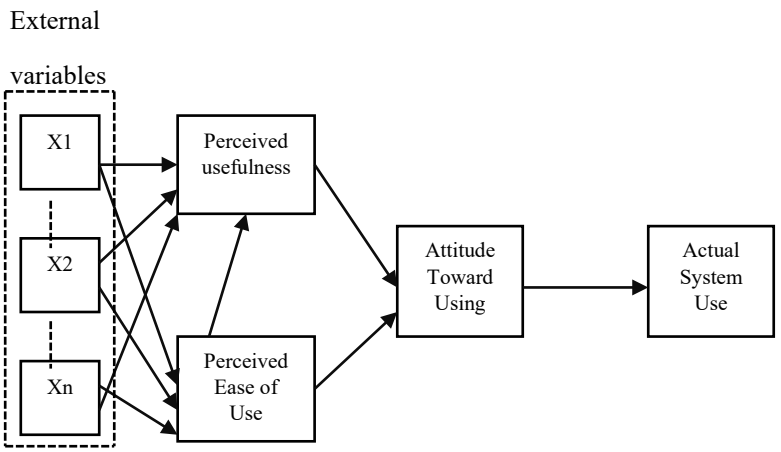


Figure 6: TAM-TOE derivative integration model [30].

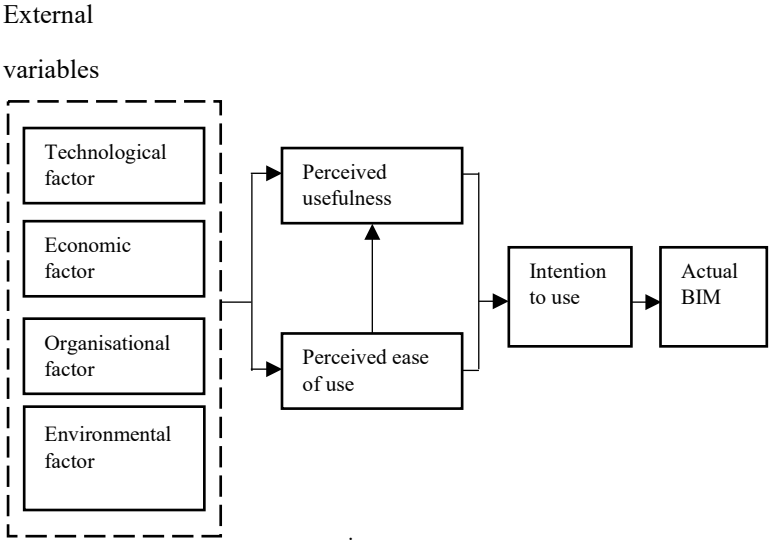


Figure 7: The derivative integrated TAM-T-O-E model of BIM technology [30].

8.1. Technical factors

According to [30] technical factors can be classified into three past research studies, namely localization, standardization, and compatibility.

- Localization: Localization as per [13], Building Information Modelling (BIM) consists of software operating environment, the adaption of ventures, and management process.
- Standardization: Specifies milestones, deliverables, and objectives to accomplish.
- Compatibility: This describes the strength and challenges of different application software from a different corporation, their integration, and are compatibility with each other.

8.2. Economic factors

According to [30] there two economic factors: cost and return in investment.

- Cost: Cost normally consist of hardware, software, training and consulting fee for professionals [59].
- Return on investment: Refers is a new technology that would have the potential and opportunities for new revenue income.

Table 2: A summary of the literature on external variables in the Building Information Modelling (BIM) TAM-T-O-E model.

Category	Code	External Variables	References																					sum
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
Technological Context	EV1	Localization of BIM				●										●	●	●	●	●			●	7
	EV2	Standardization of BIM	●	●	●		●	●	●	●	●	●	●	●	●	●	●		●	●	●	●	●	19
	EV3	Compatibility of BIM	●		●	●	●		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	19
Economic Context	EV4	Cost of using BIM	●	●	●		●	●	●		●	●	●			●	●		●	●	●	●		15
	EV5	Return on investment		●				●	●	●		●							●				●	7
Organisational Context	EV6	Organisational pattern and workflow	●	●	●		●	●		●	●	●		●	●	●	●		●	●			●	15
	EV7	Traditional thinking mode	●	●			●	●	●	●	●	●		●		●			●	●			●	13
	EV8	Executive support	●		●	●	●	●		●		●	●	●			●	●	●	●		●	●	15
	EV9	Number of BIM experts and technical staff	●	●	●		●		●			●	●			●	●		●	●	●		●	13
Environmental Context	EV10	Requirement from in the industry		●	●			●	●						●		●	●		●	●	●		10
	EV11	Popularity of BIM in the industry		●	●	●			●	●		●		●	●			●		●	●	●		12
	EV12	Competitions from other companies			●	●			●					●	●		●	●				●		8

Notes. A = [44] B = [19] C = [77]; D = [45]; E = [6]; F = [18] G = [26]; H = [46]; I = [24] J = [23]; K = [34]; L = [34]; M = [47]; N = [72]; O = [59]; P = [23]; Q = [24]; R = [48]; S = [8]; T = [26]; U = [22].

8.3. Organisational factors

As per [30], the four organisational factors can be listed as organisational mode and workflow, traditional thinking mode, support from senior management, and several experts and technicians.

- Organisational model and Workflow: It is normally fixed and is difficult to change the organisational model and workflow. Roles and responsibilities and work context frustrate application extension to a degree [30].
- Traditional Thinking mode: People are hesitant to change and handle issues in the same way [30], [24].
- Top Management Support: top management duty is one of the significant achievement factors for receiving BIM innovations [33].
- A number of BIM experts and technicians: The requirement of professionals with related technical experience and insights.

8.4. Environmental factors.

As per [30], environmental factors are shown at national, industrial, and enterprise levels for perceptions and attitudes in adoption and have been classified as national policy requirements, popularity in the BIM industry, and competition from other companies.

- National policy requirements: As per [15], national policies is crucial to the development of information technology. For instance, the British government has assumed a significant part in advancing the application and the improvement of BIM [18].
- Popularity in the BIM industry: With greater acceptance, BIM adoption is becoming more popular.
- Competition from other companies: Completion can be clarified in two ways of competition namely, cost and differentiation.

The study merges TAM and TOE and proposes a Building Information Modelling (BIM) TAM-TOE model that consists of 12 external variables. An informative questionnaire was derived from the 12 external variables and 3 internal variables. A span number DEMATEL strategy was utilized to figure the impact degree, affected degree, centrality degree, and causal level of every factor. By ascertaining the centrality and reason for each factor, the key components influencing the selection of BIM are acquired, an 'impact and cause chart' was made dependent on the absolute relations between factors.

The study significant discoveries can be summed up as:

Firstly, the Requirement from public arrangements (EV10) is the most grounded driving variable, followed by Traditional thinking modes (EV7), Standardization of BIM (EV2), Compatibility of BIM (EV3), Competitions from other companies (EV12), Popularity of BIM in the industry (EV11), Localization of BIM and Return on investment (EV5).

Secondly, the re-enactments and emphases of the eight propulsive variables affirm that the objective variable Intention to use (IV3) fluctuated essentially and is responsive with changes of firstly, The Requirement from national policies (EV10), and then the Standardization of BIM (EV2).

9. Conclusion

TAM is a powerful adoption theory in the IT sector and has been used by numerous research studies. The model looks at an individual's goal when utilizing a system or application and recommend PU and PEOU as the central reason for IT adoption. Over the years TAM has shown great acceptance, utilization, and reproduction but the model failed to give essential data on the user's assumption to only PU and PEOU. Based on this limitation it was necessary for the model by expanding or integrating with other IT acceptance models [9]. The T-O-E framework proposed a generic set of factors of technology adoption. The T-O-E framework looks at three contexts namely Technology, Organisation, and Environmental. TAM and its all-encompassing forms have a high ability to clarify the innovation reception while the meaning of the T-O-E system is likewise perceived in clarifying technology adoption. The marriage of these two models brings a new and unique developed redesign that takes TAM and T-O-E models to a more extensive level to advance and encourage improved informative and prescient focal points of IT adoption.

The TAM and T-O-E framework can be used separately or as a hybrid depending on the situation at hand. In future it important to harmonize the so many factors of the models that have been suggested and used in literature.

This review paper doesn't overrule the acknowledgment of other reception models, yet it investigates the writing to create a selection model for study comparable data advances as referenced before.

Acknowledgment

This work is supported by the Vaal University of Technology under the supervision of Professor Tranos Zuva in Johannesburg (Vanderbijlpark), South Africa.

References

[1] C. Low, Y. Chen, M. Wu, "Understanding the determinants of cloud computing adoption," Industrial Management and Data Systems, **111**(7), 1006–1023, 2011, doi:10.1108/02635571111161262.

[2] K. Zhu, K. Kraemer, S. Xu, "Electronic business adoption by European firms: A cross-country assessment of the facilitators and inhibitors," European Journal of Information Systems, **12**(4), 251–268, 2003, doi:10.1057/palgrave.ejis.3000475.

[3] T.S.H. Teo, S. Lin, K. hung Lai, "Adopters and non-adopters of e-procurement in Singapore: An empirical study," Omega, **37**(5), 972–987, 2009, doi:10.1016/j.omega.2008.11.001.

[4] M.I. Salwani, G. Marthandan, M.D. Norzaidi, S.C. Chong, "E-commerce usage and business performance in the Malaysian tourism sector: Empirical analysis," Information Management and Computer Security, **17**(2), 66–185, 2009, doi:10.1108/09685220910964027.

[5] V. Venkatesh, F.D. Davis, M.G. Morris, "Dead or alive? The development, trajectory and future of technology adoption research," Journal of the Association for Information Systems, **8**(4), 267–286, 2007, doi:10.17705/1jais.00120.

[6] S. Xu, K. Zhu, J. Gibbs, "Global Technology, Local Adoption: A Cross-Country Investigation of Internet Adoption by Companies in the United States and China," Electronic Markets, **14**(1), 13–24, 2004, doi:10.1080/1019678042000175261.

[7] K. Zhu, K.L. Kraemer, V. Gurbaxani, S.X. Xu, Migration to open-standard interorganizational systems: Network effects, switching costs, and path dependency, MIS Quarterly: Management Information Systems, 2006, doi:10.2307/25148771.

[8] J. Xu, "Research on Application of BIM 5D Technology in Central Grand Project," in Procedia Engineering, 174 600-610, 2017, doi:10.1016/j.proeng.2017.01.194.

[9] H. Okorie Awa, O. Ukoha, B. Emecheta, "Integrating TAM and TOE Frameworks and Expanding their Characteristic Constructs for E-Commerce Adoption by SMEs," in Proceedings of the 2012 InSITE Conference, 571–588, 2012, doi:10.28945/1676.

[10] C. Carnaghan, K. Klassen, "Exploring the determinants of web-based E-business evolution in Canada," in Association for Information Systems - 13th Americas Conference on Information Systems, AMCIS 2007: Reaching New Heights, 5 3532-3543, 2007.

[11] S. Alwahaishi, V. Snášel, "Modeling the Determinants Affecting Consumers' Acceptance and Use of Information and Communications Technology," International Journal of E-Adoption, **5**(2), 25–39, 2013, doi:10.4018/jea.2013040103.

[12] Y.M. Wang, Y.S. Wang, Y.F. Yang, "Understanding the determinants of RFID adoption in the manufacturing industry," Technological Forecasting and Social Change, **77**(5), 803–815, 2010, doi:10.1016/j.techfore.2010.03.006.

[13] Y.H. Li, "An empirical investigation on the determinants of E-procurement adoption in Chinese manufacturing enterprises," in 2008 International Conference on Management Science and Engineering 15th Annual Conference Proceedings, ICMSE, 32–37, 2008, doi:10.1109/ICMSE.2008.4668890.

[14] Y. a Au, H. Zafar, A Multi-Country Assessment of Mobile Payment Adoption, 2008, doi:10.1111/1467-8616.00112.

[15] N. Schillewaert, M.J. Ahearne, R.T. Frambach, R.K. Moenaert, "The adoption of information technology in the sales force," Industrial Marketing Management, **34**(2), 323–336, 2005, doi:10.1016/j.indmarman.2004.09.013.

[16] M.J. Pan, W.Y. Jang, "Determinants of the adoption of enterprise resource planning within the technology-organization-environment framework: Taiwan's communications industry," Journal of Computer Information Systems, **48**(3), 94–102, 2008, doi:10.1080/08874417.2008.11646025.

[17] K. Amoako-Gyampah, A.F. Salam, "An extension of the technology acceptance model in an ERP implementation environment," Information and Management, **41**(6), 731–745, 2004, doi:10.1016/j.im.2003.08.010.

[18] W.W. Wu, "Developing an explorative model for SaaS adoption," Expert Systems with Applications, **38**(12), 15057–15064, 2011, doi:10.1016/j.eswa.2011.05.039.
[19] V. Venkatesh, M.G. Morris, G.B. Davis, F.D. Davis, "User acceptance of information technology: Toward a unified view," MIS Quarterly: Management Information Systems, **27**(3), 425–478, 2003, doi:10.2307/30036540.
[20] O. Ukoha, H.O. Awa, C.A. Nwuche, I.F. Asiegbu, "Analysis of explanatory and predictive architectures and the relevance in explaining the adoption of IT in SMEs," Interdisciplinary Journal of Information, Knowledge, and Management, (6), 217–230, 2011, doi:10.28945/1431.
[21] I. Troshani, C. Jerram, H.S. Rao, "Industrial Management & Data Systems Exploring the public sector adoption of HRIS," Industrial Management & Data Systems Personnel Review International Journal of Productivity and Performance Management Iss Kirstie S. Ball Personnel Review, **111**(3), 470–488, 2011.
[22] S.Y. Yousafzai, G.R. Foxall, J.G. Pallister, Technology acceptance: a meta-analysis of the TAM: Part 1, Journal of Modelling in Management, 2007, doi:10.1108/17465660710834453.
[23] H.D. Yang, Y. Yoo, "It's all about attitude: Revisiting the technology acceptance model," Decision Support Systems, **38**(1), 19–31, 2004, doi:10.1016/S0167-9236(03)00062-9.
[24] K. Zhu, K.L. Kraemer, S. Xu, "The Process of E-Business Assimilation in Organizations : A Technology Diffusion Perspective," Journal of the Institute for Operations Research and Management Sciences, 2005.
[25] L. Raymond, F. Bergeron, S. Blili, "The Assimilation of E-business in Manufacturing SMEs: Determinants and Effects on Growth and Internationalization," Electronic Markets, 2005, doi:10.1080/10196780500083761.
[26] M.D. Williams, Y.K. Dwivedi, B. Lal, A. Schwarz, "Contemporary trends and issues in IT adoption and diffusion research," Journal of Information Technology, **24**(1), 1–10, 2009, doi:10.1057/jit.2008.30.
[27] I. Arpaci, O. Turetken, Y.C. Yardimci, S. Ozkan, "Organizational Adoption of Information Technologies: A Literature Review Telework View project Sentiment Analysis View project ORGANIZATIONAL ADOPTION OF INFORMATION TECHNOLOGIES: A LITERATURE REVIEW," INTERNATIONAL JOURNAL OF EBUSINESS AND EGOVERNMENT STUDIES, **4**(2), 731–745, 2012.
[28] L. Silva, Post-positivist review of Technology Acceptance Model, Journal of the Association for Information Systems, 2007, doi:10.17705/1jais.00121.
[29] J.W. Moon, Y.G. Kim, "Extending the TAM for a World-Wide-Web context," Information and Management, **38**(4), 217–230, 2001, doi:10.1016/S0378-7206(00)00061-6.
[30] X. Qin, Y. Shi, K. Lyu, Y. Mo, "USING A TAM-TOE MODEL TO EXPLORE FACTORS OF BUILDING INFORMATION MODELLING (BIM) ADOPTION IN THE CONSTRUCTION INDUSTRY," JOURNAL OF CIVIL ENGINEERING AND MANAGEMENT, **26**(3), 463–468, 2020, doi:10.3846/jcem.2020.12176.
[31] P.A. Pavlou, "Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model," International Journal of Electronic Commerce, **7**(3), 101–134, 2003, doi:10.1080/10864415.2003.11044275.
[32] B. Ramdani, P. Kawalek, O. Lorenzo, "Predicting SMEs' adoption of enterprise systems," Journal of Enterprise Information Management, (22), 10–24, 2009, doi:10.1108/17410390910922796.
[33] M.S. Musawa, E. Wahab, "The adoption of electronic data interchange (EDI) technology by Nigerian SMEs: A conceptual framework," Journal of Business Management and Economics, 2012.
[34] H. Xu, J. Feng, S. Li, "Users-orientated evaluation of building information model in the Chinese construction industry," Automation in Construction, (39), 32-46, 2014, doi:10.1016/j.autcon.2013.12.004.
[35] K.W. Wen, Y. Chen, "E-business value creation in Small and Medium Enterprises: a US study using the TOE framework," International Journal of Electronic Business, **8**(1), 80, 2010, doi:10.1504/ijeb.2010.030717.
[36] T. Oliveira, M.F. Martins, U.N. De Lisboa, "Literature Review of Information Technology Adoption Models at Firm Level," Review of Economics Studies, 2011.
[37] L. Da Chen, J. Tan, "Technology adaptation in E-commerce: Key determinants of virtual stores acceptance," European Management Journal, **22**(1), 74–86, 2004, doi:10.1016/j.emj.2003.11.014.
[38] S. Priyanka, "Technology Acceptance Model: A Survey of Literature," International Journal of Business and Social Research (IJBSR), 2015.
[39] A.Y.L. Chong, B. Lin, K.B. Ooi, M. Raman, "Factors affecting the adoption level of c-commerce: An empirical study," Journal of Computer Information Systems, **50**(2), 13–22, 2009, doi:10.1080/08874417.2009.11645380.
[40] W. Hong, K. Zhu, "Migrating to internet-based e-commerce: Factors affecting e-commerce adoption and migration at the firm level," Information and Management, **43**(3), 204–221, 2006, doi:10.1016/j.im.2005.06.003.
[41] C. Zhang, L. Cui, L. Huang, C. Zhang, "Exploring the role of government in information technology diffusion: An empirical study of IT usage in Shanghai firms," in IFIP International Federation for Information Processing, 393–407, 2007, doi:10.1007/978-0-387-72804-9_26.
[42] A.N. Mishra, P. Konana, A. Barua, "Antecedents and consequences of Internet use in procurement: An empirical investigation of U.S. manufacturing firms," Information Systems Research, **18**(1), 103–120, 2007, doi:10.1287/isre.1070.0115.
[43] N. Rao Kowtha, T. Whai Ip Choon, "Determinants of website development: A study of electronic commerce in Singapore," Information and Management, **39**(3), 227–242, 2001, doi:10.1016/S0378-7206(01)00092-1.
[44] D. Mehran, "Exploring the Adoption of BIM in the UAE Construction Industry for AEC Firms," in Procedia Engineering, 2016, doi:10.1016/j.proeng.2016.04.144.
[45] L. Zhou, L. Dai, D. Zhang, "ONLINE SHOPPING ACCEPTANCE MODEL-A CRITICAL SURVEY OF CONSUMER FACTORS IN ONLINE SHOPPING（自己构造的模型，列表写出了已存在的文献中用于衡量online shopping acceptance的factors，分为两方面：online和shopping）," Journal of Electronic Commerce Research, 2007.
[46] Z. Zakaria, N. Mohamed Ali, A.T. Haron, A.M. Ponting, Z. Abd. Hamid, "Exploring the barriers and driving factors inimplementing Building Information Modeling (BIM) in the Malaysian Construction Industry: A Preliminary Study," The Journal of The Institution of Engineers, Malaysia, **58**(12), 7250–7257, 2014.
[47] J. Won, G. Lee, C. Dossick, J. Messner, "Where to Focus for Successful Adoption of Building Information Modeling within Organization," Journal of Construction Engineering and Management, **139**(11), 04013014, 2013, doi:10.1061/(asce)co.1943-7862.0000731.
[48] R. Bose, X. Luo, "Integrative framework for assessing firms' potential to undertake Green IT initiatives via virtualization - A theoretical perspective," Journal of Strategic Information Systems, **110**(2), 269–283, 2011, doi:10.1016/j.jsis.2011.01.003.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 139-153 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Visual Saliency Detection using Seam and Color Cues

Sk. Md. Masudul Ahsan[1], Aminul Islam[*,2]

[1]*Department of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh*

[2]*Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 14 February, 2021
Online: 10 March, 2021

Keywords:
Salient Object Detection
Seam Importance Map
Color Importance Map
Saliency Optimization
Background Extraction
Visual Attention
Computer Vision

ABSTRACT

Human have the god gifted ability to focus on the essential part of a visual scenery irrespective of its background. This important area is called the salient region of an image. Computationally achieving this natural human quality is an attractive goal of today's scientific world. Saliency detection is the technique of finding the salient region of a digital image. The color contrast between the foreground and background present in an image is usually used to extract this region. Seam Map is computed from the cumulative sum of energy values of an image. The proposed method uses seam importance map along with the weighted average of various color channels of Lab color space namely boundary aware color map to extract the saliency map. These two maps are combined and further optimized to get the final saliency output using the optimization technique proposed in a previous study. Some intermediate combinations which are closer to the proposed optimized version but differ in the optimization technique are also presented in this paper. Several standard benchmark datasets including the famous MSRA 10k dataset are used to evaluate performance of the suggested procedure. Precision-recall curve and F-beta values found from the experiments on those datasets and comparison with other state of the art techniques prove the superiority of the proposed method.

1 Introduction

Computer vision techniques are usually complex and time consuming, since every pixel of the input image needs to be processed through them. Saliency detection is the technique of finding the salient or important region in an image. So, researches in this field have become popular to reduce this type of computational complexities and shrink the search space of an input image in various vision related applications including object detection, action recognition, image segmentation, video surveillance, medical data processing, robotics and many more. Color contrast present between the salient region and the background is the most popular cue for saliency detection. Another prior that produces good result is the boundary prior which assumes that the boundary regions are mostly backgrounds. These are two popular and successful cues in saliency detection solutions. But, those cues are challenged when the salient regions have less color dissimilarities with the background, they are closer to the border or touch the boundary regions. In those complex conditions, relying only on contrast and boundary priors may not produce good results. So, a different method which does not exclusively rely on contrast based dissimilarities is presented in this study. Experimental results with popular standard datasets and comparison with other state of the art techniques proves the superiority of the proposed system. Evaluation is done based on PR curves and F-Beta measures using publicly available state of the art benchmark results of different studies.

2 Related Works

Since there is usually a visible contrast between the salient region and its background, many researchers mostly relied on contrast prior. Moreover, boundary prior offered promising results in most of the cases.

Early saliency based researches emphasized on center prior which assumes that the salient regions are usually at the center of an image. Those methods usually biased the center of the image with higher saliency values. It was not always true and there are many examples where the saliency is not at the center of an image. So, it was a fragile assumption and nowadays researchers are not much interested in that cue.

On the other hand, boundary prior is an opposite assumption to center prior. Unlike its counterpart, it works better for objects which are not strictly at the center of an image. Although this assumption has some drawbacks, it resulted in better detection in many researches. This prior does not consider the condition where an

[*]Corresponding Author: Aminul Islam, CSE Discipline, Khulna University, Khulna-9208, Bangladesh, Contact No: +880 1911013641 & Email: aminul@ku.ac.bd

object contains some of the boundary parts. In [1] the shortest-path distance from an image region to its boundary is used as a cue for saliency detection. It shows that the boundary pixels can easily be connected to the image background instead of the foreground. As a result, it results in incorrect detection if an object is touching the boundary region.

In [2],the authors have taken into consideration the problems of boundary prior and introduced weighted boundary connectivity to successfully solve them and presented a better boundary prior based implementation. They also have proposed a robust optimization technique to combine multiple saliency cues to achieve cleaner saliency maps.

In [3], the authors have used contrast based spatial features along with boundary prior to generate their state of the art saliency detection result. They measured region based color dissimilarity of a superpixelized image found by applying the SLIC algorithm on an input image and suppressed the background using boundary prior as a cue to it. They also have combined the contrast based salient region with boundary based output and fine-tuned the combined saliency map using Gestalt thresholding technique to generate the final salient region.

In [4], the authors have presented a color map based saliency detection technique. It is the weighted average of various color channels. They combined it with border map to acquire the final saliency output. This color map is used successfully in [5] to calculate salient region in an image.

Seam is an optimal 8-connected path of pixels in an image. It is introduced by Avidan et al. in [6] for content aware image resizing and suggested for image enhancement too. Yijun et al. has used seam information in their saliency detection technique [7] with success.

Many methods including [8]–[12] have combined multiple saliency maps to produce better optimized versions successfully. Some of them became complex and their implementation and performance in real world situations are still to be tested. In [13], saliency detection problem has been solved using graph based manifold ranking. In [14], a color and spatial contrast based saliency model is presented which uses four corners of an image to detect the image background. That method was inspired by reverse-management methods and it also applied energy function to further optimize foreground and background.

In [15], a regional contrast based saliency detection method is presented. They show that there are local and global contrast based saliency detection methods. Local contrast based methods commonly produce higher saliency values near object edges and fail to highlight the entire salient region. On the other hand, they criticized some global contrast based methods for being insufficient to analyse common variations in natural images. Inspired from biological vision they propose a histogram-based contrast method and a region based contrast method including spatial information for saliency detection.

In [16], the authors have proposed a saliency detection approach based on maximum symmetric surrounds. They used the low level features of luminance and color. They narrow down the search area for symmetric surrounds as they approach the borders. The disadvantage of their method is that if the salient region touches the border or is cut by the border, it cannot detect it and is treated as the background.

So, there are biologically inspired methods and computationally optimized methods, local and global contrast based methods. There are also combinations of biological and computational models. In [17], a biology based method is used in [18] to extract the feature map and they used graph based approach to normalize the output. Some of the methods computed the entire feature map or some combined multiple maps to produce the final saliency model.

According to the previous discussion, it is obvious that the main challenges of saliency detection problem are low color dissimilarity of objects with its surrounding regions, size and position of a salient region in an image, computational complexity and usability of the method in real life applications. In this study, a combination of seam and color map based salient region detection technique is presented. The proposed method is easy to implement and does better than the other compared methods in response to the mentioned challenges. Details of the proposed method are presented in the following section.

3 Proposed Method

In this study, a novel approach is presented that combines seam and color map based saliency detection procedures in addition to traditional contrast and boundary prior based methods. Moreover, introduction of the quadratic equation based optimization method has improved the result significantly.

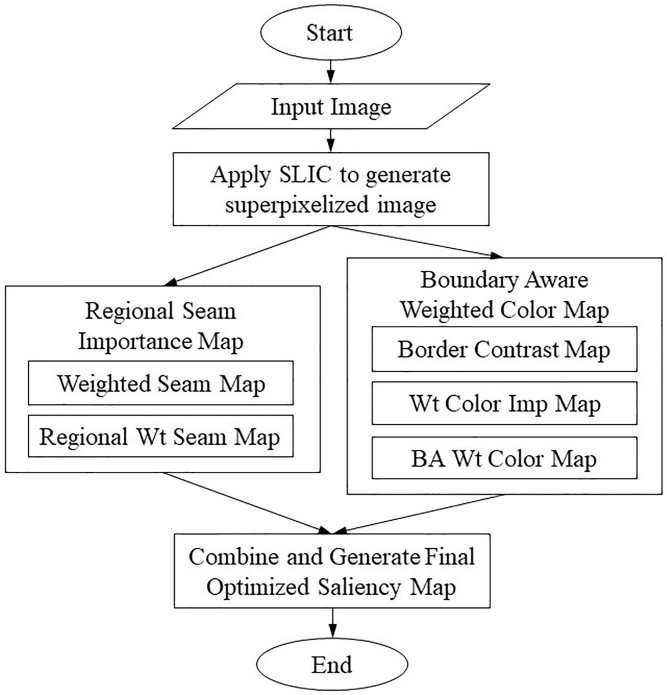
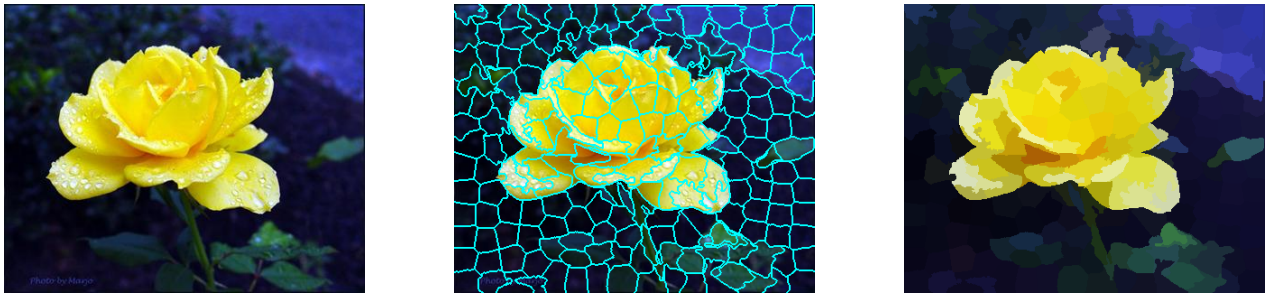


Figure 1: Flow chart of the proposed method.

Structure of the proposed method is shown in the flowchart of Figure 1. The raw input image is transformed into superpixel image. Seam and color importance maps are generated from this superpixelized image. These two saliency maps are then combined and optimized using a suitable optimization procedure to generate the



(a) Input Image (b) Suoerpixel Boundaries (c) Superpixelized Image

Figure 2: The SLIC approach to transform the input image into superpixels.

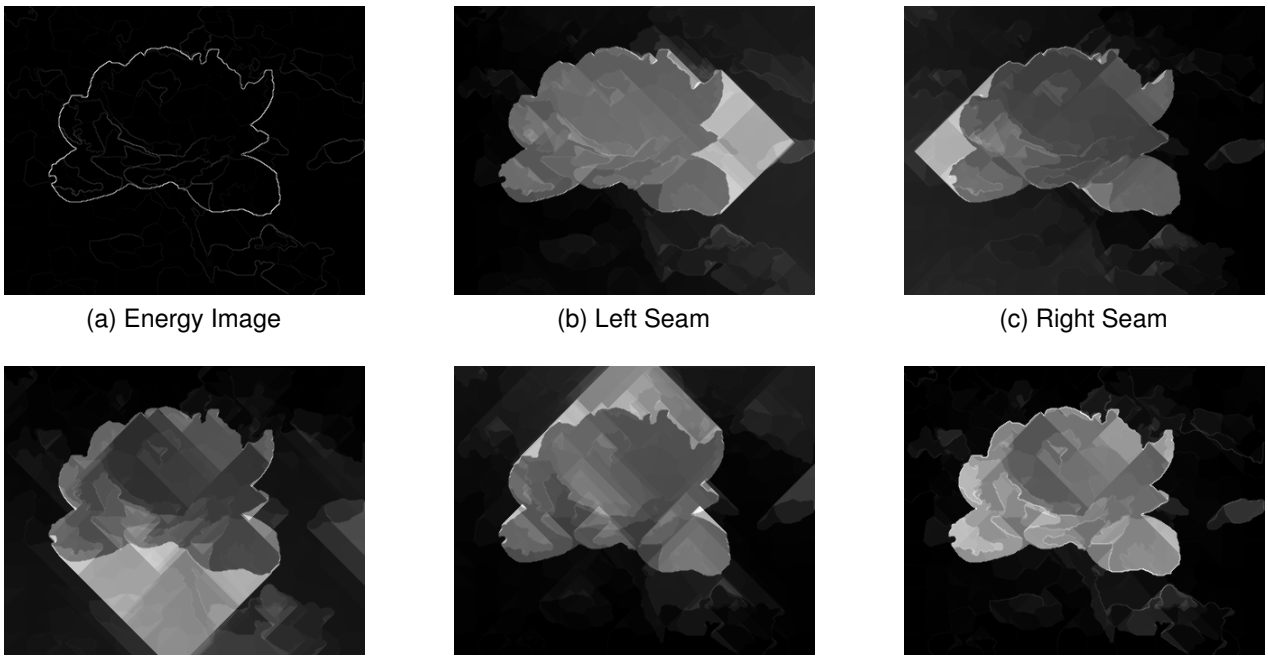
final saliency map. There are more internal steps seen in the flow chart. In this section the components of this flowchart are discussed in details.

3.1 Superpixel Image

At first, the input image is converted into superpixels using the SLIC method [19]. It groups similar pixels together using some clustering methods around randomly chosen pixel centers. Each region contains the averaged pixel intensity value in that region. It is done to speed up the calculation procedure. Since, working in pixel level will require more time to complete the calculation, it assigns a region ID to each pixel and computation can be continued using those regions only. By doing that the calculation becomes faster and quality is not much considered. It can be converted into the full size image anytime. Our main goal is to detect the salient region only. So, pixel level perfection is not necessary. Figure 2 depicts how an input image is transformed into its superpixel form. Figure 2(b) shows the cluster boundaries which are generated from Figure 2(a). The final superpixelized image will look like the region base image of Figure 2(c).

3.2 Seam Importance Map

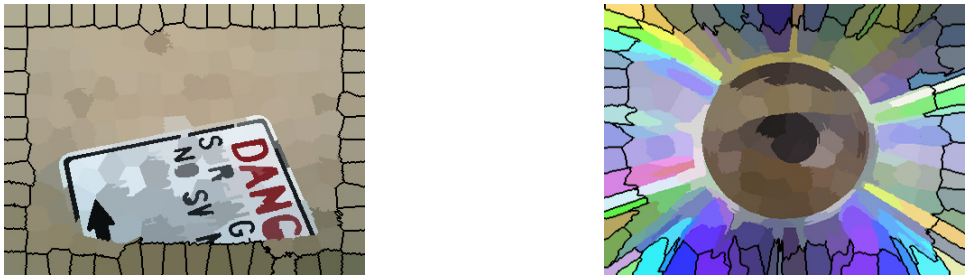
Seam map is computed from the energy map of a given image. It is the cumulative sum of minimum energy values from all four directions of an image namely top to bottom, left to right, bottom to top and right to left. Unlike [6] which computed a seam for the whole image, we follow [7] to get the seam for each pixel of the superpixelized image. The energy function is defined by Sobel operator to get the gradient image using eq. 1 which minimizes the cost. Here, the seam is an optimal 8-connected path of pixels. The superpixelized input image taken as the input to the seam map generation process following [5]. It produces better results as compared



(a) Energy Image (b) Left Seam (c) Right Seam

(d) Top Seam (e) Bottom Seam (f) Combined Seam

Figure 3: Various stages of seam map calculation method.



(a) (b)

Figure 4: Example of simple and complex border regions.

to [8] which uses the raw input image in the seam map generation procedure. The reason for this enhancement is the smoothness and stronger edges of the superpixelized image. The computation process is as follows: to compute the seam in the direction of top to bottom, each pixel follows eq. 2. The minimum cumulative sum of energy values of its previous row's 8-connected neighbours are summed up with the current pixel's energy value according to eq. 2 to calculate the current pixel's seam value. Other seam maps are generated in a similar fashion. Since, the background pixels take lower values in any of the four maps as seen in Figure 3, the minimum seam value at each pixel is taken in eq. 3 which suppresses the background well. This way the combined map is found by selecting the minimum of four values for each pixel by eq. 3. The seam map generation process of this study is presented in Figure 3 which shows that Figure 3(b), 3(c), 3(d) and 3(e) are computed from the energy image 3(a); and 3(f) is the combination of seams from Figs. 3(b), 3(c), 3(d) and 3(e). It removes the disadvantage of boundary prior and successfully enumerates the salient object even if it cuts the boundary of an image. The region level seam map is found by averaging seam values in each region as seen in eq. 4. The final seam map is generated by further down-weighting distant regions using average spatial distance as in eq. 5. Pictorial representation of the seam map generation process is visually presented in Figure 3. In this study this seam importance map is combined and further optimized intelligently to produce the final saliency model which does better in conditions where the salient region cuts the boundary.

$$\text{Energy Image}, e(I) = \left|\frac{\delta}{\delta x}I\right| + \left|\frac{\delta}{\delta y}I\right| \tag{1}$$

$$M_T(x,y) = \begin{cases} e(x,y); & \text{if y = 0} \\ e(x,y) + min[M_T(x-1,y-1), & \\ M_T(x-1,y), M_T(x-1,y+1)]; & \text{otherwise} \end{cases} \tag{2}$$

$$Imp_{seam}(x,y) = min[M_L(x,y), M_T(x,y), M_R(x,y), M_B(x,y)] \tag{3}$$

$$Imp_{seam}(r_i) = \frac{1}{|r_i|} \sum_{\forall I_{cst}(x,y)=i} Imp_{seam}(x,y), where$$
$$|r_i| = \text{Number of pixels in region } r_i \tag{4}$$

$$S_{seam,s}(r_i) = \frac{Imp_{seam}(r_i)}{d_s^2(r_i)}$$
$$d_s = \left\|p_i - p_j\right\|, \text{Euclidian region distance where} \tag{5}$$
$$p_i, p_j \text{ is mean position vector for } r_i, r_j$$

3.3 Border Contrast Map

Boundary prior is a very successful cue in case of saliency detection. Many researchers have found this cue useful in generating state of the art results. It mainly assumes that boundary regions are mostly backgrounds. This can be easily understood by examining the simple examples and their surroundings in Figure 4. In that figure, the backgrounds of both 4(a) and 4(b) have similar colors like the border regions. Although, Figure 4(a) is simple and have a background similar to the border regions and Figure 4(b)'s boundary regions have several color intensities but all of them constructs the boundary and significantly differ from the salient object or region. To get advantage of this simple but significant characteristic of a digital image, every region is compared with the border regions to find the dissimilarity of that region with the image boundary. In eq. 6 these contrast dissimilarities are computed and the border contrast map is found by averaging the dissimilarities of each region with all boundary regions using eq. 7.

$$\rho(r_i, b_j) = 1 - exp\left(\frac{-d_c^2(r_i, b_j)}{\sigma^2}\right)$$
$$d_c = \text{Euclidian color distance} \tag{6}$$
$$\sigma = \text{constant to control strength}$$
$$b_j = j^{th} \text{ region of total P border regions}$$

$$S_B(r_i) = \frac{1}{\alpha P} \sum_{k=1}^{\alpha P} B_i(k)$$
$$B_i(k) = \rho(r_i, b_j), \text{Dissimilarity vector for } r_i \tag{7}$$
$$\alpha = \text{constant } 0 < \alpha < 1$$

3.4 Color Importance Map

A color map in this study consists of the weighted average of different color channels in CIE Lab color space. In a digital image most of the image pixels are backgrounds. So, the main idea of



(a) L channel (b) a channel (c) b channel

(d) diff. L (e) diff. a (f) diff. b

(g) G_{map} (h) N_{map} (i) Imp_{clr}

Figure 5: Various stages of computing color map.

color importance map is to suppress those background pixels by applying some weighting techniques so that the salient regions become prominent. The device independent Lab color space is used to calculate the weighted color map. Because, it provides us more information than RGB color space where all three channels contain similar information. The average L, a and b channels are calculated using eq. 8 and the color difference images are found by subtracting each channel's average from color intensity values of each pixel as of eq. 9. The visual model of different Lab channels are given in Figure 5. In that figure, the color difference images of Figure 5(d), 5(e) and 5(f) shows remarkable contrasts than its corresponding raw channel images of Figures 5(a), 5(b) and 5(c). The average calculation of the L channel is a bit different. Because, in this channel the intensity values can differ in a range of 0 to 255. Thus the mean and mode can have a large difference and because of frequency, the mode may become insignificant. For this reason, the minimum of $(L_{mean} + L_{mode})/2$ and L_{mean} is taken as the weighted average in eq. 8. The Grey map (G_{map}) in eq. 10 sums up all three channels and Normalized map (N_{map}) in eq. 11 produces the magnitude value of the color space. These two maps are combined into a single saliency map using eq. 12. Though, gamma correction improves the result very slightly, it is performed on the grey map before the combination in eq. 12 as seen in Figure 5(i). The pixel level color map is transformed into a region level map using the region average in eq. 13. The combined map highlights the salient region significantly, but it also fails at image boundary. So, the border contrast map is introduced with weighted color map in eq. 15 to produce the boundary aware color importance map. 6 shows the boundary aware version of the color importance map. As seen in this figure, the boundary aware color map of Figure 6(c) is found by merging border contrast map of Figure 6(a) and seam map of Figure 6(b). The resulting combination is a better candidate for saliency detection.



Figure 6: Boundary aware color importance map.

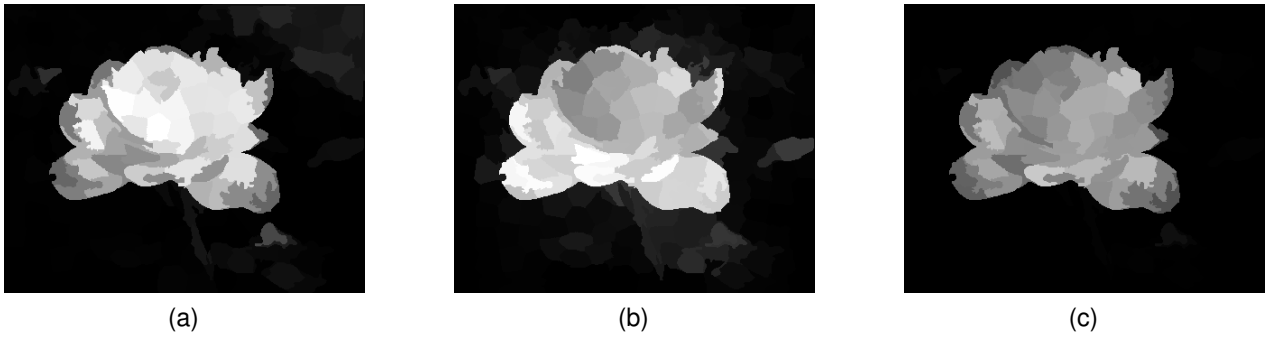


Figure 7: Combination of boundary aware color map and seam map.

$$
\begin{aligned}
C_{avg} &= [L_{avg}, a_{avg}, b_{avg}] \\
L_{avg} &= min\left(\frac{L_{mean} + L_{mode}}{2}, L_{mean}\right) \\
a_{avg} &= \frac{a_{mean} + a_{mode}}{2} \\
b_{avg} &= \frac{b_{mean} + b_{mode}}{2}
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
dI_{Lab}(x,y) &= \left|I_{Lab}(x,y) - C_{avg}\right| \\
I_{Lab}(x,y) &= \text{Color intensity of pixel(x,y) Lab image}
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
G_{map}(x,y) &= w_L d_L(x,y) + w_a d_a(x,y) + w_b d_b(x,y) \\
where, w_L &= 0.1 \text{ and } w_a = w_b = 0.45
\end{aligned} \tag{10}
$$

$$
N_{map}(x,y) = \sqrt{d_L^2(x,y) + d_a^2(x,y) + d_b^2(x,y)} \tag{11}
$$

$$
\begin{aligned}
Imp_{clr}(x,y) &= G_{map}^{\gamma}(x,y) \times N_{map}(x,y) \\
where, \gamma &= 1.5^{-1}
\end{aligned} \tag{12}
$$

$$
\begin{aligned}
Imp_{clr}(r_i) &= \frac{1}{|r_i|} \sum_{\forall I_{clr}(x,y)=i} Imp_{clr}(x,y), where \\
|r_i| &= \text{Number of pixels in region } r_i
\end{aligned} \tag{13}
$$

$$
\begin{aligned}
S_{clr,s}(r_i) &= \frac{Imp_{clr}(r_i)}{\frac{1}{N}\sum_{k=1}^{N} d_s(r_i, r_k)} \\
d_s &= \text{Euclidian region distance as in eq. 5} \\
N &= \text{Number of regions}
\end{aligned} \tag{14}
$$

$$
\begin{aligned}
S_{bClr}(r_i) &= S_B(r_i) \times [1 + S_{clr,s}(r_i)] \\
S_B(r_i) &= \text{Border contrast map from eq. 7}
\end{aligned} \tag{15}
$$

3.5 Combining Seam and Color Importance Maps

Following the traditional way of combining saliency maps, the boundary aware color importance map is combined with seam importance map by eq. 16 using multiplication as seen in Figure 7. The combined saliency map of Figure 7(c) is found by merging boundary aware color map of Figure 7(a) and seam map of Figure 7(b). The samples in Figure 7(a) and Figure 7(b) both contain highlighted portions of backgrounds. Though the combined result in Figure 7(c) successfully suppresses the background, it does not enumerate the salient result sufficiently. So, an optimization procedure is required to do the rest of the job.

$$
Sal'(r_i) = S_{bClr}(r_i) \times S_{seam,s}(r_i) \tag{16}
$$

3.6 Color Distance and Contrast Dissimilarity

The existing color contrast between the regions in an image is one of the main attractions of saliency detection studies. The euclidean color distance in the Lab color space of the superpixels is used to measure the contrast difference between the superpixels as in eq. 17. The global contrast map is found by using eq. 18. This is used in the optimization equation as the smoothness weight.

$$\begin{aligned} \rho(r_i, r_j) &= exp\left(\frac{-d_c^2(r_i, r_j)}{\sigma^2}\right) \\ d_c(r_i, r_j) &= \sqrt{(c_i - c_j)^2} \\ c_i, c_j &= \text{Mean color vector of } r_i, r_j \\ \sigma &= \text{constant to control strength} \end{aligned} \tag{17}$$

$$\begin{aligned} S_G(r_i) &= \frac{1}{N}\sum_{k=1}^{N}\rho(r_i, r_j) \\ N &= \text{Number of regions} \end{aligned} \tag{18}$$

3.7 Optimization

Most of the previous researchers use multiplication or weighted summation to combine multiple saliency cues. The resulting combination does not do well in case of generalization. The main objective of saliency detection techniques is to generate a map closer to the binary ground truth image. So, some authors developed an optimization technique towards this goal. In this research, optimization is done by minimizing their quadratic cost function 19 proposed in [2]. All three terms of this cost function are square errors and it can be solved by using least square method. The first term of the cost function represents the background which forces a superpixel p_i with greater background weight w_i^{bg} to take a small value s_i closer to 0. In this research, the background weight w_i^{bg} is $1 - Saliency$. The second term represents the foreground which forces a superpixel p_i with greater foreground weight w_i^{fg} to take a value closer to 1. Here, the foreground weight is one of the saliency maps found in this study. The third term is the smoothness term which smooths both the background and foreground by removing small noises. The color dissimilarity based global contrast map from eq. 18 is used as the weight for the smoothness term. Since the target of this weight is to smooth the combined saliency map, connecting two levels of nearest neighbours and the boundary regions before dissimilarity measures produces better results. In this study, different combinations of saliency measures are used as the foreground and background weight in eq. 19 as shown in table 2 and a balanced combination is proposed as the optimal combination. The optimal saliency map of Figure 8(c) is found by solving the quadratic equation proposed in [2] using 8(a) as background map and 8(b) as foreground map.

$$\sum_{i=0}^{N} w_i^{bg} s_i^2 + \sum_{i=0}^{N} w_i^{fg}(s_i - 1)^2 + \sum_{i,j} w_{i,j}(s_i - s_j)^2 \tag{19}$$

4 Experiments and Results

4.1 Datasets

Standard datasets like MSRA 1k, MSRA 10k, CSSD and ECSSD are used to evaluate the performance of this study. Here, MSRA 1k and 10k are subsets of widely used MSRA original dataset which contains 20,000 images along with their hand labelled rectangular ground truth images which are manually annotated by 3-9 users. In [20] and [21], the authors has claimed that the original MSRA dataset has limitations in fine grained evaluation because of being too coarse in some cases. In [21] the authors presented the MSRA 1k or ASD dataset by manually segmenting each salient region which overcomes the limitations of the original MSRA dataset. An extended version of MSRA 1k dataset which contains pixel level binary ground truth images of 10,000 images from the original MSRA dataset is presented in [15]. They named this dataset MSRA10k which was previously called THUS10000. The first row of Figure 9 shows some bounding box annotation from original MSRA dataset and the second row shows the pixel level ground truth annotation from MSRA10k dataset. Almost all modern saliency detection techniques use this dataset as a benchmark for evaluation of saliency detection methods.

MSRA based datasets are good for saliency detection evaluation and contain a large variety in contents, but their backgrounds are relatively simple and smooth. To evaluate the complex situations in saliency detection, where objects and backgrounds have relatively low contrast and object surroundings are comparatively complex, in [12] the authors have proposed CSSD dataset with pixel level binary ground truths in 2013 which is a more challenging dataset with 200 images containing complex situations and they extended the dataset by increasing the images to 1,000 in [22] and named it Extended Complex Scene Saliency Dataset (ECSSD). Some members of CSSD and ECSSD datasets with complex backgrounds and low contrast surroundings are presented in Figure 10.

Table 1: Datasets and compared methods.

Dataset	Compared Methods
MSRA 1k	CA[23], FT[24], HC[15], LC[25], MSS[16], GS[1], SF[10], SWD[26], BARC[3], BARCS[8], BACS[5]
MSRA 10k	GS[1], SF[10], BARC[3], BARCS[8], BACS[5]
CSSD	GS[1], SF[10], BARC[3], BARCS[8], BACS[5]
ECSSD	GR[27], IT[17], SR[28], MSS[16], GS[1], SF[10], BARC[3], BARCS[8], BACS[5]

4.2 Evaluation Metrics

There are usually two types of evaluation strategies for every computer vision technique - qualitative analysis and quantitative analysis. Results found in this study are compared with several state of the art methods which are labelled as HC[15], MSS[16], GB[18], SWD[26], RC[29], AC[30], IT[17], GR[27], LC[25], FT[24], CA[23], AIM[31], GS[1], SF[10], SR[28], BARC[3], BARCS[8] and BACS[5]. Most of these labels are used in previous researches like [21], [15] and [32]. Other labels are constructed by taking some meaningful letters from the titles of the related publications.



Figure 8: Combination of foreground and background maps to generate the optimal saliency model.

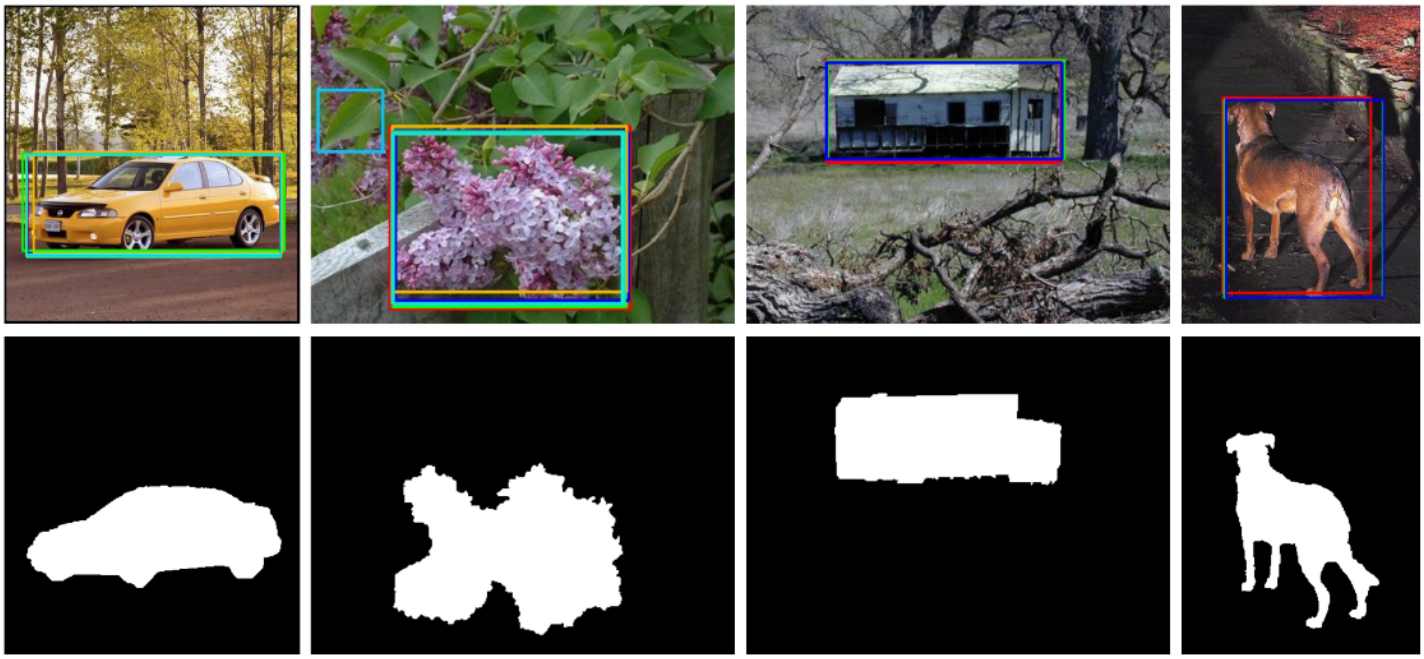


Figure 9: Rectangular annotation of original MSRA dataset and the pixel level ground truth annotation from MSRA10k dataset.

Based on publicly available results for specific dataset, the proposed method has been compared with other state of the art methods as shown in Table 1.

Qualitative analysis of this study are presented in Figure 15 and 16 with some of the mentioned methods containing comparative samples from respective technologies. For quantitative analysis precision recall curves are shown for different datasets along with other techniques that clearly proves the superiority of the proposed method. To draw this precision recall curve simple binarization is performed by varying the threshold in the range from 0 to 255 and the curves are plotted according to the ground truth data available for each input image in the corresponding dataset. Since the only consideration of precision-recall curve is whether the foreground region saliency is higher than the background saliency, to evaluate the overall performance F-beta measures in Figure 14 are also presented in this study to better understand the performance of the compared methods.

4.3 Experimental Setup

Standard datasets are used to evaluate the proposed method. An initial 600 pixel region size is considered for converting the input image into the superpixelized version using the SLIC method. Other static variables are considered as $\alpha = 0.25$ and $compactness = 20$.

4.4 Optimization Experiments

Some promising combinations of saliency maps are explored in this study which are briefly discussed in this section. In the proposed optimization method, if the border contrast map is used as the background weight and the seam importance map is used as the foreground weight, we find closer PR curves as seen in Figure 12. The f-beta score is also closer to the proposed method as depicted in that figure. Similarly, different combinations are experimented for this study. Some of them are presented in table 2 and their qualitative and quantitative results are shown in Figure 11 and Figure 12 respectively. The runtime for these three combinations are given



Figure 10: Some members of CSSD and ECSSD datasets with complex backgrounds and low contrast surroundings.



(a) Comb1 (b) Comb2 (c) Comb3 (Proposed)

Figure 11: Optimization combinations.

in table 3. In all three combinations, the smoothness weight is kept the same which is the global contrast map between the superpixels from eq. 18.

From the results in Figure 11, 12 and runtimes in table 3 it is seen that all three compared methods are closer to each other. But the second combination is the slowest since it introduces more computational complexities and the first combination is the fastest in terms of runtime comparison. On the other hand, for the MSRA 1k dataset the second combination does well in case of PR curve and F-beta test. But, the third one does well in complex situations like the ECSSD dataset. Its performance is also competitive in MSRA 1k dataset and the runtime is quite acceptable in comparison to the fastest method and its excellent performance on the complex dataset. Thus the third combination is chosen as the proposed method in this study and it will be termed as *Our* method onwards.

Table 2: Proposed optimization combinations.

Method	Foreground	Background
Comb1	Border Contrast Map from eq. 7	Seam Map from eq. 4
Comb2	Border Contrast Map from eq. 7	Combined Seam and Boundary Aware Color Map from eq. 16
Comb3 (Proposed)	Boundary Aware Color Map from eq. 15	Weighted Seam Map from eq. 5

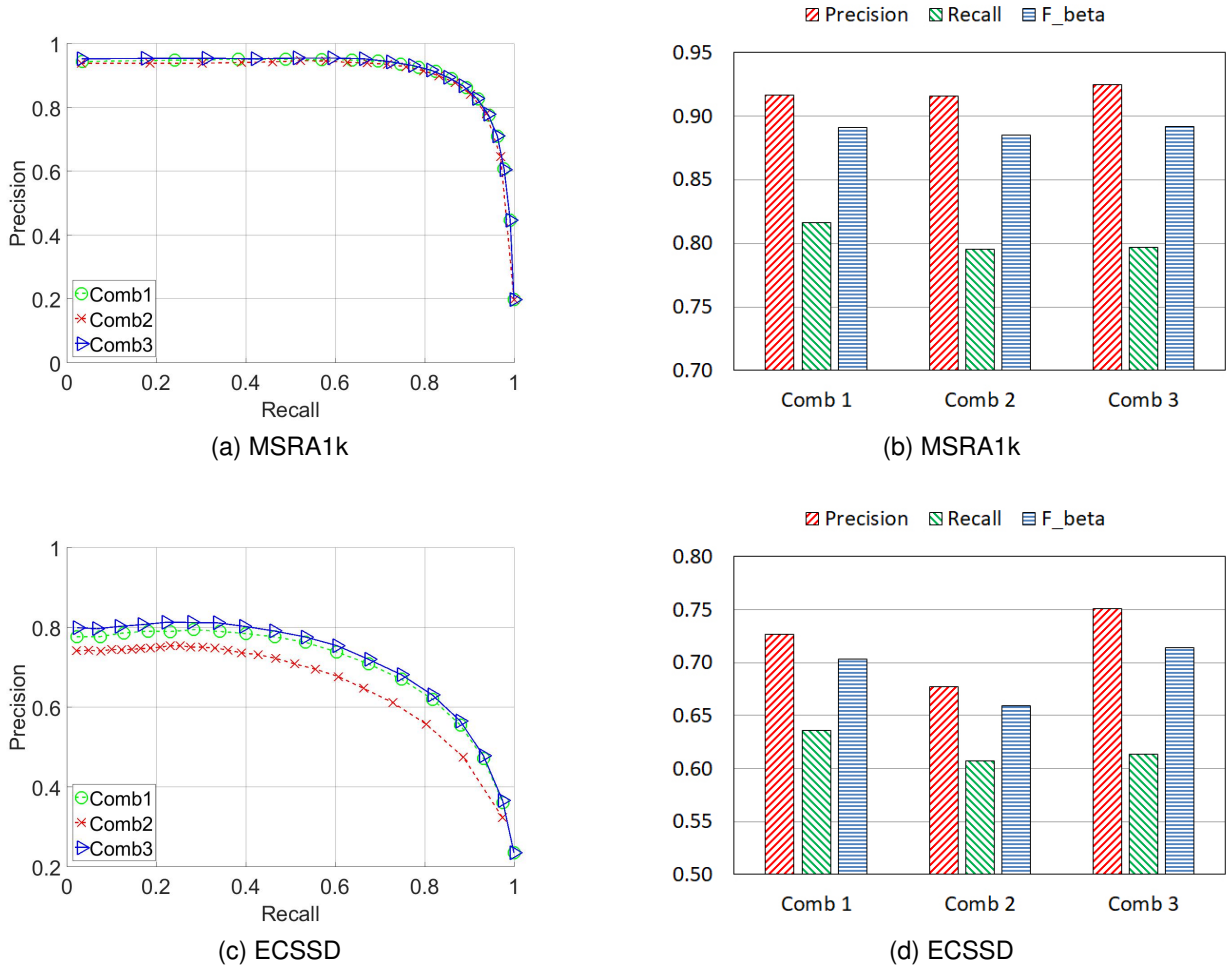


Figure 12: Precision recall curves for different combinations.

Table 3: Runtime comparison of proposed combinations.

Method	Comb1	Comb2	Comb3 (Proposed)
Time (sec)	1.63	2.65	1.745
Code	Matlab	Matlab	Matlab

4.5 Results

The qualitative analysis in Figure 15 and 16 visually demonstrates the performance of the presented method compared to other state of the art techniques. The first row shows the input images and the last row contains the target pixel level ground truths for corresponding input images. From those figures it is seen that the proposed method done quite well with respect to other compared techniques.

In case of quantitative comparison, the precision-recall curve in Figure 13 clearly demonstrates the superiority of the proposed method. As pointed in [15], all methods have the same precision and recall values at maximum recall for threshold = 0 and all pixels are considered to be in the foreground. On the other hand, if the curves are observed from recall value 0.5 to 1, it is realized that for all datasets the proposed method exhibits higher precision values due to smoother saliency containing more pixels in the salient region with saliency value 255 until it reaches to the common minimum precision value at recall = 1 and distinctly remains at the top of other curves until that. When a PR curve shows higher precision values for higher recalls, that is considered better. So, for the given datasets the proposed method clearly outperforms all other methods in terms of PR curve analysis.

$$F_\beta = \frac{(1+\beta^2)Precision \times Recall}{\beta^2 \times Precision + Recall} \tag{20}$$

To measure the overall quantitative performance F_β measures are also calculated. In doing so, $\beta = 0.25$ is considered in eq. 20 like [8] and [3] to draw F-beta measures in Figure 14. The bars in this figure demonstrate that the presented method clearly surpasses other state of the art techniques.

The average runtime of different methods are given in table 4 as presented in [15]. The given times are dependent on the programming platform and different machine specifications. It may vary on other computer or programming platforms. So, this is mentioned here only for references. The average runtime of the proposed method is also mentioned in that table. The configuration of the

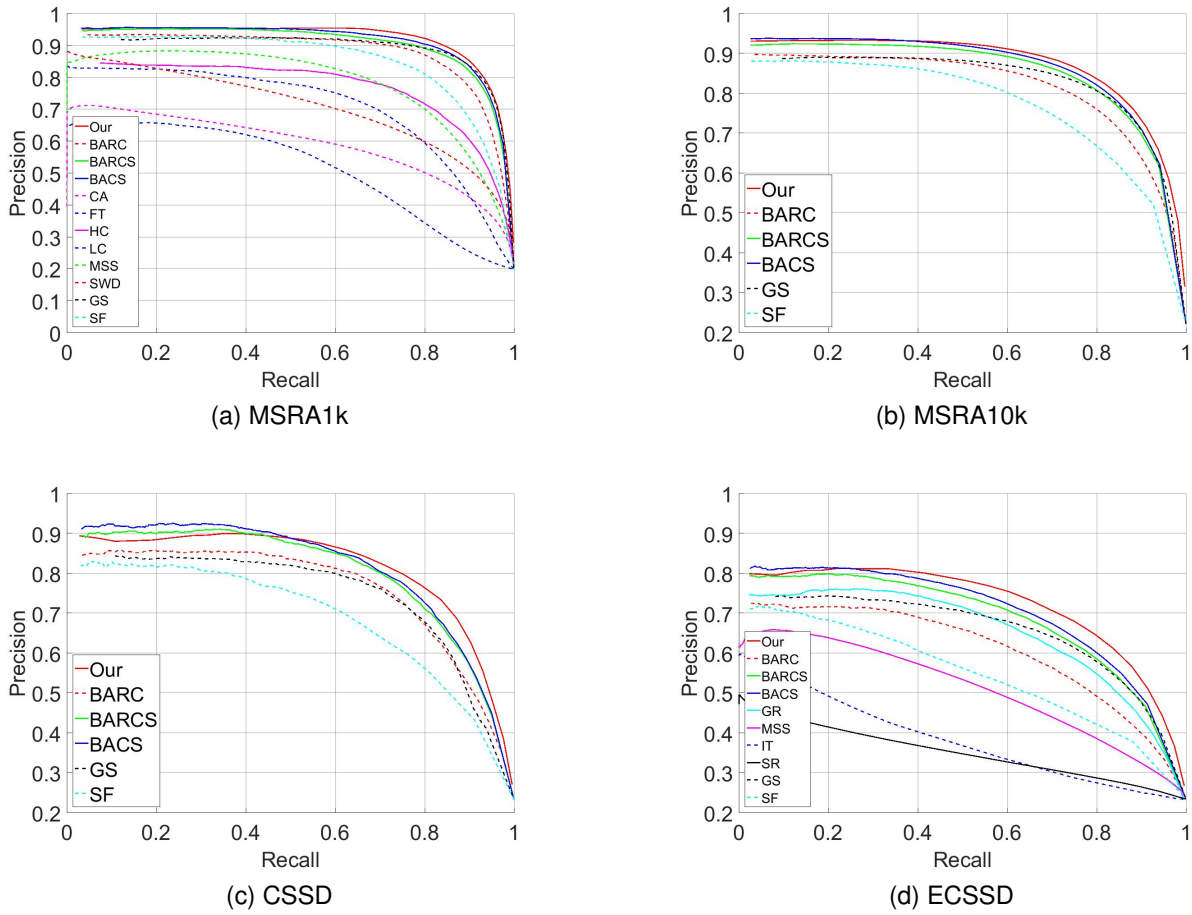


Figure 13: Precision recall curves for different datasets.

windows machine for this study is a second generation Intel core i5 CPU with 8 Gb of DDR3 RAM.

Table 4: Time comparison of different methods.

Method	Our	AIM	CA	CB	FT	GB
Time (s)	1.745	4.288	53.1	5.568	0.102	1.614
Code	Matlab	Matlab	Matlab	M&C	C++	Matlab
Method	HC	IT	LC	MSS	RC	SWD
Time (s)	0.019	0.611	0.018	0.106	0.254	0.100
Code	C++	Matlab	C++	Matlab	C++	Matlab

5 Conclusion

A novel method of saliency detection is presented in this paper that combines seam and color map based saliency detection procedures. The background and foreground weights proposed for the quadratic eq. 19 have successfully outperformed other state of the art methods and those have never been explored before. Some intermediate combinations are also presented with their detailed comparison. Though, one of the combinations is marked as the proposed solution balancing the runtime and performance, any of them can be used with respect to the application demand. Moreover, the presented results in the experimental section of this study clearly demonstrates the superiority of the proposed method over other state of the art techniques in terms of qualitative and quantitative analysis. The use of standard benchmark datasets strengthens the acceptability of this study and its publicly available results make an opportunity for other researchers to compare their findings with the proposed method.

Computational devices are improving day by day. So, in future exactness of saliency detection technique will be the main concern. Moreover, a fine tuned version of the proposed method is possible and its implementation in an application level programming language like C++ would certainly improve its runtime. Besides that, different saliency weight calculation procedures of this study like border contrast map, color map and contrast based dissimilarity can be parallelized and this way runtime can come to its minimum. So, the future plan is to improve the saliency detection technique performance and apply it into various computer vision related applications like pedestrian detection, action recognition, object detection and recognition etc. Experimenting the proposed model with more complex datasets can be another interesting scope of study.

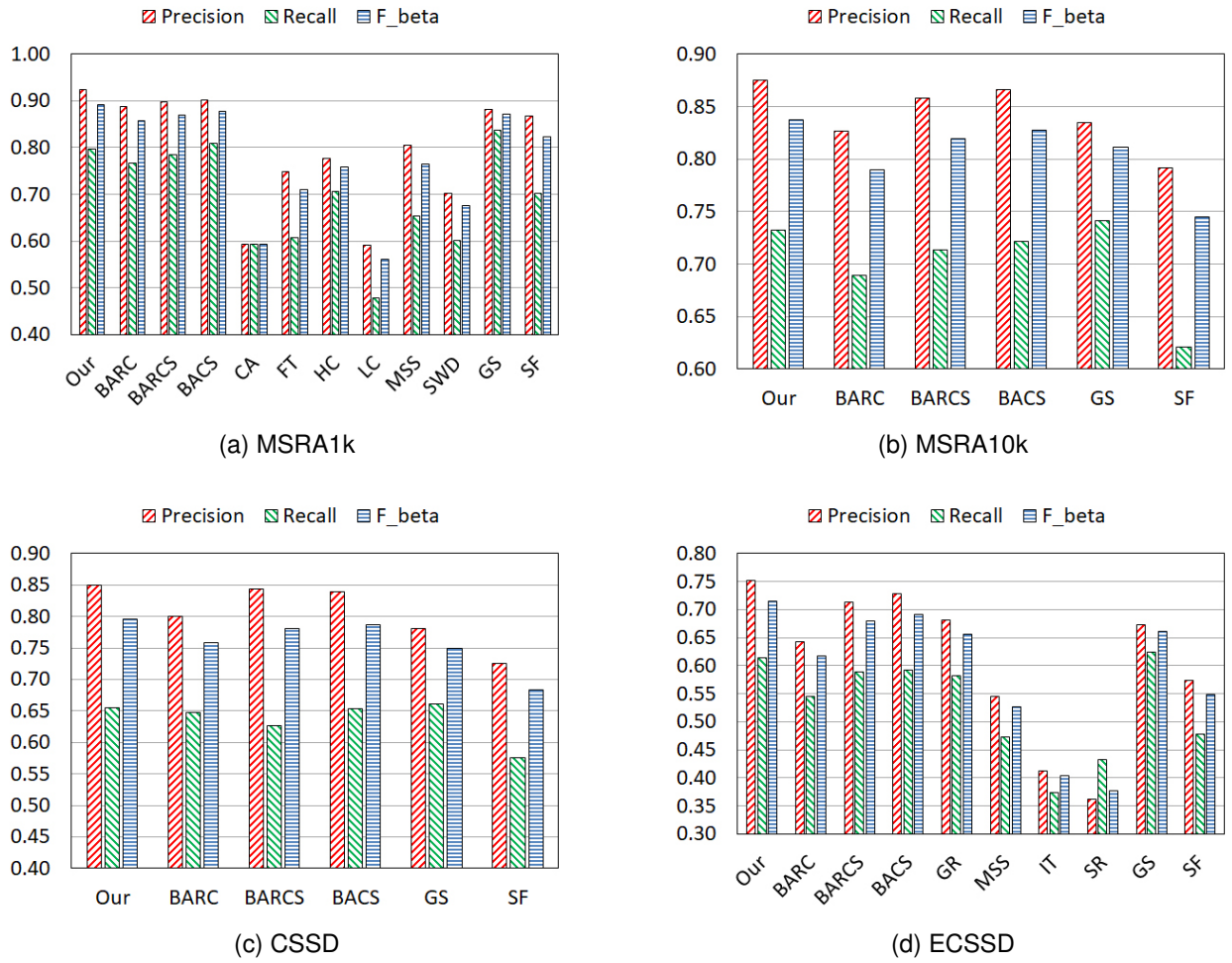


Figure 14: F-beta curves for different datasets.

6 Declaration

6.1 *Availability of data and materials*

The experimental results of this study are available for download at (`https://aminulislam.net/#research`). The famous MSRA 10k dataset is freely available at (`https://mmcheng.net/msra10k`), the CSSD and ECSSD datasets are available at (`http://www.cse.cuhk.edu.hk/~leojia/projects/hsaliency/dataset.html`). Other datasets used and/or analyzed during this study are available from the authors upon reasonable request.

6.2 *Competing interests*

The authors declare that they have no competing interests.

6.3 *Authors' contributions*

AI executed the experiments, analyzed results, and wrote the initial draft of the manuscript. SMMA contributed to the concept, edited the manuscript, supervised the project and provided technical support and conceptual advice. The authors reviewed and approved the final manuscript.

6.4 *Acknowledgements*

We thank all the people who gave us various perceptive and constructive comments.

References

[1] Y. Wei, F. Wen, W. Zhu, J. Sun, "Geodesic Saliency Using Background Priors," in 12th European Conf. on Computer Vision (ECCV), Proc., Part III, volume 7574, 29–42, Springer, 2012, doi:10.1007/978-3-642-33712-3_3.

[2] W. Zhu, S. Liang, Y. Wei, J. Sun, "Saliency Optimization from Robust Background Detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2814–2821, 2014, doi:10.1109/CVPR.2014.360.

[3] S. M. M. Ahsan, J. K. Tan, H. Kim, S. Ishikawa, "Boundary Aware Regional Contrast Based Visual Saliency Detection," in The Twenty-First Int. Symposium on Artificial Life and Robotics, 258–262, 2016.

[4] T. G. Joy, M. M. Hossan, "A Study on Salient Region Detection using Bounday and Color Cue," Khulna University of Engineering & Thechnology, 2018, thesis no. CSER-18-26.

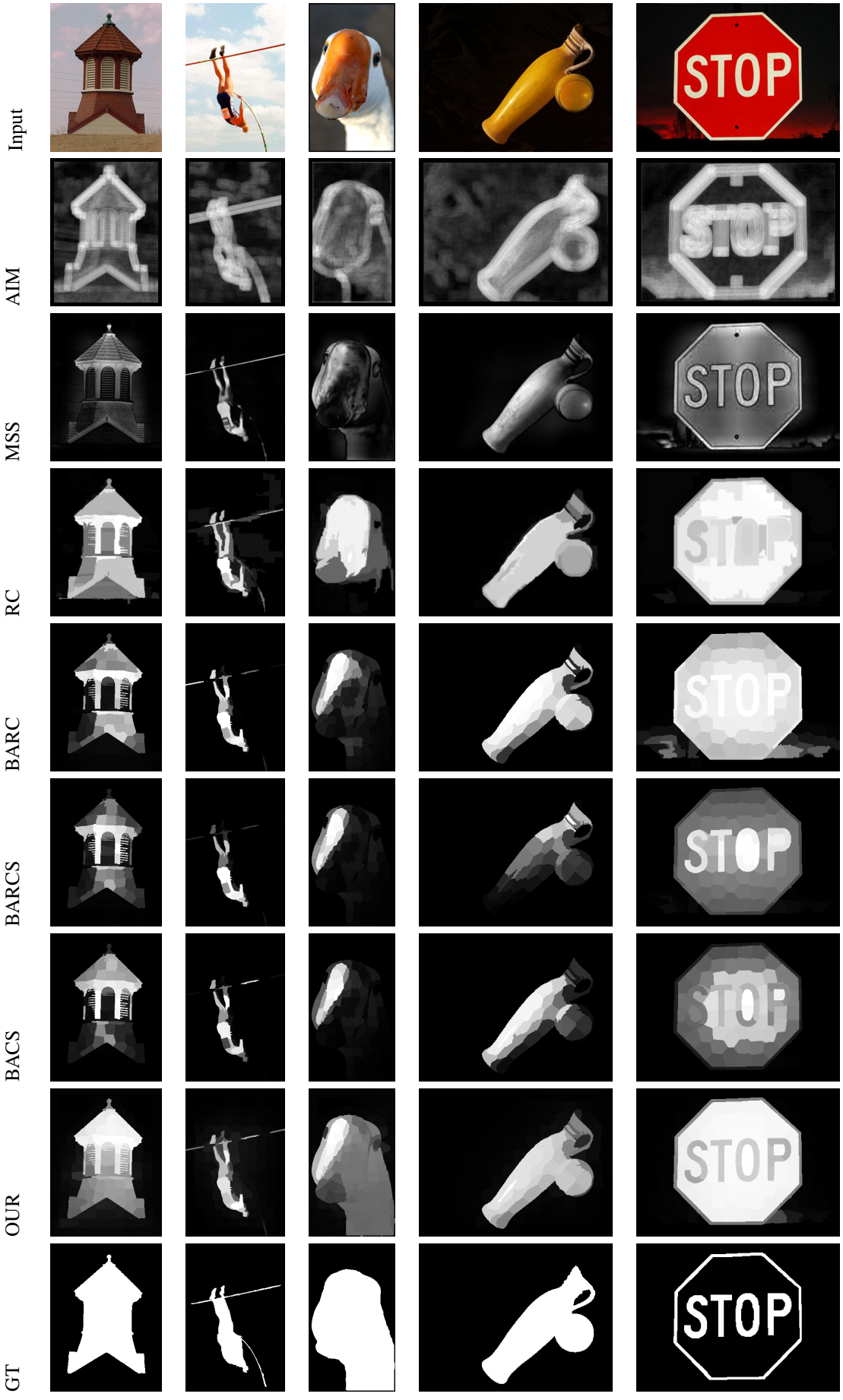


Figure 15: Qualitative comparison of our saliency output with other state of the art methods on MSRA 10K dataset.



Figure 16: Qualitative comparison of the proposed saliency map with other state of the art methods on ECSSD dataset.

[5] A. Islam, S. M. M. Ahsan, J. K. Tan, "Saliency Detection using the Combination of Boundary Aware Color-map and Seam-map," in International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), 2019.

[6] S. Avidan, A. Shamir, "Seam Carving for Content-aware Image Resizing," ACM Trans. Graph., **26**(3), 2007, doi:10.1145/1276377.1276390.

[7] Y. Li, K. Fu, L. Zhou, Y. Qiao, J. Yang, "Saliency detection via foreground rendering and background exclusion," in Proc. IEEE Int. Conf. Image Processing (ICIP), 3263–3267, 2014, doi:10.1109/ICIP.2014.7025660.

[8] A. Islam, S. M. M. Ahsan, J. K. Tan, "Saliency Detection using Boundary Aware Regional Contrast Based Seam-map," in Proc. Int. Conf. Innovation in Engineering and Technology (ICIET), 2018, doi:10.1109/CIET.2018.8660825.

[9] T. Liu, J. Sun, N. Zheng, X. Tang, H. Shum, "Learning to Detect A Salient Object," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1–8, 2007, doi:10.1109/CVPR.2007.383047.

[10] F. Perazzi, P. Krähenbühl, Y. Pritch, A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 733–740, 2012, doi:10.1109/CVPR.2012.6247743.

[11] A. Borji, M. Cheng, H. Jiang, J. Li, "Salient Object Detection: A Benchmark," IEEE Transactions on Image Processing, **24**(12), 5706–5722, 2015, doi:10.1109/TIP.2015.2487833.

[12] Q. Yan, L. Xu, J. Shi, J. Jia, "Hierarchical Saliency Detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1155–1162, 2013, doi:10.1109/CVPR.2013.153.

[13] C. Yang, L. Zhang, H. Lu, X. Ruan, M. Yang, "Saliency Detection via Graph-Based Manifold Ranking," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 3166–3173, 2013, doi:10.1109/CVPR.2013.407.

[14] H. Zhang, C. Xia, "Saliency Detection combining Multi-layer Integration algorithm with background prior and energy function," arXiv preprint arXiv:1603.01684, 2016.

[15] M. Cheng, G. Zhang, N. J. Mitra, X. Huang, S. Hu, "Global contrast based salient region detection," in Proc. CVPR 2011, 409–416, 2011, doi:10.1109/CVPR.2011.5995344.

[16] R. Achanta, S. Süsstrunk, "Saliency detection using maximum symmetric surround," in Proc. IEEE Int. Conf. Image Processing, 2653–2656, 2010, doi:10.1109/ICIP.2010.5652636.

[17] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, **20**(11), 1254–1259, 1998, doi:10.1109/34.730558.

[18] J. Harel, C. Koch, P. Perona, "Graph-based visual saliency," in Advances in neural information processing systems, 545–552, 2007.

[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, **34**(11), 2274–2282, 2012, doi:10.1109/TPAMI.2012.120.

[20] Z. Wang, B. Li, "A two-stage approach to saliency detection in images," in 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, 965–968, IEEE, 2008.

[21] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, "Frequency-tuned salient region detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1597–1604, 2009, doi:10.1109/CVPR.2009.5206596.

[22] J. Shi, Q. Yan, L. Xu, J. Jia, "Hierarchical Image Saliency Detection on Extended CSSD," IEEE Transactions on Pattern Analysis and Machine Intelligence, **38**(4), 717–729, 2016, doi:10.1109/TPAMI.2015.2465960.

[23] S. Goferman, L. Zelnik-Manor, A. Tal, "Context-Aware Saliency Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, **34**(10), 1915–1926, 2012, doi:10.1109/TPAMI.2011.272.

[24] R. Achanta, S. Süsstrunk, "Saliency detection for content-aware image resizing," in Proc. 16th IEEE Int. Conf. Image Processing (ICIP), 1005–1008, 2009, doi:10.1109/ICIP.2009.5413815.

[25] Y. Zhai, M. Shah, "Visual Attention Detection in Video Sequences Using Spatiotemporal Cues," in Proc. of the 14th ACM Int. Conf. on Multimedia, MM '06, 815–824, ACM, 2006, doi:10.1145/1180639.1180824.

[26] L. Duan, C. Wu, J. Miao, L. Qing, Y. Fu, "Visual saliency detection by spatially weighted dissimilarity," in Proc. CVPR, 473–480, 2011, doi:10.1109/CVPR.2011.5995676.

[27] C. Yang, L. Zhang, H. Lu, "Graph-regularized saliency detection with convex-hull-based center prior," IEEE Signal Processing Letters, **20**(7), 637–640, 2013.

[28] X. Hou, L. Zhang, "Saliency detection: A spectral residual approach," in 2007 IEEE Conference on computer vision and pattern recognition, 1–8, Ieee, 2007.

[29] M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, S. Hu, "Global Contrast Based Salient Region Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, **37**(3), 569–582, 2015, doi:10.1109/TPAMI.2014.2345401.

[30] R. Achanta, F. Estrada, P. Wils, S. Süsstrunk, "Salient region detection and segmentation," in International conference on computer vision systems, 66–75, Springer, 2008.

[31] N. D. B. Bruce, J. K. Tsotsos, "Saliency, attention, and visual search: An information theoretic approach," Journal of Vision, **9**, 5–5, 2009, doi:10.1167/9.3.5.

[32] H. Chen, P. Wang, M. Liu, "From co-saliency detection to object co-segmentation: A unified multi-stage low-rank matrix recovery approach," in Proc. IEEE Int. Conf. Robotics and Biomimetics (ROBIO), 1602–1607, 2015, doi:10.1109/ROBIO.2015.7419000.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 3, 146-154 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Synchronization in a Class of Fractional-order Chaotic Systems via Feedback Controllers: a Comparative Study

Juan Luis Mata-Machuca*

Instituto Politecnico Nacional, UPIITA, Department of Advanced Technologies, Mexico City, 07340, Mexico

ARTICLE INFO

Article history:
Received: 03 December, 2020
Accepted: 04 May, 2021
Online: 20 May, 2021

Keywords:
Fractional-order chaotic system
Proportional control
Nonlinear fractional-order PI control
Active control
Master-slave synchronization
Equivalent electronic circuit

ABSTRACT

In this paper a synchronization methodology of two fractional-order chaotic oscillators under the framework of identical synchronization and master-slave configuration is introduced. The proposed methodology is based on a fractional-order feedback control design under the frame of control theory, the feedback controllers provide synchronization convergence. A comparative study between a proportional control, a nonlinear fractional-order proportional-integral control and an active control is presented. The above is showed via an analysis of the dynamic of the called synchronization error. Numerical experiments using the mathematical model of the fractional-order unified chaotic system and its equivalent electronic circuit corroborate the satisfactory results of the proposed schemes.

1 Introduction

A significant development of chaos occurred after the publication by Lorenz [1] of a work related with the existence of non-periodic solutions in a model formed by nonlinear differential equations. Lorenz's contribution provided the interest in studying deterministic systems that generate dynamic trajectories strongly influenced by sensitivity to initial conditions. However, it was until 1975 when the term chaos was introduced [2], since then chaotic systems have been studied intensively, and this is one important area of current research in physics, mathematics, and engineering. The analysis of nonlinear complex dynamic behaviors has led to important technological application. In particular the analysis of chaotic oscillators has had important advances in process engineering, cryptography of images, secure data transmission, life science, and information processing, between others [3]–[5].

There are several definitions about chaotic systems. However, in general terms, a chaotic system is a deterministic dynamic system that exhibits irregular behavior, similar to random behavior [6, 7]. A very important aspect in relation to chaos theory is the synchronization of chaotic systems. Since Pecora and Carrol [8] showed that two seemingly random and unpredictable chaotic behaviors can converge in a single trajectory, new expectations arose around chaos theory, trying to solve its control, understanding and prediction.

The synchronization of chaotic systems consists of coupling two or more chaotic systems, after a transitory time, they exhibit identical chaotic oscillations. Synchronization can be solved from the point of view of control theory, designing a slave system which follows the variables of the master system. In the open literature has been presented a lot of important contributions for the design of control schemes for synchronization purposes in chaotic oscillators [9]–[11], observer based controllers [12, 13], adaptive algorithms [14, 15], optimal controllers [16, 17] and so on [18]–[20].

From the above, control theory has been looking for advanced techniques in order to reach a high performance in the operation of processes, in particular with dynamic systems with complex behavior, such as the nonlinear fractional-order systems [21]–[23]. The fractional calculus for the study of dynamic systems with chaotic behavior has emerged as an alternative framework, since it has been proved that it is possible that a class of systems exhibits chaos despite its order is fractional [24]–[27]. Generally, nonlinear systems display complex dynamic behavior as steady state multiplicity, in-

*Corresponding Author: Juan Luis Mata-Machuca, Email: jmatam@ipn.mx

stabilities, complex oscillations and so on under different initial conditions, external disturbances and time-varying parameters, leading to chaotic dynamic behaviors. However, besides of the scientific interest on the study and analysis of nonlinear system with exotic dynamic behaviors, the applications for engineering purposes have been growing importantly [28]–[31].

There are many engineering applications based on dynamical systems with fractional operators, for instance: fractional modelling of the human arm dynamics [32], modelling and identification of viscoelastic mechanical systems [33], modelling of electrical systems [34], analysis and control of economics and financial systems [35, 36], vibration and acoustics phenomena [37], modeling of friction in electric machines [38], problems in electrochemistry, biophysics and bioengineering [39]–[41], methods for signal and image processing [42, 43], applications in automatic control, mechatronics and robotics [44]–[46], among others [47]–[51]. However, there are a few practical applications in this field to prove the feasibility of the physical realization of the proposed techniques [52]–[54]. Some works deal with the master-slave synchronization of fractional-order systems via unidirectional linear coupling [55]–[58], where the proposed methods are applied in the synchronization of fractional-order Rössler, Chua, Chen, unified systems and so on.

The main contribution of this work is to achieve the synchronization of a class of fractional-order systems and its verification via the equivalent electronic circuit by means of a proportional control, a nonlinear fractional-order proportional-integral control and an active control. The synchronization implies that all trajectories of the slave system tend to follow all trajectories of the master system. In this sense, the main problem is to find a control input such that the synchronization is possible. The suitable conditions for the synchronization are obtained by solving the dynamical equation of the synchronization error. Furthermore, to verify the effectiveness of the suggested approach we have implemented an electronic circuit using Matlab-Simulink's Simscape library to obtain the synchronization regime between two fractional-order unified chaotic systems for the same order.

2 Fractional-order feedback controllers

The master system of fractional order α can be represented by,

$$D^{\alpha} x_i = f_i(x_1, x_2, \ldots, x_n), \tag{1}$$

with $x_i(0) = x_{i0}$, $i = 1, 2, \ldots, n$, $x = [x_1,\ x_2,\ \ldots,\ x_n]^T \in \Omega \subset \mathbb{R}^n$ is the state vector, and $f : \Omega \to \mathbb{R}^n$ is a nonlinear continuous function.

In (1), the fractional derivative of x_i is defined as [**?**],

$$\begin{aligned} D^{\alpha} x_i = \ {}_{t_0}D_t^{\alpha} x_i(t) &= \frac{d^{\alpha} x_i(t)}{dt^{\alpha}} \\ &= \frac{1}{\Gamma(m-\alpha)} \int_{t_0}^{t} \frac{d^m x_i(\tau)}{d\tau^m} (t-\tau)^{m-\alpha-1}\, d\tau, \end{aligned} \tag{2}$$

where: $\alpha \in \mathbb{R}^+$, $m - 1 \leq \alpha < m$, $m \in \mathbb{N}$, $\frac{d^m x_i(\tau)}{d\tau^m}$ is the m-th time derivative of x_i and Γ is the gamma function.

2.1 Proportional control

Let us propose the controlled slave system of fractional order α as follows,

$$D^{\alpha} z_i = f_i(z_1, z_2, \ldots, z_n) + u_i \tag{3}$$

$$u_i = k_i(x_i - z_i) \tag{4}$$

with $z_i(0) = z_{i0}$, $i = 1, 2, \ldots, n$, $z = [z_1,\ z_2,\ \ldots,\ z_n]^T \in \Omega \subset \mathbb{R}^n$ is the state vector of the slave system, and k_i is the proportional gain.

Let us define the synchronization error as,

$$\varepsilon_i = x_i - z_i. \tag{5}$$

The synchronization performance is obtained by means of the analysis of the dynamics of the synchronization error.

First, we take the time derivative of order α of (5),

$$D^{\alpha} \varepsilon_i = D^{\alpha} x_i - D^{\alpha} z_i \tag{6}$$

Now, replacing the fractional dynamics of equations (1) and (3) into (6),

$$D^{\alpha} \varepsilon_i + k_i \varepsilon_i = f_i(x) - f_i(z) \tag{7}$$

Then, the solution of (7) is given by,

$$\begin{aligned} \varepsilon_i(t) = \ & \varepsilon_{i0} E_{\alpha}(-k_i t^{\alpha}) + \\ & + \int_0^t (t-\tau)^{\alpha-1} E_{\alpha,\alpha}(k_i (t-\tau)^{\alpha}) \left[f_i(x) - f_i(z) \right] \mathrm{d}\tau \end{aligned} \tag{8}$$

where $\varepsilon_{i0} = x_{i0} - z_{i0}$.

In (8), the Mittag-Leffler function is defined as [**?**],

$$E_{\alpha}(q) = \sum_{j=0}^{\infty} \frac{q^j}{\Gamma(\alpha j + 1)}, \quad q \in \mathbb{C},\ \mathbb{R}e(\alpha) > 0 \tag{9}$$

$$E_{\alpha,\alpha}(q) = \sum_{j=0}^{\infty} \frac{q^i}{\Gamma(\alpha j + \alpha)}, \quad q,\ \alpha \in \mathbb{C},\ \mathbb{R}e(\alpha) > 0 \tag{10}$$

To guarantee the convergence of the synchronization error in (8), Assumption 1 and 2 are introduced,

Assumption 1 $|f_i(x) - f_i(z)| \leq L_i < \infty$

Assumption 2 *The proportional gain is positive.*

Applying Assumption 1 in (8) we have,

$$\begin{aligned} |\varepsilon_i(t)| \leq \ & |\varepsilon_{i0} E_{\alpha}(-k_i t^{\alpha})| + \\ & + L_i \int_0^t |(t-\tau)^{\alpha-1} E_{\alpha,\alpha}(-k_i (t-\tau)^{\alpha})| \mathrm{d}\tau \end{aligned}$$

Finally, by Assumption 2,

$$\lim_{t \to \infty} |\varepsilon_i(t)| \leq \frac{L_i}{k_i} \tag{11}$$

From the previous analysis, the following proposition is established.

Proposition 1 *If Assumptions 1 and 2 are fulfilled, then the synchronization of the master system (1) and the controlled slave system (3)-(4) is achieved, where, the trajectories of ε_i belongs to a closed ball with radius proportional to L_i/k_i.* □

2.2 Nonlinear fractional-order PI control

Let us consider (1) as the master system. In this case, the slave system is described by,

$$\begin{aligned} D^{\alpha}z_i &= f_i(z_1, z_2, \ldots, z_n) + u_i, \quad i = 1, 2, \ldots, n && (12)\\ u_i &= f_i(x) - f_i(z) + k_i\varepsilon_i + K_{I_i}e_i && (13)\\ D^{\alpha}e_i &= \varepsilon_i && (14) \end{aligned}$$

where, $\varepsilon_i = x_i - z_i$, k_i is the proportional gain and K_{I_i} is the integral gain.

Property 1 (see [59]) *Let I^{α} be the fractional integral of order α. Then,*

$$I^{\alpha}(D^{\alpha}e_i) = e_i \tag{15}$$

Applying Property 1 in both sides of (14),

$$e_i = I^{\alpha}(D^{\alpha}e_i) = I^{\alpha}\varepsilon_i \tag{16}$$

We analyze the dynamics of the synchronization error,

$$\begin{aligned} D^{\alpha}\varepsilon_i &= f_i(x_1, x_2, \ldots, x_n) - f_i(z_1, z_2, \ldots, z_n) - u_i && (17)\\ u_i &= f_i(x) - f_i(z) + k_i\varepsilon_i + K_{I_i}e_i && (18)\\ D^{\alpha}e_i &= \varepsilon_i && (19) \end{aligned}$$

Proposition 2 *Let P be a matrix defined by,*

$$P = \begin{bmatrix} -k_i & -k_{I_i} \\ 1 & 0 \end{bmatrix}$$

For $0 < \alpha < 1$, the interconnected system (17)-(19) is asymptotically stable if,

$$|arg(eig(P))| > \frac{\alpha\pi}{2} \tag{20}$$

where, $eig(P)$ are the eigenvalues of P.

Proof. The closed-loop system could be expressed as,

$$\begin{bmatrix} D^{\alpha}\varepsilon_i \\ D^{\alpha}e_i \end{bmatrix} = \underbrace{\begin{bmatrix} -k_i & -k_{I_i} \\ 1 & 0 \end{bmatrix}}_{P} \begin{bmatrix} \varepsilon_i \\ e_i \end{bmatrix} \tag{21}$$

The eigenvalues of P are obtained in terms of proportional and integral gains,

$$eig(P) = \frac{-k_i \pm \sqrt{k_i^2 - 4K_{I_i}}}{2}$$

Hence, (21) is asymptotically stable if $|arg(eig(P))| > \frac{\alpha\pi}{2}$. □

2.3 Active control

Let us represent the master system (1) as,

$$D^{\alpha}x = Ax + g(x), \tag{22}$$

In (22), $g : \Omega \to \mathbb{R}^n$ is a nonlinear continuous function and the matrix $A \in \mathbb{R}^{n\times n}$ is defined by,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \tag{23}$$

where, a_{ij} are constant coefficients and $A_i = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$ denotes the i-th row of matrix A, for $i = 1, 2, \ldots, n$.

Then, master system (22) can be rewritten as,

$$\begin{aligned} D^{\alpha}x_i &= A_i x + g_i(x) \\ &= \sum_{j=1}^{n} a_{ij}x_j + g_i(x), \quad i = 1, 2, \ldots, n \end{aligned} \tag{24}$$

The corresponding slave system is proposed as,

$$D^{\alpha}z_i = \sum_{j=1}^{n} a_{ij}z_j + g_i(z) + u_i, \quad i = 1, 2, \ldots, n \tag{25}$$

Then, the dynamics of the synchronization error is obtained by (24) and (25),

$$D^{\alpha}\varepsilon_i = \sum_{j=1}^{n} a_{ij}\varepsilon_j + g_i(x) - g_i(z) - u_i, \quad i = 1, 2, \ldots, n \tag{26}$$

Proposition 3 *Equation (26) is asymptotically stable if the signal u_i is active,*

$$u_i = \sum_{j=1,\ j\neq i}^{n} a_{ij}\varepsilon_j + g_i(x) - g_i(z) + k_i\varepsilon_i, \quad i = 1, 2, \ldots, n \tag{27}$$

and the following condition is satisfied,

$$k_i > a_{ii} \tag{28}$$

Proof. Substituting 27 into (26),

$$D^{\alpha}\varepsilon_i = (-k_i + a_{ii})\varepsilon_j, \quad i = 1, 2, \ldots, n \tag{29}$$

or, in matrix form

$$D^{\alpha}\varepsilon = \underbrace{\begin{bmatrix} -k_1 + a_{11} & 0 & \ldots & 0 \\ 0 & -k_2 + a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & -k_n + a_{nn} \end{bmatrix}}_{\bar{A}} \varepsilon \tag{30}$$

Matrix $\bar{A}$ is Hurwitz stable if $k_i > a_{ii}$, $i = 1, 2, \ldots, n$. Then (30) is asymptotically stable. □

3 The fractional-order unified chaotic system

Consider the fractional-order master system [54],

$$\begin{aligned} D^{\alpha}x_1 &= (25\beta + 10)(x_2 - x_1) \\ D^{\alpha}x_2 &= 28x_1 - x_2 + 29\beta x_2 - 35\beta x_1 - 10x_1x_3 \\ D^{\alpha}x_3 &= 10x_1x_2 - \frac{\beta + 8}{3}x_3 \end{aligned} \tag{31}$$

With $\beta = 1$ and $\alpha = 0.9$ system (31) exhibits chaotic behaviour, in the sense of Definition 1.

Definition 1 ([6]) *An attractor is said to be chaotic if it is bounded and all the trajectories starting form it are Lyapunov unstable. Then, a dynamic system is chaotic if it has at least one chaotic attractor.*

System (31) is simulated through the Matlab-Simulink diagram shown in Figure 1, where blocks F1, F2 and F3 correspond to the approximation of the integral operator with order $\alpha = 0.9$.

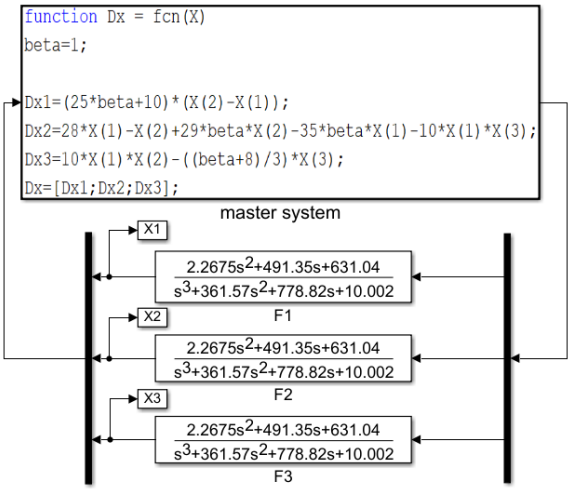


Figure 1: Simulation of system (31) using Matlab-Simulink

For simulation purposes, initial condition of block F1, was set to -0.001. Figure 2 shows the numerical simulation of system (31) via the phase portraits.

Now, in order to illustrate the behavior of system (31), we change the weight parameters (α, β). The simulation results are given in Figure 3 with parameters: (a) $\alpha = 0.9$, $\beta = 0$, (b) $\alpha = 0.9$, $\beta = -1$, (c) $\alpha = 0.8$, $\beta = 0.5$ and (d) $\alpha = 0.7$, $\beta = 1$. Clearly, system (31) is not chaotic with these set of parameters, where the integral operator with orders $\alpha = 0.8$ and $\alpha = 0.7$ is defined by [24],

$$\frac{1}{s^{0.8}} \approx \frac{5.235s^3 + 1453s^2 + 5306s + 254.9}{s^4 + 658.1s^3 + 5700s^2 + 658.2s + 1} \tag{32}$$

and

$$\frac{1}{s^{0.7}} \approx \frac{5.406s^4 + 177.6s^3 + 209.6s^2 + 9.197s + 0.01450}{s^5 + 88.12s^4 + 279.2s^3 + 33.30s^2 + 1.927s + 0.0002276} \tag{33}$$

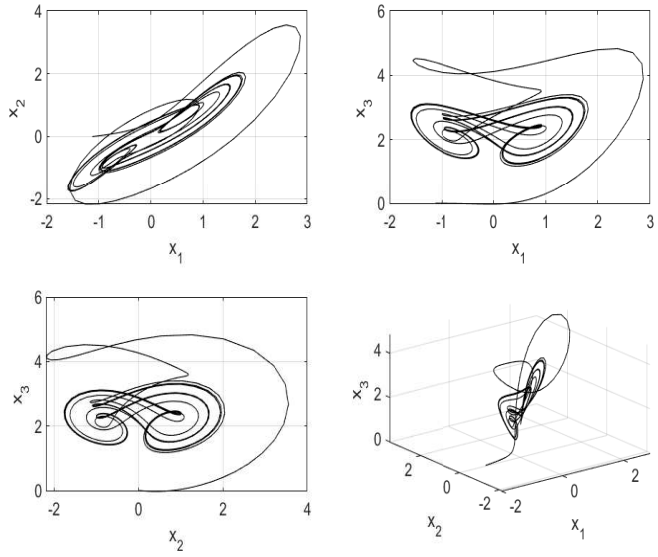


Figure 2: Simulation results of system (31), $\beta = 1$ and $\alpha = 0.9$.

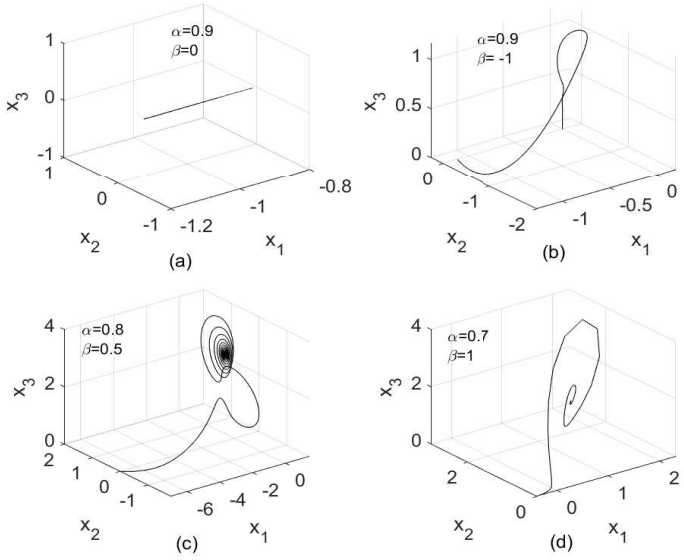


Figure 3: Phase portraits of system (31) with different values of (α, β)..

4 Synchronization results and discussion

The fractional-order slave system that synchronizes with system (31) is represented as,

$$\begin{aligned} D^{\alpha}z_1 &= (25\beta + 10)(z_2 - z_1) + u_1 \\ D^{\alpha}z_2 &= 28z_1 - z_2 + 29\beta z_2 - 35\beta z_1 - 10z_1z_3 + u_2 \\ D^{\alpha}z_3 &= 10z_1z_2 - \frac{\beta + 8}{3}z_3 + u_3 \end{aligned} \tag{34}$$

where $\beta = 1$, $\alpha = 0.9$ and u_1, u_2 and u_3 are the control signals.

It is necessary to calculate the control gains values. First, for the proportional control, we obtain the constants L_i over the bounded set $\Omega = \{x \in \mathbb{R}^n \mid |x_i| \leq M_i\}$,

$$L_i = \max_{1 \leq i \leq n} \|f_i'(x)\| \tag{35}$$

where $f_i'(x)$ is the Jacobian matrix.

Then, considering that $M_1 \approx 2$, $M_2 \approx 3$ and $M_3 \approx 4$, we have that $L_1 = 35$, $L_2 = 47$, and $L_3 = 30$.

According to Proposition 1, the maximum synchronization errors are defined as

$$\varepsilon_{i_{\max}} = \frac{L_i}{k_i} = 0.05 \tag{36}$$

In Equation (36) we consider the synchronization errors are bounded by 0.05, they are acceptable for complete practical synchronization. Equation (36) yields to $k_1 = 700$, $k_2 = 940$, and $k_3 = 600$.

For the nonlinear PI controller, the gain values are obtained in order to guarantee a settling time of 0.02s and an overshoot of 1% in the three coordinates, then $ki = 400$ and $K_{I_i} = 58613$, for $i = 1, 2, 3$. And for the active controller we have the following conditions,

$$k_1 > -25\beta + 10 \tag{37}$$

$$k_2 > 29\beta - 1 \tag{38}$$

$$k_3 > -\frac{\beta + 8}{3} \tag{39}$$

In this case, we choose $k_1 = 200$, $k_2 = 100$ and $k3 = 100$.

Figure 4 shows the phase portraits of the master and slave systems when the control signals are off, clearly all systems are unsynchronized due to the initial conditions are different. However, when the feedback controllers are activated the synchronization is achieved since the trajectories of the controlled slave system tend to follow the trajectories of the master system.

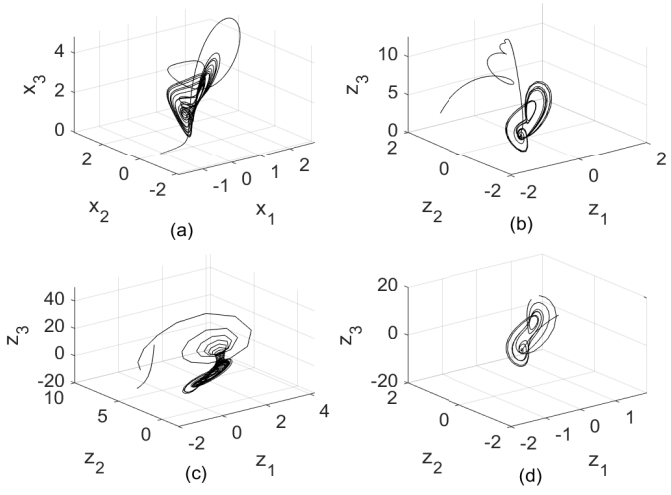


Figure 4: Unsynchronized systems.

In Figure 5 the synchronization performance is verified via the synchronization errors. We observe that the synchronization errors converge to zero when the feedback control is activated. As we can note, the corresponding slave trajectory reach almost immediately the master trajectory for the proportional control, also, when the nonlinear PI control acts, the corresponding trajectory has a higher oscillatory overshoot and a settling time of 0.02s, and for the active control, the synchronization errors converge exponentially to zero at 0.02 s.

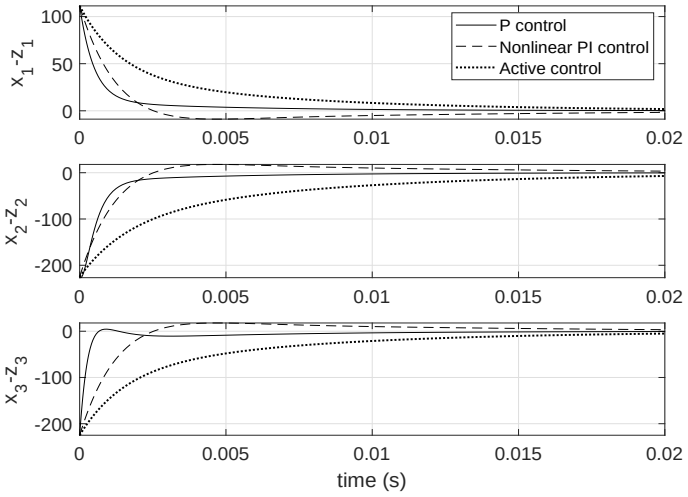


Figure 5: Synchronization errors.

The above-mentioned behaviors for the state variables under feedback controllers can be explained for the performance of the three controllers under comparison. Figure 6 is related to the control effort performance, it can be observed that the control effort for the proportional controller is very high due to the magnitude presented at the startup and transient condition. On the other hand, PI and active controllers show a smooth behavior, leading the trajectory to the controlled slave system to the required master trajectory in a larger settling time.

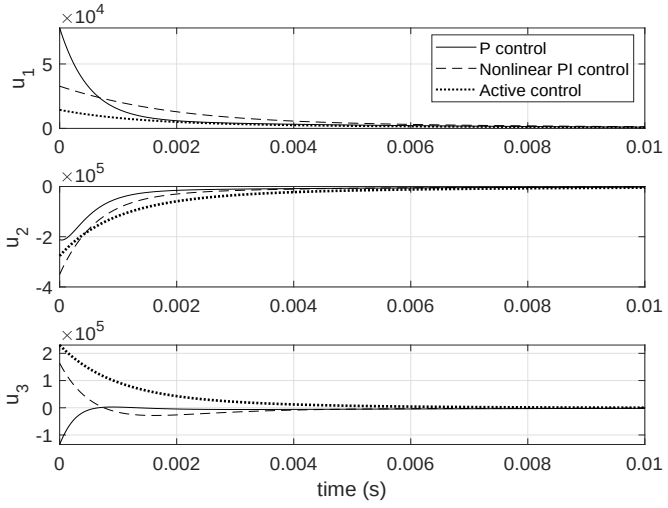


Figure 6: Control signals.

5 Equivalent electronic circuit

System (31) is translated to the electric circuit shown in Figure 7.

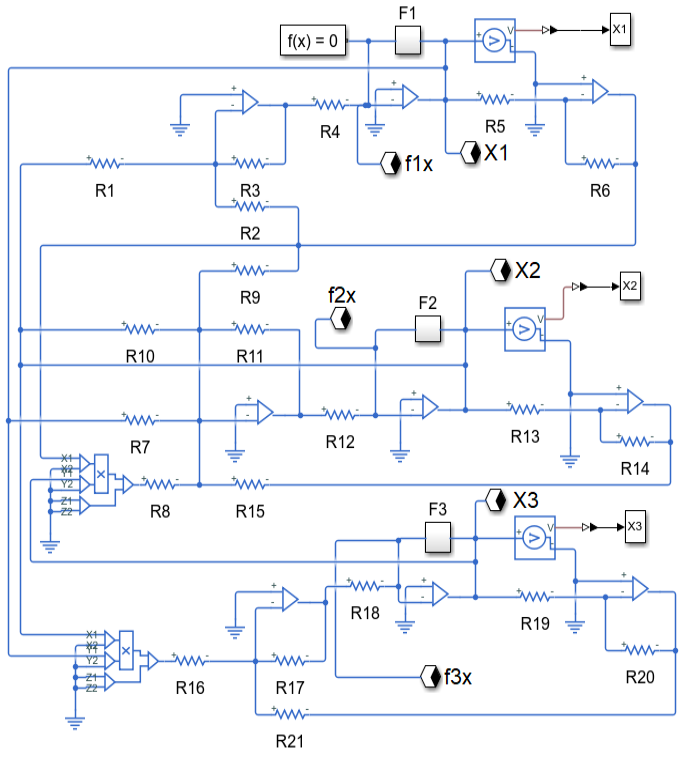


Figure 7: Electrical model of master system (31) using Simulink's Simscape library.

Component values are taken as: $R_4 = R_{13} = R_{18} = 100\,k\Omega$, $R_1 = R_5 = R_7 = R_8 = R_{10} = R_{14} = R_{16} = R_{20} = 10\,k\Omega$, $R_2 = R_3 = R_6 = R_{21} = 33\,k\Omega$, $R_9 = R_{11} = R_{15} = 27\,k\Omega$ and $R_{12} = R_{17} = R_{19} = 100\,k\Omega$.

Fractional integrators blocks F1, F2 and F3 are represented as $\frac{1}{s^{0.9}}$. Figure 8 shows the approximation of $\frac{1}{s^{0.9}}$ as an electric circuit using the Simscape libraries.

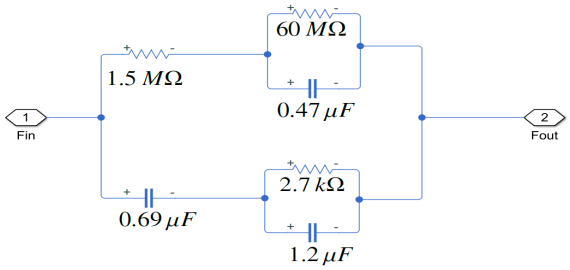


Figure 8: Integrator blocks F1, F2,..., F12.

6 Verification of the master-slave synchronization

The synchronization is verified via the equivalent electronic circuits of master and slave systems. The controlled slave electronic circuit is given in Figure 9, where u_1, u_2 and u_3 are the control inputs introduced in Section 2. In order to obtain $-z_1$, $-z_2$ and $-z_3$, in this circuit, for each signal z_1, z_2 and z_3 is added an inverting operational amplifier with unity gain, the value of R is equal to 10kΩ. To avoid the crossing of physical lines we employ connection labels. The label block makes connections between two blocks without using lines.

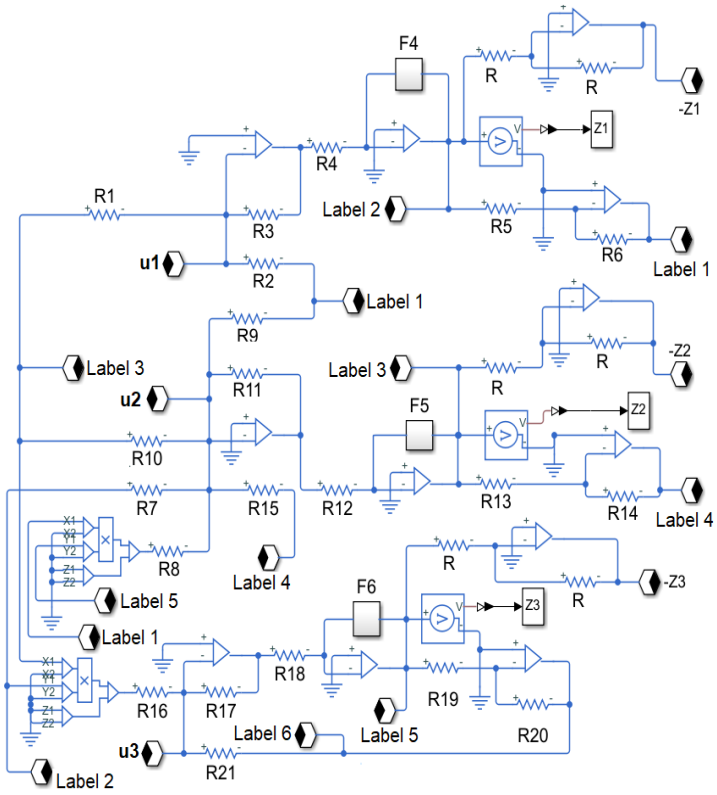


Figure 9: Electrical circuit schematic for the slave system (34).

Master circuit of Figure 7 and slave circuit of Figure 9 are coupled since the control inputs require information signals from both circuits.

The equivalent circuit for the proportional control (4) is shown in Figure 10. Resistors R_{P1}, R_{P2} and R_{P3} are related with the control gains of Section 4 as follows,

$$R_{P1} = 10\frac{R_3}{k_1} = 471\ \Omega \tag{40}$$

$$R_{P2} = 10\frac{R_{11}}{k_2} = 287\ \Omega \tag{41}$$

$$R_{P3} = 10\frac{R_{17}}{k_3} = 1.67\ \text{k}\Omega \tag{42}$$

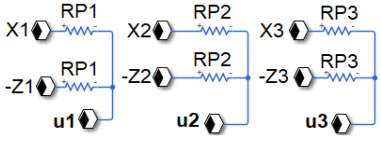


Figure 10: Proportional control.

The nonlinear fractional-order PI control (13)-(14) is proposed by means of the circuit of Figure 11.

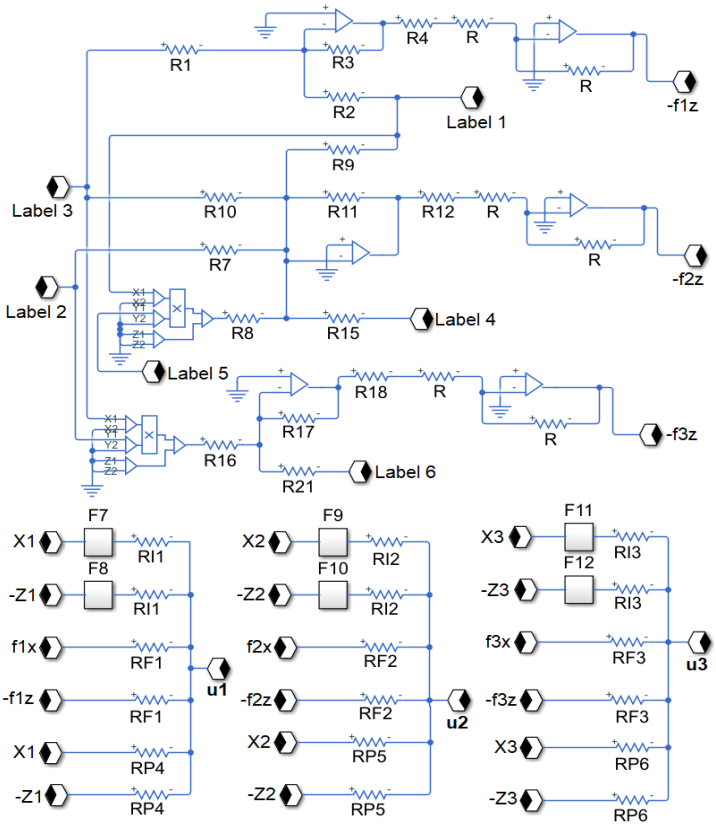


Figure 11: Nonlinear fractional-order PI control.

The parameter values are obtained as,

$$R_{P4} = 10\frac{R_3}{k_1} = 825\ \Omega \quad (43)$$

$$R_{P5} = 10\frac{R_{11}}{k_2} = 675\ \Omega \quad (44)$$

$$R_{P6} = 10\frac{R_{17}}{k_3} = 250\ \Omega \quad (45)$$

$$R_{I1} = 10\frac{R_3}{K_{I_1}} = 5.63\ \Omega \quad (46)$$

$$R_{I2} = 10\frac{R_{11}}{K_{I_2}} = 4.6\ \Omega \quad (47)$$

$$R_{I3} = 10\frac{R_{17}}{K_{I_3}} = 17\ \Omega \quad (48)$$

$$R_{F1} = 10\ R_3 = 330\ \text{k}\Omega \quad (49)$$

$$R_{F2} = 10\ R_{11} = 270\ \text{k}\Omega \quad (50)$$

$$R_{F3} = 10\ R_{17} = 1\ \text{M}\Omega \quad (51)$$

Finally, the equivalent circuit for the active control (27) is given by Figure 12. The active control (27) has been rewritten as,

$$u_i = F_i(x) - F_i(z) + k_i\varepsilon_i, \quad i = 1, 2, \ldots, n \quad (52)$$

where, $F_i(x) = \sum_{j=1,\ j\neq i}^{n} a_{ij}x_j + g_i(x)$ and $F_i(z) = \sum_{j=1,\ j\neq i}^{n} a_{ij}z_j + g_i(z)$.

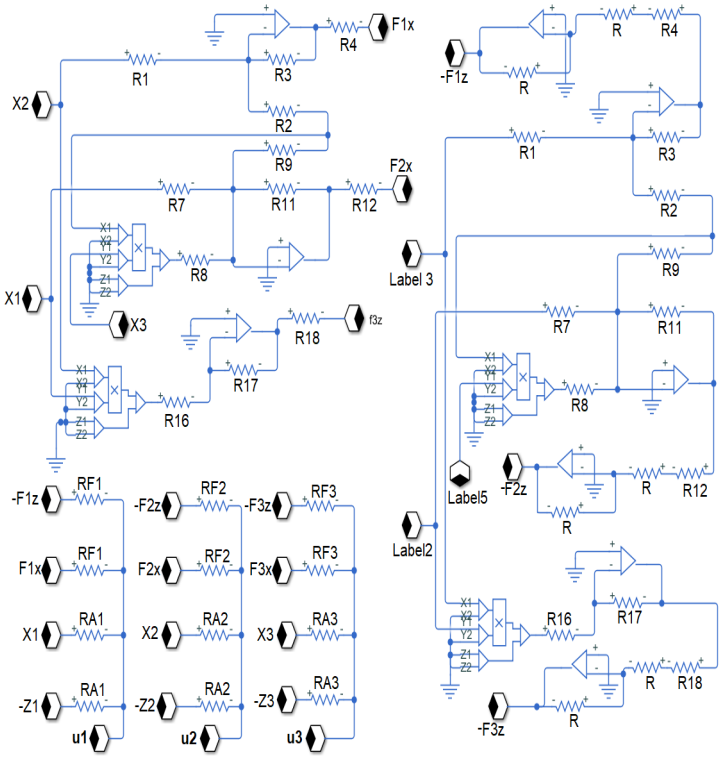


Figure 12: Active control.

The corresponding parameter values are: $R_{F1} = 330$ kΩ, $R_{F2} = 270$ kΩ, $R_{F3} = 10\ R_{17} = 1$ MΩ and,

$$R_{A1} = 10\frac{R_3}{k_1} = 1.65\ \text{k}\Omega \quad (53)$$

$$R_{A2} = 10\frac{R_{11}}{k_2} = 2.7\ \text{k}\Omega \quad (54)$$

$$R_{A3} = 10\frac{R_{17}}{k_3} = 10\ \text{k}\Omega \quad (55)$$

Figure 13 shows the phase portraits of the master-slave synchronization obtained by means of the equivalent electronic circuits given by Figures 7 and 9 along with the control inputs of Figures 10, 11 and 12. For all simulations, initial voltage on capacitor of $0.69\,\mu F$ takes several arbitrary values: 1 V (block F1), -2 V (block F4), 1 V (block F5), 1 V (block F6), while the rest of capacitors remain at 0 V. Notably, the whole phase portrait of the controlled slave system is synchronized with the corresponding phase portrait of the master system using the proposed controllers, for arbitrary initial conditions. However, an important characteristic is that the nonlinear fractional-order PI control and the active control lead to complex circuit structures.

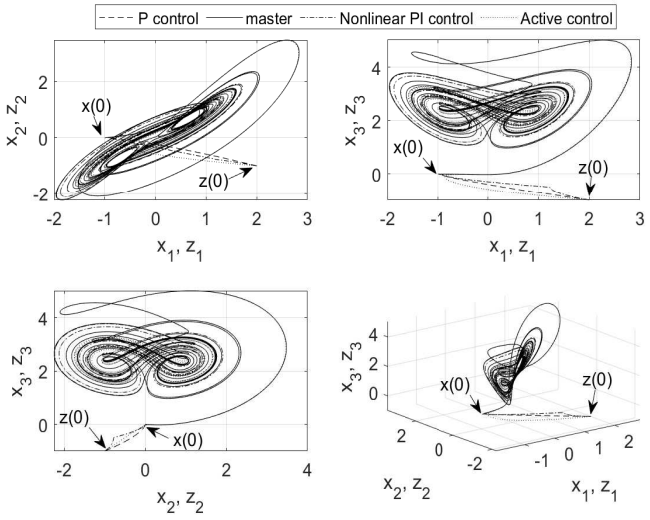


Figure 13: Master-slave synchronization.

7 Conclusions

This work studied the master-slave synchronization in a class of chaotic systems with fractional-order via feedback control. First, the feedback controllers were designed. In this sense, a proportional control was proposed, it guarantees that the synchronization error is bounded, as well as, a nonlinear fractional-order PI control and an active control were introduced, where the stability of the closed-loop system depends on the control gains, here for both controllers the synchronization error is asymptotically stable. Next, as application example, the fractional-order unified chaotic system was considered as master system. In fact, a set of numerical simulations based on the mathematical models of master and slave system showed that the synchronization is fulfilled by means of the proposed feedback controllers. It is important to mention that the proportional control exhibited the higher control effort, which could be a restriction in another class of systems, for instance, in those involving actuators. On other hand, in the transient condition, the active control is slower than the other ones. Finally, the synchronization was verified with success using the equivalent electronic circuits. As can be noticed, the proportional control has a simple circuit structure.

Conflict of interest The author declares no conflict of interest.

Acknowledgment This paper was supported by the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional (SIP-IPN) under the research grant SIP20211509.

References

[1] E. N. Lorenz, "Deterministic Nonperiodic Flow," Journal of the Atmospheric Sciences, **20**, 1963, doi:10.1175/1520-0469(1963)020⟨0130:dnf⟩2.0.co;2.

[2] T.-Y. Li, J. A. Yorke, "Period Three Implies Chaos," The American Mathematical Monthly, **82**, 1975, doi:10.2307/2318254.

[3] G. Li, B. Zhang, "A Novel Weak Signal Detection Method via Chaotic Synchronization Using Chua's Circuit," IEEE Transactions on Industrial Electronics, **64**, 2017, doi:10.1109/TIE.2016.2620103.

[4] M. H. Abd, F. R. Tahir, G. A. Al-Suhail, V. T. Pham, "An adaptive observer synchronization using chaotic time-delay system for secure communication," Nonlinear Dynamics, **90**, 2017, doi:10.1007/s11071-017-3825-2.

[5] H. Li, X. Liao, M. Luo, "A novel non-equilibrium fractional-order chaotic system and its complete synchronization by circuit implementation," Nonlinear Dynamics, **68**, 2012, doi:10.1007/s11071-011-0210-4.

[6] A. L. Fradkov, Cybernetical Physics, From Control of Chaos to Quantum Control, 2007.

[7] R. Femat, G. Solis-Perales, "Robust synchronization of chaotic systems via feedback," 2008, doi:10.1007/978-3-540-69307-9.

[8] L. M. Pecora, T. L. Carroll, "Synchronization in chaotic systems," Physical Review Letters, **64**, 1990, doi:10.1103/PhysRevLett.64.821.

[9] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, C. Zhou, "The synchronization of chaotic systems," Physics reports, **366**(1-2), 1–101, 2002.

[10] L. Huang, R. Feng, M. Wang, "Synchronization of chaotic systems via nonlinear control," Physics Letters, Section A: General, Atomic and Solid State Physics, **320**, 2004, doi:10.1016/j.physleta.2003.11.027.

[11] L. M. Pecora, T. L. Carroll, "Synchronization of chaotic systems," Chaos: An Interdisciplinary Journal of Nonlinear Science, **25**(9), 097611, 2015.

[12] Y. Zhao, W. Zhang, H. Su, J. Yang, "Observer-Based Synchronization of Chaotic Systems Satisfying Incremental Quadratic Constraints and Its Application in Secure Communication," IEEE Transactions on Systems, Man, and Cybernetics: Systems, **50**(12), 5221–5232, 2020, doi:10.1109/TSMC.2018.2868482.

[13] S. Mohammadpour, T. Binazadeh, "Observer-based synchronization of uncertain chaotic systems subject to input saturation," Transactions of the Institute of Measurement and Control, **40**, 2018, doi:10.1177/0142331217705435.

[14] A. Das, F. L. Lewis, "Distributed adaptive control for synchronization of unknown nonlinear networked systems," Automatica, **46**, 2010, doi:10.1016/j.automatica.2010.08.008.

[15] S. Wen, Z. Zeng, T. Huang, "Adaptive synchronization of memristor-based Chua's circuits," Physics Letters, Section A: General, Atomic and Solid State Physics, **376**, 2012, doi:10.1016/j.physleta.2012.08.021.

[16] B. Naderi, H. Kheiri, A. Heydari, R. Mahini, "Optimal Synchronization of Complex Chaotic T-Systems and Its Application in Secure Communication," Journal of Control, Automation and Electrical Systems, **27**, 2016, doi:10.1007/s40313-016-0245-3.

[17] W. D. Chang, "PID control for chaotic synchronization using particle swarm optimization," Chaos, Solitons and Fractals, **39**, 2009, doi:10.1016/j.chaos.2007.01.064.

[18] Z. Liu, "Design of nonlinear optimal control for chaotic synchronization of coupled stochastic neural networks via Hamilton–Jacobi–Bellman equation," Neural Networks, **99**, 2018, doi:10.1016/j.neunet.2018.01.003.

[19] C. Huang, J. Cao, "Active control strategy for synchronization and anti-synchronization of a fractional chaotic financial system," Physica A: Statistical Mechanics and its Applications, **473**, 2017, doi:10.1016/j.physa.2017.01.009.

[20] R. Martínez-Guerra, J. L. Mata-Machuca, "Fractional generalized synchronization in a class of nonlinear fractional order systems," Nonlinear Dynamics, **77**(4), 1237–1244, 2014.

[21] X. J. Wen, Z. M. Wu, J. G. Lu, "Stability analysis of a class of nonlinear fractional-order systems," IEEE Transactions on Circuits and Systems II: Express Briefs, **55**, 2008, doi:10.1109/TCSII.2008.2002571.

[22] Y. Li, Y. Q. Chen, I. Podlubny, "Mittag-Leffler stability of fractional order nonlinear dynamic systems," Automatica, **45**, 2009, doi:10.1016/j.automatica.2009.04.003.

[23] S. Liu, W. Jiang, X. Li, X. F. Zhou, "Lyapunov stability analysis of fractional nonlinear systems," Applied Mathematics Letters, **51**, 2016, doi:10.1016/j.aml.2015.06.018.

[24] T. T. Hartley, C. F. Lorenzo, H. K. Qammer, "Chaos in a Fractional Order Chua's System," IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, **42**, 1995, doi:10.1109/81.404062.

[25] C. Li, G. Peng, "Chaos in Chen's system with a fractional order," Chaos, Solitons and Fractals, **22**, 2004, doi:10.1016/j.chaos.2004.02.013.

[26] V. Daftardar-Gejji, S. Bhalekar, "Chaos in fractional ordered Liu system," Computers and Mathematics with Applications, **59**, 2010, doi:10.1016/j.camwa.2009.07.003.

[27] Z. Wang, X. Huang, G. Shi, "Analysis of nonlinear dynamics and chaos in a fractional order financial system with time delay," Computers and Mathematics with Applications, **62**, 2011, doi:10.1016/j.camwa.2011.04.057.

[28] H. G. Sun, Y. Zhang, D. Baleanu, W. Chen, Y. Q. Chen, "A new collection of real world applications of fractional calculus in science and engineering," Communications in Nonlinear Science and Numerical Simulation, **64**, 2018, doi:10.1016/j.cnsns.2018.04.019.

[29] J. A. T. MacHado, M. F. Silva, R. S. Barbosa, I. S. Jesus, C. M. Reis, M. G. Marcos, A. F. Galhano, "Some applications of fractional calculus in engineering," Mathematical Problems in Engineering, **2010**, 2010, doi:10.1155/2010/639801.

[30] J. Zarei, M. Tabatabaei, "Fractional order unknown input filter design for fault detection of discrete fractional order linear systems," Transactions of the Institute of Measurement and Control, **40**, 2018, doi:10.1177/0142331217746493.

[31] G. C. Wu, Z. G. Deng, D. Baleanu, D. Q. Zeng, "New variable-order fractional chaotic systems for fast image encryption," Chaos, **29**, 2019, doi:10.1063/1.5096645.

[32] I. Tejado, D. Valério, P. Pires, J. Martins, "Fractional order human arm dynamics with variability analyses," Mechatronics, **23**, 2013, doi:10.1016/j.mechatronics.2013.05.003.

[33] J. J. D. Espíndola, J. M. D. S. Neto, E. M. Lopes, "A generalised fractional derivative approach to viscoelastic material properties measurement," volume 164, 2005, doi:10.1016/j.amc.2004.06.099.

[34] J. Sabatier, O. P. Agrawal, J. A. T. Machado, Advances in fractional calculus: Theoretical developments and applications in physics and engineering, 2007, doi:10.1007/978-1-4020-6042-7.

[35] W.-C. Chen, "Nonlinear dynamics and chaos in a fractional-order financial system," Chaos, Solitons and Fractals, **36**(5), 1305–1314, 2008, doi:https://doi.org/10.1016/j.chaos.2006.07.051.

[36] S. Dadras, H. R. Momeni, "Control of a fractional-order economical system via sliding mode," Physica A: Statistical Mechanics and its Applications, **389**, 2010, doi:10.1016/j.physa.2010.02.025.

[37] D. Klatt, R. L. Magin, F. Mainardi, T. J. Royston, N. C. Perkins, I. Y. Shen, "Special section: Fractional calculus in vibration and acoustics," 2014, doi:10.1115/1.4027482.

[38] R. Cipin, C. Ondrusek, R. Huzlík, "Fractional-order model of DC motor," 2014, doi:10.1007/978-3-319-02294-9_46.

[39] R. L. Magin, "Fractional calculus in bioengineering," 2004, doi:10.1615/CritRevBiomedEng.v32.10.

[40] K. B. Oldham, "Fractional differential equations in electrochemistry," Advances in Engineering Software, **41**, 2010, doi:10.1016/j.advengsoft.2008.12.012.

[41] R. L. Magin, "Fractional calculus models of complex dynamics in biological tissues," Computers and Mathematics with Applications, **59**, 2010, doi:10.1016/j.camwa.2009.08.039.

[42] P. Ghamisi, M. S. Couceiro, J. A. Benediktsson, N. M. Ferreira, "An efficient method for segmentation of images based on fractional calculus and natural selection," Expert Systems with Applications, **39**, 2012, doi:10.1016/j.eswa.2012.04.078.

[43] H. Sheng, Y. Chen, T. Qiu, Fractional processes and fractional-order signal processing: techniques and applications, Springer Science & Business Media, 2011.

[44] S. Ghasemi, A. Tabesh, J. Askari-Marnani, "Application of fractional calculus theory to robust controller design for wind turbine generators," IEEE Transactions on Energy Conversion, **29**, 2014, doi:10.1109/TEC.2014.2321792.

[45] R. E. Gutiérrez-Carvajal, L. F. D. Melo, J. M. Rosário, J. A. MacHado, "Condition-based diagnosis of mechatronic systems using a fractional calculus approach," International Journal of Systems Science, **47**, 2016, doi:10.1080/00207721.2014.978833.

[46] L. B. G. Bozzini, "Fractional-order derivatives and their application to the position control of robots," International Journal of Mechanics and Control, 39, 2009.

[47] C. Li, G. Chen, "Chaos and hyperchaos in the fractional-order Rössler equations," Physica A: Statistical Mechanics and its Applications, **341**, 2004, doi:10.1016/j.physa.2004.04.113.

[48] W. H. Deng, C. P. Li, "Chaos synchronization of the fractional Lü system," Physica A: Statistical Mechanics and its Applications, **353**, 2005, doi:10.1016/j.physa.2005.01.021.

[49] C. Li, X. Liao, J. Yu, "Synchronization of fractional order chaotic systems," Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, **68**, 2003, doi:10.1103/PhysRevE.68.067203.

[50] R. Martínez-Martínez, J. Mata-Machuca, R. Martínez-Guerra, J. León, G. Fernández-Anaya, "Synchronization of nonlinear fractional order systems," Applied Mathematics and Computation, **218**, 2011, doi:10.1016/j.amc.2011.08.075.

[51] J. D. Gabano, T. Poinot, "Fractional modelling and identification of thermal systems," Signal Processing, **91**, 2011, doi:10.1016/j.sigpro.2010.02.005.

[52] M. F. Tolba, A. M. AbdelAty, N. S. Soliman, L. A. Said, A. H. Madian, A. T. Azar, A. G. Radwan, "FPGA implementation of two fractional order chaotic systems," AEU - International Journal of Electronics and Communications, **78**, 2017, doi:10.1016/j.aeue.2017.04.028.

[53] K. Rajagopal, A. Karthikeyan, A. K. Srinivasan, "FPGA implementation of novel fractional-order chaotic systems with two equilibriums and no equilibrium and its adaptive sliding mode synchronization," Nonlinear Dynamics, **87**, 2017, doi:10.1007/s11071-016-3189-z.

[54] X. R. Chen, C. X. Liu, F. Q. Wang, "Circuit realization of the fractional-order unified chaotic system," Chinese Physics B, **17**, 2008, doi:10.1088/1674-1056/17/5/022.

[55] L. G. Yuan, Q. G. Yang, "Parameter identification and synchronization of fractional-order chaotic systems," Communications in Nonlinear Science and Numerical Simulation, **17**, 2012, doi:10.1016/j.cnsns.2011.04.005.

[56] J. G. Lu, "Chaotic dynamics of the fractional-order Lü system and its synchronization," Physics Letters, Section A: General, Atomic and Solid State Physics, **354**, 2006, doi:10.1016/j.physleta.2006.01.068.

[57] P. Zhou, "Chaotic synchronization for a class of fractional-order chaotic systems," Chinese Physics, **16**, 2007, doi:10.1088/1009-1963/16/5/016.

[58] X. Wu, J. Li, G. Chen, "Chaos in the fractional order unified system and its synchronization," Journal of the Franklin Institute, **345**, 2008, doi:10.1016/j.jfranklin.2007.11.003.

[59] E. Demirci, N. Ozalp, "A method for solving differential equations of fractional order," Journal of Computational and Applied Mathematics, **236**, 2012, doi:10.1016/j.cam.2012.01.005.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 153-163 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Predicting Student Academic Performance Using Data Mining Techniques

Lonia Masangu[1], Ashwini Jadhav[2,*], Ritesh Ajoodha[3]

[1]*School of Computer Science and Applied Mathematics, DSI-NICIS National e-Science Postgraduate Teaching and Training Platform (NEPTTP), The University of the Witwatersrand, Johannesburg, 2106, South Africa*

[2]*Faculty of Science, The University of the Witwatersrand, Johannesburg, 2106, South Africa*

[3]*School of Computer Science and Applied Mathematics, The University of the Witwatersrand, Johannesburg, 2106, South Africa*

ARTICLE INFO

Article history:
Received: 31 August, 2020
Accepted: 23 December, 2020
Online: 15 January, 2021

Keywords:
predicting student academic performance
Data mining techniques
Machine learning algorithms
Learner attributes

ABSTRACT

There is a crisis in basic education during this pandemic which affected everyone worldwide, we see that teaching and learning have gone online which has effected student performance. Student's academic performance needs to be predicted to help an instructor identify struggling students more easily and giving teachers a proactive chance to come up with supplementary resources to learners to improve their chances of increasing their grades. Data is collected on KAGGLE and the study is focusing on student's engagement, how often they check their announcements, number of raised hands, number of accessed forum and number of accessed resources to predict student academic performance. Various machine learning models such as Support vector machine, Decision tree, Perceptron classifier, Logistic regression and Random forest classifier is used. From the results, it was proven that Support vector machine algorithm is most appropriate for predicting student academic performance. Support vector machine gives 70.8% prediction which is relatively higher than other algorithms.

1 Introduction

Student performance prediction is a crucial job due to the large volume of data in educational databases. A lot of data has become available describing student's online behaviour and student engagement [1]. Online data has been used by a great number of researchers. Online learning and teaching is making a significant impact on the fabric of basic education.

Predicting the performance of a student is a great concern to the basic education management. The scope of this paper is to identify the factors influencing the performance of students in different grades and to find out the best machine learning model to predicting student academic performance and helps to identify students with poor grades and then be evaluated and provided with new materials and methods to improve their results.

There is a crisis in basic education during this pandemic which affected everyone worldwide, we see that teaching and learning have gone online which has effected student performance/grades. Parents are now supposed to sign whether their children must attend traditional class or online classes due to this pandemic. Online learning platforms is used to track student performance, provides unlimited access to online materials, organises online content in one location though many things that need to be learned do not add themselves to online learning. Many courses require the acquisition of soft skills, which cannot be easily learned or tested online.

They have been many attempts on predicting student academic performance using data collected from students activity logs and student demographic information. They have used several of data mining techniques. Decision trees were primarily used to find the predictor variables to the predicted variable and shows the targeted discrete value [2, 3]. Logistic Regression- was used to describe data and describe the relationship among the variables and Random forest was primarily used to maintain accuracy between the variable [4]–[5].

This research aim to predict student performance from data collected on student's activity logs on an online learning and student's demographic information. We want to find the accuracy of some classification models for predicting student academic performance. The major objectives of this study are:

*Corresponding Author: Ashwini Jadhav, The University of the Witwatersrand, Johannesburg, South Africa, Ashwini.Jadhav@wits.ac.za

- Finding the best classification model on student data set.
- Predict student academic performance using student's demographically data and student's online logging's data.

We would like to find out, does the online engagement have an influence on physical engagement? Does the parent association with the learners improve the learner's engagement?.

This study will be using various Machine Learning methods to predict student's academic performance. Five popular classification methods (Support vector machines, Decision trees and Perceptron classification, Logistic regression and Random forest) are built and compared to each other making use of their predictive accuracy on the given data samples to predict student academic performance. Data is collected on KAGGLE and we will be focusing on student's engagement, how often they check their announcements, number of raised hands, number of accessed forum and number of accessed resources to predict student success. The process of predicting student performance using online logging's data is performed using various data mining techniques. [6, 7].

This study proposes the following assumptions:

- Every student must have undergone some training on how to utilise online learning platform to predict correct results.
- Every student have an equal chance of receiving all course materials.
- If a student fails to submit an assessment or attend any forums it assumes the student has obtained 0% in that particular assessment.

2 Related Work

This chapter provides an overview of previous research on predicting student academic performance. This section illustrates other similar work-related to predicting student academic performance.

2.1 Data

Several researchers have used machine learning algorithms in predicting student's academic performance. 323 students doing undergraduate who enrol in dynamics in 4 consecutive semesters: in semester one there were 128 students, in semester two there were 58 students, in semester three there were 53 students and semester four there were 84 students [8]. Other authors have used undergraduate student's information from four different sources which included student's demographic data, course resources data, forums data and Sakai's grade book data [4].

Other authors have used electronic data from the universities to improve the quality of educational processes. They studied using machine learning techniques to study the performance of undergraduate students [9]. We also learnt that other researchers have conducted their study about the student academic performance in high schools and in Iran for bachelor degree students. They analysed data of 1100 learners where 500 where the high school students and 600 was that of the bachelor degree level [10].

2.2 Features

Several researchers have used student's demographic information and student online loggings data to predict student academic performance. The authors have done predictions for 4 class, 3 class and 2 class. The authors applied oversampling in other to balance the dataset and also used 3 different machine learning classifiers to predict student academic performance [8].

Several researchers collected data from student materials, assignments, tests, forums, and discussions on the online learning platform database to predict student success. Some researchers predicted student academic performance and they collected data from total login frequency on the system and a total number of attended forums [4].

2.3 Model

Various machine learning methods have been used by several researchers to predict student academic performance using student's demographic information and student online logging's data. Two machine learning algorithms were used which are Bayesian networks and Decision trees. The Decision trees was more accurate than Bayesian network by 3-12% [8].

Other authors have used C4.5 decision trees, Naïve bayes classifier, Random forests, neural networks and Meta-classifiers. After the analysis they found that C4.5 decision trees performed better [9]. The analysis shows that parent educational level, past examination results and gender has an impact on the presentation [11].

They also used decision trees and clustering data mining algorithms. They focused upon two type of student performance. They first predicted student academic performance at the end of the study programme [12]. Later predicted the student progressions and add them with the predicted results. The results shows good and bad performance of students and assists with identifying struggling students and giving teachers a proactive chance to come up with supplementary measures to improve their chances of passing. The decision trees gave out the best results [4].

3 Research Contribution

This study contributes to the current body on literature by predicting student academic performance and helps to identify students with poor grades can then be evaluated and provided with new materials and methods to improve their grades. Predicting students performance allow an instructor to spot non-engagement students based on their actions and activities from online learning platform. It also assists with identifying struggling students and giving teachers a proactive chance to come up with supplementary measures to improve their chances of passing during the course of their study programme.

From the literature that was reviewed various machine learning methods have been used by several researchers to predict student academic performance, they predicted student academic performance at the end of the study programme and they were unable able to detect which students may need immediate attention so that they lower the chances of them failing.

Table 1: Related Works.

Authors	Features	Models	Accuracy
Shaobo Huang and Ning Fang in (2012)	323 students doing undergraduate who enrol in dynamics in 4 consecutive semesters	Decision trees and Bayesian networks	The decision trees consistently 3 - 12% more accurate than the Bayesian network.
M. Lauria et al (2012)	Used undergraduate student's data	Logistic regression, support vector machines and C4.5 decision trees	Support vector machines and the logistic regression gave out higher
Oskouei Askari et al (2014)	Total number of 1100 and 500 where the high school students and 600 was that of the bachelor degree level	C4.5 decision trees, Naïve bayes classifier, Random forests, neural networks and Meta-classifiers	The best prediction was found using the C4.5 decision trees.
Raheela Asif et al (2017)	Used electronic data from the universities to improve the quality of educational processes.	Decision trees and clustering.	The decision trees gave out the best results.

4 Research Methodology

This chapter presents the methods that is proposed in this study to predict performance of students using student's demographical information and student's online logging's data.

4.1 Research design

Based on the literature the study proposed using data mining techniques to find whether the total spent time on online learning affect the student's performance and how have learning management system features affect the student performance.

4.2 Data description:

The data set consists a total of 480 records and 16 features. We found that there are 3 major categories of features: first being student demographic information such as student nationality and student gender. Secondly being student academic information such as student grade, student section and student educational stage. Thirdly being student behavioural information such as number of access resources, number of raised hands in class and school satisfaction.

The data set consists a total of 305 males and 175 females where 179 are from Kuwait, 28 are from Palestine, 172 are from Jordan, 22 are from Iraq, 17 are from Lebanon, 12 are from Tunis, 11 are from Saudi Arabia, 9 are from Egypt, 6 are from USA, 7 are from Syria, 6 are from Iran, 6 are from Libya, 4 are from Morocco and 1 is from Venezuela.

The data set is collected in two academic semesters, on first semester they were 245 student records and 235 student records on semester two.

The data set also covers the features of the school attendance days. There are two categories based on their number of absence days. We found that there were 289 students their absence days were under 7 days and 191 students were absent for more than 7 days.

Finally, the data set also contains parent's participation on their children academic process. There are two categories: first being the parent answering survey and secondly being parent school satisfaction. We found that 270 parents managed to answer the survey and 210 parents did not answer the survey. We also found that 292 parents are happy and satisfied with the school and 188 parents are not satisfied.

4.3 Classification field

There are three numerical intervals of student grades. The first interval is for students who obtained a failing percentage (L), the interval includes values from 0% to 69%. The second interval is for students who obtained low passing percentage (M), the interval includes values from 70% to 89%. Lastly is for students who got high passing percentage (H), the interval includes values from 90% to 100%.

4.4 Methods

Predictive Models: Five popular classification methods (Decision trees and Perceptron classification and Support vector machines, Logistic regression and Random forest) are built and compared to each other making use of their predictive accuracy on the given data samples. Brief description of the predictive models that will be used in this study.

Support Vector Machines: Support vector machines (SVMs) [13] helps in detecting the outliers on the data set and it also perform classification. SVMs are set of supervised learning methods. SVMs used kernel trick to modify data and use the modified data to find the difference between the possible end results. SVM finds the optimal solution by computing on each feature by using partial differentiation after employing the Lagrange multiplier [14]. The model decreases convolution of the training data consequential subset of support vectors.

Table 2: This table depicts the student related variables and their description.

Data Field	Description
Gender	The student's gender.
Nationality	The student's nationality
Place of Birth	The student's country of birth.
Stage ID	Educational level student belongs to (Elementary, Middle, or High School
Grade ID	The grade year of the student.
Section ID	The classroom the student is in.
Topic	The topic of the course.
Semester	The semester of the school year. (F for Fall, S for Spring)
Relation	The parent responsible for student.
raised hands	How many times the student raises his/her hand on classroom
Visited Resources	How many times the student visits a course content
Announcements View	How many times the student checks the new announcements
Discussion	How many times the student participate on discussion groups
Parent Answering Survey	Parent answered the surveys which are provided from school or not
Parent School Satisfaction	Whether or not the parents were satisfied. "Good" or "Bad". Oddly this was not null for parents who did not answer the survey. It is unclear how this value was filled in.
Student Absence Days	Whether or not a student was absent for more than 7 days
Class	Our classification field. 'L' is for students who got a failing percentage (Less than 69%), 'M' for students who got a low passing grade (Between 70% and 89%), and 'H' for students who achieved high marks in their course (90% to 100%)

Given a data set containing a training set of N data points, $\{x_k, y_k\}_{k=1}^{N}$ and input data, which is an n-dimensional data vector ($x_k \epsilon R\hat{} N$) and output, which is the one-dimensional vector space ($y_k \epsilon r$); SVM create the classifier as shown below in this equation:

$$y(x) = sign\Big[\sum_{k=1}^{N} \alpha_k y_k \Psi(x, x_k) + b\Big]$$

where α_k are positive real constants and b is a real constant.

Decision trees (DTs): In this study Decision trees is used to find the predictor variables to the predicted variable and shows the targeted discrete value. "Decision trees uses variable values to create a structure that has nodes and edges" [15]. A DT has internal nodes and leaves, rectangles represents nodes and ovals represents leaves. Data set features are represented by the internal node and it contain two or more child. The value of these features is found at the branches. Each leaf contains a classification label [16].

Decision trees are established from a training set. A tree is called the hierarchy and a node is called segment. The entire data set is contained at the original segment called the node of the tree. The branches are formed by the node with its successors that created it and the leaves are final nodes. The decision is made on each leaf and it is applied to all the observations in the leaf. The decision is the predicted value.

Perceptron classification: The Perceptron classifier [17] is a set of supervised learning, the classification field of a sample can be predicted using Perceptron classifier. Perceptron classifier accept numerous input and if the number of inputs is more than the specified condition, it does not return the output, it output the massage for corrections. In the Perceptron algorithm features on the data set are taken as inputs and it is represented by x1, x2, x3, x4,...,xn where features value is indicated by x and the total occurrences is represented by n.

The required features to be trained is stored as input in the first layer. Now the total inputs and weights will be multiplied and add their outcome. The weights are the values obtained through the training of the model and are denoted by w1, w2, w3,..,wn. The output function will be shifted by the bias value and this value will later be presented to the activation function then the output value is obtained after receiving the value on the last step.

Logistic regression (LR): Logistic regression describe the association among variables and it was used to predict student academic performance by estimating the probability of an event occurring [18]. It also shows the probability of two categories by fitting the explanatory variables and log odds to model using this equation.

$$log(\frac{(P(Y=1|X))}{1-P(Y=1|X)}) = \beta_0 + \beta_1 X_1 +, ..., \beta_N X_N$$

where Y=(0,1) is the binary variable; 1 if it is higher than the reference level and 0 if not, $X = (X_1, ..., X_n)$ are n explanatory variables and $\beta = (\beta_0, ...\beta_n)$ are the estimated regression coefficient.

Random forest (RF): Random forest uses begging method to generate trees in which its prediction is more accurate than that of any individual tree [19]. Random forest was also used to avoid over fitting on the training set and limiting errors due to bias hence yield accurate and useful results. RF can handle outliers and noise in the data and gains high classification accuracy. RF generate numerous decision trees in the training phase and output class labels [20]. RF

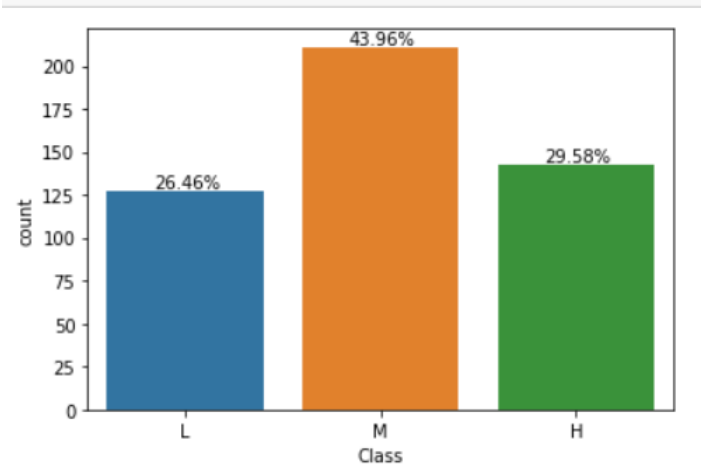


Figure 1: This figure depicts the proportion of student academic performance of the classification field.

is used in this study since it is permissible to less over-fitting and it has proofed to be good classification results previously.

RF is a theoretical framework grounded on mixture of decision trees; $\{T_1X, ..., T_BX\}$. The ensemble produces B outputs $\{\check{Y}_1 = T_1(X), ..., \check{Y}_B = T_B(X)$, where $\check{Y}_b, b = 1, ..., B$ is the predicted grades by the bth tree. Output of all trees are aggregated to produce one final prediction $\check{Y}$, which is the class predicted by majority of trees.

4.5 Analysis

The data set consist a total of 480 records. Machine learning algorithms are applied to predict student academic performance and we found that Support vector machine algorithm is best suited to predict student academic performance. We achieve a total accuracy of 70.8% which shows the potential accuracy of Support Vector Machine technique, followed by Random forest with 69.7% accuracy, Logistic regression with 67.7% accuracy, Perceptron with 64.5% and lastly Decision tree with 46.8%.

5 Results and Discussion

5.1 Pre-processing

Typically in machine learning before processing and running a test on a data set, it is necessary to prepare the data and select the targeted attribute. Selecting attributes requires putting all the matching combination of attributes in the data set in order to find which combination is suitable in predicting student academic performance.

Our goal with pre-processing was to change our numerical fields that have a value like Grade ID to a numerical only value in a way that we preserve that distance in a meaningful way. We also assign our three classes to numerical outcomes with a preserved distance and set setting L = -1, M = 0, and H = 1. We chose to preserve the distance between the categorical values and scale our numerical fields so that they would be more meaningful when compared together.

The five machine learning models are used to evaluate the student's academic performance and to check which model best predict student performance.

5.2 Data visualisation

The data set consist a total of 480 records. In this study, the purpose of selecting an attributes was to find the attributes that contain numerical data, attributes that contain categorical data and classification label. Our goal with data visualisation is to get an idea of the shape of the data set and to see if we can easily identify any possible outliers and also look to see if any of data is unclear or redundant.

5.3 Summary of results

In the previous section, student academic performance was discussed and how it will be evaluated. In this section we discuss the performance of five machine learning models that was specified in section 4. First of all, we performed data visualisation after performing data pre-processing by generating simple plots of data distributions to get an idea of the shape of the data set and then 5 main machine learning techniques was evaluated and also describe the variables contained on the data set. The evaluation was done on the full data set that consist of 480 features [21].

From the results, it was proven that Support vector machine algorithm is most suitable in predicting the performance of students. SVM is relative higher than other algorithms and it has 70.8% prediction accuracy, followed by Random forest with 69.7% accuracy, Logistic regression with 67.7% accuracy, Perceptron with 64.5% and lastly Decision tree with 46.8%. We found the overall percentage of passing rate using class variable. There were 26.46% students who got a failing percent (less than 69%), 43.96% students who got a low passing grades (between 70% and 89%) and 29.58% who achieved high marks in their course (90% to 100%).

Student absence days seems to have a strong correlation with class variable. Very few students who missed more than 7 days managed to achieve high marks and very few students who missed less than 7 days failed their course. From grade 2 to grade 12 we found that grade 5, 9, and grade 10 have very few counts. No 5th grade students pass and no 9th grade students achieve high marks.

In figure 4, the bar plot shows the accuracy of five popular

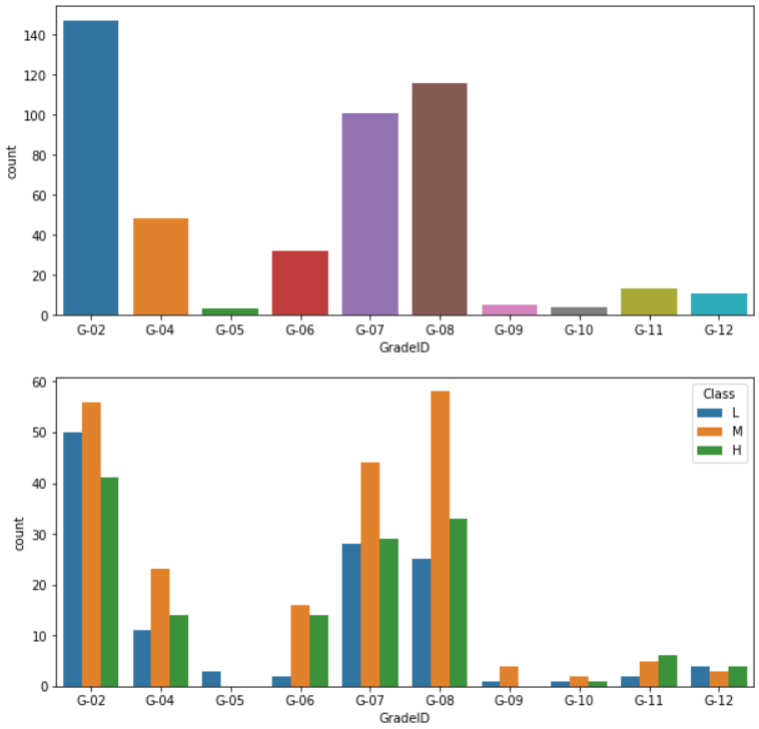


Figure 2: This figure depicts the grade the student is in and it also shows the average student performance for each grade.

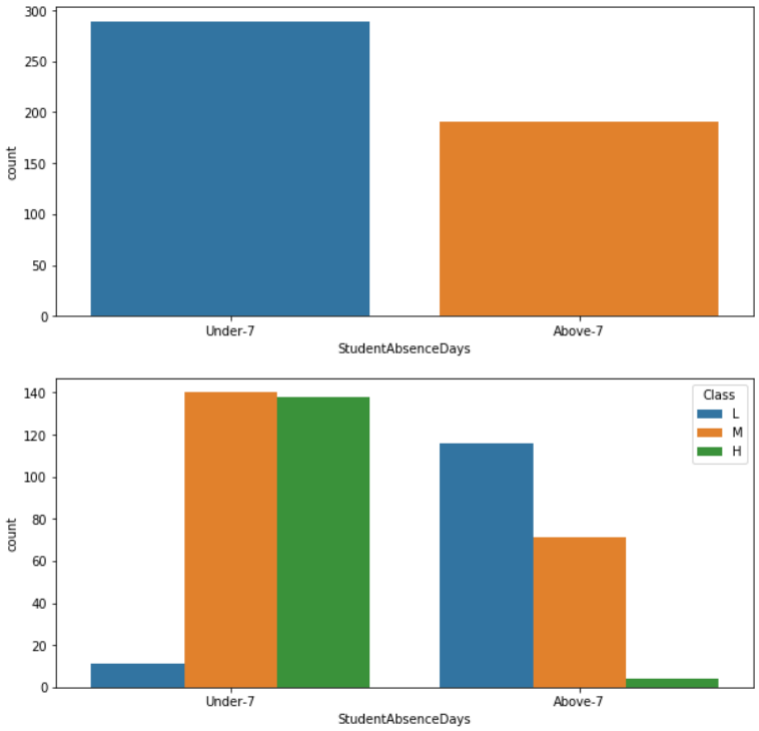


Figure 3: This figure depicts the proportion of student absence days and it also shows whether or not a student was absent for more than 7 days.

Table 3: This table shows the confusion matrices that describe the performance of the classification models.

Support vector machine			
Actual Values	**Predicted Values**		
	High	Low	Medium
High	23	1	11
Low	0	19	3
Medium	5	8	26
Decision tree			
Actual Values	**Predicted Values**		
	High	Low	Medium
High)	20	1	14
Low	1	15	6
Medium	7	8	24
Perceptron			
Actual Values	**Predicted Values**		
	High	Low	Medium
High	8	4	25
Low	0	21	1
Medium	2	19	18
logistic regression			
Actual Values	**Predicted Values**		
	High	Low	Medium
High	8	4	25
Low	0	21	1
Medium	2	19	18
Random Forest			
Actual Values	**Predicted Values**		
	High	Low	Medium
High	22	1	12
Low	0	16	6
Medium	4	6	29

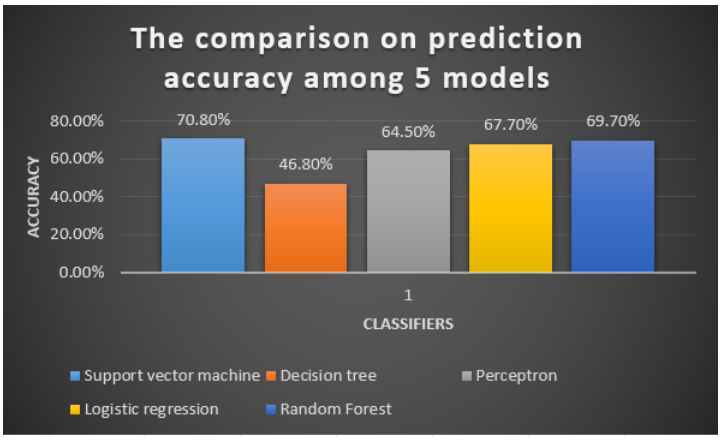


Figure 4: The comparison on prediction accuracy among 5 models.

Table 4: This table depicts the machine learning algorithms used in this study and their performance accuracy.

Classifier	Accuracy
Support vector machine	70.8%
Decision tree	46.8 %
Perceptron	64.5%
Logistic regression	67.7%
Random forest	69.7%

Table 5: This table depicts detailed accuracy of Logistic regression.

Features	Weighted Average
Correctly classified Instances	73.75%
Incorrectly classified Instances	26.25%
Mean absolute error	21.33%
Root mean squared error	37.53%
Relative absolute error	49.27%
Root relative squared error	80.6%
Precision	73.8%
Recall	73.8%
F-Measure	73.8%
Roc Area	83.8%
Total number of Instances	480

Table 6: This table depicts detailed accuracy of Perceptron.

Features	Weighted Average
Correctly classified Instances	79.37%
Incorrectly classified Instances	20.62%
Mean absolute error	14.88%
Root mean squared error	34.59%
Relative absolute error	34.36%
Root relative squared error	74.35%
Precision	79.3%
Recall	79.4%
F-Measure	79.3%
Roc Area	89.3%
Total number of Instances	480

Table 7: This table depicts detailed accuracy of Random Forest.

Features	Weighted Average
Correctly classified Instances	76.66%
Incorrectly classified Instances	23.33%
Mean absolute error	24.28%
Root mean squared error	33.37%
Relative absolute error	56.09%
Root relative squared error	71.72%
Precision	76.6%
Recall	76.7%
F-Measure	76.6%
Roc Area	89.7%
Total number of Instances	480

Table 8: This table depicts detailed accuracy of Decision tree.

Features	Weighted Average
Correctly classified Instances	72.70%
Incorrectly classified Instances	27.29%
Mean absolute error	29.53%
Root mean squared error	37.18%
Relative absolute error	68.19%
Root relative squared error	79.92%
Precision	72.8%
Recall	72.7%
F-Measure	72.7%
Roc Area	84.2%
Total number of Instances	480

Table 9: This table depicts detailed accuracy of Support Vector Machine.

Features	Weighted Average
Correctly classified Instances	78.75%
Incorrectly classified Instances	21.25%
Mean absolute error	27.22%
Root mean squared error	35.22%
Relative absolute error	62.87%
Root relative squared error	75.71%
Precision	78.8%
Recall	78.8%
F-Measure	78.7%
Roc Area	86%
Total number of Instances	480

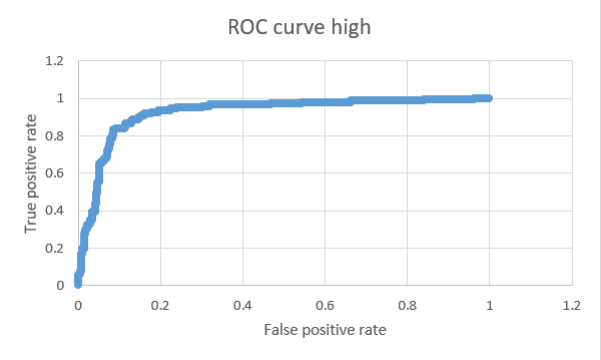


Figure 5: Receiver operating characteristic curve (class high).

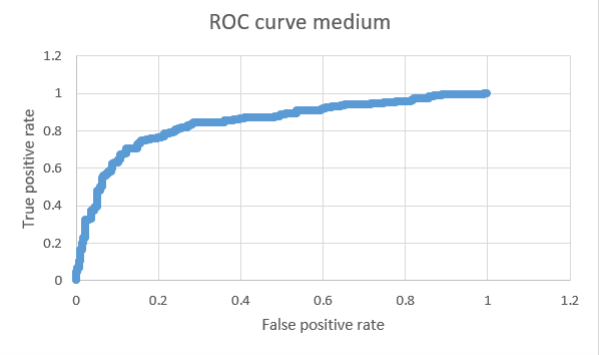


Figure 6: Receiver operating characteristic curve (class medium).

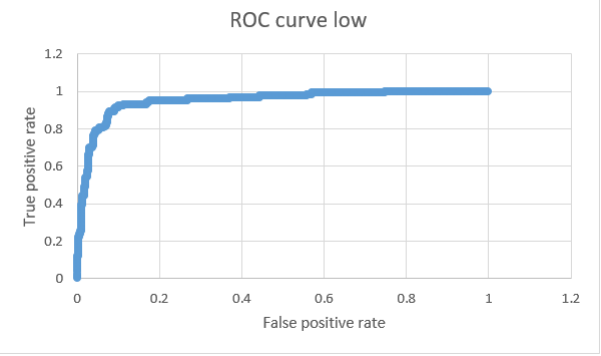


Figure 7: Receiver operating characteristic curve (class low).

machine learning models that is used to evaluate the student performance. The legend of the plot indicate that blue colour is Support vector machine, Yellow colour indicate Logistic regression, orange colour indicate Decision tree, light blue indicate Random forest and grey colour indicate Perceptron classifier.

Support vector machine has performed well when compared to other machine models. 78.75% instances was correctly classified and 21.25% instances was incorrectly classified. Another way of representing accuracy of the machine learning models that are in figure 4 is through confusion matrices. Table 3 describe how each algorithm has performed.

The graphical representation of receiver operating characteristic curve shows the performance of our best classification model, Support vector machine at all classification thresholds [22]. The ROC curve figure 5, figure 6 and figure 7 shows the classification performance of our best classifiers which is Support vector machine. There are three numerical intervals of student grades (low, medium and high). As can be seen from figure 5, figure 6 and figure 7. The ROC curve occupies the upper left corner, which means the classifier (SVM) used in this paper indicate a better performance and the prediction of positive value is specific in some degree, with AUC of 86%.

6 Conclusion and Future work

Companies and educational institutions uses learning management systems to create and manage lessons, courses, quizzes and other training materials [23]. Student's success needs to be predicted to help an instructor identify academic performance and helps with identifying struggling students more easily and giving teachers a proactive chance to come up with supplementary resources to learners to improve their chances of increasing their grades. It may be difficult for students to learn virtually than in a traditional class hence the student's performance vary due to difference methods of delivering the course materials [24].

Various machine learning models were used to predict student success using the learning management system. Each model indicate different percentage of accuracy when is tested with different features that are associated in an online learning platforms [25].

Student's performance was evaluated by five machine learning techniques which is Perceptron classifier, Support vector machine and Decision trees, Logistic regression and Random forest. Support vector machine ends up handling the data the best with 70.8% accuracy. The obtained results shows that the student absence days influence student academic performance on the other hand student class grades is not influencing student academic performance.

We aim to extend the study by collecting more additional features such as encouraging and motivational strategies taken by facilitators and teachers and considering more materials available for students in an E-learning platforms. We will also consider features such as psychological factors available which affect student's performance. We also intent to use more interesting and detailed data set to predict student academic performance in our future studies.

7 Acknowledgement

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant number: 121835).

References

[1] M. D. Dixson, "Measuring student engagement in the online course: the Online Student Engagement scale (OSE).(Section II: Faculty Attitudes and Student Engagement)(Report)," Online Learning Journal (OLJ), **19**(4), 143, 2015, doi: 10.3102/00346543074001059.

[2] S. Huang, N. Fang, "Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models," Computers and Education, **61**(1), 133–145, 2013, doi: 10.1016/j.compedu.2012.08.015.

[3] R. Ajoodha, R. Klein, B. Rosman, "Single-labelled music genre classification using content-based features," in 2015 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), 66–71, IEEE, 2015.

[4] E. J. Lauría, J. D. Baron, M. Devireddy, V. Sundararaju, S. M. Jayaprakash, "Mining academic data to improve college student retention: An open source perspective," in Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, 139–142, 2012.

[5] R. Ajoodha, B. S. Rosman, "Learning the influence structure between partially observed stochastic processes using iot sensor data," AAAI, 2018.

[6] K. F. Li, D. Rusk, F. Song, "Predicting student academic performance," in 2013 Seventh International Conference on Complex, Intelligent, and Software Intensive Systems, 27–33, IEEE, 2013.

[7] R. Ajoodha, A. Jadhav, S. Dukhan, "Forecasting learner attrition for student success at a south african university," in Conference of the South African Institute of Computer Scientists and Information Technologists 2020, 19–28, 2020.

[8] S. Huang, N. Fang, "Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models," Computers & Education, **61**, 133–145, 2013, doi:https://doi.org/10.1016/j.compedu.2012.08.015.

[9] R. J. Oskouei, M. Askari, "Predicting academic performance with applying data mining techniques (generalizing the results of two different case studies)," Computer Engineering and Applications Journal, **3**(2), 79–88, 2014.

[10] R. Asif, A. Merceron, S. A. Ali, N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," Computers & Education, **113**, 177–194, 2017.

[11] Y. Gerhana, I. Fallah, W. Zulfikar, D. Maylawati, M. Ramdhani, "Comparison of naive Bayes classifier and C4. 5 algorithms in predicting student study period," in Journal of Physics: Conference Series, volume 1280, 022022, IOP Publishing, 2019.

[12] Y. Salal, S. Abdullaev, M. Kumar, "Educational Data Mining: Student Performance Prediction in Academic," IJ of Engineering and Advanced Tech, **8**(4C), 54–59, 2019, doi:10.1016/6895.125478.

[13] S. Hua, Z. Sun, "Support vector machine approach for protein subcellular localization prediction," Bioinformatics, **17**(8), 721–728, 2001.

[14] I. Burman, S. Som, "Predicting students academic performance using support vector machine," in 2019 Amity International Conference on Artificial Intelligence (AICAI), 756–759, IEEE, 2019.

[15] F. Esposito, D. Malerba, G. Semeraro, J. Kay, "A comparative analysis of methods for pruning decision trees," IEEE transactions on pattern analysis and machine intelligence, **19**(5), 476–491, 1997.

[16] F. Jauhari, A. A. Supianto, "Building student's performance decision tree classifier using boosting algorithm," Indonesian Journal of Electrical Engineering and Computer Science, **14**(3), 1298–1304, 2019, doi:10.1016/j.compedu.2012.08.015.

[17] B. Chaudhuri, U. Bhattacharya, "Efficient training and improved performance of multilayer perceptron in pattern classification," Neurocomputing, **34**(1-4), 11–27, 2000.

[18] D. W. Hosmer Jr, S. Lemeshow, R. X. Sturdivant, Applied logistic regression, volume 398, John Wiley & Sons, 2013.

[19] A. Liaw, M. Wiener, et al., "Classification and regression by randomForest," R news, **2**(3), 18–22, 2002.

[20] K. Deepika, N. Sathyanarayana, "Relief-F and Budget Tree Random Forest Based Feature Selection for Student Academic Performance Prediction," International Journal of Intelligent Engineering and Systems, **12**(1), 30–39, 2019, doi:10.1016/j.5896408.015.

[21] C. Sievert, Interactive Web-Based Data Visualization with R, plotly, and shiny, CRC Press, 2020.

[22] R. Kannan, V. Vasanthi, "Machine learning algorithms with ROC curve for predicting and diagnosing the heart disease," in Soft Computing and Medical Bioinformatics, 63–72, Springer, 2019, doi:10.1016/j.compedu.2012.08.015.

[23] C. Hodges, S. Moore, B. Lockee, T. Trust, A. Bond, "The difference between emergency remote teaching and online learning," Educause Review, **27**, 2020.

[24] H. M. Niemi, P. Kousa, et al., "A case study of students' and teachers' perceptions in a Finnish high school during the COVID pandemic," International journal of technology in education and science., 2020.

[25] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, S. Ali, "Using machine learning to predict student difficulties from learning session data," Artificial Intelligence Review, **52**(1), 381–407, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 154-160 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Multi-Objective Design of Current Conveyor using Optimization Algorithms

Abdelaziz Lberni[*,1], Malika Alami Marktani[2], Abdelaziz Ahaitouf[3], Ali Ahaitouf[1]

[1]*Faculty of Sciences and Technology, SIGER Laboratory, University of Sidi Mohamed Ben Abdellah, Fez, 2202, Morocco*

[2]*National School of Applied Sciences, SIGER Laboratory, University of Sidi Mohamed Ben Abdellah, Fez, 2202, Morocco*

[3]*SI Laboratory, Polydisciplinary Faculty of Taza, University of Sidi Mohamed Ben Abdellah, Taza-Gare, 1223, Morocco*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 15 February, 2021
Online: 10 March, 2021

Keywords:
Multi-objective optimization
evolutionary algorithms
Metaheuristics
Pareto front
Analog design
CMOS Current conveyor

ABSTRACT

*The design of microelectronic systems is often complex, therefore metaheuristics can be of a great interest, because in most cases these systems have conflicting objectives and constraints. In this paper, we demonstrate the application of multi-criteria design strategies to a CMOS current conveyor. This provides designers with the ability to develop solutions that can meet several objectives respecting the design constraints. Therefore, three evolutionary algorithms well-known for their best performance in the resolution of more difficult multi-objective problems are proposed. They are first applied to the well-known benchmark functions and then for the optimal design of the current conveyor transistors in the framework of the 0.18*µm *CMOS technology. The aim is to maximize the bandwidth and minimize the parasitic input resistance respecting the technological constraints of the circuit. The obtained results are integrated in Cadence tool to show their validities. Final performances obtained by the three methods are in agreement and are better compared to the state-of-art-results.*

1 Introduction

Today, with the complex growth of VLSI technology, it is very difficult to hand design analog integrated circuits with multiple parameters and purposes. The characterization of complicated trade-offs between conflicting and nonlinear performances while ensuring the required design specifications makes the design of analog circuits a tedious and time-consuming process. In this regard, due to their design difficulty and complexity, analog circuits have been attracted a lot of optimization attention. In general, optimization is often a time-consuming process having several contradictory criteria as well as a wide variety of design parameters. However, the design of electronic circuits is carried out by optimizing the circuit parameters to be able to rapidly design high-performance circuits [1]. For instance, finding passive elements values and transistor sizes and bias currents, so it can meet output performances such as gain, frequency band, power consumption, etc.

Several metaheuristics have been developed in the literature, which can be divided into two principal categories [2]: Single solution based methods, such as Taboo Search (TS) [3], Local Search (LS) [4], Simulated Annealing (SA) [5], or population based approaches like, Ant Colony Optimization (ACO)[6], Whale Optimization Algorithms (WOA)[7], Grey Wolf Optimizer (GWO), Particle Swarms Optimization (PSO)[8], Hybrid PSO-GWO [9], Non-dominated Sorting Genetic Algorithm (NSGA II)[10], Multi-objective Genetic Algorithm (MOGA)[11], Strength Pareto Evolutionary Algorithm (SPEA2)[12], etc. Single solution based techniques do not offer good results for problems where different types of variables, objectives and constraint functions (linear or non-linear constraints) are used [8]. Moreover, their efficiency highly depends on their parameters, the search space dimension and the number of variables. The population based techniques are generally classified into two groups, Particle Swarms (PS) and Evolutionary Algorithms (EA). PS give good results for problems that are not so difficult [13], but their performance also depends on their parameters and the complexity of the problem, especially for multi-objective optimization problems (MOP) [5, 13]. However, EAs are optimization techniques based on biological evolution and natural selection of species [14], are population-based, where each individual represents a possible solution. The initial population is generated randomly. At every new generation, the population iteratively evolves by the mutation, the crossover and the selection operators on each individual, and only non-dominated solutions meeting the constraints, will survive. Hence, such algorithms are well-known for their efficiency when

[*]Corresponding Author: Abdelaziz Lberni, abdelaziz.lberni@usmba.ac.ma

solving complex MOP, Unlike PS, EAs do not need big parameter adjustments.

In this paper, we apply EAs for the optimal design of the current conveyor (CCII). The used EAs are NSGA II, MOGA and SPEA2, which allows the simultaneous optimization of multiple conflicting targets resulting into a set of Pareto Front (PF) solutions. Hence, the main objective to optimize the size of the MOS transistors transistors, using these algorithms, to achieve the high performances of the CCII. The CCII is one of the best known current mode circuits, making it the objective of several applications, such as filters, oscillators, etc., [15]. As far as we know, few works have been reported until now on CCII optimization by the EAs. In [16], the MOPSO with Crowding Distance (CD) was used for the optimization of the CCII, the differential CCII and the current feedback operational amplifier (CFOA) for low voltage low power applications, the MOPSO-CD was used as a part of a simulation-based tool to find the optimal sizing transistors that operate in weak inversion. In [17], the NSGA II and the decomposition-based multi-objective EA (MOEA/D) were used for the optimization of other purposes of the CCII, i.e., current gain and offset. However, the high-performance CCII design requires that the input parasitic resistance be small and its cut-off frequency be high. For this reason, we have chosen in this work these two characteristics as objectives to be optimized. All EAs generate Pareto fronts and simulations are carried out on Cadence using 0.18µm CMOS process. The simulation results are conform to those obtained by the optimization.

This paper is structured as follows: Section 2 gives an overview of the EAs, Section 3 is dedicated to the EAs validation using usual test functions. Section 4 presents the CMOS CCII. Section 5, is devoted to the results and discussion. A conclusion is given at the end.

2 Evolutionary Algorithms

2.1 MOGA

The MOGA was introduced in [11], as a new variant of the Golberg approach [18]. It uses the concept of dominance and a random-based fitness assignment. The non-dominanted solutions are ranked into groups, which are assigned the same rank in each group.

The pseudo-code of MOGA is given by algorithm 1.

Algorithm 1: MOGA Pseudo Code

```
Generate the initial population P_i;
Evaluate of P_i;
while the stopping criteria is not met do
    i=i+1;
    Selection of P_i;
    Mutation and Crossover of P_i;
    Evaluate P_i;
end
Result: non-dominated solutions
```

2.2 NSGA II

NSGA-II, which was proposed in [10] as a modified version of NSGA, is among the most commonly used and effective EAs due to its simplicity and effectiveness. The basic operation of the NSGA-II is : A random population is created. This generated population is sorted using the notion of dominance. A fitness function is assigned to each solution. Therefore, it is assumed that physical fitness is minimized. At first, selection, mutation and crossover operators are used to create a new population from the first. The NSGA II algorithm uses the notion of elitism, to compare the current population with previously found best non-dominated solutions.

NSGA II relies on two major procedures: crowding distance and fast non-dominated sorting. Both procedures ensure elitism and feasibility of solutions. Algorithm 2 represents the NSGA II pseudo-code.

Algorithm 2: NSGA II Pseudo Code

```
Generate randomly the initial population P_i
while the stopping criteria is not satisfied do
    while population in not classified do
        Search for non-dominated individuals;
        Fitness calculation;
        Sharing;
    end
    Mutation;
    Crossover;
    Selection;
end
```

2.3 SPEA2

SPEA2 is a MO algorithm introduced in [12] as an improved version of SPEA. It is based on the notion of dominating fitness evaluation to to generate the PF. SPEA2 uses elitist concept maintaining an external archive of non-dominated solutions. It also uses a nearest neighbor density estimation method, and an improved archive Truncation approach.

The SPEA2 is given by algorithm 3.

Algorithm 3: SPEA2 Pseudo Code

```
Generate randomly population P_i;
Evaluate objective and create external Archive A_i ;
for i = 1, i ≤ Max iterations do
    i=i+1;
    Evaluate objectives and update A_i by Truncation
      operator;
    Perform Binary Tournament Selection ;
    Crossover of P_i ;
    mutation of P_i ;
end
Result: External Archive with non dominated solutions.
```

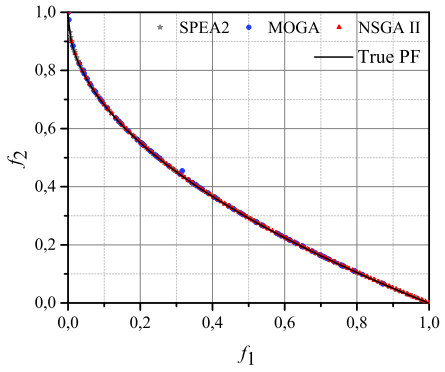
3 EA robustness

Before using the proposed algorithms, we evaluated performances using multi-objective standards ZDT functions [19].

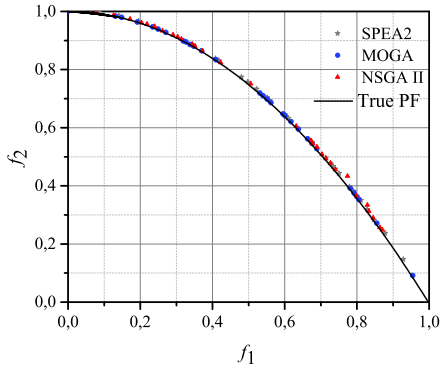
Each ZDT function includes two objectives f_1 and f_2 with 30 variables, which demonstrates the high complexity of such test functions. All tests are carried out with the algorithms parameters of

1000 iterations and a population size of 50 and a crossover probability of 0.8 and a mutation probability of 0.1. The ZDT benchmark functions and their expressions are given in the appendix.

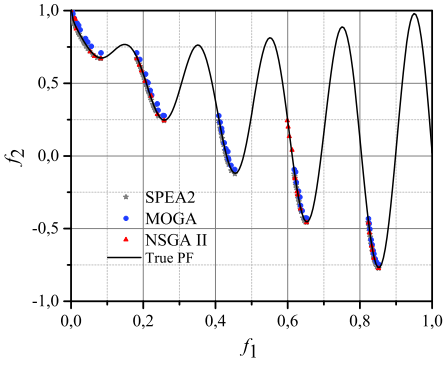
Figure 1 shows the PFs obtained for the three chosen algorithms. As can be clearly seen the generated PFs with the all proposed methods achieve good approximations to the exact benchmark functions PFs. Therefore, we can confidently use them for the optimization problems of the CCII optimal design with guaranteed results.



(a) ZDT1



(b) ZDT2



(c) ZDT3

Figure 1: Evaluation of algorithms by ZDT benchmark functions

4 CMOS Current Conveyor

In this section, we studied the current conveyor shown in Figure 2, it has three active ports X, Y and Z, and its main function consists of

- Current follower between ports X and Z, which can be provided by the translinear loop formed by transistors M1- M4.
- Voltage follower between ports X and Y, which can be provided by M5-M6 and M7-M8 current mirrors.

The present current conveyor topology is the most used one for its good performances and the interest of using a translinear loop [20]. Therefore, the design optimization of the CCII is performed considering its main objective functions : the parasitic resistance at the port X (R_X) and the cut-off frequency ($f_{\text{-3dB}}$). Recall that the aim is to minimize the first objective to obtain low input resistance and to maximize the second objective to get high bandwidth.

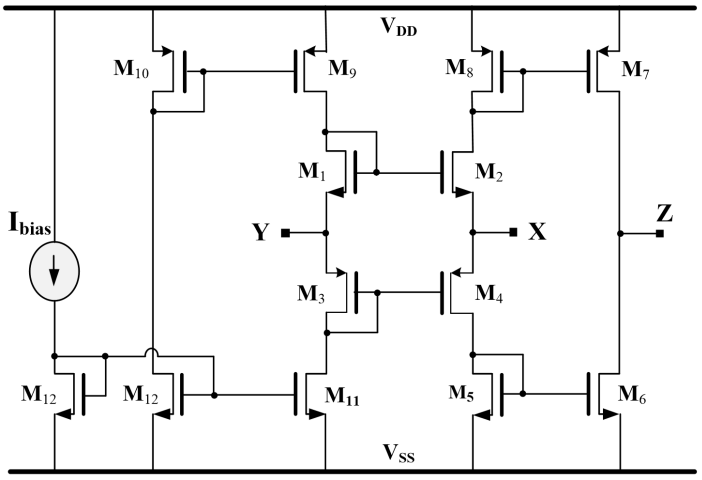


Figure 2: CMOS Current conveyor.

The CCII sizing is performed as in [14], using CMOS 0.18 µm process, and with the following conditions $V_{DD} = V_{SS} = \pm 1.8V$. All the transistors are characterized by their geometrical parameters: the channel length L and the gate width W.

The CCII optimal sizing issue is treated as a MOP. The aim is to find the best trade-off between a small R_X and a high $f_{\text{-3dB}}$ as a function of the transistors parameters. The problem constraints are given by Eqs. 4 and 5 bellow corresponding to the saturation transistor regime [15].

The design problem can be formulated as:

$$\begin{aligned} &\text{Minimize} \quad R_X(x) \text{ and } -f_{\text{-3dB}}(x). \\ &\qquad\qquad x = \{W_n, W_p, L_n, L_p, I_0\}. \end{aligned} \tag{1}$$

subject to

$$g_{1,2}(x) \leq 0,$$

where

- The resistance R_X is

$$R_X = \frac{1}{g_{mn} + g_{mp}} = \frac{1}{\sqrt{2\mu_n C_{ox}\frac{W_n}{L_n} I_0} + \sqrt{2\mu_p C_{ox}\frac{W_p}{L_p} I_0}} \quad (2)$$

$g_{mn(p)}$ is the transconductance for N(P) channel transistor. μ_n (μ_p) and C_{ox} are the electrons (holes) mobility and the gate oxide capacitance per unit area, respectively. I_0 is the bias current.

- The cut-off frequency $f_{\text{-3dB}}$ is

$$f_{\text{-3dB}} = \frac{\omega_{\text{-3dB}}}{2\pi} \quad (3)$$

The saturation constraints g_1 and g_2 are given by :

- The constraint of M_2 and M_8 transistors:

$$g_1 = V_{SS} - V_x(min) + V_{tn} + \sqrt{\frac{2I_0}{\mu_n C_{ox}\frac{W_n}{L_n}}} + \sqrt{\frac{2I_0}{\mu_p C_{ox}\frac{W_p}{L_p}}} \quad (4)$$

- The constraint of M_4 and M_5 transistors:

$$g_2 = V_x(max) - V_{DD} - V_{tp} + \sqrt{\frac{2I_0}{\mu_n C_{ox}\frac{W_n}{L_n}}} + \sqrt{\frac{2I_0}{\mu_p C_{ox}\frac{W_p}{L_p}}} \quad (5)$$

where $V_{tn}(V_{tp})$, $V_{\text{DD(SS)}}$ are the threshold voltage for NMOS(PMOS) and the supply voltage, respectively. W_n (W_p), L_n (L_p) are the gate width and the channel length for n-channel (p-channel) transistors, respectively.

5 Results and discussion

All the CCII transistors with the same channel type have the same parameters (W_n, L_n for NMOS and W_p, L_p for PMOS). To respect the industrial design constraints, we also used identical channel length (L) for all transistors.

5.1 CCII optimization results

The optimal MOS transistors sizes are reached using EAs by minimizing R_X and maximizing $f_{\text{-3dB}}$ in two ways:

First, the optimization process is performed by MOGA, SPEA2 and NSGA II using three I_0 values, i.e., 20µA, 40µA and 80µA. For these experiments, we use the minimum channel length $L_n = L_p = L_{min}$. The generated PFs (R_X and $-f_{\text{-3dB}}$) by the algorithms for different bias currents are shown in Figure 3. From this figure, we can see that the best trade-off between R_X and $-f_{\text{-3dB}}$ is obtained for the PF with the large bias current I_0=80µA. Therefore, designers have to select the best trade-off solution depending on the circuit design and its application requirements.

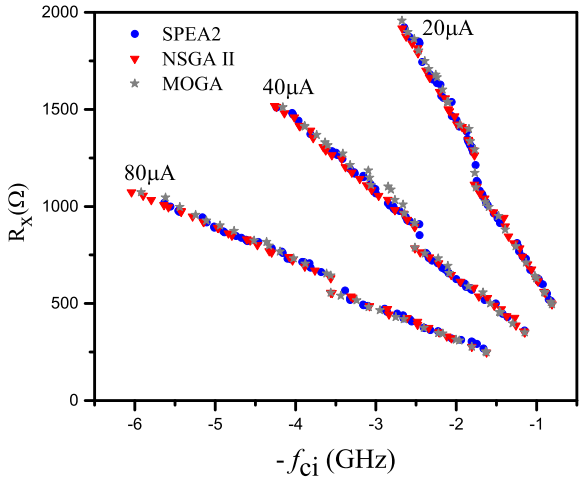


Figure 3: Pareto fronts obtained by the three algorithms for three bias currents

Second, the optimization process is performed by the same algorithms using three channel lengths L_{min}, $2L_{min}$ and $3L_{min}$. For these experiments, we use I_0 = 80µA. The generated PFs for these cases are shown in Figure 4. From this figure, we can see that the best trade-off between performances is achieved for the PF with a minimum channel length L_{min}. Unlike digital circuits, the channel length for analog circuits is usually at least $3L_{min}$, which is why we tested all three channel lengths. This relatively large channel lengths minimizes the effects of channel modulation.

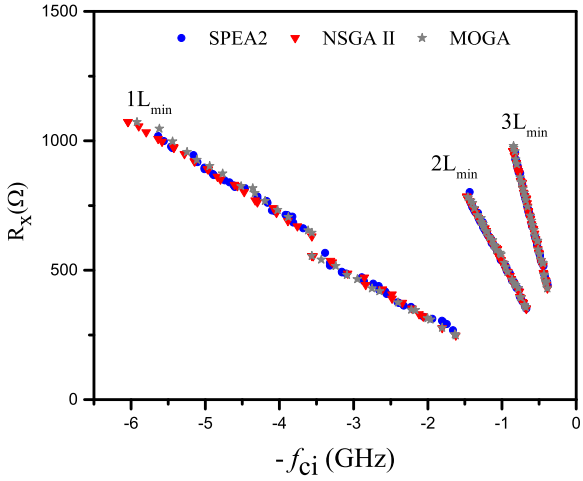


Figure 4: Pareto fronts obtained by the three algorithms for three channel lengths.

Based on the results in both scenarios, we can see that the best performances are obtained with a smaller channel length and a higher bias current. Considering the effects of channel modulation, power consumption and the circuit design and application requirements, IC designers should select the appropriate solution.

We run all algorithms 10 times, to evaluate the weakness and strength of each algorithm in all the performed experiments. To evaluate the diversity and the distribution of Pareto solutions, we

use the hyper volume metric [21]. This metric shows the covered area by the Pareto front. In case of minimization, of the problem, the lager hyper-volume value, the good is the quality of Pareto front solutions.

The hypervolume results are shown in Table 1. The optimization procedure required a mean CPU time of approximately 6.707 s, 67.25 s, and 148.412 s for NSGA II, MOGA and SPEA2, respectively.

Compared to the others algorithms, the NSGA II has the rapid CPU time and a good convergence rate confirmed by the higher hypervolume value.

Table 1: The Hypervolume Values.

		SPEA2	MOGA	NSGA II
L_{min}	20 µA	0.750	**0.752**	**0.752**
	40 µA	0.825	0.824	**0.826**
	80 µA	0.873	0.875	**0.876**
80 µA	1 L_{min}	0.873	0.875	**0.876**
	2 L_{min}	0.824	0.823	**0.825**
	3 L_{min}	0.782	0.784	**0.785**

5.2 *CCII validation results*

Table 2 presents the solutions to be validated by the CADENCE tool, they are defined by their cutoff frequencies, their values of R_X and the corresponding transistors parameters. They are randomly chosen from the PF that corresponds to I_0 = 80µA and L_n=L_p=L_{min}.

Table 2: Parameters to be validated by Cadence, obtained in the PF that gives the best trade-off (R_X,f_{-3dB}).

	Test	W_n(µm)	W_p(µm)	f_{-3dB}(GHz)	R_X(Ω)
MOGA	1	1.08	4.26	5.11	926.5
	2	3.77	13.38	3.08	481.26
SPEA2	1	0.86	3.54	5.63	1018.6
	2	4.30	17.85	2.66	439.07
NSGA II	1	0.80	3.08	6.04	1073.3
	2	4.65	15.77	2.64	443.26

Figures 5–8 show the simulation results of the selected solutions in table 2. The maximum (minimum) deviation between the simulation and the theoretical results is 10.1% (1%) and 9.4% (4.7%) for R_X and f_{-3dB}, respectively. This interval of variations is quite narrow and makes it possible to consider that the simulation results are in good agreement with the theoretical ones, obtained by the algorithms.

Table 3 presents a qualitative comparison of the obtained CCII performance with works previously published but with 0.35µm process. From the data given in this table, it is clear that the achieved performances are higher than reported, i.e, very higher frequency and much lower power consumption and good X-port resistance.

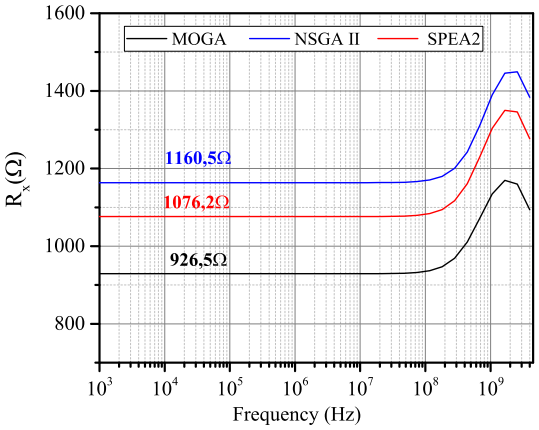


Figure 5: Cadence simulation results of parasitic resistance for the used algorithms (Tests 1 in table 2).

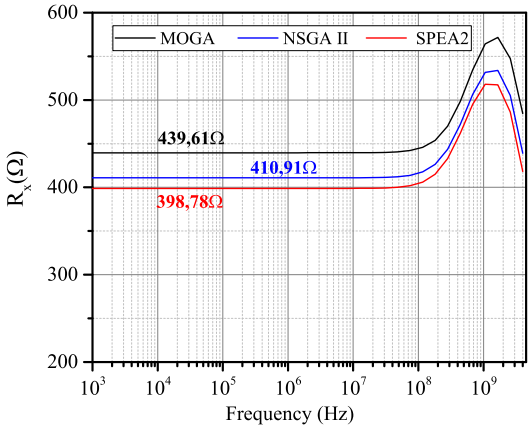


Figure 6: Cadence simulation results of parasitic resistance for the used algorithms (Tests 2 in table 2).

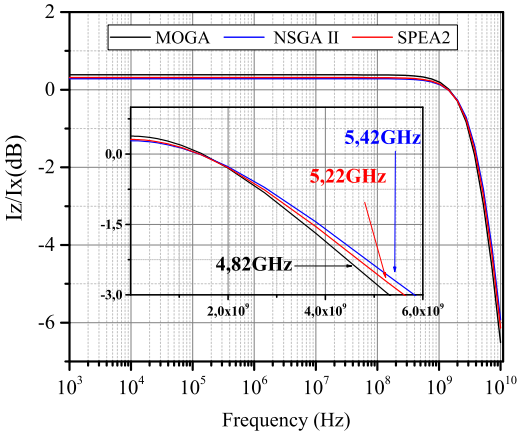


Figure 7: Cadence simulation results of cut-off frequency for the used algorithms (Tests 1 in table 2).

Table 3: CCII performance comparison with works previously published.

	This work	[22]	[23]	[24]
Algorithms	NSGA II	PSO-2S	MODE	EGO-PEI
Technology (µm)	CMOS 0.18	CMOS 0.35	CMOS 0.35	AMS 0.35
V_{DD}/V_{SS} (V)	±1.8	±2.5	±2.5	–
I_{bias} (µA)	80	300	300	–
Power (µW)	288	1500	1500	–
$f_{\text{-3dB}}$(max) (GHz)	6.04	2.121	2.132	1.2255
R_X(min) (Ω)	247	225	221	259.53

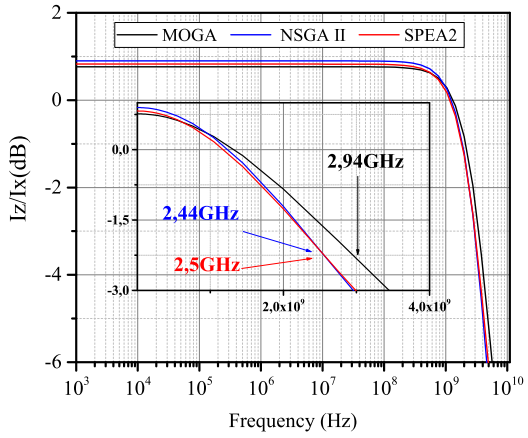


Figure 8: Cadence simulation results of cut-off frequency for the used algorithms (Tests 2 in table 2).

6 Conclusion

In this paper, we presented the usefulness of applying EAs, namely MOGA, SPEA2 and NSGA II, for the automatic optimization of high performances CCII. Several optimization experiments were carried out with three bias currents and three channel lengths, minimizing the parasitic resistance and maximizing the cut-off frequency. In all experiments, the achieved results show that these methods can provide Pareto Fronts with greater solutions diversity. The simulations are performed by Cadence using the CMOS 0.18µm process, showing good accuracy with the theoretical results. The best performances achieved with EAs in this work can be summarized in a power consumption of 288µW, a minimum parasitic resistance of about 247Ω and a maximum frequency of about 6.04GHz.

Appendix

- ZDT1 function:

$$\begin{aligned} &g(\vec{x}) = 1 + 9\sum_{i=2}^{Dim}\frac{x_i}{30-1} \quad \text{and} \quad h(\vec{x}, f_1, g) = 1 - \sqrt{\frac{f_1(\vec{x})}{g(\vec{x})}} \\ &f_1(\vec{x}) = x_1 \\ &f_2(\vec{x}) = g(\vec{x}).h(\vec{x}, f_1, g) \end{aligned} \tag{6}$$

- ZDT2 function:

$$\begin{aligned} &g(\vec{x}) = 1 + 9\sum_{i=2}^{Dim}\frac{x_i}{30-1} \quad \text{and} \quad h(\vec{x}, f_1, g) = 1 - \left(\frac{f_1(\vec{x})}{g(\vec{x})}\right)^2 \\ &f_1(\vec{x}) = x_1 \\ &f_2(\vec{x}) = g(\vec{x}).h(\vec{x}, f_1, g) \end{aligned} \tag{7}$$

- ZDT3 function:

$$\begin{aligned} &g(\vec{x}) = 1 + 9\sum_{i=2}^{Dim}\frac{x_i}{30-1} \\ &h(\vec{x}, f_1, g) = 1 - \left(\frac{f_1(\vec{x})}{g(\vec{x})}\right)\sin(10\pi f_1(\vec{x})) - \sqrt{\frac{f_1(\vec{x})}{g(\vec{x})}} \\ &f_1(\vec{x}) = x_1 \\ &f_2(\vec{x}) = g(\vec{x}).h(\vec{x}, f_1, g) \end{aligned} \tag{8}$$

References

[1] B. Liu, G. Gielen, F. V. Fernández, Automated design of analog and high-frequency circuits, Springer, 2014.

[2] M. Fakhfakh, A. Sallem, M. Boughariou, S. Bennour, E. Bradai, E. Gaddour, M. Loulou, "Analogue circuit optimization through a hybrid approach," in Intelligent Computational Optimization in Engineering, 297–327, Springer, 2011.

[3] A. Gharbi, M. Benrejeb, P. Borne, "A taboo search optimization of the control law of nonlinear systems with bounded uncertainties," International Journal of Computers Communications & Control, **11**, 224–232, 2016, doi: 10.15837/ijccc.2016.2.2010.

[4] C. Dai, Y. Wang, L. Hu, "An improved α-dominance strategy for many-objective optimization problems," Soft Computing, **20**, 1105–1111, 2016, doi:10.1007/s00500-014-1570-8.

[5] F. Y. Vincent, A. P. Redi, Y. A. Hidayat, O. J. Wibowo, "A simulated annealing heuristic for the hybrid vehicle routing problem," Applied Soft Computing, **53**, 119–132, 2017, doi:10.1016/j.asoc.2016.12.027.

[6] Q. Yang, W. Chen, Z. Yu, T. Gu, Y. Li, H. Zhang, J. Zhang, "Adaptive Multimodal Continuous Ant Colony Optimization," IEEE Transactions on Evolutionary Computation, **21**(2), 191–205, 2017, doi:10.1109/TEVC.2016.2591064.

[7] A. Lberni, M. A. Marktani, A. Ahaitouf, A. Ahaitouf, "Adaptation of the Whale Optimization Algorithm to the Optimal Sizing of Analog Integrated Circuit: Low Voltage Amplifier Performances," in 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 1–6, 2020, doi:10.1109/ICECOCS50124.2020.9314428.

[8] M. Fakhfakh, Y. Cooren, A. Sallem, M. Loulou, P. Siarry, "Analog circuit design optimization through the particle swarm optimization technique," Analog integrated circuits and signal processing, **63**, 71–82, 2010, doi:10.1007/s10470-009-9361-3.

[9] A. Lberni, M. A. Marktani, A. Ahaitouf, A. Ahaitouf, "Application of hpsgwo to the optimal sizing of analog active filter," in International Conference on Electronic Engineering and Renewable Energy, 309–315, Springer, 2020.

[10] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE Transactions on Evolutionary Computation, **6**(2), 182–197, 2002, doi:10.1109/4235.996017.

[11] T. Murata, H. Ishibuchi, "MOGA: multi-objective genetic algorithms," in Proceedings of 1995 IEEE International Conference on Evolutionary Computation, volume 1, 289–, 1995, doi:10.1109/ICEC.1995.489161.

[12] E. Zitzler, M. Laumanns, L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," TIK-report, **103**, 2001, doi:10.3929/ethz-a-004284029.

[13] A. Sallem, B. Benhala, M. Kotti, M. Fakhfakh, A. Ahaitouf, M. Loulou, "Application of swarm intelligence techniques to the design of analog circuits: evaluation and comparison," Analog Integrated Circuits and Signal Processing, **75**, 499–516, 2013, doi:10.1007/s10470-013-0054-6.

[14] A. Lberni, A. Ahaitouf, M. A. Marktani, A. Ahaitouf, "Sizing of second generation current conveyor using evolutionary algorithms," in 2019 International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS), 1–5, 2019, doi:10.1109/ISACS48493.2019.9068896.

[15] A. Lberni, M. A. Marktani, A. Ahaitouf, A. Ahaitouf(in press), "International Journal of Computer Aided Engineering and Technology," .

[16] A. Lberni, A. Sallem, M. A. Marktani, A. Ahaitouf, N. Masmoudi, A. Ahaitouf, "An efficient Multi-Objective Simulation-based Approach for Low Voltage Low Power Analog ICs," in 2020 IEEE International Conference on Design Test of Integrated Micro Nano-Systems (DTS), 1–5, 2020, doi:10.1109/DTS48731.2020.9196166.

[17] I. Guerra-Gómez, E. Tlelo-Cuautle, T. McConaghy, G. Gielen, "Optimizing current conveyors by evolutionary algorithms including differential evolution," in 2009 16th IEEE International Conference on Electronics, Circuits and Systems - (ICECS 2009), 259–262, 2009, doi:10.1109/ICECS.2009.5410989.

[18] D. Golberg, S. Mahfoud, "A Genetic algorithm for parallel simulated annealing," in Proceedings of international conference on parallel problem solving from nature. Netherland, 1992.

[19] E. Zitzler, K. Deb, L. Thiele, "Comparison of multi-objective evolutionary algorithms: Empirical result," Evolutionary computation, **8**, 173–195, 2000, doi:10.1162/106365600568202.

[20] M. A. Yakout, T. A. Alawadi, "A novel simple BiCMOS Current Controlled Current Conveyor for RF applications," International Journal of Electronics Letters, **8**, 1–16, 2019, doi:10.1080/21681724.2019.1600726.

[21] G. Chen, J. Qian, Z. Zhang, Z. Sun, "Applications of novel hybrid bat algorithm with constrained Pareto fuzzy dominant rule on multi-objective optimal power flow problems," IEEE Access, **7**, 52060–52084, 2019, doi:10.1109/ACCESS.2019.2912643.

[22] A. El Dor, M. Fakhfakh, P. Siarry, "Performance optimization of CMOS second generation current conveyors using a multi-swarm algorithm," AEU-International Journal of Electronics and Communications, **68**, 496–503, 2014, doi:10.1016/j.aeue.2013.12.007.

[23] A. El Dor, M. Fakhfakh, P. Siarry, "Multiobjective differential evolution algorithm using crowding distance for the optimal design of analog circuits," Journal of Electrical Systems, **12**(3), 612–622, 2016.

[24] N. Drira, M. Kotti, M. Fakhfakh, P. Siarry, E. Tlelo-Cuautle, "Pseudo Expected Improvement Based-Optimization for CMOS Analog Circuit Design," in 2019 16th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), 137–140, 2019, doi:10.1109/SMACD.2019.8795259.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 155-159 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Inferring Topics within Social Networking Big Data, Towards an Alternative for Socio-Political Measurement

Khalid Ait Hadi*,[1], Rafik Lasri[1], Abdellatif El Abderrahmani[2]

[1]*Department of Computer Sciences, Polydisciplinary Faculty of Larache, Abdelmalek Essaadi University, 93000, Morocco*

[2]*Department of Computer Sciences, Dhar El Mahraz Sciences Faculty, Sidi Mohamed Ben Abdellah University, 30050, Morocco*

ARTICLE INFO

Article history:
Received: 03 September, 2020
Accepted: 28 October, 2020
Online: 08 November, 2020

Keywords:
Language processing
Data Mining
Topics detection

ABSTRACT

This research sought to measure some socio-political indicators using millions of opinionated messages from social network sourced big data. Thus, and using an enhanced mixed method for sentiment analysis and a fusion model algorithm to infer topics from short text, this study attempted to demonstrate the value of computational approaches in measuring some phenomena in the real social world and quantifying public opinion fluctuations in response to certain socio-political issues. The validity of the experimental results was examined by comparing them with data obtained from representative surveys, thus providing a better understanding of the relationships between online and offline opinion dynamics. This contribution is intended to be multidisciplinary, both useful for policymakers and opinion analysts to explore public trends and to inquire into socio-political issues.

1. Introduction

Nowadays, in the communication fields, market research and public opinion analysis, as in many other disciplinary domains, social media are mobilized as new societal trends observatories [1, 2]. The evolution of these networking platforms in the first decade of the 21st century has enabled Internet users without specific technical skills to easily publish and share content on their concerns [3]. The participation architecture on which these services are based has facilitated the information co-production, and has offered spaces for socio-political engagement [4].

In this context, millions of people interact daily on social networks, producing several hundreds of millions of messages of all kinds and of any content. These messages represent valuable clues concerning practices, representations and Internet users' opinions. As such, they form a particularly interesting material to investigate when looking at citizen public behavior studies. Furthermore, due to the development of technical and computational tools that have made it possible to collect, archive and analyze huge data volumes, namely big data, these numerical phenomena can be objectified and quantified, on a large scale, for the investigation's needs. Social media analytics open up a promising way for the quantitative study of certain substantial subjects of declarative surveys, such as quantifying a social object, capturing behavioral representations and uncovering political intentions.

In this work, we seek to exploit data extracted from Facebook and use computational methods to measure some socio-political indicators. This article does not claim to provide affirmative answers about the effective correlations between the phenomena observed on social media and those perceived in the real world. On the contrary, it aims to help support the observation indicating that social networking data constitute a particularly interesting material to investigate when trying to study citizens' behavior, public opinion's fluctuations and social trends, particularly where declarative surveys and others conventional approaches turn out to be often costly, time-consuming and labor-intensive [5].

This paper is an extension of work originally presented in 2019 International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS) [6]. An initial work, which was interested in measuring of notoriety capital on Facebook of a political figure. Here, we will not be limited to measuring political popularity, but we will improve this research to be able to evaluate other socio-political indicators. Likewise, we will not settle for the use and study of messages formulated only in one language, as was the French case in [6], but we will extend the prospecting to include also social networking data expressed in classical Arabic and dialect as well.

*Corresponding Author: Khalid Ait Hadi, khalid.aithadi@gmail.com

Another improvement of the present work is linked this time to the observation relating to the fact that topic models, like Latent Dirichlet Allocation (LDA) and its derivatives, are not efficient in the analysis of short texts due to data sparsity [7]. Alas, it is clear that the overwhelming majority of data extracted from social networks consists of short texts. Here, we will overcome this inconvenience by opting for an improved Biterm Topic Model (BTM) model based on the heat matrix [8].

The rest part of the paper is structured as follows. Section 2 presents an overview of related work. Section 3 is dedicated to the data collection process description and the suggested approach implementation. An experimental evaluation of the method is also presented in the same section. Finally, discussions, conclusions and suggestions for future research openings are made in the last section.

2. Related Work

Studies seeking to examine predictive capacity of social media data and to establish existence of possible correlations between the facts perceived on social media and those observed in the effective world have been subject to a rich and growing literature since the beginning of the 21st century [2].

Many recent studies have used social media data as a new sensor of societal trends and a new predictor of economic or political phenomena, ranging from stock market volatility to box office performance or, in a certain perspective that interests us more directly, public opinion fluctuations and social dynamics [9, 10]. Research in this area has been driven by a positivist approach that is certainly strengthened owing to the unprecedented possibilities offered by these new social platforms, in the sense that they allow unheard-of access to the data shared daily by people as well as to their networks, regardless of any temporal or geographic consideration.

In this context, a review article published in 2014 [11] has listed more than a hundred studies devoted solely to political messages published on the social network Twitter. The penchant of researchers for Twitter can be explained by easier access to data, where, unlike Facebook, the vast majority of accounts are “accessible” and by the brevity of posts on Twitter, limited to 280 characters.

Always on Twitter, a meta-analysis developed by Gayo-Avello [12] examined the predictive usefulness of data published on this social network, leading to the conclusion that such data have relatively predictive power and provide some information referring to electoral consultations results.

It should be noted on the socio-political aspect that the most commonly studied subject is that concerning electoral outcomes prediction using social media data [13]. In the opposite, studies relating to the political popularity evaluation, for example, or to the measurement of social indicators are not numerous, and here too Twitter has often been called upon to make forecasts. In addition, and almost globally, these studies have not examined the validity of the predictions by making comparisons to conventionally more admissible results types, such as those from censuses and surveys [14].

It should also be emphasized, according to the existing scientific literature [1], that most results in the field of forecasting with social media data claim to have produced predictive results to some extent, registering a positive statistical correlation between offline and online data. However, several scientists have criticized these results, which draw up a fundamental limitation to these works that is linked to the lack of sociological representativeness of the populations recorded on social networks [15]. Though, dismissing social media data on this ground would fail to capture the opinion forming dynamics. Some studies have even argued that representativeness search is no longer consubstantial with the need to sample populations in the big data era. The arguments put forward in this wake point to fact that debates led by certain politically active groups prevail over those who develop in society at large [16], and that individuals active on social networks are thought leaders, more politicized than average, influential in their entourage and acquaintances and whose opinions count more than those of "ordinary" individuals. To poll these opinion leaders would indirectly mean polling their entourage and, ultimately, the entire population. In addition, discussions on certain social platforms would above all reflect the concerns and themes put on the agenda by the mainstream media [17]. In many ways, the social network appears to be an echo chamber for the media field, and the associated media agenda would indirectly influence citizens' concerns [18].

Regarding the diversity of methods used in measuring public opinion, whether sentiment-based or social network-based approaches, it should be underlined that there is no unanimity about the most efficient methods in terms of prediction, and several researchers have reported some performance with different approaches and implementations.

Thus, and particularly in the category of sentiment-based studies, where opinions tone is used as a behavior’s indicator, two lines of research stand out in this context, one is based on pre-established lexicas [19], while the other relies on new sentiment models specifically for political messages [20].

Although concerning the dictionary-based sentiment classifiers, some studies have criticized the reliability of these approaches for predictions [21, 22], advancing in this regard some deficiencies such as the incorrect classification of the word in the lexicon, the lack of words disambiguation, and the neglect of contextual inference. Imperfections that are more accentuated in the socio-political context, with the emergence of difficulties especially linked to the lack of subtleties of socio-political language [23].

Through this literature review, it is seen that the majority of research has focused on the electoral and public opinion forecasts, without identifying and analyzing the reasons behind these opinions development, problematic that the present work is concerned with. In addition, there are two more obstacles that we will overcome in this paper, firstly by using a data source other than Twitter, Facebook in this case, and secondly by comparing our results with those of a declarative survey [24].

3. Methodology and Experimental Results

As previously announced, this research aims precisely to support the predictive interest of data extracted from Facebook. This involves, on the one hand, developing algorithm to analyze a very large amount of messages with socio-political connotation in

order to identify their polarity and target and, on the other hand, to compare these information to opinion poll statistics so as to better grasp similarity degrees between online and offline opinion dynamics. For this, we will consider the same period covered by the public opinion survey 'The Arab Barometer' [24] conducted in Morocco from May 7 – June 11, 2016.

Data collection has been made using '*Rfacebook*' (a package that provides a series of functions that allow R users to access Facebook's API to collect public status updates that mention specific tags). Therefore, and in order to cover a very large sample of socio-politically concerned Moroccan Facebook users, the choice has been made so as to extract both messages and comments written in French or Arabic, even dialect. On this register, we admit that debate around socio-politically charged issues is also, if not largely, done in the mother tongue and that disregarded educational level, people remain always sensitive to socio-political themes and use informal language to take part in the debate.

Therefore, and to test our approach, we have looked at two socio-political issues: (I1) 'opinions about politics' and (I2) 'perceptions of freedoms'. For the first theme, we have collected data (posts and comments) corresponding to the following queries: 'politique', 'gouvernement', 'politique gouvernementale', 'سياسة', 'حكومة', and 'السياسة الحكومية'. The second theme was examined using three tags: 'libertés', 'حرية', and 'الحريات'.

3.1. Sentiment analysis

It would be necessary before approaching the opinion mining phase, and in the ultimate objective to recover pure personal opinion, to operate some text preprocessing and cleaning steps: stemming, removal of URLs, punctuations, stopwords, screen names, special caracters and duplicate comments.

For sentiment analysis, two approaches will be used, one reserved for data written in French and the other for those formulated in Arabic or dialect language. For the first category, the choice focused on the use of Carousel greedy algorithm with Cat Swarm Optimization based Functional Link Artificial Neural Networks (CSO-FLANN), a technique for sentiment mining, enhanced in term of accuracy and computational effort [25]. For the other category of data, we used 'SentiArabic', a sentiment lexicon package for standard Arabic [26] coupled with the 'MADAR' corpus, a collection of parallel sentences covering the dialects of 25 cities from the Arab World [27], and for our case we choose the package relating to the Moroccan city 'Fez'. Using these tools, we found 1.15 million positive, 1.48 million negative, and 3.37 million neutral posts and comments for the first issue (I1) 'opinions about politics', and 2.58 million positive, 6.02 million negative, and 4.4 million neutral posts and comments for the second issue (I2) 'perceptions of freedoms'.

3.2. Inferring Topics

Our approach focuses on detecting the main reasons behind a positive or negative impression expressed by Facebook users on a given theme.

Unlike the Latent Dirichlet Allocation (LDA)-based models, which are relatively inefficient in short text processing, where they generate a high dimensional and sparse data problem [7], we opt here for a model called HMBTM (Heat Matrix based Biterm Topic Model), based on Biterm Topic Model (BTM) and improved by introducing the heat matrix. This is also merged with VSM (Vector Space Model) for counting similarities, thus inducing accuracy improvement and dimensionality reduction [8]. Applying an implementation of VSM- HMBTM to reveal the topics from the set of negative posts and comments for the first issue (I1), and from the set of positive posts and comments for the second issue (I2), and using thereafter Maximum Likelihood Estimation (MLE) [28], 906 topics were selected as the optimum number of topics for the first issue (I1), and 798 for the second issue (I2). We admitted here that people express their agreement with positive sentiments as they express their critical positions with negative sentiments.

By examining the most used terms, and despite overlaps and similarities between several themes, we try to assign a unifying label in order to bring together the most relevant subjects. For example, we have assigned for the second issue (I2) 'perceptions of freedoms', the label "freedom of association" to topics containing: joint associations, civil associations, civil organizations… and the label "political participation" to a set including the terms: political parties, free election, partisan activities…

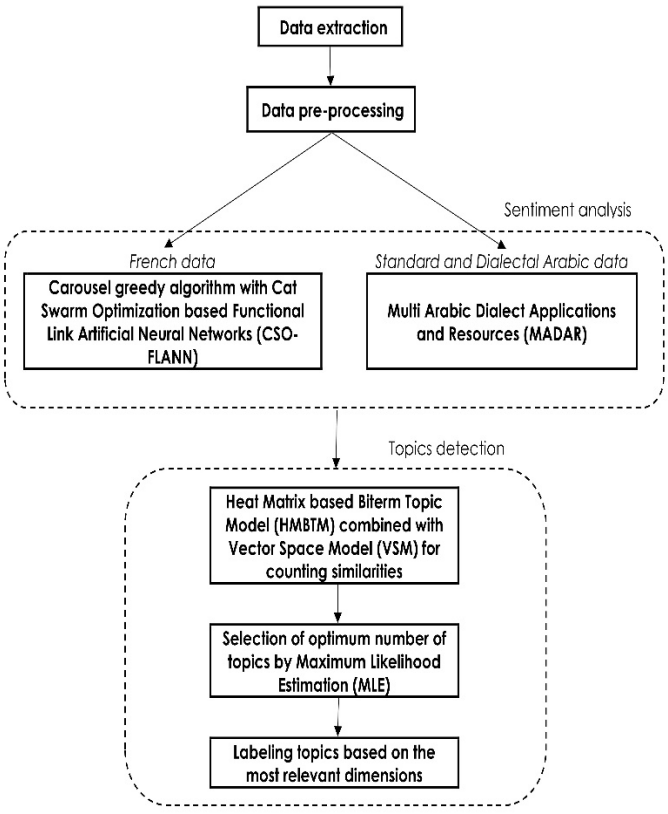


Figure 1: The construction process of our approach

We adopt the same process for the first issue (I1) 'opinions about politics. By examining the main related words such as those in Table 1, we agreed to assign a label based on the most relevant dimensions. For example, we reserved the label "complexity of politics" to topics containing: blurry political landscape, multiplicity of political concepts, number of political parties...

Table 1: A Sample of First Issue's Topics

Complexity of politics	Government not concerned with citizens
multiplicité innombrable ضبابية مبهم	opportunisme détachement لامبالاة ابتعاد
Freedom to criticize the government	Obligation to support the government
reproches critiquable انتقاد خوف	soutien supporter تأييد مساندة

The construction process of our approach is illustrated in Figure 1.

Subsequently, we explore the distribution of labels to get an idea of the topics' distribution and weight for Facebook users. Therefore, Figure 2 represents the rates corresponding to the four relevant subjects that emerge from the analysis, compared to those established by of a declarative survey [24]. Obviously, the values have been scaled given the difference in measurements and scales adopted in each situation.

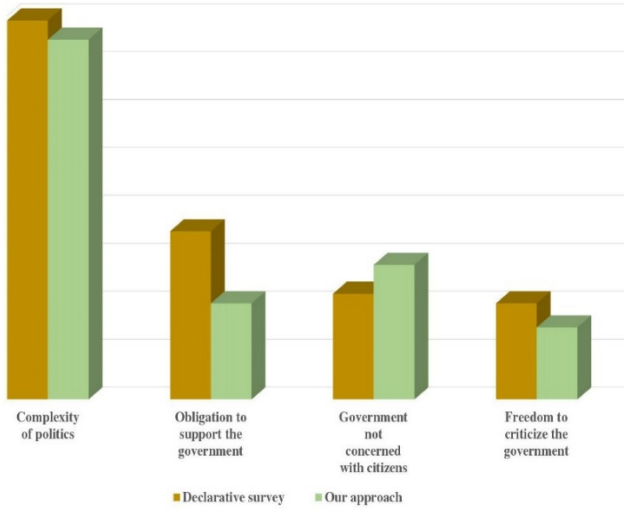


Figure 2: Distribution of First Issue's Topics

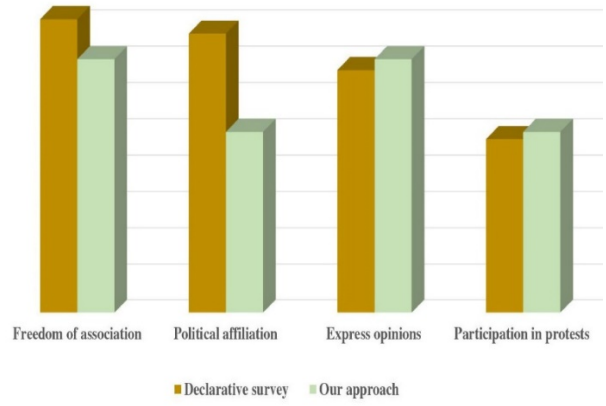


Figure 3: Distribution of Second Issue's Topics

For the second issue (I2) 'perceptions of freedoms', four topics were the most salient: freedom of association, political affiliation, express opinions and participation in protests. Figure 3 compares the weight of these topics to that resulting from the declarative survey [24].

4. Discussion and Conclusions

We have presented a research approach whose objective is to compare the results obtained via two techniques for measuring opinions: the declarative survey, on the one hand, and the analysis of messages published on Facebook, on the other hand. The work sought to situate the data produced by computational methods in relation to the results obtained by surveys. Although the current results do not release certainties, but it appears that, our approach opens up promising horizons, in the sense that it provides very indicative results about the configuration of future declarative surveys results, particularly when the latter are costly in time and even materially. The results presented through this research demonstrate the potential of data extracted from social networks to provide basic and essential information about public opinion dynamics.

However, the results as they stand confirm the observation so much raised by several researchers, indicating that the analysis of social networks would not replace, at least now, public opinion studies based on conventional polls [14]. Nevertheless, it offers indicator data and informative tools to refine our assimilation of public opinion and socio-political behavior. It also opens up perspectives for better capturing indicators relating to engagement and trends measurement, and assessing by the way, the feasibility of many promises in terms of renewing mobilization practices in socio-political field.

One of the possible openings of the present work is to diversify the platforms of data extraction, something that could play, in our point of view, a crucial role to improve the analyses precision. Broadening the base of opinionated data would undoubtedly amount to increasing the representativeness index, so much criticized in this area. One of the concerns that are shaping up recently is the future unavailability of social networks data when the confidentiality of private life is increasingly questioned. End-to-end encrypted messaging applications are attracting new users every day. Social platforms further restrict access to user data to ensure confidentiality of privacy or simply for purely commercial reasons.

One of the possible directions of research is also the one linked to the following question: does not reasoning in terms of positive or negative tones lead to neglecting an essential dimension of the analysis of socio-political messages on social networks, namely the strong majority of messages considered neutral (something that can be seen from the results above)? This leads us to question the silence or neutrality of users and to apprehend its socio-political significance, which would constitute a major issue for future promising research.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. F. Schober, J. Pasek, L. Guggenheim, C. Lampe, F. G. Conrad, "Research Synthesis: Social Media Analyses for Social Measurement" Public Opinion Quarterly, **80**(1), 180–211, 2016. https://doi.org/10.1093/poq/nfv048

[2] H. Schoen, D. Gayo-Avello, P. Takis Metaxas, E. Mustafaraj, M. Strohmaier, P. Gloor, "The Power of Prediction with Social Media" Internet Research, **23**(5), 528–543, 2013. https://doi.org/10.1108/IntR-06-2013-0115
[3] D. Cardon, V. Jeanne-Perrier, F. Le Cam, N. Pélissier, "Présentation" Réseaux, **24**(137), 9–25, 2006.
[4] N. A. Jackson, D. G. Lilleker, "Building an Architecture of Participation? Political Parties and Web 2.0 in Britain" Journal of Information Technology & Politics, **6**(3-4), 232–250, 2009.
[5] W. Wang, D. Rothschild, S. Goel, A. Gelman, "Forecasting Elections with Non-representative Polls" International Journal of Forecasting, **31**(3), 980–991, 2015. https://doi.org/10.1016/j.ijforecast.2014.06.001
[6] K. Ait Hadi, R. Lasri, A. El Abderrahmani, "Social Media Reputation and Political Popularity: Study of a Moroccan Case" in 2019 IEEE International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS), Taza, Morocco, 2019. https://doi.org/10.1109/ISACS48493.2019.9068920
[7] L. Jiang, H. Lu, M. Xu, C. Wang, "Biterm Pseudo Document Topic Model for Short Text" in 2016 IEEE International Conference on Tools with Artificial Intelligence (ICTAI), San Jose, CA, USA, 2016. https://doi.org/10.1109/ICTAI.2016.0134
[8] Q. Liqing, J. Wei, L. Haiyan, F. Xin, "Microblog Hot Topics Detection Based on VSM and HMBTM Model Fusion" IEEE Access, **7**, 120273–120281, 2019. https://doi.org/10.1109/ACCESS.2019.2932458
[9] E. Kalampokis, E. Tambouris, K. Tarabanis, "Understanding the Predictive Power of Social Media" Internet Research, **23**(5), 544–559, 2013.
[10] D. Rousidis, P. Koukaras, C. Tjortjis, "Social Media Prediction: a Literature Review" Multimed Tools Appl, **79**, 6279–6311, 2020. https://doi.org/10.1007/s11042-019-08291-9
[11] A. Jungherr, "Twitter in Politics: A Comprehensive Literature Review" Social Science Research Network, 2014. https://doi.org/10.2139/ssrn.2402443
[12] D. Gayo-Avello, "A Meta-analysis of State-of-the-art Electoral Prediction from Twitter Data" Social Science Computer Review, **31**(6), 649–679, 2013. https://doi.org/10.1177/0894439313493979
[13] S. C. McGregor, R. R. Mourão, L. Molyneux, "Twitter as a Tool for and Object of Political and Electoral Activity: Considering electoral context and variance among actors" Journal of Information Technology & Politics, **14**, 154–167, 2017. https://doi.org/10.1080/19331681.2017.1308289
[14] M. M. Skoric, J. Liu, K. Jaidka, "Electoral and Public Opinion Forecasts with Social Media Data: A Meta-Analysis" Information, **11**(4), 187, 2020. https://doi.org/10.3390/info11040187
[15] T. H. McCormick, H. Lee, N. Cesare, A. Shojaie, E. S. Spiro, "Using Twitter for Demographic and Social Science Research: Tools for Data Collection and Processing" Sociol Methods Res. **46**(3), 390–421, 2017. https://doi.org/10.1177/0049124115605339
[16] A. Dür, "How Interest Groups Influence Public Opinion: Arguments Matter more than the Sources" European Journal of Political Research, **58**(2), 514–535, 2019. https://doi.org/10.1111/1475-6765.12298
[17] J. T. Feezell, "Agenda Setting through Social Media: The Importance of Incidental News Exposure and Social Filtering in the Digital Era" Political Research Quarterly, **71**(2), 1–13, 2018. https://doi.org/10.1177/1065912917744895
[18] Y. Tsfati, J. Cohen, Perceptions of Media and Media Effects, The International Encyclopedia of Media Studies: Media Effects/Media Psychology, 1st Edition, Blackwell Publishing Ltd, 2013.
[19] M. Ibrahim, O. Abdillah, A. F. Wicaksono, M. Adriani, "Buzzer Detection and Sentiment Analysis for Predicting Presidential Election Results in a Twitter Nation" in 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, 2015. https://doi.org/10.1109/ICDMW.2015.113
[20] D. Contractor, T. A. Faruquie, "Understanding Election Candidate Approval Ratings using Social Media Data" in 2013 International Conference on World Wide Web (www '13 Companion), Rio de Janeiro, Brazil, 2013. https://doi.org/10.1145/2487788.2487883
[21] J. E. Chung, E. Mustafaraj, "Can Collective Sentiment Expressed on Twitter Predict Political Elections?" in 2011 AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 2011.
[22] B. O'Connor, R. Balasubramanyan, B. R. Routledge, N. A. Smith, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series" in 2010 AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 2010.
[23] D. Gayo-Avello, "Don't Turn Social Media into Another 'Literary Digest' Poll" Communications of the ACM, **54**, 121–128, 2011. https://doi.org/10.1145/2001269.2001297
[24] The Arab Barometer, "Morocco Five Years after the Arab Uprisings", 2017. https://www.arabbarometer.org/wp-content/uploads/Morocco_Public_Opinion_Survey_2016.pdf
[25] K. Ait Hadi, R. Lasri, A. El Abderrahmani, "An efficient Approach for Sentiment Analysis in a Big Data Environment" International Journal of Engineering and Advanced Technology (IJEAT), **8**(4), 263–266, 2019.
[26] R. Eskander, "SentiArabic: A Sentiment Analyzer for Standard Arabic" in 2018 Language Resources and Evaluation Conference (LREC'18), Miyazaki, Japan, 2018.
[27] H. Bouamor, et al., "The MADAR Arabic Dialect Corpus and Lexicon" in 2018 Language Resources and Evaluation Conference (LREC'18), Miyazaki, Japan, 2018.
[28] W. H. Greene, Econometric Analysis, 7th Edition, Pearson Education, Prentice Hall, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 3, 155-163 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Realization and Energy Optimization of a Recharging Station for Electric Vehicles with Fixed Storage and Photovoltaic Panels

David Roszczypala*, Christophe Batard, Nicolas Ginot, Frédéric Poitiers

Communicating Systems team, I.E.T.R, Polytech, rue Christian Pauc Nantes, 44300, France

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 29 April, 2021
Online: 20 May, 2021

Keywords:
Electric vehicle
Smart-charging
Micro-grid

ABSTRACT

During the past years, a lot of research work have been done on the topic of smart grids and more specifically on the charging of electric vehicles (EVs), which will become an essential aspect in the coming years. The various works carried out on these themes have allowed the development of efficient tools to organize energy exchanges within these networks and to make this energy available to electric vehicles on certain time intervals. However, the problems related to the compatibility between the different elements of these networks seem to be largely underestimated. The collaborative work between IETR and Dropbird highlights the technological challenges that significantly hinder the deployment of relevant charging algorithms and experiments with dynamic programming-based algorithms to circumvent these obstacles.

1 Introduction

The advent of many electric vehicles implies some environmental changes, however, the conversion of a number of motorists to electric vehicles implies a growth in electricity demand and adds additional constraints on the power grid. Figure 1 shows the growth (estimated by the french grid operator) in the number of EVs and their demand for electricity on the French electrical grid in the coming years

Year	2010	2016	2030
Number of EVs	a few thousands	100 000	5 millions
Cumul. power demand (3.7kW)	a few MW	400 MW	20 GW
Cumul. power demand (50kW)	100 MW	5 GW	250 GW
Needed energy per year	a few GWh	a few 100 GWh	10 TWh
Cumul. storage capacity	a few 10 MWh	2.5 GWh	125 GWh

Figure 1: Evolution of EV energy needs (France)

The need for electricity corresponds to the virtual demand (in the scenario where all electric vehicles charge simultaneously), the calculation of the annual energy needed is based on a scenario of 15kWh/100 km and 13,000 km/year (source: ENT2008, datas in perpetual variation) and on a storage possibility of 20kWh of on-board batteries (again, these data are constantly changing). On the basis of these elements, we note that the energy required to recharge the vehicle is always moderate compared to the total annual electricity consumption (consumption of electric vehicles in 2030: 10 TWh, -total grid consumption in 2016: 470 TWh), while the electricity needs can be very large (several tens to several hundreds of GW, compared to the total installed power in 2016 of 130 GW). The development of EVs seems feasible without significantly increasing electricity production, but the punctual need for energy can be problematic if judicious recharging methods are not planned. A research work carried out in Spain in 2011 [1], shows the interest of smart charging planning methods. In this work, a simple decrease in the concurrency factor can decrease the additional costs of EVs by 60-70%. On the other hand, the application of strategies that allow charging during off-peak hours could also avoid up to 35% of the necessary additional expenses. In conjunction with this limited increase in electricity consumption, it is important to highlight the storage opportunities that several million EVs will provide, and the new possibilities for controlling an increasingly intermittent electricity grid. In this paper, we will develop the implementation of a smart parking with renewable energy generation and fixed storage taking into account the constraints of the grid, while showing the difficulties encountered due to interoperability issues. The load management of an EV fleet is a subject that is widely covered in the literature, around five main axes: bi-directional charging [2], decentralised [3], mobility management [4], [5], integration of renewable energies [6], adding ancillary services to the grid [7].

Many algorithms based on different methods have been devel-

*Corresponding Author: David Roszczypala, Email: david.roszczypala@univ-nantes.fr

oped and tested in simulation for the intelligent management of energy flows in a network with multiple sources and loads. Multi-agent systems (MAS) allow the decentralisation of the energy management system based on the collaboration of different actors. Initially designed to model the collective behaviour of groups of individuals evolving together while being able to make their own choices, MAS are today frequently used for energy management systems in micro-grids with several sources [8]. The energy management of a system is usually carried out by an EMS (Energy Management System) with centralised control. The use of multi-agent systems is tantamount to decentralising energy control by giving decision-making power to different actors. Another application example is given by Taesic Kim et al. in their paper [9]. It is an application on a micro-grid composed of a set of dwellings, each with production, storage and load devices (domestic, electric car ...). An SMA solution is proposed to optimise the energy flows within the network in order to minimise peak load, reduce the energy costs of the inhabitants and maximise the energy efficiency of the storage system. Despite their wide use, the use of SMAs is controversial in the scientific community, as the results of the interaction between agents are sometimes unpredictable, therefore predicting the behaviour of the overall system from these constituent entities can become very complex.

Fuzzy logic is an extension of classical logic which allows the imprecision of data to be taken into account and thus seeks to come closer (to some extent) to the flexibility of human reasoning. It is based on Lotfi Zadeh's mathematical theory of fuzzy sets, which presents an extension of classical set theory to imprecisely defined sets. As opposed to Boolean logic, fuzzy logic allows to add a degree of truth to a condition, which allows this condition to be in a state different from true or false. In their paper [10], Kyriakarakos et al. give an example of the use of fuzzy logic for energy management in a microgrid with multi-source power supply for remote areas. The network in question consists of an AC bus on which are connected production elements : PV panels with their inverter, wind generator, hydrogen fuel cell, loads : seawater desalination module, various domestic loads, an electrolyser to create hydrogen from water, and an electrochemical storage battery connected through an AC/DC converter. Simulations have shown that the use of fuzzy logic for energy control makes it possible to respect the constraints imposed for a much smaller system dimensioning and therefore to save a lot of energy and money. Another example of fuzzy logic energy management including intelligent load management of electric vehicles is given in [11]. The performance of the algorithm is evaluated from the point of view of peak consumption and charging cost. Fuzzy logic makes it possible to implement inference systems with seamless, flexible and non-linear decision making, closer to human behaviour than conventional logic. Moreover, the rules are expressed in natural language. This has many advantages, such as including the knowledge of a non-computer expert at the heart of a decision-making system or even modelling certain aspects of natural language more finely. However, fuzzy logic has some drawbacks, such as not being able to predict that the system will behave optimally. Performance will therefore have to be measured a posteriori and adjustments will be made by trial and error. For these reasons, fuzzy logic hasn't been chosen for this specific experimentation.

Artificial neural networks (ANN) are a family of bio-inspired algorithms : their principle is borrowed from the functioning of biological neurons (in a simplistic way). They are designed to reproduce certain characteristics of biological memories by the fact that they are parallel, capable of learning, able to store information in the connections between neurons, able to process incomplete information. An example of the use of neural networks for the control of a generator in a micro-grid is given in [12]. The system is a micro-grid for the production of energy and drinking water for regions isolated from the power grid; it is composed of solar photovoltaic panels, a diesel generator, an electrochemical storage battery, a desalination unit, and conventional domestic loads. A neural network is used to control the diesel generator (On/Off and intensity) in order to minimise dependence on the generator, greenhouse gas emissions and engine wear due to incomplete combustion. Another example is given in [13], where the design and validation of an innovative control system based on an artificial neural network for a hybrid micro-grid is proposed. The use of neural networks for energy flow management is however complex to implement and may require a large computing capacity.

In this paper, we will develop a supervision method for the intelligent charging of electric vehicles based on dynamic programming. This optimization method, developed by Richard Bellman in 1957, consists in solving a global problem by breaking it down into subproblems, each with its own solution. After presenting the concept of the microgrid, this method will be developed in section III of this paper. The results obtained will be presented and analysed

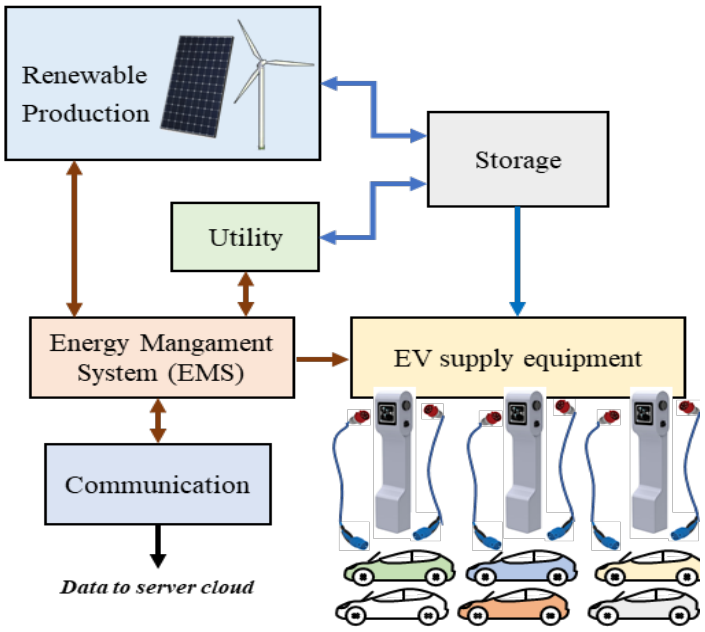


Figure 2: Experimental bench structure

2 Context and Experimental Bench

Dropbird is working on EV charging station management solutions and also on a planned charging organization in order to improve the performance of charging sites. The objective is to ensure max-

imum access to energy for EV users and to reduce the negative consequences for parking lot/network managers. This objective can manifest itself in two ways: with a given set of constraints, ensure that a maximum number of EVs can be charged or, with a given amount of EVs, minimize the strain on the power grid.

In order to carry out our studies, a flexible experimental environment has been created. It consists of an energy management system connected to multiple elements: EVSEs, DN, ESS, PV, remote control as shown in figure 2. Despite the fact that the environment is experimental, the users are various real local workers so the user satisfaction has to be respected the owned data is limited. The only data gathered on the PEV is the vehicle characteristics (battery capacity) and the initial SoC. Communication is also allowed between the EMS and the DN to receive requests of load demand response.

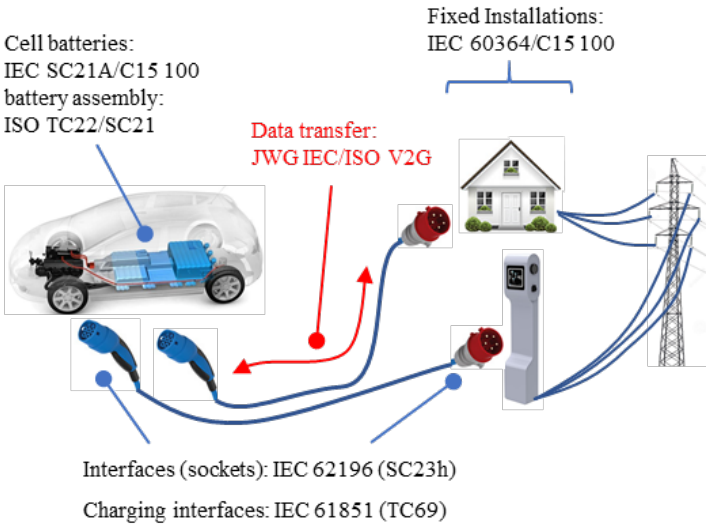


Figure 3: Normative context

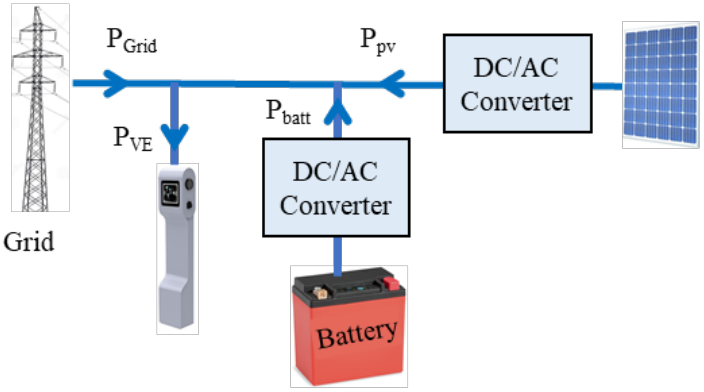


Figure 4: Energy flow within the experimental micro-grid

The development of such an experimental platform requires to meet normative constraints and standards. The main international standards associated to EV charging can be seen in figure 3.

Figure 4 sums up the energy flows exchanged by the different equipments connected to the micro-grid.

3 Algorithm structure

The energy management algorithms developed here are primarily EV recharge control algorithms; however, they are intended to control not only the EVSEs but also all elements of the system. We can speak of an EMS (energy management system) that will interact with the different elements described in figure 2: the EMS communicates with the EVSEs of course, but also with the inverter of the solar production, the fixed storage system, the distribution grid manager. The figure 5 shows the exchanges of measurements or control between the EMS and the other elements: an instantaneous measurement of the photovoltaic production, of the battery state of charge, of the charge power of each connected EV is transmitted to the EMS at each time step of the algorithm.

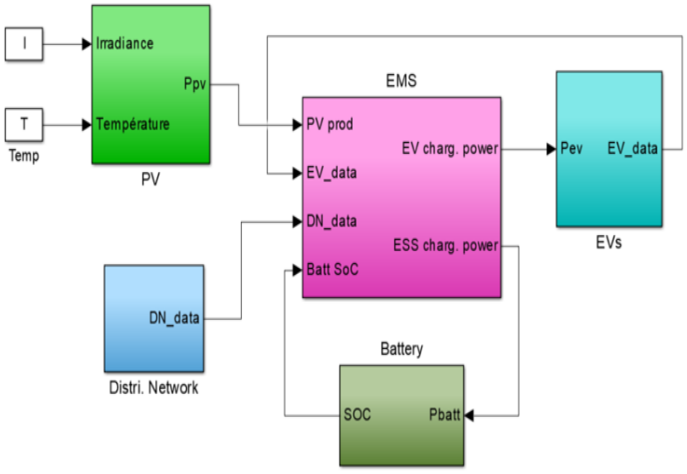


Figure 5: Information and control links

4 Dynamic Programming

Energy management in a parking lot with renewable energy sources consists in making a sequence of decisions spread over a day. The behavior of the users not being modeled, a simulation on an indeterminate day (working week or weekend) is rational. The concepts of artificial intelligence and more particularly of decision support are perfectly adapted to our problem. Enumeration (the simplest principle) is used [14].

4.1 Choosing the right energy management tool

In order to develop an independent intelligent energy management system, different strategies based on artificial intelligence were analyzed in simulation: multi-agent systems, fuzzy logic, artificial neural networks and dynamic programming.

4.2 Energy management uses

The method, defined by Bellman and used for sequential optimization problems, was developed in the 1980s to manage water reservoirs in hydroelectric installations [15]. Since the beginning of the

2000s, it has been used for the optimized use of energy in electric vehicles (such as the simultaneous management of several motors) [16]–[17], and, recently, for the management of electric energy storage from wind power and the management of a fleet of electric vehicles in [18], [19].

4.3 Principle

Bellman's principle of optimality forms the basis of the concept of dynamic programming. This principle states that "every optimal policy is composed of optimal sub-policies". The simplest (and most time-consuming) way to construct an optimal decision-making scenario is to list all possible choices at each step time and determine the least costly path. The number of possible paths can be considerably reduced using Bellman's principle by considering only the optimal subpaths at each computational step. As we elaborate the different possible paths, at each time step k, we keep only the best path that allows us to reach this state at this time: x(k). To use this method, it is necessary to discretize the time in time steps k, with 1¡k¡N, as well as each of the state variables x_1, x_2, x_3, x_4, x_5.

Dynamic programming is both a mathematical optimization method and a computer programming method. It is an algorithmic method for solving optimization problems.
Dynamic programming is used in particular for problems requiring a sequential decision-making sequence. It allows to determine the optimal trajectory by considering the problem as an optimal path to compute, and by decomposing it into sub-paths.

For example, if you want to solve the following optimization problem :

$$\min_{u_0,\dots,u_{T-1}} \sum_{t=0}^{T-1} C_t(x_t, u_t) + K(x_T) \tag{1}$$

$$x_{t+1} = f_t(x_t, u_t), \text{ for } x_0 \text{ given} \tag{2}$$

$$u_t \in U_t(x_t)$$

where

- $C_t(x, u)$ is the cost to go from t to $t + 1$ starting from the state x by applying the u command;
- $K(x)$ is the final cost for the final state x;
- f_t is the dynamic function of the system;
- $U_t(x_t)$ is the set of possible commands at time t starting from state x.

Problem can be written as follows :

$$\min_{u_0,\dots,u_{T-1}} \left\{ L_0(x_0, u_0) + \min_{u_1,\dots,u_{T-1}} \sum_{t=1}^{T-1} C_t(x_t, u_t) + K(x_T) \right\} \tag{3}$$

$$x_{t+1} = f_t(x_t, u_t) \tag{4}$$

$$x_1 = f_0(x_0, u_0) \tag{5}$$

$$u_t \in U_t(x_t)$$

Or again :

$$\min_{u_0} C_0(x_0, u_0) + V_1(f_0(x_0, u_0)) \tag{6}$$

Where $V_1(x)$ is the value for the problem starting from the state $x_1 = x$ at time $t = 1$.

The optimal value $V_{t_0}(x)$ starting from the state x at time step t can thus be written :

$$V_{t_0}(x) = \min_{u_0,\dots,u_{T-1}} \sum_{t=t_0}^{T-1} C_t(x_t, u_t) + K(x_T) \tag{7}$$

$$with \quad x_{t+1} = f_t(x_t, u_t), x_{t_0} = x$$

$$u_t \in U_t(x_t)$$

The Bellman equation appears:

$$V_T(x) = K(x) \qquad \forall x \in \mathbb{X}_T \tag{8}$$

$$V_t(x) = \min_{u_0,\dots,u_{T-1}} C_t(x_t, u_t) + V_{t+1} \circ \underbrace{f_t(x, u_t)}_{x_{t+1}} \qquad \forall x \in \mathbb{X}_T \tag{9}$$

And the optimal strategy is given by :

$$\pi_t^*(x) \in arg\ min \left\{ C_t(x_t, u_t) + V_{t+1} \circ \underbrace{f_t(x_t, u_t)}_{x_{t+1}} \right\} \forall x \in \mathbb{X}_T \tag{10}$$

5 Simulations and Results

5.1 Validation of the operation of the dynamic programming algorithm

Before carrying out complex simulations using dynamic programming algorithms, a first series of tests was performed on a simple case that is easy to grasp in order to verify the behavior of the system controlled by the dynamic programming algorithms, and to validate that its behavior seems intuitively coherent. For this purpose, a simulation with a single EV (always connected) and a fixed battery was carried out, with the parameter to minimize the cost of charging. The simulation does not take into account photovoltaic production. The electricity tariff is variable depending on the time of day and defined so that the EV is not able to charge 100% while the cost is at its lowest (consistent with the values of nominal charging power and battery capacity). Thus, to enable the EV to charge at minimum cost, the system must necessarily be controlled so that the storage battery also charges during the low cost period, and provides the EV with this low-cost energy purchased from the EV outside the off-peak period. The state of charge of the battery must also be the same at the beginning and end of the simulation.

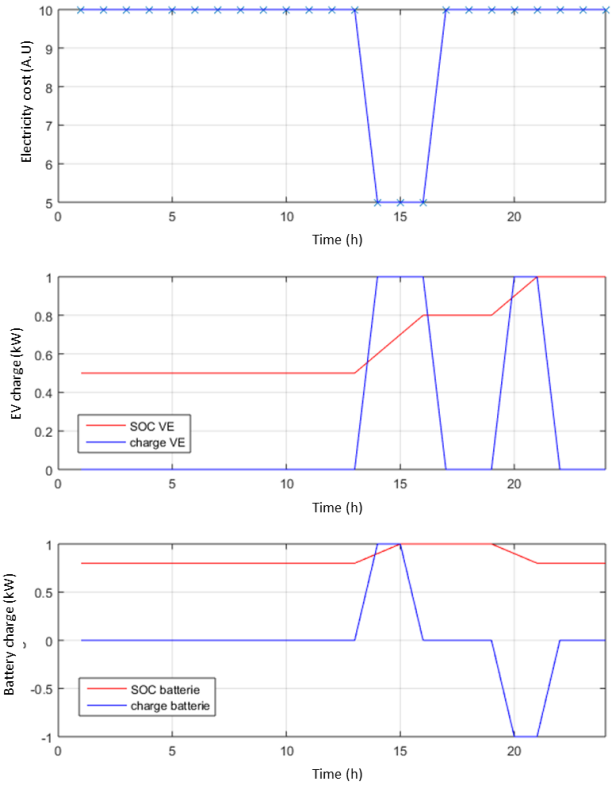


Figure 6: Test of the dynamic algorithm : after DP algorithm has been implemented, a simple intuitive case has been used to check its correct behaviour. The figure shows the battery is charged when electricity price is low according to cost function as long as it is possible

The results of the simulation are given in figure 6. The first graph describes the variation of the hourly rate of electricity for our simulation, the second graph represents the charge of the electric vehicle in blue (1 for *in charge* and 0 for *no charge*) and the variations of the state of charge of its internal battery in red. The third graph represents the charge and the variation of the state of charge for the fixed storage battery. It is immediately apparent that the algorithm controlled the system logically: the EV was charged when energy was cheapest (cost function to be minimized: total price of recharge). Since this was not enough time for the EV to receive enough energy to fully charge, the fixed battery was used well: it was also recharged when electricity was cheap, so that enough energy was stored to complete the charge of the EV outside the period of low price, and end the day with exactly the same level of charge as at the beginning. The timing of the energy delivery is random, from a cost minimization point of view, so it makes no difference whether the EV is charged with battery energy before or after the off-peak period, since the battery starts the day with a fairly high state of charge (70 percent), and no setpoint has been implemented for this action. This result allows to validate this first test of consistency of the dynamic programming algorithms.

In this paper, a simulation based on elementary rules makes it possible to judge the efficiency of the dynamic programming algorithms implemented: cases 1 and 2 are designed by implementing simply-expressed rules to control the system.

- 1. In case 1, the EV is always charged at nominal power as long as the demand exists (presence of EV and SoC less than 1), the power taken from the grid corresponds to the difference between the demand and what can be produced by the solar panels.
- 2. In case 2, the EV is also charged at constant nominal power, the fixed batteries tend to compensate for solar production (they produce if the power demanded Pve is higher than the power of the solar panels Ppv and are charged otherwise). The energy in the fixed batteries must be the same at the beginning and end of the day (battery state of charge at 70%) to justify the consistency of the simulation, and to compare with the case without fixed storage.
- 3. In the third case, a decision-making process is based on a single decision at each time step: charging EVs at rated power or not charging, which allows the cost matrix to be filled quickly (no use of the fixed battery). The cost function chosen is to minimize the total energy consumed by the network, as modeled by equation14.

$$minJ = min \sum_{k=1}^{N-1} P_{grid}(k) * cost(k) \tag{11}$$

$$= min \sum_{k=1}^{N-1} (P_{EV}(k) - P_{PV}(k)) * cost(k) \tag{12}$$

- 4. A fourth case is also added for comparison purposes, similar to case 3, but this time to minimize the cost of energy over a day, as modeled by equation12.

$$minJ = min \sum_{k=1}^{N-1} P_{grid}(k) \tag{13}$$

$$= min \sum_{k=1}^{N-1} P_{EV}(k) - P_{PV}(k) \tag{14}$$

5.2 *Simply-expressed rules (cases 1 and 2)*

For the first case, the results are given in the figure 7 : in blue, the photovoltaic production over the day, in cyan the hourly energy pricing, which was defined in collaboration with the distribution network manager in order to limit the consumption of the EVs over 3 periods corresponding to morning, midday, and evening peaks. All the electric vehicles arrive at 6:00 am and are automatically charged without planning as soon as they arrive (red curve on the figure). The charging is not optimized and does not take into account the pricing or the photovoltaic production.

For the second case, the fixed storage battery is added, allowing to store the energy produced by the solar panels in order to increase the self-consumption of renewable energy. The first graph of the figure 8 shows the use of the battery (in red, the power supplied by

the battery, and in green the state of charge). On the second graph, we find the behavior of case 1, since it is still a simulation without load planning intelligence.

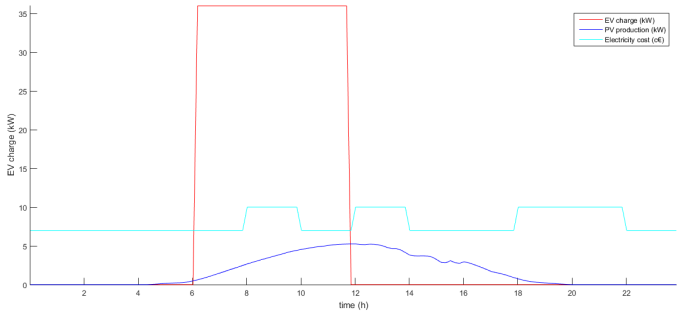


Figure 7: Reference case 1 : without battery

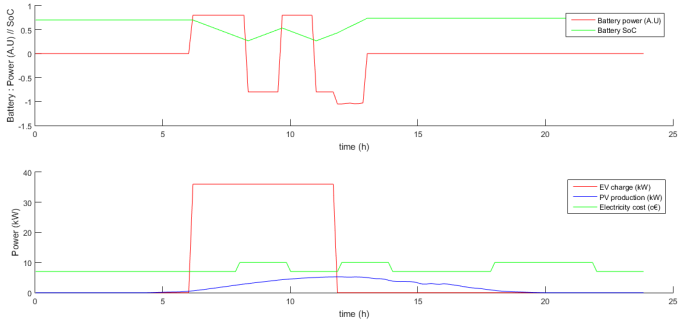


Figure 8: Reference case 2 : with battery

5.3 Minimizing energy consumption

We have now integrated the dynamic scheduling algorithm to plan the charging in an intelligent way. In the figure 9, we can see on the top graph the hourly charging, and on the bottom one, we find in green the PV production, in blue the charging power, and in red, the state of charge of the EV batteries. In this case, the algorithm is designed to minimize energy consumption, so EV charging is favored during the time slots when renewable production is at its maximum.

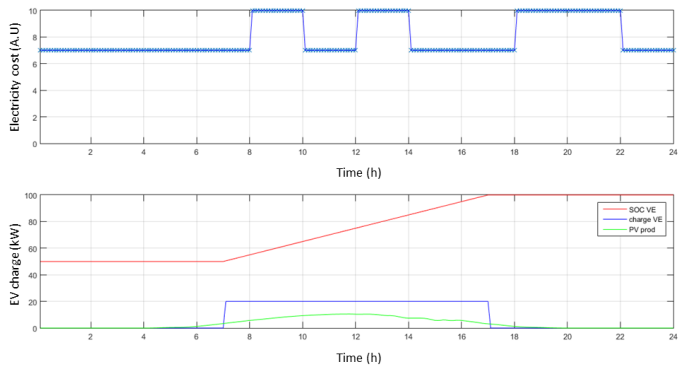


Figure 9: Results for minimal energy consumption

5.4 Minimizing charging cost

For the last case, the cost function to minimize is the total cost. Thus, we can see on the figure 10 that EV charging is planned on periods when electricity is cheaper and renewable generation is also available.

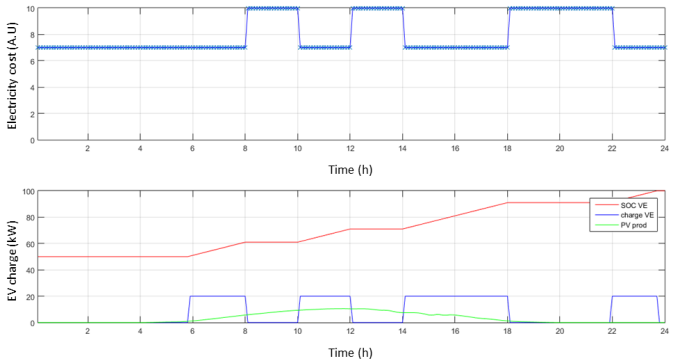


Figure 10: Results for minimal charging cost

5.5 Experimental test

The developed dynamic programming algorithm has been implemented and tested for minimizing charging cost in a real parking lot with 22kW AC EVSEs, which is illustrated on figure 11. The experimentation environment has been created and developed on the experimental site of the French distribution network operator. We also developed a supervision interface to remotely monitor the system and energy supply strategies. Figure 12 shows the web monitoring interface. Experimental data is confidential but the first experimental results are consistent with the simulations.



Figure 11: Experimental environment : part of the EVs, EVSEs and PV panels used for our study

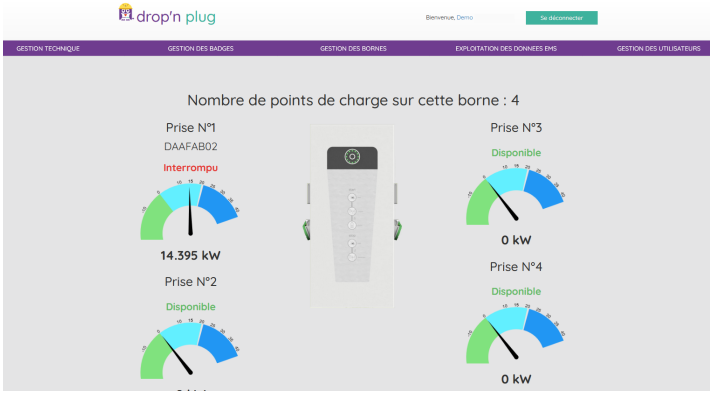


Figure 12: Screen capture of the web monitoring interface used for experimental data

6 Limits of deterministic model

Difficulties in using a deterministic model for inconsistencies with reality and planning is due to a lack of access to system data. Unlike the EV driver, the operator responsible for managing the charging stations is unable to easily visualize the SoC visible on the dashboard. Data exchanges between the EV and the EV charging infrastructure (IRVE) are governed by the IEC 61851 standard (and soon by the IEC 15118 standard). However, the access to the SoC through the EVSE will not be facilitated by EV manufacturers. This decision, political and commercial above all, limits the development of smart charging stations. The relevance of the criteria to be used for efficient load planning is discussed in [20]. The need to have access to certain data (arrival / departure of VE, SoC, etc.) for robust management algorithms is also mentioned. The unpredictable behaviour of people using ESVEs is considered a non-deterministic aspect.

6.1 Data gathering

In order to compensate for the lack of information described above, a system for collecting data from users can be set up. In the case of a fleet of electric vehicles of a company or that of people having a subscription to a public or private parking, the collaboration of the people in charge of the recharging of the electric vehicles and their agreement to provide the data necessary for the correct sequencing of the algorithm is the main drawback of this method. One solution to retrieve the data is that each user declares the EV's expected departure time, model and state of charge. Without going through a declarative process, an experiment was carried out to recover useful data for recharging EVs. The company Dropbird was tested a communication system between the EV and the EVSE through on-board diagnostics (OBD). The first tests carried out seem very promising.

6.2 Users behaviour modelling

Modelling user behavior is a key issue, whether access to the SoC is available or not. So far, all users have been modelled as arriving at the same time to charge their EV. This particular case is not realistic. Also, an optimization taking into account more realistic user behaviour was carried out. An optimized management law for a family of random cases (in the sense of "average") is preferred over the search for a list of optimal choices for a given case. An identification of the recurring behaviours of the users will be carried out from real behaviour data, which will make it possible to create a matrix of transitions between the different possible states. Each transition will be associated with a probability of passing from one state to another (present/absent) depending on the arrival time, presence time and/or state of charge.

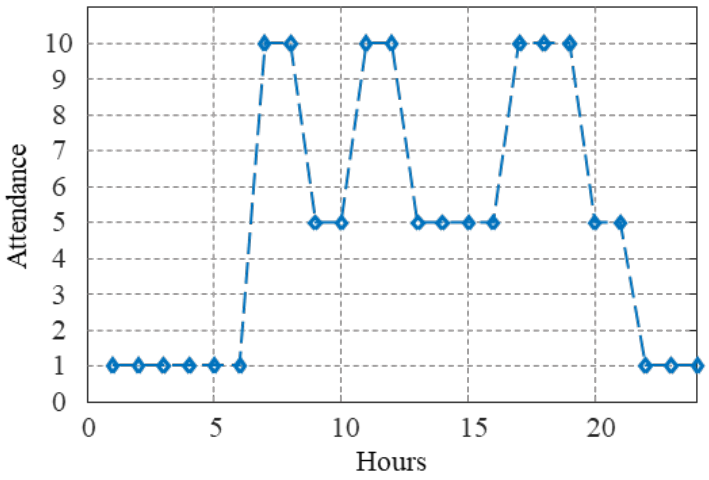


Figure 13: Attendance factor added to model to take into account the probability of a new incoming vehicle to the parking, depending on the time of the day

6.3 Users behaviour quick approximation

A study of fictitious cases was carried out with the aim of making a first approximation of user behavior. Fictitious users whose behaviour is modelled by several random variables (initial state of charge, battery capacity, arrival time) were simulated.

Fictitious users whose behaviour is modelled by several random variables were simulated: initial state of charge, battery capacity, arrival time. These variables are initialized for each EV when it arrives by picking a random value from a range of realistic values. An "attendance" variable was defined on the basis of our observations. This variable gives the probability that an EV will arrive to recharge, depending on the time of day, as illustrated to the Figure 13.

In order to observe the impact of user behaviour on the load planning and to compare the planning in cases where access to the state of charge is possible or not, a simulation was carried out on a parking lot. not having renewable energies. Figure 14 shows, for different scenarios, the distribution of the final state of charge and the final energy on board for all the electric vehicles that have used the parking lot during the day. For the same user behaviour but with different management methods, 3 simulations were performed. In Fig. 14a), the state of charge of the incoming vehicles is unknown to the parking manager, the energy is allocated first to the earliest arrived EV.

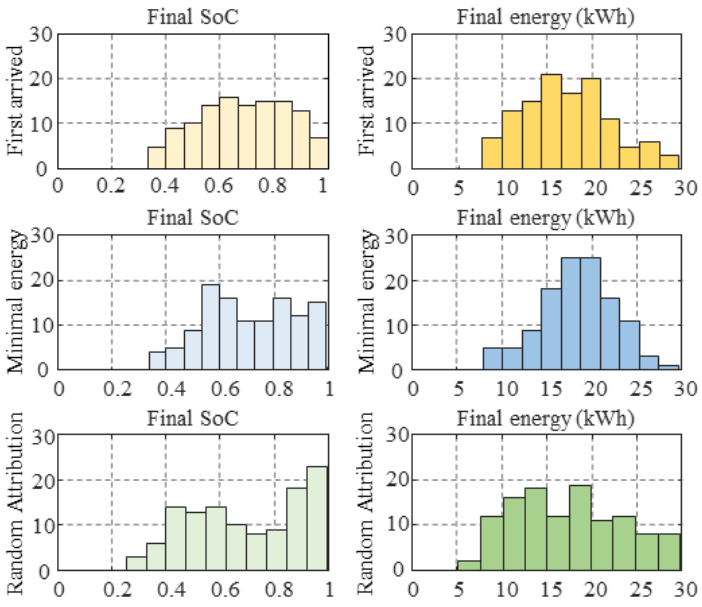


Figure 14: Number of EVs for each value of final energy and final SOC at departure time, for three different strategies of allocation of available energy : first come first served (a), minimum energy (b), random (c)

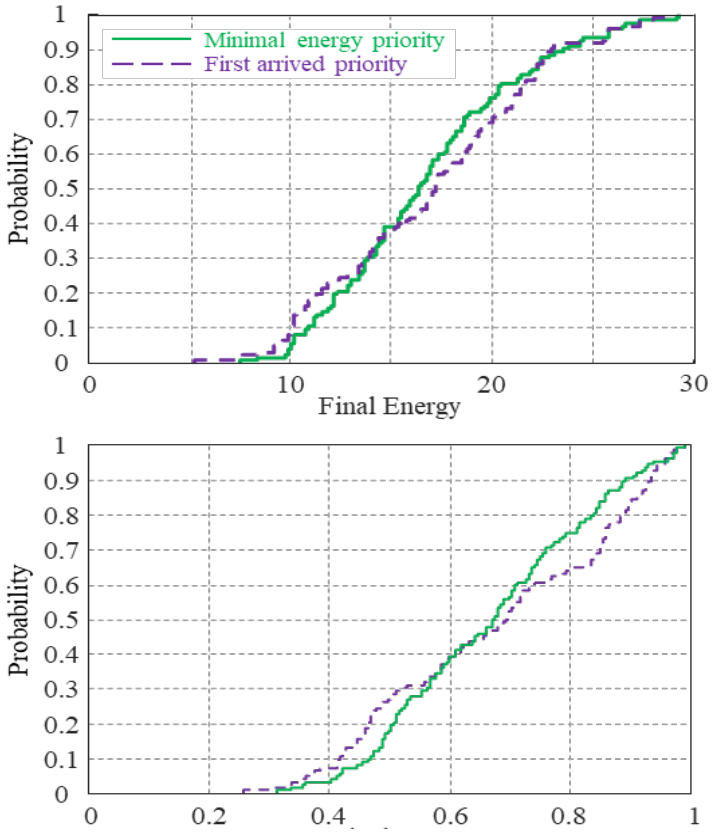


Figure 15: Cumulative probability of EV final Energy / final SOC at departure time. Cumulative distribution is given for 2 strategies : priority to first arrived or priority to lowest energy

In Figure 14 (b), it is considered that the SoC and the battery capacity of each EV are accessible. The energy is then allocated to the EV with the lowest autonomy, that is to say to the EV with the minimum stored energy. In Fig. 14c), the SoC of the incoming vehicles is unknown. Energy is assigned to the EV when it comes to connect. To allow another waiting EV to start charging, we must wait for an EV to reach full charge or for an EV start. The choice is made randomly if several EVs are on standby. The simulation of figure 14b) shows that it is possible to significantly reduce the number of EVs with final energy less than 15 kWh (the distribution is shifted to the right). Figure 15 compares more explicitly the cases illustrated in Figures 14 (a and b).

In figure 13, the criterion used is the energy available in the EVs from the car park (at the end of charging). The distribution of this data is used to qualify our load planning algorithm. This distribution is given for 2 planning algorithms: in red for the case where the energy allocation is given to the EV with the minimum energy (the SoC is accessible) and in blue for the case where the energy allocation is given on first arrival (the SoC is not accessible). A criterion corresponding to a range of approximately 100 km is set, i.e. approximately a minimum energy at the end of charging of 15kWh. For the case without access to the SoC, this value is guaranteed for 60% of the EVs, while for the case with access to the Soc, this value is guaranteed for 70% of the EVs. To ensure a minimum energy level for each EV at the end of charging, knowledge of SoC is essential.

7 Conclusion and perspectives

In this paper, the optimization of energy flows within a network of micro-parking lots with EV charging stations and energy production was carried out using dynamic programming algorithms. Their proper functioning has been first validated in simulation by considering a predetermined use. The implementation of the algorithms in a real experimental demonstrator has been made. The first results obtained validate the effectiveness of the algorithms. This is the first step in a long process, the final objective of which is the creation of an energy management system. The final objective is to propose an optimal management law for all types of car parks equipped with recharging infrastructures, taking into account their specificity. Among the work to be developed, we can cite the increasing complexity of dynamic simulation by adding degrees of freedom, but also the strengthening of the study of user behavior in order to define random cases based on reality.

References

[1] L. Pieltain Fernandez, T. Gomez San Roman, R. Cossent, C. Mateo Domingo, P. Frias, "Assessment of the Impact of Plug-in Electric Vehicles on Distribution Networks," **26**(1), 206–213, doi:10.1109/TPWRS.2010.2049133.

[2] G. B. Shrestha, B. C. Chew, "Study on the optimization of charge-discharge cycle of electric vehicle batteries in the context of Singapore," 1–7, IEEE, 2007, doi:10.1109/AUPEC.2007.4548023.

[3] Z. Ma, D. Callaway, I. Hiskens, "Decentralized charging control for large populations of plug-in electric vehicles: Application of the Nash certainty equivalence principle," 191–195, IEEE, 2010, doi:10.1109/CCA.2010.5611184.

[4] S. Deilami, A. S. Masoum, P. S. Moses, M. A. S. Masoum, "Real-Time Coordination of Plug-In Electric Vehicle Charging in Smart Grids to Minimize Power Losses and Improve Voltage Profile," **2**(3), 456–467, doi: 10.1109/TSG.2011.2159816.

[5] D. Dallinger, D. Krampe, M. Wietschel, "Vehicle-to-Grid Regulation Reserves Based on a Dynamic Simulation of Mobility Behavior," IEEE Transactions on Smart Grid, **2**(2), 302–313, 2011, doi:10.1109/TSG.2011.2131692.

[6] A. Y. Saber, G. K. Venayagamoorthy, "Resource Scheduling Under Uncertainty in a Smart Grid With Renewables and Plug-in Vehicles," **6**(1), 103–109, doi:10.1109/JSYST.2011.2163012.

[7] E. Sortomme, M. A. El-Sharkawi, "Optimal Scheduling of Vehicle-to-Grid Energy and Ancillary Services," IEEE Transactions on Smart Grid, **3**(1), 351–359, 2012, doi:10.1109/TSG.2011.2164099.

[8] J. Lagorse, D. Paire, A. Miraoui, "A multi-agent system for energy management of distributed power sources," **35**(1), 174–182, doi:10.1016/j.renene.2009.02.029.

[9] T. Kim, J. Yun, W. Qiao, "A multiagent system for residential DC microgrids," in 2015 IEEE Power & Energy Society General Meeting, 1–5, IEEE.

[10] G. Kyriakarakos, A. I. Dounis, K. G. Arvanitis, G. Papadakis, "A fuzzy logic energy management system for polygeneration microgrids," **41**, 315–327, doi: 10.1016/j.renene.2011.11.019.

[11] N. J. Schouten, M. A. Salman, N. A. Kheir, "Energy management strategies for parallel hybrid vehicles using fuzzy logic," **11**(2), 171–177.

[12] A. Al-Alawi, S. M Al-Alawi, S. M Islam, "Predictive control of an integrated PV-diesel water and power supply system using an artificial neural network," **32**(8), 1426–1439, doi:10.1016/j.renene.2006.05.003.

[13] Y. Ates, O. Erdinc, M. Uzunoglu, B. Vural, "Energy management of an FC-UC hybrid vehicular power system using a combined neural network-wavelet transform based strategy," **35**(2), 774–783, doi:10.1016/j.ijhydene.2009.11.021.

[14] Louis Frecon, Manuel d'intelligence artificielle, presses polytechniques et universitaires romandes edition, 2009.

[15] Yakowitz, Dynamic programming applications in water resources ». Water Resources Research, water resources research edition, 1982.

[16] X. Zeng, J. Wang, "A Parallel Hybrid Electric Vehicle Energy Management Strategy Using Stochastic Model Predictive Control With Road Grade Preview," IEEE Transactions on Control Systems Technology, **23**(6), 2416–2423, 2015, doi:10.1109/TCST.2015.2409235.

[17] D. Scaradozzi, M. Fanesi, "Advanced Control Strategies to Improve Nonlinear Automotive Dynamical Systems Consumption," Axioms, **8**, 123, 2019, doi:10.3390/axioms8040123.

[18] P. Haessig, Dimensionnement et gestion d'un stockage d'énergie pour l'atténuation des incertitudes de production éolienne, Ph.D. thesis, École normale supérieure de Cachan-ENS Cachan, 2014.

[19] R. L. G. Latimier, Gestion et dimensionnement d'une flotte de véhicules électriques associée à une centrale photovoltaïque: co-optimisation stochastique et distribuée, Ph.D. thesis, Université Paris-Saclay, 2016.

[20] K. N. Kumar, B. Sivaneasan, P. L. So, "Impact of Priority Criteria on Electric Vehicle Charge Scheduling," **1**(3), 200–210, doi:10.1109/TTE.2015.2465293.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 161-174 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Large Empirical Study on Automatically Classifying Software Maintainability Concerns from Issue Summaries

Celia Chen[*,1], Michael Shoga[2]

[1]*Department of Computer Science, Occidental College, Los Angeles, 90041, United States of America*

[2]*Center for Systems and Software Engineering, Department of Computer Science, University of Southern California, Los Angeles, 90089, United States of America*

ARTICLE INFO

Article history:
Received: 22 December, 2020
Accepted: 19 February, 2021
Online: 10 March, 2021

Keywords:
Software Maintainability
Open Source Software
Empirical Study
Linguistic Analysis

ABSTRACT

Software maintenance contributes the majority of software system life cycle costs. However, existing approaches with automated code analysis are limited by accuracy and scope. Using human-assessed methods or implementing quality standards are more comprehensive alternatives, but they are much more costly for smaller organizations, especially in open-source software projects. Instead, bugs are generally used to assess software quality, such as using bug fixing time as an estimate of maintenance effort. Although associated bug reports contain useful information that describe software faults, the content of these bug reports are rarely used. In this paper, we incorporate quality standards with natural language processing techniques to provide insight into software maintainability using the content of bug reports and feature requests. These issues are classified with an automated approach into various maintainability concerns whose generalizability has been validated against over 6000 issue summaries extracted from nine open source projects in previous works. Using this approach, we perform a large empirical study of 229,329 issue summaries from 61 different projects. We evaluate the differences in expressed maintainability concerns between domains, ecosystems, and types of issues. We have found differences in relative proportions across ecosystem, domain and issue severity. Further, we evaluate the evolution of maintainability across several versions in a case study of Apache Tomcat, identifying some trends within different versions and over time. In summary, our contributions include a refinement of definitions from the original empirical study on maintainability related issues, an automated approach and associated rules for identifying maintainability related quality concerns, identification of trends in the characteristics of maintainability related issue summaries through a large-scale empirical study across two major open source ecosystems, and a case study on changes in maintainability over versions in Apache Tomcat.

1 Introduction

Software maintainability measurements provide organizations with a greater understanding of how difficult it is to repair or enhance their software. The importance of having this understanding is underscored in [1], which reported that 75-90% of business and command&control software and 50-80% of cyber-physical system software costs are incurred during maintenance. In addition, maintainability serves as a crucial link to other quality characteristics. In [2], the author lists maintainability as a contributing quality to life cycle efficiency, changeability and dependability. In [3], the maintainability is a key quality in understanding software quality interrelationships. Thus, having comprehensive knowledge of software maintainability is significant in the software development and maintenance process.

A number of metrics and approaches have been developed to provide ways to measure and evaluate software maintainability. In this study, they are classified into the following categories:

- Automated analysis: Automatic analysis involves analyzing source code or other software artifacts and quantifying software maintainability into numeric results. This includes static code analysis such as measuring Maintainability Index, technical debt, code smells, and other Object-Oriented metrics [4, 5], as well as bug-focused metrics such as bug fixing time [6] and accumulated defect density [7].

[*]Corresponding Author: Celia Chen, Email: qchen2@oxy.edu

- Human-assessed analysis: Human-assessed analysis includes reuse cost models that estimate maintainability of potentially reusable components based on human-assessed maintainability aspects, such as code understandability, documentation, and developer self-reported surveys [8, 9].
- Software ontology, standards and frameworks: These introduce immense high-level knowledge, which are mostly coming from consensus wisdom, professional discipline and expert sources. This includes standards such as the ISO/IEC 25010 and Software Engineering Body of Knowledge.

While the automated analysis metrics are easy to use and often require relatively low human effort, in [10] author points out that the effective use of accuracy measures for these metrics has not been observed and there is a need to further validate maintainability prediction models. Moreover, despite having the advantage of identifying the particular parts of the software most needing maintainability improvement at the module and method level, they do not provide an overall quality status for the current version of the software.

Although bug fixing times may reflect maintenance effort [6], these bug-focused metrics also do not provide a systematic understanding of software maintainability. Furthermore, these metrics do not utilize the information provided by the natural language descriptions due to their unstructured nature.

On the other hand, human-assessed analyses are able to more accurately reflect maintenance effort, yet they are limited in use due to cost and subjectivity based on developers' skills and experience [8, 11].

Software ontology, standards and frameworks tend to be used in larger organizations as guidelines during the development process. They provide insightful knowledge in understanding, evaluating, and improving a system's maintainability planning, staffing, and preparation of technology for cost-effective maintenance. However, it is very difficult to enforce standards on actual program behavior. Moreover, while standardizing the process can help make sure that no steps are skipped, standardizing to an inappropriate process can reduce productivity, and thus leave less time for quality assurance. Especially in smaller organizations and open source ecosystems, it is extremely difficult to apply and enforce these paradigms due to their limited resources and functionality-focused nature.

To provide a way to effectively measure and keep track of the overall maintainability while involving relatively low human effort, we utilize bug report information in conjunction with a software maintainability ontology to assess software maintainability at the system level in an initial empirical study [12]. By manually mapping over 6000 bug reports to maintainability subgroup software qualities (SQs) in the ontology, we validated the approach to evaluate overall system maintainability. However, this approach is limited by the amount of manual effort needed for mapping the bug reports. To overcome the high effort requirements, we incorporate natural language processing techniques to automatically classify "issue summaries," which include the descriptions of bug reports and feature requests, to the maintainability subgroup SQs. In this paper, we provide a refinement of definitions from the original empirical study on maintainability related issues and the rule set. We expand upon the scale of the analysis done in [12], made possible by the fuzzy classifier, to identify trends in maintainability related issue summaries from two major open-source software ecosystems. We further perform an in-depth case study on the maintainability changes over versions and time in Apache Tomcat. In total we classify 229,329 issue summaries from 61 projects and the trends over 7 versions and 20 years in Apache Tomcat.

The rest of this paper is organized as follows. Section 2 summarizes related work and presents the differences of those compared to our study. Section 3 describes the background of the automated approach and introduces the research questions and design of an empirical study on maintainability trends in two open-source software ecosystems. Section 4 discusses the results, analysis and implications. Section 5 concludes the study.

2 Related Work

2.1 Software Maintainability Measurement

Maintainability Index (MI) is the most widely used metric to quantify maintainability in software projects. Since its introduction in 1992 [13], several variations have been developed [14, 15]. While it is widely used, the metric's effectiveness has been brought into question and several shortcomings identified [16].

Other approaches to measuring maintainability have incorporated other metrics as well as frameworks and ontology. In [17], the author provided an overview of an approach that uses a standardized measurement model based on the ISO/IEC 9126 definition of maintainability and source code metrics. These metrics include volume, redundancy, complexity and more.

In [18], the author investigated 11 different types of source code metrics in an empirical study to develop a maintainability prediction model for Service-Oriented software and compare their model with the Multivariate Linear Regression (MLR) and Support Vector Machine (SVM) approaches. They found that using a smaller set of source code metrics performed better than when they used all of the available metrics.

Approaches utilizing machine learning have also been proposed. In [19], the author conducted a comparative study on using machine learning algorithms for predicting software maintainability on two commercial ADA datasets. They examined Group Method of Data Handling, Genetic Algorithms, and Probabilistic Neural Network with Gaussian activation function for predicting a surrogate maintenance effort measure, the number of lines of code changed per class over a three year maintenance period. Their results showed improvement over previously reported models.

In [20], the author proposed an LSTM algorithm for software maintainability metrics prediction. They considered 29 OO metrics and applied their approach on a large number of open source projects. In addition to comparing against other machine learning algorithms, they also used FSS to determine which metrics are most relevant for maintainability prediction.

In [21], the author presented a study using several classifiers to evaluate maintainability at the class level using the output of different static analysis tools. In their approach, ConQAT, Teamscale, and Sonarqube are used to extract metrics such as SLOC, average method length, clone coverage, etc. The classifiers are trained using

expert-labeled data from three different systems. Their best results provided a classification accuracy of 81% and a precision of 80%.

2.2 Bug Characteristics Analysis with Natural Language Processing

Several studies have investigated the characteristics of bugs and bug reports through the use of natural language processing.

In [22], the author collected 709 bugs including security related and concurrency bugs. They analyzed the characteristics of those bugs in terms of root causes, impacts and software components. Their findings reveal characteristics of memory bugs, semantic bugs, security bugs, GUI bugs, and concurrency bugs. They verified their analysis results on the automatic classification results by using text classification and information retrieval techniques.

In [23], the author proposed an approach to binary classification of bug reports into 'bug' and 'nonbug' by leveraging text mining and data mining techniques. Analyzing the summary and some structured features including severity, priority, component, and operating systems, they use Bayesian Net Classifier as the machine learner. They performed an empirical study of 10 open source projects to validate their method and provide a MyLyn plugin prototype system that will classify given reports.

In [24], the author analyzed bug reports from nine systems and found that a large percentage of bug reports lack Steps to Reproduce (S2R) and Expected Behavior (EB) information. They in turn developed an automated approach to detect missing S2R and EB from bug reports. They produced three versions using regular expressions, heuristics and natural language processing, and machine learning. They found their machine learning version to be the most accurate with respect to F1 score, but the regular expressions and heuristics and natural language processing approaches had similar accuracy results without training.

In [25], the author constructed models for identifying security and performance related bug reports utilizing feature selection, random under-sampling, and Naive Bayes Multinominal approach. They evaluated their approach on datasets of bug reports from four software projects, achieving average AUC values of 0.67 and 0.71 for their security and performance models respectively.

Summing up, here is how our work differs from the existing studies: with regard to measurement of maintainability, our work enables study of maintainability evolution with relatively low cost. By using issues, preexisting software artifacts, it allows for expert knowledge to be applied to open source software systems wherein there is less control over development and maintenance tasks.

3 Empirical Study

3.1 Background

This section presents the software maintainability ontology used and an extension of the SQ definitions provided in [12]. It also provides a brief summary on our previous works and the overall automated approach.

3.1.1 Software Maintainability Ontology Background

The ontology provided in [2] presents maintainability as depending on two alternative SQs, repairability and modifiability, which handle defects and changes respectively. These SQs are further enabled by several subgroups. The automated approach focuses on maintainability in the context of these mean-ends SQs as shown in Figure 1.

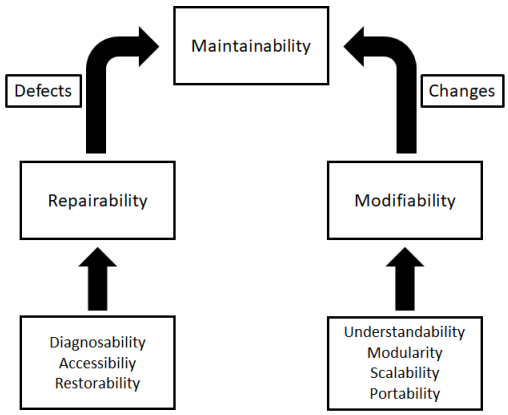


Figure 1: Software maintainability ontology hierarchy

The following are the refined definitions for each subgroup SQ to better capture the scope of these quality concerns.

Repairability involves handling of defects in software. It is enabled by the following SQs:

- ***Diagnosability***:

  Diagnosability is the characteristic of being diagnosable. It is the property of a partially observable system with a given set of potential faults, which can be detected with the certainty given finite observation. Issues that affect this SQ involve problems with lack of logging and diagnosability management, faulty error messages and the process of tracing where they originate, failure of tests, and insufficient information provided for accurate assessments [26]–[28].

- ***Accessibility***:

  Accessibility [29] generally describes the ability of a software system to accommodate people with special needs. This requires a software system to be suitable for most of the potential users without any modifications and be easily adaptable to different users with adaptable and customized user interfaces.

  Another definition for accessibility is at the architecture level. The JCIDS manual [30] defines the Accessibility of Architectures as the ability to grant access to authorized users in a timely fashion in order to "support architecture-based analysis and decision making processes." In this paper, accessibility is defined as the quality of being available and reachable, which involves whether the intended areas of a software system can be accessed as desired. Issues that affect this SQ prevent authorized users from accessing data or functions due to things such as redirects to unintended locations, broken links to intended areas, and incorrect user permission and authorization.

- ***Restorability***:

  Restorability describes the ability of a software system to restore to a previous state. Issues that affect this SQ include activities such as clearing of caches, refreshing settings, proper removal of data and backups of the current system.

Modifiability involves handling of software changes. It is enabled by the following SQs:

- ***Understandability***:

  Software understandability can be considered in the context of source code as well as non-source code artifacts and further depends on the person assessing the software. This may include the level of experience and familiarity with the software's code base if considering a developer's perspective or whether or not the software is clear in its usage and applicability if considering an end user's perspective. Understandability can have an impact on maintenance tasks especially in cases where the original developers are not the ones responsible for maintaining the system. Further explanation of software understandability is provided in [31].

  Issues that affect this SQ involve activities such as system enhancement, lack of explanations and comments, confusing or inaccurate descriptions, presence of deprecated software and more.

- ***Modularity***:

  Modularity involves separation of code into modules. It indicates the degree to which a system's components are made up of relatively independent components or parts which can be combined [32, 33].

  Issues that affect this SQ involve unwanted interactions between different modules and separation of one module into multiple modules.

- ***Scalability***:

  Scalability is the ability of a system to continue to meet its response time or throughput objectives as the demand for the software functions increases [34, 35]. Issues that affect this SQ involve latency in functionality, hangs, and insufficient resources for functionality to scale up or down.

- ***Portability***:

  Portability refers to the ability of a software unit to be ported to a given environment and being independent of hardware, OS, middle-ware, and databases [36, 37]. Issues that affect this SQ prevent proper interfacing between software components and external platforms.

3.1.2 Background Studies

In our previous empirical study [12], we manually analyzed 6372 bugs found in the Mozilla community. By categorizing them into one of the subgroup SQs described above, we identified various trends in maintainability changes as software evolves and the relationships between these subgroup SQs. The findings were valuable but it was difficult to scale up the study due to the large amount of manual effort required to produce such mappings between bug reports and subgroup SQs.

Thus, a manual analysis on the ground-truth dataset[1] was first performed, and we identified three types of linguistic patterns from bug reports: lexical patterns, syntax patterns and semantic patterns. These patterns illustrate the recurrent linguistic rules that users are likely to use when reporting bugs or requesting new features. Motivated by these heuristic linguistic patterns, we proposed a fuzzy classifier [31, 38] that aims to identify the maintainability subgroup SQ concerns expressed in issue summaries. Based on the definitions of these patterns, a set of 24 initial fuzzy rules was generated by heuristically identifying them from subgroup SQ definitions and practice guidelines. To improve this initial fuzzy rule set, an incremental approach was constructed to identify potential new rules from issue summaries mined from four open-source projects. The rule performance was used to determine whether the existing rule set should be updated. As a result, we obtained a final set of 99 rules[2]. To evaluate the generalizability of the obtained rule set, we evaluated it on projects that were not used in generating the rules. All metrics (accuracy, precision, recall, and f-measure) had an average above 0.8, indicating that the rule set is able to perform well in classifying issue summaries with all of the subgroup SQs. Thus, with such an automated classifier that can identify maintainability concerns expressed in issue summaries, we conduct a large empirical study to investigate the trends of maintainability across 61 open-source software projects and over 200,000 issue summaries.

3.2 Research Questions

To explore the characteristics of maintainability, we look to answer the following research questions:

- RQ1: How are software maintainability concerns expressed in different domains and ecosystems? For this RQ, each project is classified as one of the following:
  - Applications: these projects are designed to have some sort of direct interaction with general users [39]. Examples of these projects include web browsers, email clients, and office suites.
  - Infrastructure: these projects are not designed to interact with users directly. Instead, they provide facilities and services for other software to run [39]. Examples of these projects include build tools, web servers, and libraries.

  To answer this RQ, we analyze the differences for subgroup SQs between these domains as well as between Apache and Mozilla projects.

- RQ2: How are software maintainability concerns expressed across different types of issues? For this empirical study, we report on the following characteristics:

[1] Dataset can be found: https://bit.ly/2WhtWJx

[2] The final rule set can be found: http://bit.ly/3az2CRy

- Won't Fix: Issues whose resolutions are WONTFIX have been classified such that they are not planned to be fixed. This can be for a variety of reasons, such as when the issue involves an unsupported method or tool, or when the issue is not worth the cost [40].
- Reopened: Issues that have been previously closed can be reopened in cases such as when new reproducibility information is reported, previous root causes are identified as misunderstood, reports with insufficient information are updated, or the priority of the issue has been increased [41].
- Unresolved: Some issues are left unresolved without updates. To investigate these unresolved issues, we filter the lists to identify issues whose status is not resolved or closed, and whose last changed or updated date is more than a year from June 30, 2020. This date is based on the point up to when the issues were collected.
- Severity: Issues are often classified depending on their impact, with most of the projects defining them as blocker, critical, major, etc. The MozillaWiki defines severity in terms of levels: S1, S2, S3, S4 for catastrophic, serious, normal, and small/trivial respectively; however, the previously mentioned descriptors are used more commonly. Thus for this RQ, the severities are defined as follows:
  - Blocker: Blocker, S1
  - Critical: Critical
  - Major: Major, S2
  - Others: All other categories

- RQ3: How does software maintainability change as software evolves?

  To gain a better understanding of how maintainability changes as software evolves, we look to the issues of Apache Tomcat. This project has been selected as it has a long history: the Apache Bugzilla contains issues from Tomcat 3 to Tomcat 9, and it has versions separated to the patch level of granularity. To answer this RQ, we look at the data in three ways: by major version, within major version by year, and by year overall. The last updated date in the Apache Tomcat Archive is used to map each patch to a year. Evaluation by changes in minor version is not done as many versions have at most 2 minor versions for a given major version. In cases where versions have patch variants, such as release candidates or betas, the versions are combined. For example, Tomcat 4 has issues of version 4.0 Beta 1, Beta 2, etc. These are combined with the release candidates, milestones, and final version issues into a single 4.0.0 category.

3.3 Study Design

3.3.1 Study Subjects

This empirical study focuses on projects found within the Mozilla and Apache ecosystems. Table 1 provides the characteristics of the projects chosen for this study. Some projects are filtered out of the study subjects. From Mozilla, projects from other and graveyard are excluded from the study as they contain many projects that do not focus on software. From Apache, projects that contain fewer than 100 issues are excluded. Apache OpenOffice and Apache SpamAssassin have their own Bugzilla repositories which are included with the other Apache projects.

3.3.2 Data Extraction and Analysis

Issue summaries from the selected projects are downloaded from their respective Bugzilla repositories along with the issue characteristics such as version, Open Date, etc. The issues are then classified as described in Section 3.1.2. They are then separated according to the criteria described in the RQs. Issues that are identified as invalid or duplicates are filtered out to avoid over-counting. For each SQ, the overall proportion is calculated from the number of expressing issues over the total number of issues to correct for differences in the number of issues reported between groups. Relative proportion is calculated from the number of expressing issues over the total number of issues that express any maintainability concern to compare how much each subgroup SQ contributes to overall maintainability.

In total, 229,329 issues are analyzed and classified as relating to one of the maintainability subgroup SQs or as non-maintainability. Of these, 82,577 (36%) are maintainability related and 146,752 (64%) are non-maintainability related. Figure 2 shows the relative and overall proportions of each of the maintainability subgroup SQs. The most prevalent maintainability subgroup SQs are understandability, portability, and accessibility.

4 Results and Discussion

4.1 RQ1

Of the 229,329 issues analyzed and classified, 180,706 come from Mozilla systems and 48,623 from Apache systems.

Table 2 complies the number of issues expressing each SQ, the overall proportion of each SQ over the total number of issues, and the relative proportion of each SQ over the total number of maintainability issues.

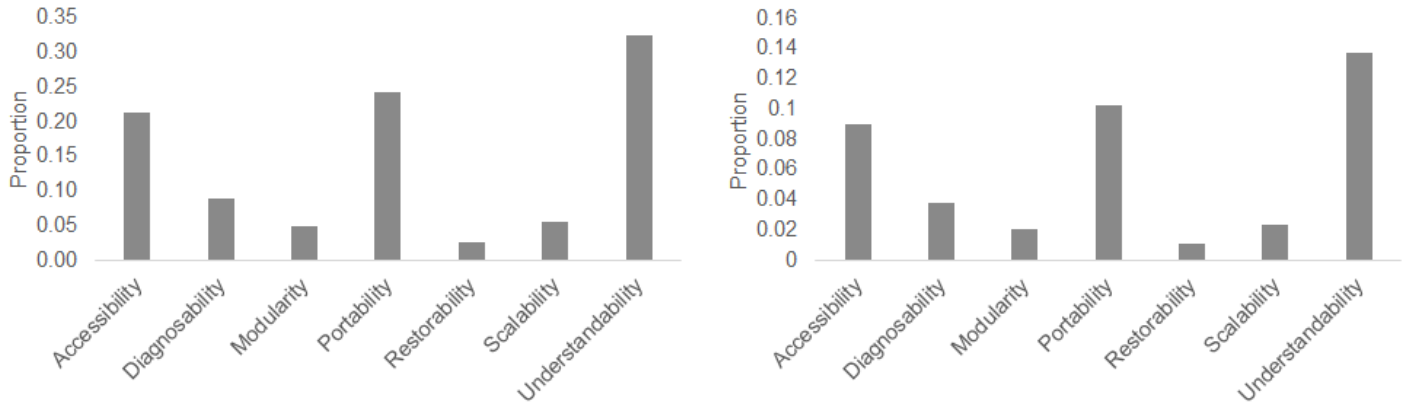
Figures 3a and 3b show the proportion of issue summaries that express maintainability concerns across the chosen ecosystems and domains respectively. There is not a large difference between the proportion of maintainability issues when considering ecosystem or domain.

4.1.1 Overall Proportions

A MANOVA is performed to examine whether there is a significant association between the overall proportion of maintainability and each maintainability subgroup SQ with ecosystems or domains. More specifically, whether ecosystems or domains have a significant effect on the overall proportions of maintainability and each maintainability subgroup SQ. While there is a significant association found between ecosystem and the overall proportions of maintainability and its subgroup SQs, $F = 19.136$, $df = 8$, $p=0.014$ (<0.05), univariate analysis does not identify statistically significant SQs that contribute to the differences between ecosystems. In addition,

Table 1: Study subject characteristics

Apache			Mozilla		
Project	Number of Issues	Domain	Project	Number of Issues	Domain
Ant	6144	Infrastructure	Bugzilla	10000	Infrastructure
Apache httpd-1.3	898	Infrastructure	bugzilla.mozilla.org	10000	Application
Apache httpd-2	8288	Infrastructure	Calendar	10000	Application
APR	818	Infrastructure	Chat Core	1295	Infrastructure
Batik	1029	Infrastructure	Cloud Services	9932	Application
BCEL	168	Infrastructure	Conduit	1713	Infrastructure
Fop	2170	Application	Core	10000	Infrastructure
JMeter	4609	Infrastructure	Data Platform and Tools	4251	Application
Lenya	1449	Application	DevTools	10000	Infrastructure
Log4j	1387	Infrastructure	Directory	726	Infrastructure
OpenOffice	10000	Application	External Software Affecting Firefox	1567	Application
POI	4667	Infrastructure	Firefox	10000	Application
Regex	102	Infrastructure	Firefox Build System	10000	Infrastructure
Security	252	Infrastructure	Firefox for Android	10000	Application
Slide	432	Application	Firefox for iOS	7067	Application
Spamassassin	7713	Application	Gecko View	2523	Infrastructure
Taglibs	764	Infrastructure	Instantbird	1709	Application
Tomcat Connectors	804	Infrastructure	JSS	455	Infrastructure
Tomcat Modules	187	Infrastructure	MailNews Core	10000	Infrastructure
Tomcat Native	178	Infrastructure	Mozilla Localizations	10000	Application
Tomcat 3	1129	Infrastructure	NSS	10000	Infrastructure
Tomcat 4	3374	Infrastructure	Remote Protocol	650	Infrastructure
Tomcat 5	3118	Infrastructure	SeaMonkey	10000	Application
Tomcat 6	1386	Infrastructure	Socorro	8822	Infrastructure
Tomcat 7	1643	Infrastructure	Testing	10000	Infrastructure
Tomcat 8	1213	Infrastructure	Testopia	920	Infrastructure
Tomcat 9	486	Infrastructure	Thunderbird	10000	Application
XercesJ	426	Infrastructure	Toolkit	10000	Infrastructure
Xindice	163	Infrastructure	Tree Management	6673	Infrastructure
			Web Compatibility	4603	Application
			WebExtensions	8344	Infrastructure
			WebTools	5682	Infrastructure



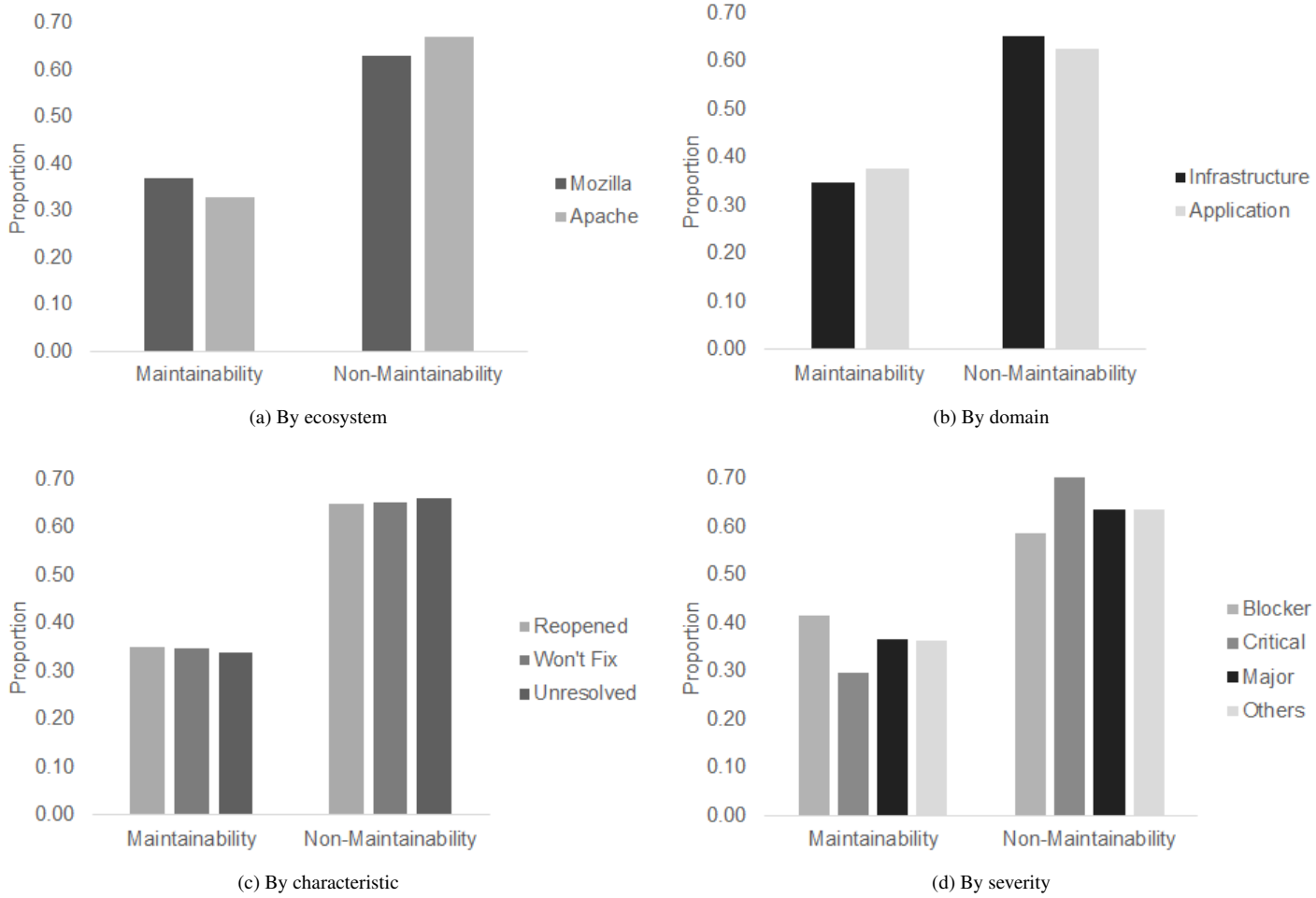
(a) Relative proportion

(b) Overall proportion

Figure 2: Proportions of subgroup SQ concerns across all issues

Table 2: SQ issue counts (#), overall proportions (O), and relative proportions (R) by various types and classifications

Type		Acc.			Dia.			Mod.			Por.			Res.			Sca.			Und.		
		#	O. %	R. %	#	O. %	R. %	#	O. %	R. %	#	O. %	R. %	#	O. %	R. %	#	O. %	R. %	#	O. %	R. %
Ecosystem	*Apache*	4506	9	25	1264	3	7	669	1	4	4021	8	22	434	1	2	1292	3	7	6143	13	34
	Mozilla	16191	9	21	7361	4	9	4079	2	5	19482	11	25	2088	1	3	4057	2	5	25201	14	32
Domain	*Application*	9586	10	22	2132	2	5	1951	2	4	12553	13	29	1116	1	3	2090	2	5	14258	14	33
	Infrastructure	11111	8	21	6493	5	12	2797	2	5	10950	8	21	1406	1	3	3259	2	6	17086	13	32
Severity	*Blocker*	572	12	24	112	2	5	117	2	5	910	19	38	43	1	2	128	3	5	507	10	21
	Critical	1550	9	24	189	1	3	369	2	6	2567	14	40	139	1	2	1072	6	17	556	3	9
	Major	2152	14	32	322	2	5	339	2	5	1936	13	29	221	1	3	581	4	9	1122	7	17
	Others	16423	9	20	8002	4	10	3923	2	5	18090	9	22	2119	1	3	3568	2	4	29159	15	36
Other Resolutions	*Reopened*	114	9	23	107	9	21	23	2	5	88	7	18	19	2	4	24	2	5	127	10	25
	Won't-Fix	1392	8	20	484	3	7	348	2	5	2017	12	29	185	1	3	305	2	4	2123	13	31
	Unresolved	1623	9	22	941	5	13	381	2	5	1450	8	20	253	1	3	492	3	7	2295	12	31



(a) By ecosystem

(b) By domain

(c) By characteristic

(d) By severity

Figure 3: Distributions of maintainability versus non-maintainability issues

no statistically significant results are found for the overall proportions of maintainability and its subgroup SQs in terms of different domains, *F = 13.036, df=8, p=0.192 (>0.05).*

4.1.2 Relative Proportions

Pearson's Chi-squared tests are performed to examine whether the ecosystems or the domains are associated with the distributions of the relative proportion of each SQ. Overall, the distributions of the relative proportion of each SQ differ significantly in domains, χ^2 (6, N = 61) = 2226.5, p < 0.001; and also in ecosystems, χ^2 (6, N = 61) = 2921.2, p < 0.001.

As shown in Table 2, understandability, portability, and accessibility are most prevalent expressed concerns. When considering ecosystem, these SQs comprise the majority for both Mozilla and Apache; however, the Apache systems tend to express more accessibility concerns than portability concerns. In considering domain, the same three SQs are the most prevalent; however, for infrastructure type software, there tend to be more diagnosability issues and fewer relating to portability when compared to application type software.

Summary of RQ1: To summarize, in this study, there is a statistically significant association between ecosystem and the overall proportions of maintainability and its subgroup SQs. The distributions of relative proportions of subgroup SQs differ significantly between ecosystems and domains; there is a trend in application software which tends to have more portability issues and fewer diagnosability issues compared to infrastructure software. As application type software is targeted toward end-users, there may be a larger variety of use cases which would necessitate compatibility with other software. In contrast, infrastructure software is targeted toward developers. In this case, they may place more importance on being able to diagnose issues with the software and may already be aware of incompatibilities with other software.

4.2 RQ2

4.2.1 RQ2a: Reopened, Won't-fix and Unresolved

Of the 229,329 issues classified, 1,251 are marked as REOPENED, 16,909 are marked as WONTFIX, and 18,809 are identified as unresolved. Figure 3c shows the proportion of issue summaries that expressed maintainability concerns across these three categories. Similar to the overall proportion, these categories are comprised of about 35% maintainability issues.

Table 2 compiles the number of issues expressing each SQ, the overall proportion that these SQs make of all tagged issues, and

the relative proportion that these SQs make of all maintainability related issues across categories. For won't-fix and unresolved issues, understandability, portability, and accessibility make up the largest percentage of expressed maintainability concerns. For reopened issues, diagnosability replaces portability of the top 3 subgroup SQ concerns. Won't-fix issues tend to express more portability concerns with relatively fewer diagnosabilty concerns compared to the other two categories.

4.2.2 RQ2b: Severity

Comparing the number of issues within each category of severity, the issues are divided into 4887, 17799, 15424, and 191219 issues associated with Blocker, Critical, Major, and Others types respectively. Figure 3d shows the proportion of issue summaries expressing maintainability concerns across these different levels of severity. For the highest severity category, blocker, maintainability issues make up 41%. Interestingly, the relative proportion of maintainabilty issues decreases for the next severity category, Critical, to 30%, while increasing again to 36% in the Major and Others categories. Table 2 compiles the number of issues expressing each SQ, the overall proportion that these SQs make of all tagged issues, and the relative proportion that these SQs make of all maintainability related issues across severities.

Portability, accessibility, and understandability remain the most prevalent expressed SQs except for the Critical category, where understandability is replaced by scalability. Blocker and Critical issues have similar proportions of portability and accessibility issues while accessibity issues have higher prevalence in Major issues. Finally, the less severe Others category is comprised largely of understandability issues.

A MANOVA is performed to examine whether there is a significant association between the overall proportion of maintainability and each maintainability subgroup SQs with different levels of severity. There is a statistically significant association found between levels of severity and the overall proportions of maintainability and its subgroup SQs, $F = 166.42$, $df = 24$, $p < 0.001$. Of the subgroup SQs, accessibility, portability, scalability, and understandability are found to have statistically significant differences across severity levels, with $p < 0.001$ after Bonferroni correction.

Pearson's Chi-squared tests are performed to examine whether levels of severity are associated with the distributions of the relative proportion of each SQ. Overall, the distributions of the relative proportion of each SQ differ significantly across severity, χ^2 (18, N = 61) = 5795.8, $p < 0.001$.

Summary of RQ2: To summarize, in this study, won't-fix and unresolved issues tend to express understandability, portability, and accessibility concerns. Reopened issues tend to express diagnosability concerns in addition to understandability and accessibility. For the case of won't-fix issues tending to express more portability concerns than the baseline-total, this result could be explained as portability issues involve factors external to the system. These types of issues are more likely to involve unsupported tools or potentially costly integrations, leading to a classification of WONTFIX.

The highest severity issues tend to have a higher proportion of maintainability issues than lower severity issues, and there is a significant association between levels of severity and the overall and relative proportions of the different subgroup SQs. This finding validates our results from the previous empirical study. As these issues have a high impact on the system, this reinforces the importance of ensuring high maintainability to avoid these types of issues.

4.3 RQ3

4.3.1 Changes between major versions

When comparing between major versions, there does not appear to be a strong trend in terms of the percentage maintainability issues make of the total. Figure 4a shows the relative proportions of each subgroup SQ. There appears to be a decreasing trend for accessibility and an increasing trend for understandability related issues for later versions. Figure 4b shows the overall proportions of each subgroup SQ. In this case, the increase in understandability related issues continues.

4.3.2 Changes within major version by year

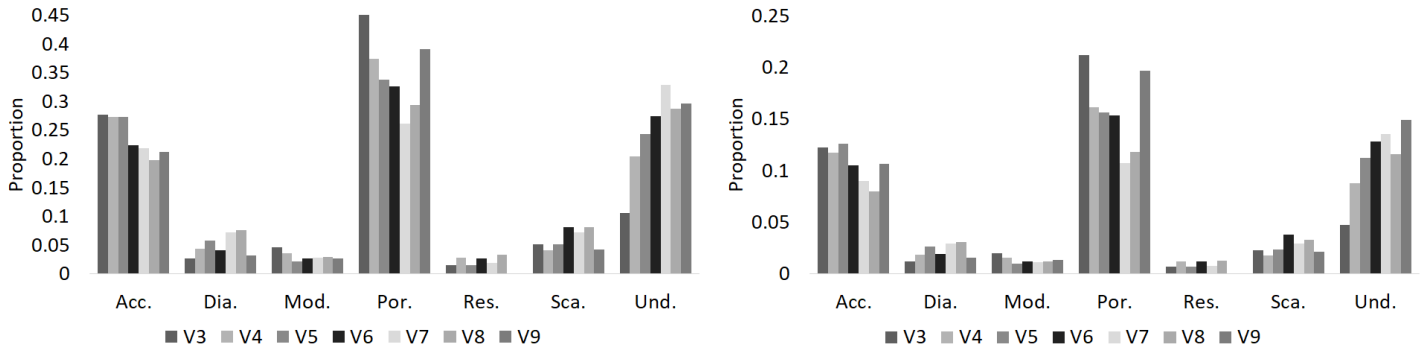
To provide an analysis of the relationship between the subgroup SQs and time, we map each patch to a year based on the last updated date in the Apache Tomcat Archive and perform linear regression within each major version. Figures 5a, and 5b show the statistically significant trends ($p < 0.05$) in relative proportions, overall proportions, and number of issues reported respectively. For clarity, the scales of the y-axes are set individually by version in Figure 6 due to relatively large differences in overall proportions and number of issues between versions.

- Version 3: Only restorability showed a statistically significant decline in number of issues reported. Relative and overall proportions did not have statistically significant relationships. We acknowledge that very few issues overall were reported in 2003 and 2004 which may contribute to this trend.

- Version 4: Other than modularity, all subgroup SQs showed statistically significant decline in number of issues reported. However, the decline in number of issues reported is also present for issues in general. In terms of relative proportions, accessibility and understandability showed declines while portability showed an increase over time. For overall proportions, accessibility, modularity, restorability, and understandability all showed declines. We acknowledge that very few issues overall were reported from 2005 to 2008 which results in the relative proportions of 0 for accessibility and understandability and the relative proportions of 1 in portability within those years. These values may skew the significance of the trends.

- Version 5: Portability and understandability showed statistically significant decline over years for number of issues reported. Those SQs showed similar declines, and accessibility showed statistically significant increase for relative and overall proportion of SQs.

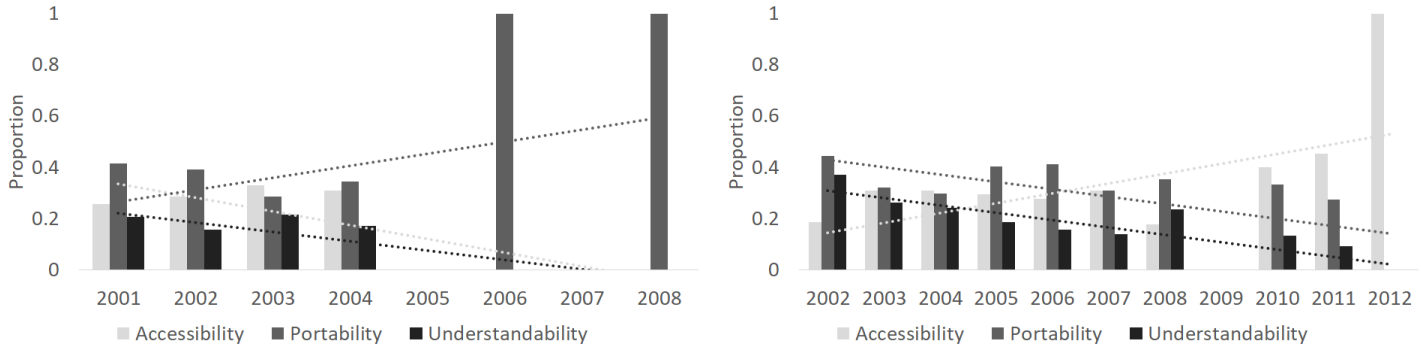
- Version 6: Accessibility and portability showed a statistically significant decrease in number of issues reported over years.



(a) Relative proportion

(b) Overall proportion

Figure 4: Proportions by major version of Apache Tomcat



(a) Version 4

(b) Version 5

Figure 5: Statistically significant relative proportions of subgroup SQ concerns across years by versions

This was shared with the number of maintainability, non-maintainability, and total issues reported. No SQs were found to have a statistically significant change in relative proportion; however, the overall proportion of portability related issues increased.

- Version 7: Accessibility, portability, scalability, and understandability all had declines over time along with decreases in numbers of maintainability, non-maintainability, and total issues reported. Overall proportions of portability and maintainability had increases over time, but no trends were found for relative proportions.
- Version 8: Accessibility, portability, and understandability all had declines over time along with decreases in number of maintainability, non-maintainability, and total issues reported. No trends were found for relative or overall proportions.
- Version 9: No significant trends were found for version 9

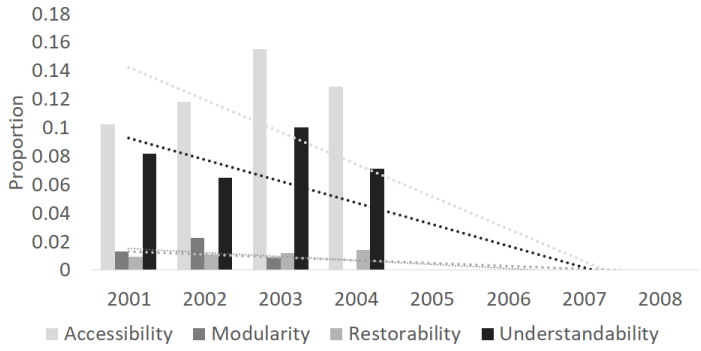
4.3.3 Changes by year overall

When considering all issues by year, there is a general decline in the number of issues reported overall which is found for all SQs. This trend is found in general for the number of issues reported; Figure 6 shows the number of maintainability and non-maintainability related issue summaries reported per year for Tomcat. However, when looking at the overall and relative proportions, there are no statistically significant trends across the SQs.

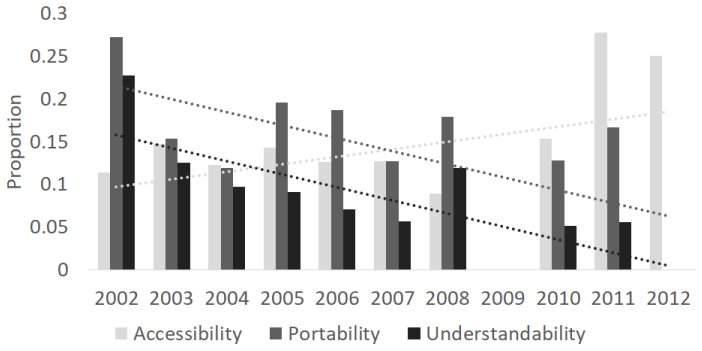
Summary of RQ3: To summarize, although there is not a significant trend in terms of the percentage of maintainability issues between major versions, there is a decreasing trend for accessibility while an increasing trend for understandability for later versions. In addition, various subgroup SQs show statistically significant trends in relative proportions, overall proportions and the number of issues reported.

While there are a number of statistically significant trends within versions, the most common is a decline in number of issues reported overall over time and within the versions themselves. Possible reasons for the decline in issues reported include that people are not reporting as many issues in general compared to in the past, the later versions are still being maintained and developed so there has been less time to report issues, and finally that the maintainability has increased over time.

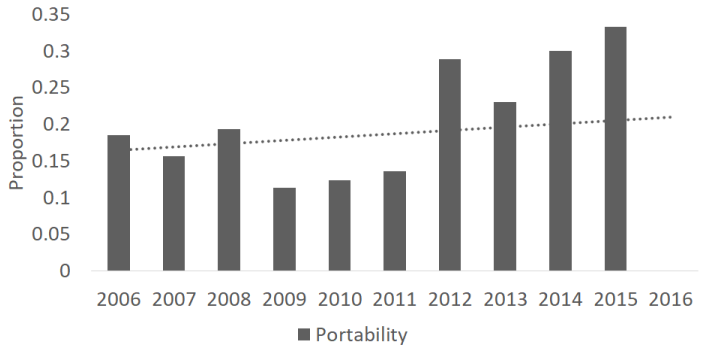
As there are no statistically significant trends in terms of relative or overall proportions of the subgroup SQs when looking at changes by year overall, this may indicate that focusing on the greatest relative proportions overall (i.e. accessibility, portability, and understandability) will be effective regardless of time in the life cycle.



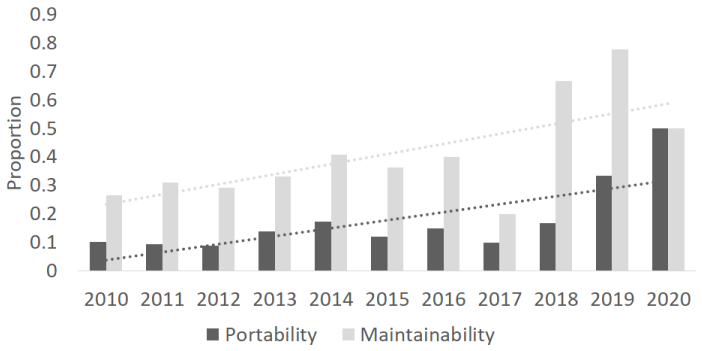
(a) Version 4



(b) Version 5



(c) Version 6



(d) Version 7

Figure 6: Statistically significant overall proportions of subgroup SQ concerns across years by version

4.4 Threats to Validity

This study depends on the model developed in [38]. The accuracy of classification is subject to the limitations and threats to validity detailed in the prior work.

Some information in issue summaries is self-reported by the developers of the different software projects such as severity, version information, etc. Validation of this information is out of the scope of this study; however, the developers reporting the issues are the most qualified to assess these metrics given their familiarity with the projects. Thus, we assume the reported information is correctly identified.

As our case study focused only on Apache Tomcat, our findings with regard to quality changes within versions and over the life of the project should not be generalized to other projects without further study.

5 Conclusion

Motivated by the lack of effective systematic measurement of maintainability in practice, we presented a novel approach to achieve automatic identification on how software maintainability and its subgroup SQs are expressed in a series of publications. Enabled by the automated approach to scale up analysis of maintainability through issue summaries, in this article, a large empirical study on 229,329 issue summaries from 61 different projects was conducted. Out of all the issue summaries, 82,577 issues were classified as expressing maintainability concerns. These issues were further analyzed to evaluate the differences between domains, ecosystems, and types. We found differences in relative proportions across ecosystems, domain and issue severity. Additional analysis was performed on Apache Tomcat to evaluate the evolution of maintainability across several versions. We identified several trends within versions and over time, such as a general decline in the number of issues reported overall in all the subgroup SQs and a statistically significant decline in portability and accessibility in multiple versions.

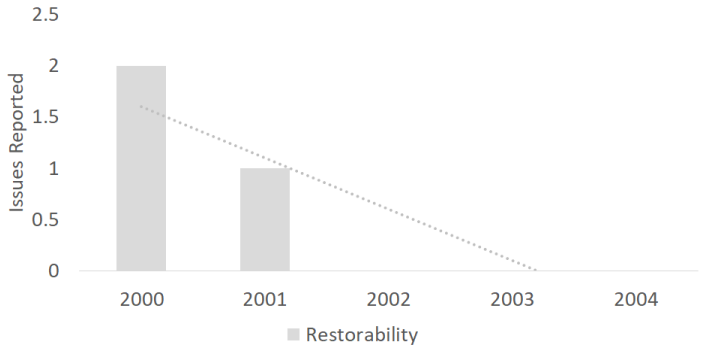
We believe that our work introduces a new angle to the area of software maintainability evaluation, encourages researchers to utilize unstructured software artifacts, and promotes automated solutions to incorporate standards and frameworks into software development process.

Conflict of Interest The authors declare no conflict of interest.

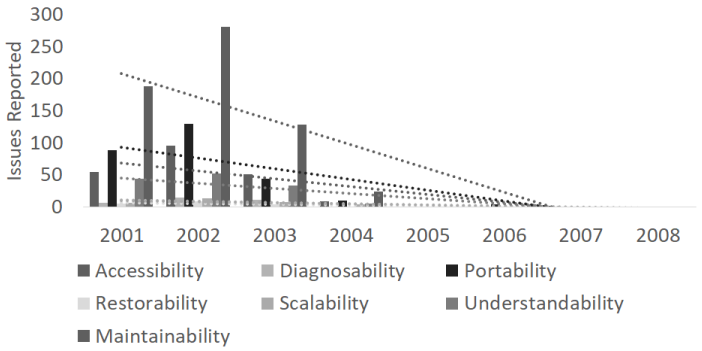
Acknowledgment This material is based upon work supported in part by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract HQ0034-13-D-0004. SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

References

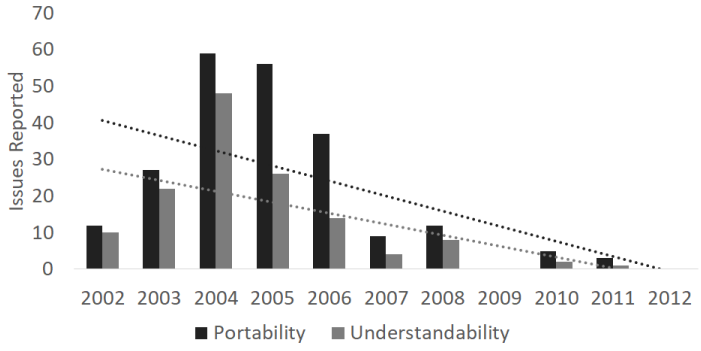
[1] J. Koskinen, "Software maintenance fundamentals," Encyclopedia of Software Engineering, P. Laplante, Ed., Taylor & Francis Group, 2009.



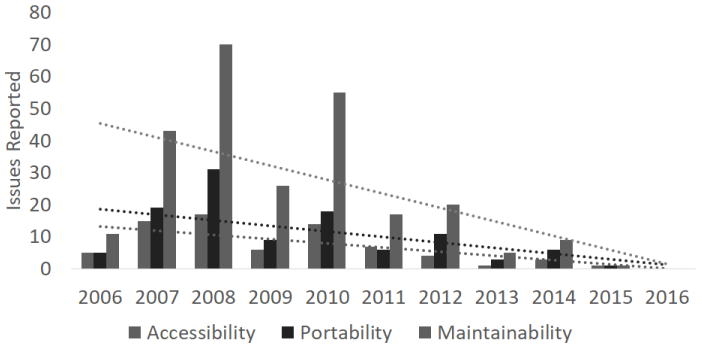
(a) Version 3



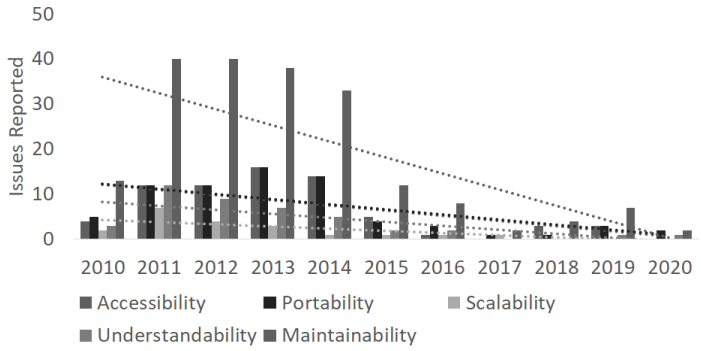
(b) Version 4



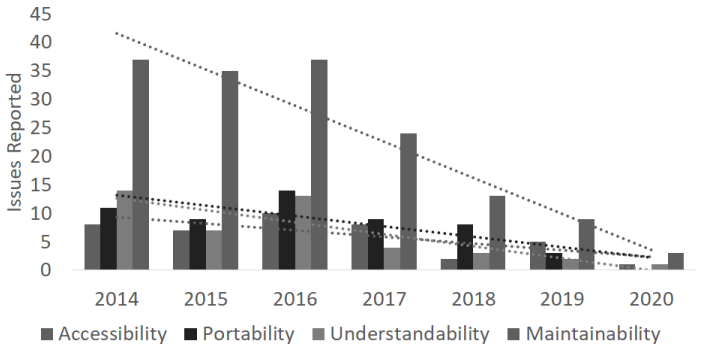
(c) Version 5



(d) Version 6



(e) Version 7



(f) Version 8

Figure 7: Statistically significant number of issues per subgroup SQ concerns across years by version

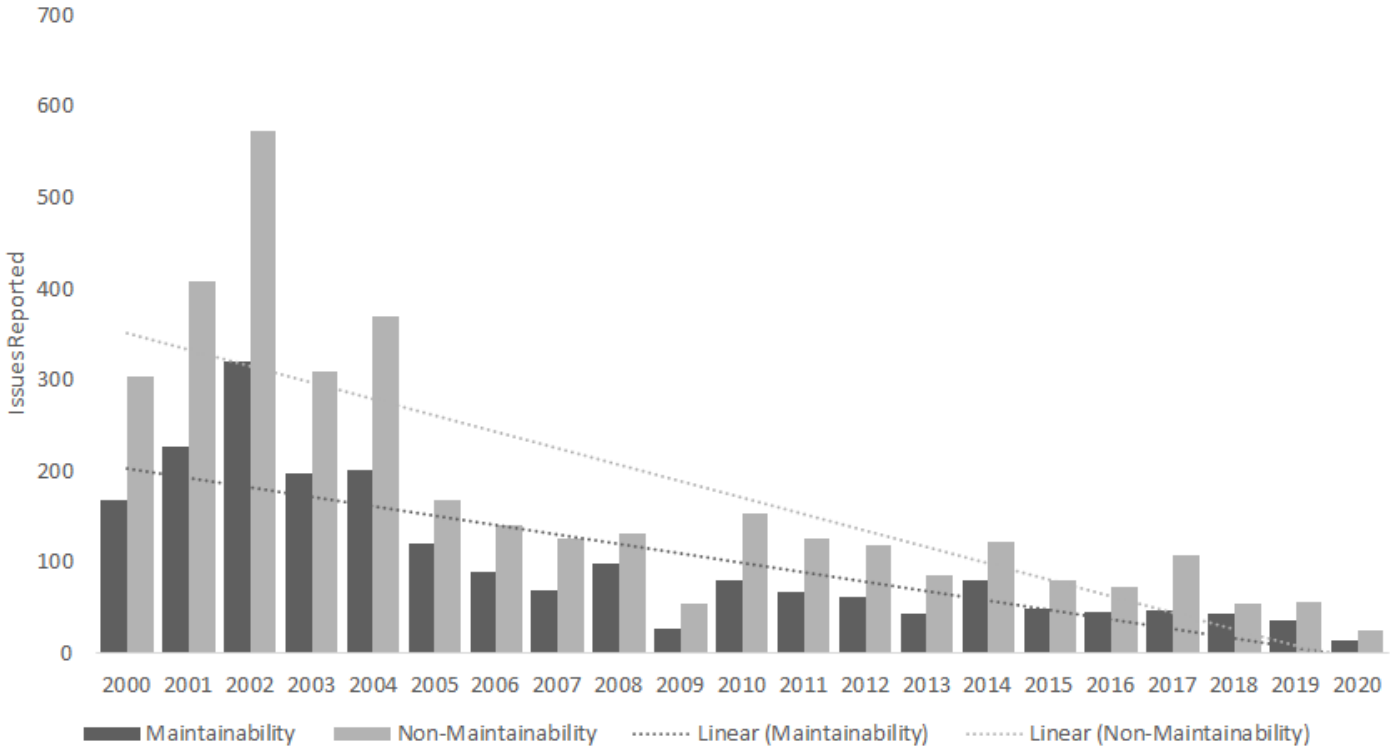


Figure 8: Number of issues reported by year

[2] B. Boehm, C. Chen, K. Srisopha, L. Shi, "The key roles of maintainability in an ontology for system qualities," in INCOSE International Symposium, **26**, 2026–2040, Wiley Online Library, 2016.

[3] M. Y. Shoga, C. Chen, B. Boehm, "Recent Trends in Software Quality Interrelationships: A Systematic Mapping Study," in 2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C), 264–271, 2020, doi:10.1109/QRS-C51114.2020.00052.

[4] S. R. Chidamber, D. P. Darcy, C. F. Kemerer, "Managerial use of metrics for object-oriented software: An exploratory analysis," IEEE Transactions on software Engineering, **24**(8), 629–639, 1998.

[5] R. Mo, Y. Cai, R. Kazman, L. Xiao, Q. Feng, "Decoupling level: a new metric for architectural maintenance complexity," in Proceedings of the 38th International Conference on Software Engineering, 499–510, ACM, 2016.

[6] H. Wu, L. Shi, C. Chen, Q. Wang, B. Boehm, "Maintenance effort estimation for open source software: A systematic literature review," in Software Maintenance and Evolution (ICSME), 2016 IEEE International Conference on, 32–43, IEEE, 2016.

[7] L. Yu, A. Mishra, "An empirical study of Lehman's law on software quality evolution," 2013.

[8] C. Chen, R. Alfayez, K. Srisopha, L. Shi, B. Boehm, "Evaluating Human-Assessed Software Maintainability Metrics," in Software Engineering and Methodology for Emerging Domains, 120–132, Springer, 2016.

[9] B. W. Boehm, R. Madachy, B. Steece, et al., Software cost estimation with Cocomo II with Cdrom, Prentice Hall PTR, 2000.

[10] M. Riaz, E. Mendes, E. Tempero, "A systematic review of software maintainability prediction and metrics," in Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement, 367–377, IEEE Computer Society, 2009.

[11] S. Scalabrino, G. Bavota, C. Vendome, M. Linaresvasquez, D. Poshyvanyk, R. Oliveto, "Automatically assessing code understandability: How far are we?" in Ieee/acm International Conference on Automated Software Engineering, 417–427, 2017.

[12] C. Chen, S. Lin, M. Shoga, Q. Wang, B. Boehm, "How do defects hurt qualities? an empirical study on characterizing a software maintainability ontology in open source software," in 2018 IEEE International Conference on Software Quality, Reliability and Security (QRS), 226–237, IEEE, 2018.

[13] P. Oman, J. Hagemeister, "Metrics for assessing a software system's maintainability," in Software Maintenance, 1992. Proceedings., Conference on, 337–344, IEEE, 1992.

[14] K. D. Welker, "The software maintainability index revisited," CrossTalk, **14**, 18–21, 2001.

[15] E. VanDoren, "Maintainability Index Technique for Measuring Program Maintainability. Software Engineering Institute," 2002.

[16] D. I. Sjøberg, B. Anda, A. Mockus, "Questioning software maintenance metrics: a comparative case study," in Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement, 107–110, ACM, 2012.

[17] R. Baggen, J. P. Correia, K. Schill, J. Visser, "Standardized code quality benchmarking for improving software maintainability," Software Quality Journal, **20**(2), 287–307, 2012.

[18] L. Kumar, S. K. Rath, A. Sureka, "Using Source Code Metrics and Multivariate Adaptive Regression Splines to Predict Maintainability of Service Oriented Software," in 2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE), 88–95, 2017, doi:10.1109/HASE.2017.11.

[19] R. Malhotra, A. Chug, "Software maintainability prediction using machine learning algorithms," Software Engineering: An International Journal (SEIJ), **2**(2), 2012.

[20] S. Jha, R. Kumar, L. Hoang Son, M. Abdel-Basset, I. Priyadarshini, R. Sharma, H. Viet Long, "Deep Learning Approach for Software Maintainability Metrics Prediction," IEEE Access, **7**, 61840–61855, 2019.

[21] M. Schnappinger, M. H. Osman, A. Pretschner, A. Fietzke, "Learning a Classifier for Prediction of Maintainability Based on Static Analysis Tools," in 2019 IEEE/ACM 27th International Conference on Program Comprehension (ICPC), 243–248, 2019, doi:10.1109/ICPC.2019.00043.

[22] Z. Li, L. Tan, X. Wang, S. Lu, Y. Zhou, C. Zhai, "Have things changed now?:an empirical study of bug characteristics in modern open source software," in The Workshop on Architectural and System Support for Improving Software Dependability, 25–33, 2006.

[23] Y. Zhou, Y. Tong, R. Gu, H. Gall, "Combining text mining and data mining for bug report classification," Journal of Software: Evolution and Process, **28**(3), 150–176, 2016, doi:10.1002/smr.1770.

[24] "Detecting Missing Information in Bug Descriptions," ESEC/FSE 2017, 396–407, New York, NY, USA, 2017, doi:10.1145/3106237.3106285, event-place: Paderborn, Germany.

[25] D. C. Das, M. R. Rahman, "Security and Performance Bug Reports Identification with Class-Imbalance Sampling and Feature Selection," in 2018 Joint 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR), 316–321, 2018.

[26] B. J. Guarraci, "Instrumenting software for enhanced diagnosability," 2012, uS Patent 8,141,052.

[27] D. Yuan, J. Zheng, S. Park, Y. Zhou, S. Savage, "Improving software diagnosability via log enhancement," in Sixteenth International Conference on Architectural Support for Programming Languages and Operating Systems, 3–14, 2011.

[28] Y. Le Traon, F. Ouabdesselam, C. Robach, "Software diagnosability," in Software Reliability Engineering, 1998. Proceedings. The Ninth International Symposium on, 257–266, IEEE, 1998.

[29] A. Kavcic, "Software Accessibility: Recommendations and Guidelines," in The International Conference on Computer As A Tool, 1024–1027, 2006.

[30] J. Manual, "Manual for the operation of the joint capabilities integration and development system," US Department of Defense. Washington. DC, 2012.

[31] C. Chen, M. Shoga, B. Li, B. Boehm, "Assessing Software Understandability in Systems by Leveraging Fuzzy Method and Linguistic Analysis," in 2019 Conference on Systems Engineering Research (CSER) (2019 CSER), Washington, USA, 2019.

[32] K. J. Sullivan, W. G. Griswold, Y. Cai, B. Hallen, "The structure and value of modularity in software design," in ACM SIGSOFT Software Engineering Notes, **26**, 99–108, ACM, 2001.

[33] R. Sanchez, J. T. Mahoney, "Modularity, flexibility, and knowledge manage-ment in product and organization design," Strategic Management Journal, **17**(S2), 63–76, 2015.

[34] C. U. Smith, L. G. Williams, Performance solutions: a practical guide to creating responsive, scalable software, 1, Addison-Wesley Reading, 2002.

[35] L. Duboc, D. Rosenblum, T. Wicks, "A framework for characterization and analysis of software system scalability," in Proceedings of the the 6th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering, 375–384, ACM, 2007.

[36] J. D. Mooney, "Issues in the specification and measurement of software portability," in 15th International Conference on Software Engineering, Baltimore, 1993.

[37] A. S. Tanenbaum, P. Klint, W. Bohm, "Guidelines for software portability," Software: Practice and Experience, **8**(6), 681–698, 1978.

[38] C. Chen, M. Shoga, B. Boehm, "Characterizing software maintainability in issue summaries using a fuzzy classifier," in 2019 IEEE 19th International Conference on Software Quality, Reliability and Security (QRS), 131–138, IEEE, 2019.

[39] R. Moazeni, D. Link, C. Chen, B. Boehm, "Software Domains in Incremental Development Productivity Decline," in Proceedings of the 2014 International Conference on Software and System Process, ICSSP 2014, 75–83, Association for Computing Machinery, New York, NY, USA, 2014, doi: 10.1145/2600821.2600830.

[40] Q. Wang, "Why Is My Bug Wontfix?" in 2020 IEEE 2nd International Work-shop on Intelligent Bug Fixing (IBF), 45–54, 2020.

[41] T. Zimmermann, N. Nagappan, P. J. Guo, B. Murphy, "Characterizing and predicting which bugs get reopened," in 2012 34th International Conference on Software Engineering (ICSE), 1074–1083, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 164-168 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Lifestyle in Nursing Students at a University of North Lima

Yanet Cruz Flores[1], Tania Retuerto-Azaña[1], Jaquelin Nuñez-Artica[1], Brian Meneses-Claudio[2,*], Hernan Matta Solis[1], Lourdes Matta-Zamudio[1]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 19 December, 2020
Online: 15 January, 2021

Keywords:
Lifestyle
Nursing students
Healthy
Unhealthy
Academic Performance

ABSTRACT

Healthy lifestyles were proposed to improve the health status of the university population, they are a set of behaviors that are reflected according to the type of situation and behavior of each person who performs it during the period of their training, the inappropriate lifestyles provides many problems, which are non-communicable diseases that manifest in the health status of each individual, which most of the time students choose to consume less nutritious foods and they are not healthy and affect their health; inadequate lifestyles do not provide good academic performance and also generate a bad psycho-emotional state. The objective of the study is to determine the Lifestyle in nursing students at a university in North Lima, 2019. As results, regarding the Lifestyle in nursing students at a university in North Lima; 51.5% have healthy lifestyle and 48.5% have an unhealthy lifestyle. Regarding the dimensions, interpersonal relationships predominate 73.8% have an unhealthy lifestyle followed by stress management, 65.8% have an unhealthy lifestyle, responsibility for health, 63.3% have a lifestyle unhealthy, Physical activity dimension, 60.8% have an unhealthy lifestyle, spiritual growth, 48.8% have an unhealthy lifestyle, healthy nutrition, 29, 5% have an unhealthy lifestyle; it is important to know those results to make decisions.

1. Introduction

In the Republic of Ireland in September 2016 they have shown that university life transition has a negative effect on student lifestyles, resulting in weight gain, increased physical inactivity and stress [1]. University students go through a variation of changes in their mental, physical, social, and economic lives during their university life [2].

The university stage is characterized by a tendency to engage in a variety of unhealthy behaviors [3]. Evidences suggest that nursing students do not engage in healthy lifestyles such negative behaviors lead nursing students to become overweight and obese, as well as other risks such as smoking, alcohol consumption, drugs, an insufficient diet and poor physical activity [4], [5].

According to the research, it has carried out a technical reform that aims to educate the entire community, regarding issues related to nutrition, promoting healthy practices for a proper life and to prevent diseases such as anemia and chronic malnutrition, as well same regarding overweight, obesity, diabetes, cardiovascular disease, high blood pressure and cancer. This technical document aims to promote the functions of exhaustive care and health promotion in the university population [6].

Unhealthy habits are factors that predominate in university students. According to studies, it is evident that they do not perform physical activity; they do not complete their sleep hours; they eat fast food [7], the negative effects are evident when they reach the adult stage, which causes health problems in older life The evidence shows that regular physical activity helps maintain good health styles and prevents diseases such as hypertension, obesity, metabolic syndrome, heart disease, diabetes, osteoporosis, some types of cancer, and produces significant benefits at the psychosocial level [8]. In this university stage, a sum of difficulties are manifested, likewise this type of qualities is linked to inappropriate health behavior; therefore, it should be noted that it is significant to determine certain health behaviors that are inadequate and have a much higher risk of suffering from non-communicable diseases [9].

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & bmeneses@uch.edu.pe

In Peru in 2018, more than 50% of the disease burden is associated with non-communicable diseases, which affect all age groups. Tobacco use, harmful use of alcohol, unhealthy diet, infrequent physical activity. Approximately 62 million people have Type 2 Diabetes. Cancer is the second leading cause of death. In 2018, it caused 1.3 million deaths, and 3.7 million new cases. Tobacco kills 7 million people each year. 62% of adults are overweight or obese. The epidemic disturbs adolescent and young children since between 20% and 25% are affected by being overweight or obese [10].

In [11], the authors showed the sample of 829 university students included 504 (60.8%) Chinese and 325 (39.2%) international students. The Chinese students had higher scores in general healthy lifestyles and the scales showed low levels in physiological and social health. Gender was associated with psychological health in both groups. Smoking was a predictor of psychological health in both groups and of social health in the international group.

In [12], the authors showed that 47% of the study participants had a moderate-level lifestyle. Regarding musculoskeletal symptoms, the most frequent were reported in the back (88,33%), knees (83,33%) and thighs (71%). Furthermore, Pearson's correlation analysis revealed significant negative correlations between eating habits and musculoskeletal symptoms and interpersonal relationships, while a significant positive correlation was found between nutrition and stress management.

In [13], the authors presented a study carried out in Turkey, showed that in nursing students there were various factors such as financial problems (70.3% of moderate income level; inadequate) and the characteristics cultural factors that may have contributed to the decrease in physical activity observed in this study; therefore it should be noted that it is significant to determine certain health behaviors that are inadequate and have a much higher risk of suffering from non-communicable diseases.

The objective of the study is to determine the Lifestyle in nursing students at a university in North Lima, 2019. This research work will be developed to improve the lifestyle of university students to reduce the risks of acquiring different diseases due to bad lifestyles in university students at a university of north lima. In this study, we suggest that we should continue to emphasize the promotion of a healthy lifestyle in nursing students; in the university stage, an important role must be fulfilled, such as prevention and health promotion; this effort must be involved with other professionals to have more significant results in the benefit of the health of the university population, they are encouraged to raise awareness.

The instrument to be used is the Health Promoting Lifestyle Profile II (HPLP-II) developed by Pender Walker and Sechrist, in 1987 and translated into Spanish in 1995, consisting of 52 items divided into six dimensions: health responsibility that is the comprising an active notion of responsibility for one's own comfort. It implies maintaining an adequate interest in one's own health, educating oneself about health [14].

All the data will be addressed to a Microsoft Excel 2013 program, then it will be sent into the SPSS24 program.

To compile the information, articles were reviewed in all the scientific databases that exist internationally and nationally, it was possible to indicate that there are a large number of articles that address the health problem with respect to a healthy lifestyle, information is evidenced found in articles, therefore, this study seeks to increase the knowledge of each of the students. The subject is important since applying it in practice would bring many benefits for the human being.

This work is structured as follows: In section II, the steps to carry out the survey are explained, how the population and sample were acquired, type of study, inclusion and exclusion criteria, analysis of variables, data collection, place and application of instruments. In section III, the results acquired according to the surveys carried out in the study are presented, through the elaboration of tables and bars. In section IV, the discussions that show the interpretation of results obtained in the research work compared to another research will be presented. Finally, in section V, the conclusions of the research work will be shown according to the results found, what measures should be taken about them.

2. Methodology

2.1. Population

There are 400 students from the first to the tenth cycle of the nursing faculty. Day and night shift.

Table 1: Number of Students from day and night shift

Study semester	Total of Students	Percentage
Semester 1	60	15.0
Semester 2	45	11.3
Semester 3	42	10.5
Semester 4	33	8.3
Semester 5	41	10.3
Semester 6	39	9.8
Semester 7	30	7.5
Semester 8	32	8.0
Semester 9	41	10.3
Semester 10	37	9.3

2.2. Inclusion criteria

- Nursing students from 1st to 10th semester.
- Nursing students who have signed the informed consent.

2.3. Exclusion Criteria

- Nursing students who do not want to participate in research.
- Nursing students who have not signed the informed consent.

2.4. Type of study

The present investigation is of a quantitative approach, its non-experimental, descriptive, and transversal design. To start with the

data collection, the necessary coordination was carried out in the Faculty of Nursing to obtain authorization for access to the classrooms of the institution [15].

2.5. Application of the instrument

The survey and the instrument used is The HPLP (Health Promoting Life Profile II) created by. Walker S., Sechrist K., Pender N. Kerr M., published in Spanish in 1995, is made up of 52 items and distributed in six dimensions, responsibility in health (9 items), physical activity (8 items), nutrition (9 items), spiritual growth (9 items), interpersonal relationships (9 items) and stress management (8 items), also a standard scale is used Likert where the punctuation varies from 1 to 4, where "1 = never", "2 = sometime", "3 = often" and "4 = usually", giving a final score of 52 to 208, ranging from " 52 to 130" is an unhealthy lifestyle and "131 to 208" is a healthy lifestyle, the higher the score the higher the healthy lifestyle the student will present [16].

2.6. Place and application of the instrument

The instrument was applied at the Universidad de Ciencias y Humanidades in Los Olivos. A permission was requested from the nursing faculty to have access to the classrooms. For this, an enrollment in the ethics committee was made for the approval of the instrument to apply.

After approval, permission was requested from teachers who teach in the classrooms to provide 15-minute access to students to fill out the questionnaire. We had some difficulties with some teachers because they made us wait until the end of their class to give us the space, and others to wait for the break to be able to enter.



Figure 1: Delivering surveys to the students at a University of North Lima

As shown in Figure 1, they were given instructions on how to complete the questionnaire. To use exclusively a pencil to mark a single alternative in the question to avoid smearing.



Figure 2: Developing the surveys by the students at a University of North Lima

As shown in Figure 2, the students carefully concentrated on reading the questionnaire; some students did not understand the questions they did not interpret it in the health field.

The doubts they had about the questions were explained to the students and they were reminded that their answer must be sincere, and the development is individual, not group to get good results.



Figure 3: Collecting student surveys.

As shown in Figure 3, it is verified that the entire questionnaire is filled out. If a question doesn't have an answer, it is returned to the participant to complete so that the questionnaire is valid to be typing to the SPSS to find questions, also if there is missing answer, the survey is eliminated.

The nursing personnel is fundamental for the health study of the university students since it offers them Psycho-emotional support for the social well-being of the university students; because many are unaware of healthy lifestyles or do not practice them and lead a routine life and acquire non-communicable diseases.

3. Results

After obtaining the surveys, the data was synthesized and statistical tables were generated for a better understanding of them, as shown in Table I.

Table 1: Sociodemographic lifestyle data in nursing students from a university of north lima, 2019 (n = 400).

Participant Information	Total	
	N	%
Total	**400**	**100**
Sex		
Female	342	85,5
Male	58	14,5
Marital Status		
Single	262	65,5
Married	54	13,5
Widowed	5	1,3
Cohabitant	56	14,0
Divorced	23	5,8
Type of Family		

Nuclear	191	47,8
Extended	43	10,8
Amplified	55	13,8
Single parent	63	15,8
Reconstituted	23	5,8
Family equivalent	25	6,3
Occupancy Condition		
Study	142	35,5
Study and work	258	64,5
Degree of Study		
Secondary	170	42,5
University	156	39,0
Technical	74	18,5
Primary Education	3	0,8

In Table 1, it has the sociodemographic data of the study participants, there were 400 participating nursing students. The minimum age was 18 years old; the maximum age was 49 years old, the mean being 24.83 years old.

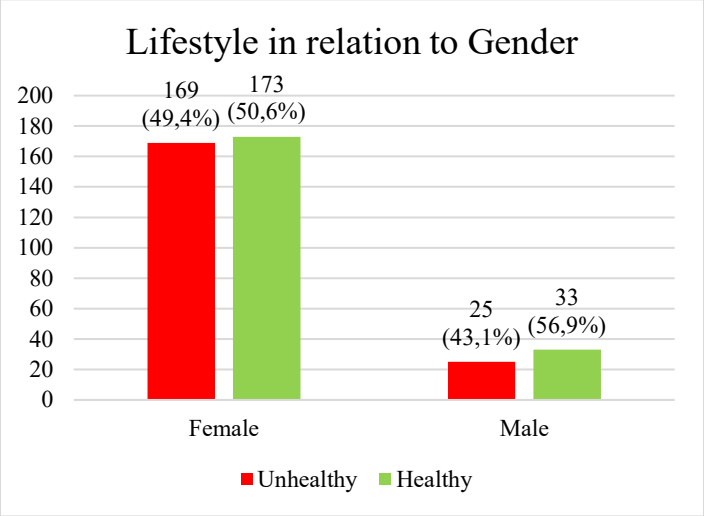


Figure 4: Lifestyle in nursing students at a university in North Lima, 2019 (N = 400)

Regarding sex, the female sex predominated with 342 students representing 85.5%, followed by the male sex with 58 students representing 14.5%. Regarding marital status, the single person predominates with 262 (65.5%) cases, followed by Cohabitant with 56 (14%) cases, married with 54 (13.5%) cases, divorced with 23 (5.8%) cases, and finally widowers with 5 (1.3%) cases.

Regarding the type of family, nuclear families predominate with 191 (47,8%) cases, followed by the single parent with 63 (15, 8%) cases, amplified with 55 (13,8%) cases, extended with 43 (10,8%) cases, family equivalent with 25 (6,3%) cases and finally reconstituted 23 (5,8%) cases.

Regarding the occupation status of nursing students, 258 study and work (64,5%) cases and only 142 with (35, 5%) cases study.

Regarding the degree of study, 238 students representing 59.5% have secondary education, 117 students representing 29.3% have technical education, 42 students representing 10.5% have university education and finally 3 students representing 0.8% have primary education.

In Figure 4, the lifestyle in nursing students from a university in Lima Norte, it can see that 206 students represent 51.5% have a healthy lifestyle and 194 students that represent 48.5% have an unhealthy lifestyle.

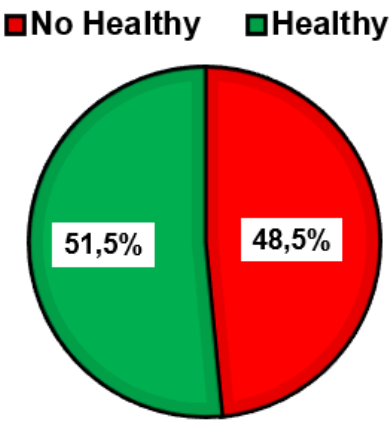


Figure 5: Lifestyle in relation to gender of nursing students from a University of Lima Norte, 2019 (N = 400).

Figure 5 shows the lifestyle in relation to gender of nursing students from a university in North Lima, where 169 (49.4%) who are female have an unhealthy lifestyle, 173 (50.6%) have a healthy lifestyle, 25 (43.1%) who are male have an unhealthy lifestyle and 33 (56.9%) have a healthy lifestyle.

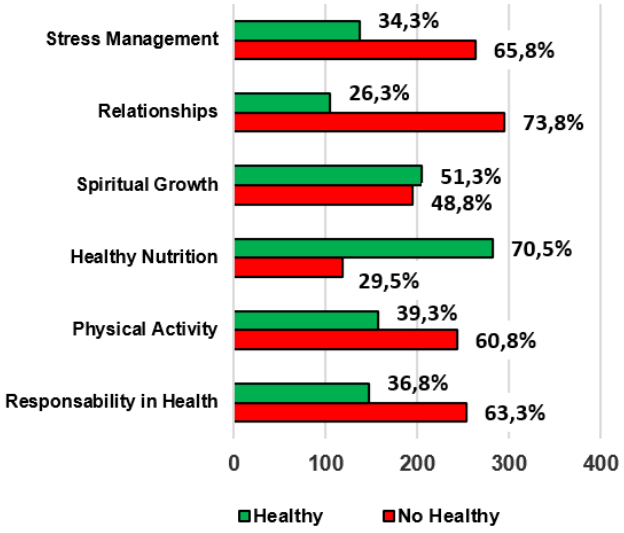


Figure 6: Lifestyle according to its general dimensions that the instrument contains according to a study in nursing students from a university in North Lima, 2019 (N = 400).

In Figure 6, Lifestyle in its dimension health responsibility, in nursing students from a university in Lima Norte, it can see that 253 students represent 63, 3%, have an unhealthy lifestyle and 147

students represent 36.8%, have a healthy lifestyle. In the physical activity dimension, in nursing students from a university in North Lima, it can see that 243 students represent 60.8%, have an unhealthy lifestyle and 157 students represent 39.3%, have a healthy lifestyle. In the healthy nutrition dimension, in nursing students from a university in North Lima, it can see that 282 students represent 70.5%, have a healthy lifestyle and 118 students represent 29.5%, have an unhealthy lifestyle. In the growth dimension spiritual, in nursing students from a university in Lima Norte, it can see that 205 students represent 51.3%, have a healthy lifestyle and 195 students represent 48.8%, have an unhealthy lifestyle. In the interpersonal relationships dimension predominates 295 students representing 73.8%, have a healthy lifestyle. In the stress management dimension, in nursing students from a university in Lima Norte, it can see that 263 students represent 65.8%, have an unhealthy lifestyle and 137 students represent 34.3%, have a healthy lifestyle.

4. Discussion

The research work confirms the unhealthy lifestyles in university students based on the articles reviewed in [12] to their study carried out on university students from Iran Shiraz where bad interpersonal relationships were evidenced in the students and based on [11] to their study carried out, low results were evidenced regarding its dimensions; according to [13] to their carried out in Turkey, students have a sedentary life and do not carry out physical activity.

In the study carried out on students from a North Lima university in terms of dimensions, it was clearly shown that a large number of students who do not have good interpersonal relationships with their peers are not responsible for their health; they do not have good stress management and do not carry out physical activity. It leads to a sedentary life. Likewise, in [5], argue that students do not have a good lifestyle and that factors such as tobacco and alcohol show an inappropriate lifestyle.

This type of studies can be applied in other university centers to help the teaching for students and also train teachers so that they can implement educational programs such as physical exercise, stress control, improving lifestyles involves commitment and will.

Promoting lifestyles is a little difficult since students do not put into practice; they always make excuses for not doing it, that's why, this research work is presented to provide guidance on the situation that university students are in.

5. Conclusion

To conclude in this regarding lifestyle, it is projected from the promotion of the health of the university students, in order to contribute to the universities so that they include programs to have an adequate healthy diet, physical activity, etc.

It is concluded that the university population is made aware of the benefits of irregular practices of physical activities that improve the lifestyle.

It is concluded that counseling is provided to university students to develop educational intervention protocols in relation to healthy life habits.

Conflicts of Interest

The authors declare no conflict of interest**.**

References

[1] E. Bastias, J. Stiepovich, "Una revisión de los estilos de vida de estudiantes universitarios iberoamericanos.," Ciencia y Enfermeria, **20**(2), 93–101, 2014.

[2] P. Mc Sharry, F. Timmins, "An evaluation of the effectiveness of a dedicated health and well being course on nursing students' health.," Nurse Education Today, **44**, 26–32, 2016, doi:10.1016/j.nedt.2016.05.004.

[3] Y. Lara, C. Quiroga, A. Jaramillo, M. Bermeo, "Estilo de vida de estudiantes en primer semestre de odontología de una universidad privada, Cali 2016.," Revista Odontológica Mexicana, **22**(3), 144–149, 2018.

[4] N. Gormley, V. Melby, "Nursing students' attitudes towards obese people, knowledge of obesity risk, and self-disclosure of own health behaviours: An exploratory. survey," Nurse Education Today, **84**, 104232, 2020, doi:10.1016/j.nedt.2019.104232.

[5] A. Montenegro, A. Ruíz, "Factores asociados a los estilos de vida en los estudiantes universitarios. Una aplicación del instrumento fantástico.," Revista Digital: Actividad Física y Deporte, **6**(1), 87–108, 2019.

[6] Ministerio de Salud, Guías Alimentarias Para La Población Peruana, 1–60, 2019.

[7] A. Gallardo, M. Alférez, E. Planells, I. López, "La etapa universitaria no favorece el estilo de vida saludable en las estudiantes granadinas.," Nutricion Hospitalaria, **31**(2), 975–979, 2015, doi:10.3305/nh.2015.31.2.8303.

[8] N. Arias, M. Solera, L. Gracia, P. Silva, V. Martínez, J. Cañete, M. Sánchez, "Levels and patterns of objectively assessed physical activity and compliance with different public health guidelines in university students.," PLoS ONE, **10**(11), 1–15, 2015, doi:10.1371/journal.pone.0141977.

[9] D. Kazemi, M. Levine, J. Dmochowski, K. Roger Van Horn, L. Qi, "Health behaviors of mandated and voluntary students in a motivational intervention program.," Preventive Medicine Reports, **2**, 423–428, 2015, doi:10.1016/j.pmedr.2015.05.004.

[10] Organización Panamericana de la Salud, Enfermedades no transmisibles y factores de riesgo, Boletín Epidemiológico, 2019.

[11] S. Lolokote, T. Hidru, X. Li, "Do socio-cultural factors influence college students' self-rated health status and health-promoting lifestyles? A cross-sectional multicenter study in Dalian, China.," BMC Public Health, **17**(1), 1–14, 2017, doi:10.1186/s12889-017-4411-8.

[12] M. Heidari, M. Borujeni, M. Khosravizad, "Health-promoting Lifestyles of Nurses and Its Association with Musculoskeletal Disorders: A Cross-Sectional Study.," Journal of Lifestyle Medicine, **8**(2), 72–78, 2018.

[13] B. Kara, B. Işcan, "Predictors of Health Behaviors in Turkish Female Nursing Students.," Asian Nursing Research, **10**(1), 75–81, 2016, doi:10.1016/j.anr.2015.12.001.

[14] M. Al-Qahtani, "Comparison of health-promoting lifestyle behaviours between female students majoring in healthcare and non-healthcare fields in KSA.," Journal of Taibah University Medical Sciences, **14**(6), 508–514, 2019, doi:10.1016/j.jtumed.2019.10.004.

[15] C. Fernández, P. Baptista, Metodología de la Investigación. 6ta ed. México: Mc Graw-Hill/Interamericana., 2015.

[16] G. Kuan, Y. Kueh, N. Abdullah, E. Tai, "Psychometric properties of the health-promoting lifestyle profile II: Cross-cultural validation of the Malay language version.," BMC Public Health, **19**(1), 1–10, 2019, doi:10.1186/s12889-019-7109-2.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 169-177 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Autonomous Robot Path Construction Prototype Using Wireless Sensor Networks

José Paulo de Almeida Amaro[1], João Manuel Leitão Pires Caldeira[2,*], Vasco Nuno da Gama de Jesus Soares[2], João Alfredo Fazendeiro Fernandes Dias[3]

[1]*Instituto Politécnico de Castelo Branco, Castelo Branco, 6000-084, Portugal*

[2]*Instituto Politécnico de Castelo Branco, Instituto de Telecomunicações, Castelo Branco, 6000-084, Portugal*

[3]*IADE–Faculdade de Design, Tecnologia e Comunicação, Lisboa, 1000-041, Portugal*

ARTICLE INFO

Article history:
Received: 25 September, 2020
Accepted: 01 December, 2020
Online: 15 January, 2021

Keywords:
Wireless Sensor Network
Received Signal Strength Indication
Robotic Path Construction
Autonomous Deployment
Prototype

ABSTRACT

The use of wireless sensor networks (WSN) can be a valuable contribution in disaster situations or life-threatening exploration. Using wireless mobile robots, it is possible to explore vast areas without human intervention. However, the wireless network coverage that can keep mobile robots connected to the base station / gateway is a major limitation. With this in mind it was created a prototype of an extensible WSN using mobile robot nodes that cooperate amongst themselves. The strategy adopted in this project proposes using three types of nodes: master node, static node, and robot node. Three different algorithms were also developed and proposed: Received Signal Strength Indication (RSSI) Request; Automovement; Robot Cooperation and Response to Static Node. The performance evaluation of the prototype was carried out using a real-world testbed with each developed algorithm. The results achieved were very promising to continue the evolution of the prototype.

1. Introduction

This paper is an extension of work originally presented in conference *2020 15th Iberian Conference on Information Systems and Technologies* (CISTI 2020) [1].

Wireless Sensor Networks (WSNs) are composed by small nodes spread in an area of interest to measure environmental conditions such as temperature, sound, humidity, although they have many other uses [2]. Each node can be seen as a small computer, having processing, sensing and communication capabilities. The nodes connect to each other wirelessly and can cooperate in collecting information and routing it to an end user [3].

The need for extensible WSNs is tied with monitoring situations that require the deployment of a WSN. Situations such as forest fires or fires is large buildings [4], determining the extent of damage of earthquakes [5], nuclear disasters as in Fukushima, Japan, rescue operations, exploratory operations, mapping flooded caves, battlefield reconnaissance or planetary exploration [6]. In some of these situations it is possible for a human operator to deploy the nodes of the WSN. But in disaster situations, or any situation where human life could be endangered, it is necessary to use alternative methods of deployment, such as the use of mobile robots.

With this purpose in mind a testbed was used, which serves as a prototype for a WSN. The testbed used was composed of three different types of nodes: master, robot and static nodes. The master node is used so an end user can receive information from the network and send commands to any node. The static nodes serve as intermediate relays that connect the master node and the robot nodes. The robot nodes serve to extend the network, having mobility besides communication capabilities. The objective of this project is to create an extensible wireless network, using intermediate static nodes that connect a gateway to a moving robot node.

This document is organized as follows. Section I introduces WSNs and the goal of the project, Section II gives a review of related literature, with examples of similarly deployed WSNs. Section III describes the testbed and the libraries used for programming the nodes. Section IV describes the algorithms developed. Section V describes four scenarios used to evaluate the performance of the testbed with the described algorithms and the

*Corresponding Author: João Caldeira, Email: jcaldeira@ipcb.pt

results achieved. Section VI discusses these results and Section VII concludes the report and proposes future avenues of exploration.

2. Related Work

There has been prior work on the problem of "as-you-go" or impromptu deployment of a WSN. In [4] and [7] are proposed algorithms for "as-you-go" deployment, in which an agent, human or robot, measures link quality at equally spaced intervals and makes placement decisions, while only moving forward along a line.

Some works consider a purely autonomous deployment of a WSN by mobile robots. In [6] is explored autonomous deployment of a WSN with the goal of human detection, in case of disaster situations, in which mobile robots perform simultaneous localization and mapping. In [5] mobile robots carry sensor nodes and measure the RSSI, to know when to deploy sensor nodes to ensure communication or restore the network.

Many of the proposed systems measure link quality, using communication metrics like RSSI or Packet Reception Ratio, and decide when to deploy a new relay when the metric satisfies predefined rules. In [8] such a system is used for dynamic deployment of a WSN through a breadcrumb system for firefighters. A breadcrumb dispenser is worn by firefighters and a link estimator uses the previous mentioned metrics to decide when to deploy breadcrumbs.

Works such as [9] also make use of mobile robots to deploy a WSN but focus on how these robots can help with problems such as coverage holes or collection of redundant sensors. They achieve this by using permanent grid-based deployment, with cluster concepts to reduce packets used in creating and maintaining the grid structure. Furthermore, they consider the communication range to be at least twice of the sensing range.

Cooperation between multiple agents is seen as an important problem to be explored. In [4] and [7] it is pointed the need for algorithms where two or more robots cooperate to deploy a WSN. While [6] calls for field test using two mobile robots and to address the communication issues that arise from coordination between multiple robots. These and more issues of multi-robot systems, such as reduction in network traffic, routing approaches and sensor data processing are pointed as areas that need appropriate operational methods by [5]. The study of communication metrics, such as RSSI, in various environments, to allow setting different parameters that could optimize the system, is called for by [8].

In the present study the aim is to use the communication metric of RSSI to know when a new node should be deployed and integrate this metric with cooperation between two mobile robot nodes. The RSSI metric has been used in studies such [10], [11] and [12] to locate the robot nodes, triangulating the robot node and static nodes for navigation and helping the robot node avoid obstacles and reach a goal. But these approaches require that the static nodes be placed before the robot node starts its action, and to locate the robot need a lot of deployed static nodes. The proposal presented in this paper gives a different approach to this problematic by promoting the static nodes deployment using mobile robots. The RSSI value was used to decide the positions to deploy static nodes.

3. Wireless Sensor Network Robot Path Construction Proposal

The goal of this project was to build a prototype of autonomous WSN deployment by a mobile robot. In light of literature requests, it was decided to consider two mobile robots and develop algorithms that ensure the coordination between them as the WSN is deployed. In order to build an extensible network, it is necessary for the robots not only to cooperate each other but also respond to the deployment of static nodes. Using static nodes allows to maintain communication between mobile robots and gateway even when the mobile robots leave the network coverage area of gateway.

In this section firstly will be described the hardware used to create the testbed, namely the master node, robot nodes and static nodes. Then will be described the libraries and methods used to achieve the WSN behavior and obtain a communication metric from each node, the RSSI. The RSSI metric was considered due to the limited resources of the hardware platform used.

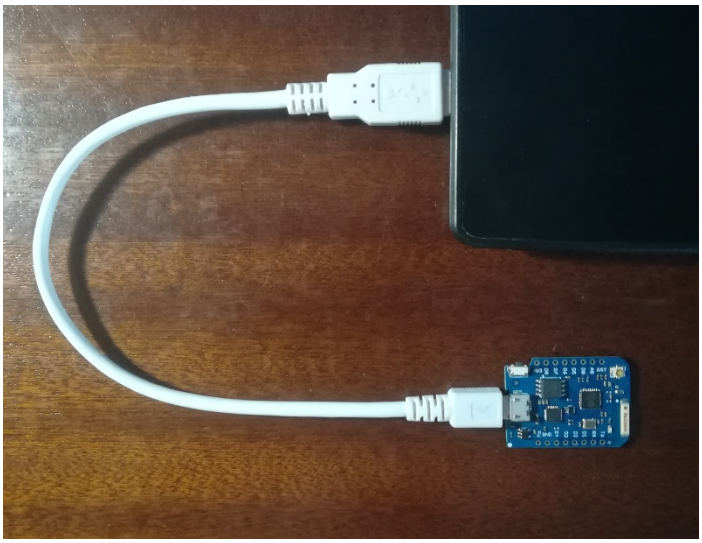


Figure 1: Photo of Master Node connected to a laptop.

3.1. Hardware Description

As mentioned before three types of nodes were used: a master node, static nodes and robot nodes. Master node and static nodes are composed by D1 mini pro boards with a wireless antenna Wi-Fi ESP8266, and a power supply. In the case of the master node the power supply was the PC it was connected to, in the case of static nodes a power banks were used.

Figure 1 is a photo of the master node used, connected to a laptop where the end user could send commands and received information from the network. Figure 2 is a photo of a static node used, which served the purpose of acting as a relay and extending the network.

The robot nodes were two wheeled mobile robots composed by a NodeMCU 1.0 board also with an ESP8266 module, a L293 Motor Driver Shield, two rubber wheels with Micro DC geared motors and a power bank, that served as a power supply. The

robots are completely similar in terms of hardware with just small differences in how the boards were mounted as can be seen in Figure 3 and Figure 4, which are photos of the first and second Robot Nodes used respectively.

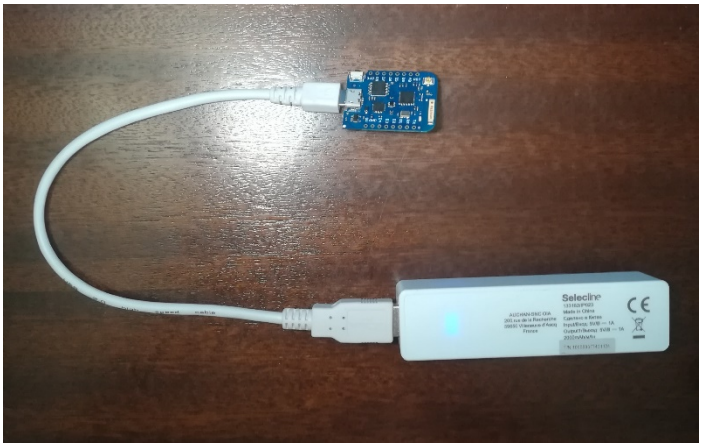


Figure 2: Photo of a Static Node.

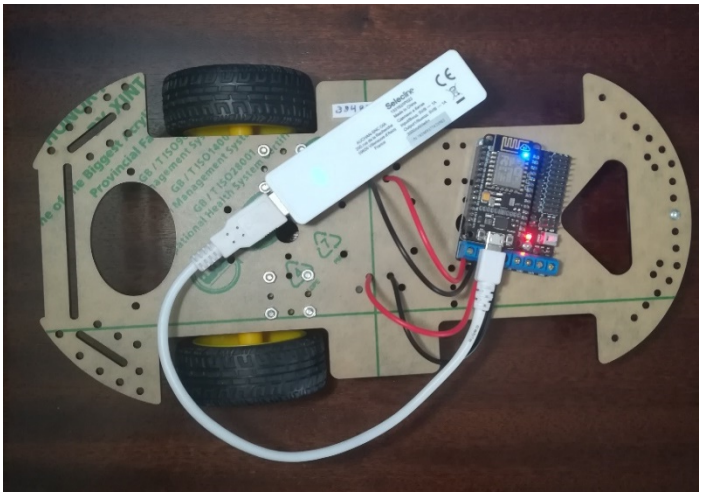


Figure 3: Photo of the first Robot Node used.

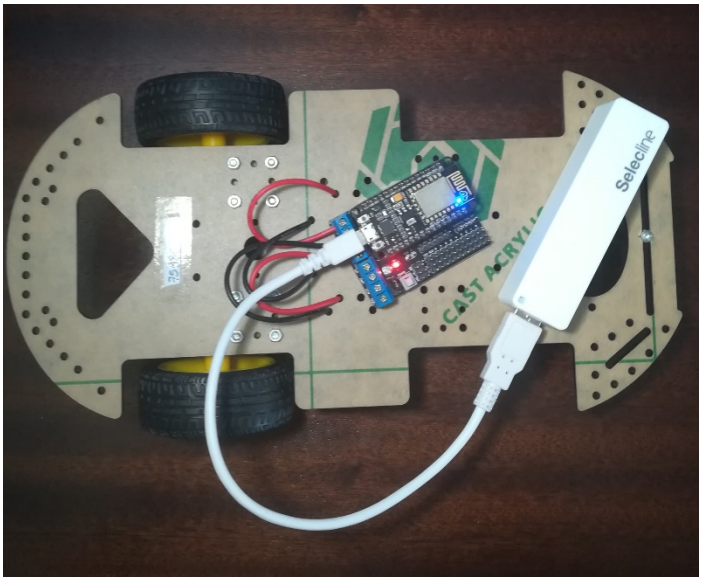


Figure 4: Photo of the second Robot Node used.

All the boards in each node were programmed using the Arduino programming language and respective Arduino IDE. The Arduino programming language makes use of two functions that are always present: setup and loop. Setup runs once when the board is connected to an energy supply. Loop runs continually after the setup, as long the board is connected to an energy supply. Future mentions to setup and loop, while explaining the code, refer to these two functions. The robots need to perform several tasks periodically, such as measuring the RSSI or moving. To achieve this behavior the Task Scheduler library was used.

3.2. Wireless Sensor Network Implementation

To manage creation of the wireless sensor network and communication between nodes the *painlessMesh* library was used [13]. This library uses the esp8266 hardware present in the boards to create a mesh network. The *painlessMesh* library does not use the Arduino WiFi libraries, due to performance problems, but the native esp8266 libraries [13]. The library allows us to make the nodes to communicate with each other, in a prototype of a WSN, without concern about how to structure or manage the network. The great advantage of using this library is that any operational nodes will automatically self-organize into a functional mesh network.

The messages between nodes are all JSON objects, which makes them human readable and facilitates integration with frontends and other applications. The mesh does not use or create a TCP/IP network, rather each node is identified by a 32bit chipId obtained through the esp8266 native SDK. As a result, every node will have a unique id, that can be used to communicate specifically to that node. Broadcast messages to all nodes in the mesh are also possible [13].

Some functions of the library are used in the setup of every node, namely init, which initializes the mesh, and the functions that set up callbacks, such as *onReceive(&receivedCallback)* and *onNewConnection(&newConnectionCallback)*. The functions given as arguments: *receivedCallback* and *newConnectionCallback*, are present in every node as well. They execute when a message is received by the node and when a new node connected to the network, respectively.

Other functions used at setup are *setRoot* for the master node and *setContainsRoot*, for all others. The purpose of these is to avoid the creation of submeshes during mesh formation. It makes sure all nodes know the gateway is the master node and a node once connected to the master node will not disconnect. This is necessary because of a limitation during mesh formation, namely that each node can only connect to another one node at any given moment. Because of this, nodes will try to randomly connect and disconnect until a full mesh is formed. This solution works but may create submeshes. The functions above aim to solve this problem. Another consequence of the “each node can only connect to one other node at a time” will be that usually only node of the mesh will be available to connect to other nodes.

A library function used in the loop of every node is the function update, which runs maintenance tasks and is required for the mesh to work. In the master node it is also used the *subConnectionJson* function, which allows to print the mesh topology in JSON format. All nodes use the *sendSingle* and

sendBroacast functions. The former allows to send a message to a node using its chipId, the latter sends a message to all nodes in the mesh.

Task Scheduler

The Task Scheduler library allows for periodic task execution. It is possible to specify execution period, in milliseconds or microseconds, and number of iterations, be they finite or infinite. The reason for needing to use this library is that the boards used cannot launch several threads at a time, hence cannot multitask with recurring to external libraries.

It works by first creating a scheduler using the Scheduler class. Then creating tasks using the Task class, and specifying the execution period, number of iterations and the actual function that will define the task behavior. Then at setup each task is added to the scheduler. From there on it is possible to enable or disable tasks as required [14]. The two main tasks used by the robot nodes are the RSSI Task and the Moving Robot Task.

The RSSI Task's main purpose is to measure the RSSI value. It will always do this if the scan for Wi-Fi networks is successful. Then depending on flags active it might just send this value to the master node or check if this value is lesser or equal to the RSSI threshold. If it is smaller, meaning more negative because the RSSI value is usually a negative integer, then it will send the RSSI value to the master node along with a request to place a new node. If the RSSI value is not smaller the robot can move further, and so it will call the appropriate movement function, disable the RSSI task and enable the Robot Moving task.

The Robot Moving Task's main purpose is to time the movement of the robot and send the RSSI value to the master node once a movement is finished. It is also used to create the robot node auto movement, as will be explained below.

The RSSI value is the most important metric in the prototype because it is based on it that the robot nodes will decide when a new node should be placed. Therefore, it regulates the extension of the WSN. It will also be used later both in the cooperation between robot nodes and in their response to deployment of new static nodes.

RSSI Request

It is useful to measure the value of the RSSI to understand how it changes, particularly due to growing distance from the master node. With this purpose in mind was created a way to simply have the master node send a request to a robot node and have the robot measure the RSSI and send its value back to the master node, to be printed and made available to the user.

The laptop connected master node expects user info from the serial monitor of the Arduino IDE. If the input received is the, case insensitive, string "RSSI" it will send a RSSI Request to all nodes. This request is simply a message that will prompt each node to enable the RSSI Task, which will make it measure the RSSI value, and then send this value back to the master node. The master node will print the value and an identifier of the node that sent it. Algorithm 1 represents the algorithm behind RSSI requests.

In terms of number of messages exchanged within the network, they are at least $2 * (n - 1)$, where n is the total number of nodes in the network. The reason for this number is that the master node will send a broadcast message upon being prompted by the user with an RSSI Request, so it will send a message to all nodes except itself, $n - 1$. Every node will then measure the RSSI and respond to the master node, so this number is doubled. Depending on the extension of the network this number will increase because the message may need to hop between nodes before reaching the master node.

Algorithm 1: RSSI Request

Result: The Master Node will print the RSSI values for all nodes in the network in its serial monitor
Initialization;
while *Master Node on* **do**
  Expects user input;
  if *User input = RSSI* **then**
    Send RSSI
    Request;
    Nodes send
    RSSI values;
    Print RSSI
    values;
  end
end

4. Testbed Deployment

With the testbed ready it is possible to start developing algorithms that will create us a prototype for an autonomous, extensible WSN. With this in mind, firstly it is necessary to develop ways of having the robot node move on its own and request new nodes for the network once a RSSI threshold is hit, which was called *AutoMOV*. Furthermore, ways of having two robot nodes cooperate with each other, to further extend the WSN, were developed, which was called *RCoop*. Finally, there was a need to develop ways for the robot nodes to know when a new static node was deployed and respond accordingly, extending the network if possible, which was called *ReStatic*. In this section the aim is to present each of these algorithms and explain how they were implemented using the testbed.

4.1. Automovement Algorithm (AutoMOV)

In the first implementation of the WSN prototype the robot node moved by commands sent from the master node by user input. But the goal was autonomous movement and deployment of an extensible WSN. To achieve this the robot nodes need to move autonomously, not prompted by user input, so an algorithm for robot node automovement was developed, hereinafter referred to as *AutoMOV*. All movement at this point occurs along a line.

An auto movement request is made if the Master Node catches a string that starts with 'A' or 'M'. This calls the function *AutoMOVRequest*. This function will print if the request is to start auto movement (A) or to stop auto movement and go back to manual (M). It will also send the given string to the robot node by method *sendMessage*. The method *sendMessage* uses the communication functions from the Painless Mesh library,

sendSingle or *sendBroadcast*, but with arguments to choose between them.

The robot node responds to this *sendMessage* with the call-back routine *receivedCallback*. In the case of 'A' it will turn the movement type to 'F', which will make the robot move forward, this is the case because all movement occurs along a line at this point. Later it should decide the movement type based on sensor information.

It will also activate an *AutoMOV* flag and enable the RSSI Task. In the case of 'M' it turns off the *AutoMOV* flag and stops movement.

Algorithm 2: Automovement (*AutoMOV*)

```
Result: Robot Node performs discrete movement until it hits
its RSSI Threshold
Initialization;
while Master Node on do
    Master Node expects input
    if User input = M do
        AutoMOV Stop Request;
        Robot Node stops;
    else if User input = A do
        AutoMOV Start Request;
        while Current RSSI <= RSSI Threshold do
            Robot moves;
        end
    end
end
```

The RSSI Task will then be responsible for making the robot always move unless the current RSSI is equal or below the threshold, which would make *AutoMOV* stop and the robot node to send a request to the master node, for placement a new static node. Otherwise it enables the Moving Robot Task. This last task has an extra 'if', to deal with auto movement, that basically says: if one movement was completed and the *AutoMOV* flag is on, then it will enable the RSSI Task again. This creates a loop between the two tasks. Algorithm 2 represents the algorithm behind *AutoMOV* requests.

4.2. Robot Cooperation Algorithm (RCoop)

Coordination between multiple robots, cooperating in deployment or exploration, is seen as an important concern by current WSN literature [3-6]. With this purpose in mind there is a necessity for algorithms that allow multiple robots to cooperate and for field tests involving more than one robot [4][6][7].

The goal then was to develop an algorithm that allows two robots to cooperate in a prototype of WSN deployment based on communication metrics, namely the RSSI. Hereinafter this algorithm will be referred as *RCoop*.

The two robot nodes were named Robot One (R1) and Robot Two (R2) to help differentiate between them. The only real distinction between robot nodes is that R1 will be the first to move. All movement described occurs along a line.

R1 will be sent a signal from the user and start movement. Each robot movement is time limited to ten seconds. After these ten seconds of movement it will measure the RSSI and based on this value decide to stop or move again. While the RSSI is above a certain threshold it will continue to perform discrete ten second movements. If the RSSI is found below the threshold it will stop, ask for the deployment of a static node and send a message to R2. Deployment of static nodes is done by human agents at this point. The reaction of robot nodes to the deployment of static nodes will discussed later. Here the main concern is about how the robots can cooperate.

R2 after receiving the message from R1 will know the first robot has reached the RSSI threshold and so commence its own movement. The message sent by R1 basically functions as R1 'calling' R2. R2 will mimic the behavior of R1 and move until it has reached the RSSI threshold. When R2 reaches the threshold, it will send a message to R1. This will prompt R1 to update its own threshold and start moving again. The goal here is to extend the network as much as possible using the two robot nodes.

The way this works in terms of code is that R1 will first be sent an *AutoMOV* request by the master node. It will normally follow *AutoMOV* until the threshold is reached. When it reaches the threshold, it will call *AutoMOVRequest*, which will send a message to R2. R2 will expect a message with this format via the callback routine *receivedCallback* and will activate its own *AutoMOV* because of it. R2 will then perform *AutoMOV*, as explained above, until it reaches its threshold. When at the threshold will send a message to R1 via *sendMessage*, informing it R2 is at the threshold. R1 will then decrement the *rssiThreshold* variable and turn its *AutoMOV* own again, and so move further until it reaches the new threshold value. Algorithm 3 represents the algorithm that deals with robot cooperation (*RCoop*).

Algorithm 3: Robot Cooperation (*RCoop*)

```
Result: The two robots will extend the network
Initialization;
R1 AutoMOV;
if R1 RSSI <= RSSI Threshold do
    R2 AutoMOV;
    if R2 RSSI <= RSSI Threshold do
        R1 expands Threshold;
        R1 AutoMOV;
    end
end
```

4.3. Response to Static Nodes Algorithm (ReStatic)

The other thing that is required to attain the goal of having an extensible WSN is that the network responds autonomously to the deployment of static nodes. To elicit this behavior, a communication metric is used again, the RSSI.

When a static node is placed it connects to the network normally, then it will measure its RSSI. It will send this value to the mobile robots connected. Then each robot will use this value to decide whether it should start movement or not. This decision is based on if the difference between its own RSSI and the static node RSSI is lesser than the value of the RSSI threshold. If the difference between RSSIs is lesser than the threshold then it means the static node is closer to the master node and hence the network can be safely extended, so the robot will move.

But if a new node connects it usually will connect as a subnode to the nodes already there, so how to solve this problem? The answer was to force a reorganization of the network. If the robot concludes by the above described method that the static node is closer to the master node then, before it initiates movement, it will reset its own board. The result of this is that when the robot reconnects to the network it will now be a subnode of the newest static node in the network topology. The value of the robot's threshold is saved in the master node, via a message sent by the robot before it resets. When the robot node reconnects the master node sends this value to the robot again, the later will update its threshold and reinitiate movement.

In terms of code this is achieved by using the RSSI Task in the static nodes as well. It will activate and measure the RSSI value as soon as the static node is turned on. Then it will send this value to the master node which will in turn send it to the robot nodes, via *sendMessage*. As usual the way to prompt behavior in a certain node, the node is made to expect a string formatted a certain way, via the *receivedCallback*, and if that string is received the desired behavior occurs. So, in that same way, when the robot nodes received the value of RSSI prefixed with the string "(mesh node)" they will first extract the RSSI value from that string. Then, they will check if the deployed static node can be considered closer to the master node, calculating if the difference between their RSSI value and the RSSI value of the static node is greater than their RSSI threshold (*currentRSSI - meshRssi > rssiThreshold*). If so, the robot node will expand its threshold and send it to the master node. The master node will save this value and send a message back to the robot node signaling the robot can now restart, with the purpose explained above. The robot node will then restart by esp8266 native function, *ESP.restart*, which will restart the board. The master node will expect a robot node connection via the *newConnectionCallback*, which is a callback routine that executes each time a new node connects the network. In case there is a saved RSSI threshold value for the robot node that connected then it will send this value to the robot that just connected. The robot will use this value to update its threshold and turn the *AutoMOV* on, thus extending the network. Algorithm 4 represents the response to static nodes algorithm described above. This algorithm will be known as *ReStatic*.

Algorithm 4: Response to Static Nodes (*ReStatic*)

Result: Robot Nodes extend network if possible
Initialization;
Static Node Deployed;
Send RSSI to Robot Nodes;
if *RSSI – Static RSSI > Threshold* **do**
  Expands Threshold and save it;
  Robot Node reset;
  Robot Node receives threshold;
  Robot Node *AutoMOV*;
end

5. Performance Evaluation

It was then time to put the algorithms to test to verify if the intended behavior of an extensible WSN could be achieved, take measurements, find potential faults or avenues for improvement. For these purposes four scenarios were developed, each progressively testing each one of the developed algorithms, first separately, then in conjunction. As the presented work is still in its initial proposal, some RSSI issues were considered out of the scope. All demonstration scenarios considered an open space area with no obstacles and, as much as possible, a linear variation of RSSI.

The measurements were conducted in an open space indoor area. The master node was connected to a laptop via the USB port and will collect RSSI data. The laptop is placed on a wooden support that measures 31 cm in height. Time and distance are collected by the testers, while RSSI values are registered by the master node.

Firstly, it was important to study the variation of RSSI based on the distance between a robot node and a master node. This would serve to understand how it naturally variates without accounting for any robot node movement and illustrate the relationship between distance and the RSSI value. It makes use of and proves the functionality of the RSSI Request.

The robot node was placed at several distances of separation to the master node, always one meter apart from 0m to 13m. At each distance, the user would elicit the RSSI value through RSSI Request. After collecting ten samples at that distance of separation an RSSI average was calculated. Figure 5 shows the RSSI averages for the various distances of separation to the master node.

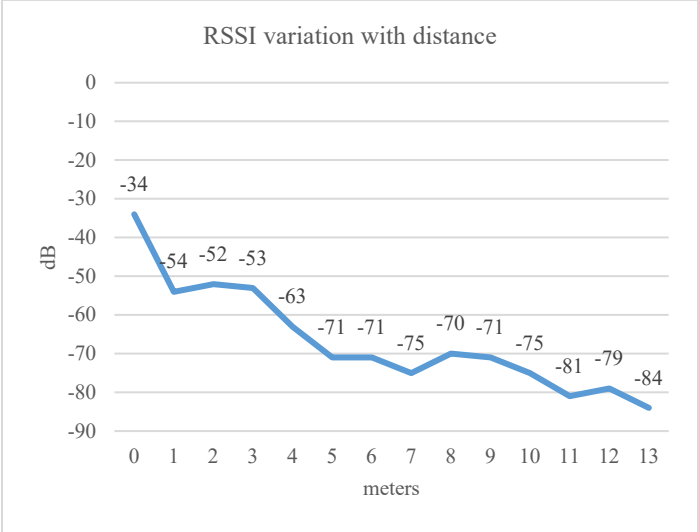


Figure 5: Robot Node RSSI variation with distance from the Master Node.

5.1. Scenario 1: One Robot Node

This first scenario intends to verify how much it was possible to extend the network simply using two nodes: the master node and a robot node. It will also illustrate the functionality of the robot node *AutoMOV* algorithm.

The robot node is place on the floor at a 0 meters distance from the master node. It will be sent a signal from the master node which will start the *AutoMOV*. Then, the robot node will perform discrete 10 second movements along a line while measuring the RSSI. When its current RSSI value is below a certain threshold it will stop and request the placement of a new static node. Figure 6 represents the end state for this scenario.

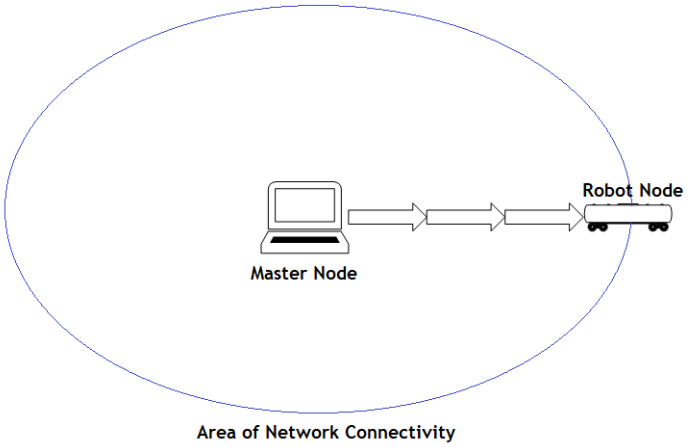


Figure 6: End state for Scenario 1 composed of Master Node and a Robot Node.

After ten trials an average of the values of time since *AutoMOV* started till robot node stoppage, distance traversed by robot node and last RSSI recorded is calculated. Table 1 holds the recorded values for this first testing using a robot node.

Table 1: Averages of time, distance and RSSI for 1 Master Node 1 Robot Node.

Time	54''
Distance	4m 83cm
RSSI	-68 dB

5.2. Scenario 2: Two Robot Nodes

The second scenario intends to extend the WSN using two robot nodes. This scenario uses two robot nodes and the master node and is meant to demonstrate the functioning of the *RCoop* algorithm.

Both robot nodes are placed next to one another at a 0m distance from the master node. Then the user enters 'A' into the serial monitor of the master node in the Arduino IDE. This is the signal that will set the *AutoMOV* of robot node named R1 on. All other network extending behavior is autonomous. Figure 7 represents the end state for this scenario.

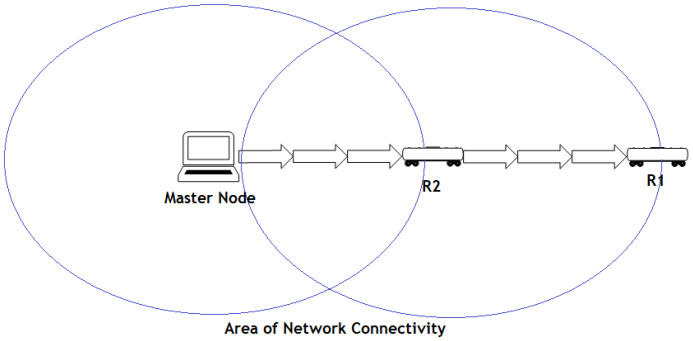


Figure 7: End state for Scenario 2 composed of Master Node and two Robot Nodes.

The last RSSI values of both R1 and R2 are recorded by the master node. The distance of both robot nodes to the master node after both robot nodes stop is recorded. The duration from the time the signal is sent until both robot nodes stop is also recorded. After ten trials an average of RSSI, distance and time was calculated. Table 2 holds the recorded values for this first test with the two robot nodes.

Table 2: Averages of time, distances and RSSIs for 1 Master Node, 2 Robot Nodes.

Time	3' 01''
R1 Distance	7m 12cm
R2 Distance	5m 11cm
R1 RSSI	-71 dB
R2 RSSI	-64 dB

5.3. Scenario 3: One Robot Node, One Static Node

This third scenario intends to extend the WSN, by guaranteeing the robot nodes will respond to the deployment of new static nodes. So, it serves to test the functionality of the *ReStatic* algorithm. It does so by using the master node, one static node and only one robot node.

The robot node will be placed at a 0m distance from the master node. Robot node *AutoMOV* will be turned on by user input and then it will move until it reaches its RSSI threshold. Then, a static node will be place at this threshold. The robot node will respond to this placement autonomously and move again till it reaches a new threshold, extending the network. Figure 8 represents the end state for this scenario.

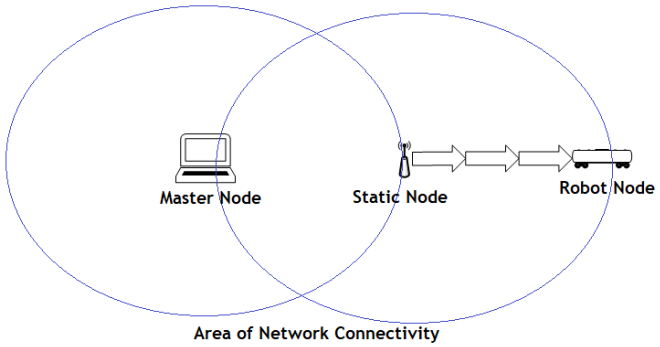


Figure 8: End state for Scenario 3 composed of Master Node, a Static Node and a Robot Node.

The last RSSI value before robot node stoppage is recorded by the master node. Time until robot node stoppage and maximum distance from static node are recorded by human users. After ten trails an average of these values is calculated. Table 3 holds the values for this first test that uses a static node.

Table 3: Averages of time, distance and RSSI for 1 Master Node, 1 Robot Nodes, 1 Static Node.

Time	3' 40''
Distance	7m 1cm
RSSI	-73 dB

5.4. Scenario 4: Two Robot Nodes, One Static Node

Finally, the fourth scenario allies the *RCoop* algorithms and the static node response algorithm. For this last scenario a master node, two robot nodes and one static were used.

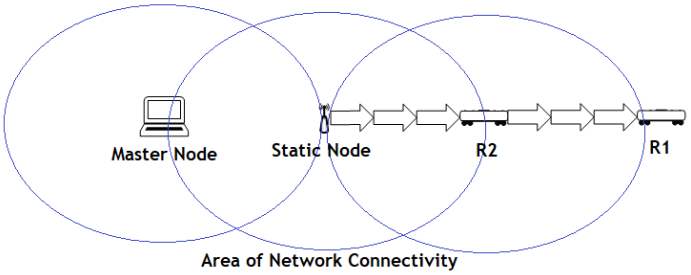


Figure 9: End state for Scenario 4 composed of Master Node, a Static Node and two Robot Nodes.

The procedure is similar to the two robot nodes test except that a static node is place at the R2 threshold when both robot nodes and stopped. So, both robot nodes are placed at 0m distance from the master node. A signal is sent to R1 by the user from the master node. Then, robot node *AutoMOV* and *RCoop* occur autonomously. Finally, a static node is placed at the R2 threshold and the robots will identify it and move again to new thresholds. Figure 9 represents the end state for this scenario.

The last RSSI values before robot node stoppage for both robot nodes are recorded by the master node. The time from signal sent to R1 till the network stops extending is recorded by a tester. Final distances of each robot node to the master node are also recorded by a tester. After ten trials an average of RSSIs, distances and time was calculated. Table 4 holds the values for this final test that incorporates all the node types: master node, robot node and static node.

Table 4: Averages of time, distances and RSSIs for 1 Master Node, 2 Robot Nodes, 1 Static Node.

Time	4’ 33’’
R1 Distance	11m 50cm
R2 Distance	6m 76cm
R1 RSSI	-82 dB
R2 RSSI	-71 dB

6. Discussion

First, it is important to have an idea of how to expect the RSSI to change with growing distance from the master node, this is the purpose of the RSSI – distance test. It is expectable that RSSI value decreases with distance from the master node, seeing as at 0m the RSSI value is -34dB and at 13m distance the RSSI is -84m. But this variation is not linear, so it must be assumed that there are other factors contributing to RSSI change besides distance.

Distance reached by the robot nodes is a key factor in knowing if the purpose of the algorithms was fulfilled and the WSN is extending. If only the number of nodes used is considered, it is expected the distance from the master node will increase with the increase in the number of nodes used, as the WSN can extend itself further. In this case the one robot test would give us the least distance, with just one node used, and the two robot nodes and one static node test would give us the furthest distance, with three nodes used. Figure 10 shows us the distances obtained according to the numbers of nodes used. It is then verified that the network is extending itself through the programmed behavior of the robot nodes.

Time was considered of less importance at this point in testing. Still it is noted a big increase, going from an average 54 seconds with one node to 4 minutes 33 seconds with three nodes. The way of reorganizing the mesh network, by resetting a robot node when a suitable static node is detected, should be reviewed. The current solution causes a lot of overhead.

Using a communication metric such as the RSSI gives rise to some problems. First off, must be said that when considering the RSSI threshold as the point where a robot node should stop and request a static node, this is a virtual limit. It is not exploring the actual hardware limit of the communication module, the wireless antenna Wi-Fi ESP8266. Explorations on RSSI variations using ESP8266 modules have used distances of up to 140m between a station and an AP and still reach acceptable values for RSSI [15]. These experiments were performed outdoors whereas the ones here were indoors, but they give an idea of the distances that could be reached if the limits of the hardware were explored. The aim here however was to develop algorithms that could generate the intended behavior on the prototype, namely the creation of an extensible WSN. Although it is legitimate to consider that testing the present proposal at the hardware limits would be interesting and closer to a live deployment scenario.

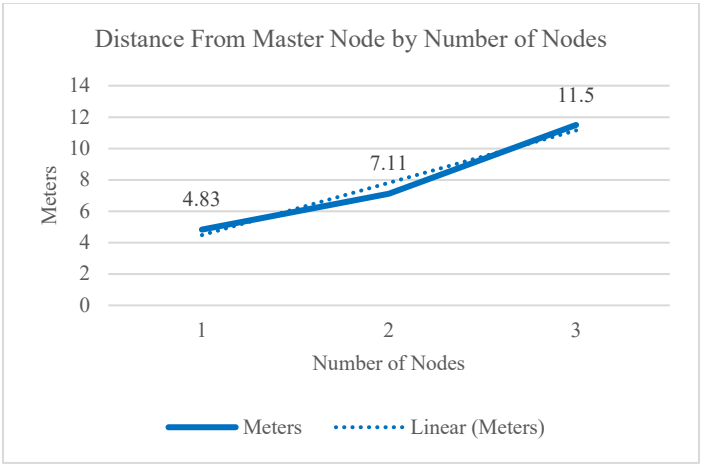


Figure 10: Variation of distance from the Master Node with number of nodes used in the WSN.

Finally, it is important to consider how the problems in expanding the RSSI threshold to make the robot nodes move further. Firstly, there is a need to find a better way to judge the distance of the static nodes when deployed. Deployment of several static nodes in proximity, or the same static node disconnecting from the network and then connecting again, might impress false information on the robot node, prompting it several times to move further. Secondly, saving the current threshold of the robot nodes in the master node when it resets could be problematic if the robot node disconnects from the network more than one time. This problem also arises from the method of reorganizing the network through resets of robot nodes. Even as it stands there might be a better way for the robot node to judge its threshold from the network topology and its own RSSI measurements, instead of receiving it from the master node.

7. Conclusion and Future Work

WSNs are ever more prevalent in modern living. Many applications, such as in disaster situations, require impromptu deployment of a WSN. Therefore, it is important to explore ways of WSN deployment. The usage of mobile robots to aid in this development is also important because it avoids the necessity of human agents. Coordination between mobile robots is essential in deployment and exploration tasks.

With this in mind a prototype of a WSN aimed at creating an extensible wireless network and proved it possible, within certain assumptions, using cooperation between mobile robots and responsiveness to the deployment of static nodes.

As for future work it should be interesting to test the solution proposed here with the actual hardware limits of the ESP8266 wireless module. Also, there is a need to develop a better way to manage network connections, even if that means foregoing the *painlessMesh* library and using the native ESP8266 SDK. Within the solution here, a better way to manage the RSSI thresholds of the robot nodes might also be a good avenue for new developments. Testing the WSN prototype with sensors able to capture environmental data would also prove interesting and open new avenues of exploring, such as how to effectively route sensing data.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work is funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/50008/2020. The authors would like to acknowledge the company InspiringSci, Lda for the interest and valuable contribution to the successful development of this work.

References

[1] J. Amaro, I.D.L.T. Diez, M.L.P. Joao, S. Vasco, J. Galan-Jimenez, "Protótipo de uma Rede de Sensores Sem Fios para Implantação Robótica de Percurso : A Prototype of a Wireless Sensor Network for Robotic Path Construction," Iberian Conference on Information Systems and Technologies, CISTI, **2020-June**, 20–25, 2020, doi:10.23919/CISTI49556.2020.9141080.

[2] S. Stage, P. Report, S. Kumar, R. No, C. Science, "Design and deployment of Wireless Sensor Networks," 313–325, 2017.

[3] Y. Sankarasubramaniam, E. Cayirci, others, I.A.. Su, "A survey on sensor networks," IEEE Communications Magazine, **40**(8), 102–116, 2002.

[4] A. Chattopadhyay, A. Ghosh, A. Kumar, "Asynchronous Stochastic Approximation Based Learning Algorithms for As-You-Go Deployment of Wireless Relay Networks Along a Line," IEEE Transactions on Mobile Computing, **17**(5), 1004–1018, 2018, doi:10.1109/TMC.2017.2750147.

[5] T. Suzuki, R. Sugizaki, K. Kawabata, Y. Hada, Y. Tobe, "Autonomous deployment and restoration of Sensor Network using mobile robots," International Journal of Advanced Robotic Systems, **7**(2), 105–114, 2010, doi:10.5772/9696.

[6] G. Tuna, V.C. Gungor, K. Gulez, "An autonomous wireless sensor network deployment system using mobile robots for human existence detection in case of disasters," Ad Hoc Networks, **13**(PART A), 54–68, 2014, doi:10.1016/j.adhoc.2012.06.006.

[7] A. Chattopadhyay, M. Coupechoux, A. Kumar, "Sequential Decision Algorithms for Measurement-Based Impromptu Deployment of a Wireless Relay Network Along a Line," IEEE/ACM Transactions on Networking, **24**(5), 2954–2968, 2016, doi:10.1109/TNET.2015.2496721.

[8] H. Liu, J. Li, Z. Xie, S. Lin, K. Whitehouse, J.A. Stankovic, D. Siu, "Automatic and robust breadcrumb system deployment for indoor firefighter applications," MobiSys'10 - Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services, 21–34, 2010, doi:10.1145/1814433.1814438.

[9] R. Soua, L. Saidane, P. Minet, "Sensors deployment enhancement by a mobile robot in wireless sensor networks," 9th International Conference on Networks, ICN 2010, 121–126, 2010, doi:10.1109/ICN.2010.29.

[10] N. Zhou, X. Zhao, M. Tan, "RSSI-based mobile robot navigation in grid-pattern wireless sensor network," Proceedings - 2013 Chinese Automation Congress, CAC 2013, 497–501, 2013, doi:10.1109/CAC.2013.6775785.

[11] O.M. Elfadil, "Navigation algorithm for mobile robots using WSN," Proceedings - 2013 International Conference on Computer, Electrical and Electronics Engineering: "Research Makes a Difference", ICCEEE 2013, 554–559, 2013, doi:10.1109/ICCEEE.2013.6634000.

[12] A.C. Jiménez, S.J. Bolaños, J.P. Anzola, "Decentralized model for Autonomous Robotic Systems based on wireless sensor networking," ARPN Journal of Engineering and Applied Sciences, **11**(19), 11378–11382, 2016.

[13] painlessMesh / painlessMesh · GitLab, Jul. 2020.

[14] GitHub - arkhipenko/TaskScheduler: Cooperative multitasking for Arduino, ESPx and STM32 microcontrollers, Aug. 2020.

[15] Yoppy, R.H. Arjadi, H. Candra, H.D. Prananto, T.A.W. Wijanarko, "RSSI Comparison of ESP8266 Modules," 2018 Electrical Power, Electronics, Communications, Controls and Informatics Seminar, EECCIS 2018, 150–153, 2018, doi:10.1109/EECCIS.2018.8692892.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 175-184 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Applications of the Heuristic Optimization Approach for Determining a Maximum Flow Problem Based on the Graphs' Theory

Simona Kirilova Filipova-Petrakieva*

Department Theory of Electrical Engineering, Faculty of Automation, Technical University of Sofia, Sofia, 1797, Bulgaria

ARTICLE INFO

Article history:
Received: 17 July, 2020
Accepted: 09 October, 2020
Online: 08 November, 2020

Keywords:
Maximum flow optimization
Graphs' theory
Power systems
Information systems
Communication systems
Transport systems
Manufacturing systems

ABSTRACT

In the present paper, a universal approach for determining the maximum flow problem in directed graph for solving different problems is applied. It can be considered as a heuristic and it can be used as a denomination for analyzing an arbitrary system with a mathematical description as a directed graph. The physical nature of the flows passing through the arcs of the system considered can be energy, information, transportation, or material. Then the studied system will be power, communication, transport, or manufacturing, respectively. In the present article five types of maximum flow problems are considered. Each of them is solved analytically by the universal heuristic method. The common between these problems is expressed in the fact that its behavior can be presented with the same mathematical model in the form of a directed graph including one initial (pending) vertex and one final (blocked) vertex, respectively. These vertexes are called either real (if they exist) or fictive (if they are introduced additionally) source and receiver depending on the topology of the associated directed graph's model. The final solution obtained with this approach for Problem 1 is compared with the similar one found through Ford-Fulkerson's, Edmonds-Karp's and Dinic's algorithms and it has been shown to be better than those determined by them.

1. Introduction

The maximum flows problem in industrial systems (energy, information or material) can be modeled by graphs' theory which adequately describe the topology of the studied system [1].

Note: For simplicity of the present analysis, it will be used only the term flow, regardless of type of the studied system - power [2, 3], information [4], communication [5, 6] or transportation [7].

Energy flows are optimized in power systems. In this sense, in [8] a review of the stochastic optimization methods, solving the optimal power flow problem. An application of the graphs' theory for the energy flows optimization in electric networks is made in [9]. The transformation of the linear model (e. g. linear goal function and linear constraints as unequations) to the directed graph's model is made there. Thus, the optimization procedure is expressed analyzing the resulting directed graph.

The definition of the flows problem in information systems is made in [10]. In [11] a review of approaches modelling information flow for organizations is presented. A new graph theoretical concept for analyzing the information flows passing through the directed graph's model, using guessing numbers is proposed in [12].

In [13], the problem of optimization the information flows passing through the uncertain networks with a given constrained capacities of the arcs is considered. The solution bases on the two NP-hard subproblems. First, it calculates the expected information flow associated to a source vertex and next it makes an optimal choice of arcs.

The flows optimization passing through the manifacturing systems is widely discussed issue. One possible solution to this problem is given by the graph's theory. For example, in [14] using the shortest path method for flows optimization in cellular manufacturing systems. In [15] an optimization for single-part flow-line configurations of reconfigurable manufacturing systems (RMS) is suggested. In [16] the other optimization approach of the flows in RMS systems is proposed. This approach also uses the graphs' theory for modeling the considered system. In [17] third model for optimizing demand period cost (fixed plus operating) of RMS flow-line configurations, based also on graphs' theory, is suggested.

A wide range of problems arise in real-world networks such as Internet, local information networks, energy systems, etc., which

*Corresponding Author: Simona Kirilova Filipova-Petrakieva, petrakievas-te@tu-sofia.bg

are modeled by directed graphs. In [18] a review of a range of theoretical primitives for generating a simple graph based on the given degree sequence is proposed. This is based on the graphs' theory and the computer networks. Several problems for which efficient algorithms for determining the theoretical solutions are proposed. Since in practice the graphs' models consists of tens of thousands of vertexes, approximation algorithms are particularly important.

The present paper is organized as follows. In a Section 2 the directed graph's model with one source – one recipient is defined, as well as the three necessary and sufficient conditions for performing its topological analysis. Five basic tasks arising in real systems, modeled with directed graphs, are also defined. In Section 3 an universal algorithm for solving all these five problems is briefly described. Solutions of these five type of problems applying the universal algorithm is proposed in Section 4. The applicability of this algorithm is visualized in Section 5 on an illustrative example describing problem 1. In the same section, a comparison is made with the obtained solutions to the same illustrative example using traditional methods based on minimum-cut the theorem, such as: Ford Fulkerson, Edmonds Carp and Dinic algorithms. The paper finishes with the conclusion remarks on the universality and advantages of the suggested algorithm.

The novelty of the proposed method consists in its universal application in determining the maximum flow passed through a arbitrary directed graph, whose model contains one initial and one final nodes. It is applicable to analysis of different systems - information, power, manufacturing, etc. The algorithm can be used for solving the five types of problems, defined in the next section

2. Problem statement

The maximum flow problem passing through the arcs of an directed graph is defined in [1].

In order to refine the considered graph's model, some basic concepts and initial hypotheses will be briefly introduced.

Let's consider a directed graph $G = < V, L >$, with $V = \{V_s\}$, $s = \overline{1,n}$ vertexes and $L = \{L_{ij}\} = \{L^k\}$, $k = \overline{1,m}$ arcs. The latter is associated to the system's channels. Let's $f_{ij} = f(V_i \to V_j) = f(V_i, V_j)$ is the flow passing through the arc L_{ij}. The capacity of this arc is q_{ij} [4].

A necessary and sufficient condition for determining the maximum flows passing through the arcs of the studied directed graph satisfies the following three assumptions.

Assumption 1

Balance condition: The sum of the incoming flows in a vertex V_i is equal to the respective sum of the outgoing ones, i.e.

$$\begin{aligned} &\sum_{V_j} f(V_i, V_j) - \sum_{V_s} f(V_s, V_i) = 0 \\ &V_j \in M^-(V_i);\ V_s \in M^+(V_i) \end{aligned} \quad . \tag{1a}$$

where: V_i is intermediate vertex for the considered directed graph;

$M^+(V_i)$ and $M^-(V_i)$ are the sets of input and output vertexes for the vertex V_i, respectively.

Assumption 2

Sequencing condition: The flow f_{ij} passing through the arc L_{ij} must not exceed its capacity q_{ij}., i.e.

$$f(V_i, V_j) \le q_{ij};\ \forall\ L_{ij} \in L. \tag{1b}$$

Assumption 3

Invariability condition: In absolute value, the flow passing through the directed graph from the source vertex to the recepient vertex must be exactly equal to that in the opposite direction.

$$\begin{aligned} &\sum_{V_j} f(V_{src}, V_j) - \sum_{V_i} f(V_i, V_s) = f_\Sigma \\ &V_j \in M^+(V_{src});\ V_i \in M^-(V_{src}) \\ &\sum_{V_j} f(V_r, V_j) - \sum_{V_i} f(V_i, V_{rec}) = -f_\Sigma \\ &V_j \in M^+(V_{rec});\ V_i \in M^-(V_{rec}) \end{aligned} \quad , \tag{1c}$$

where f_Σ is the flow passing through the directed graph.

$$\begin{aligned} &f_\Sigma = \sum_{V_j} f(V_{src}, V_j) = \sum_{V_i} f(V_i, V_{rec}) = 0 \\ &V_j \in M^+(V_{src});\ V_i \in M^-(V_{rec}) \end{aligned} \quad . \tag{2}$$

The first condition can be violated if part of the flow is dissipated in a vertex of the graph. This can be observed in the energy transfer, when each vertex is a specific physical device that consumes electricity and at the same time can dissipate it in the form of heat.

The second condition cannot be violated, since real systems are considered where it is physically impossible for a flow to pass through an arc that is greater than its capacity.

In real energy systems there are transfer losses, conversion losses, etc. In this case moving in the both directions of the graph (from the source to the recipient and vice versa) these losses cannot be equal, but they are very close.

Note: Without a loss of generality it is assumed, that the directed graph's model of the system has 1 source and 1 recipient vertexes. In general case, it can have multiple sources (N_s) and multiple recipients (N_r). Then the graph is transformed to a directed graph of the first type by adding fictive source V_{src} and recipient V_{rec} vertexes, respectively (Figure 1).

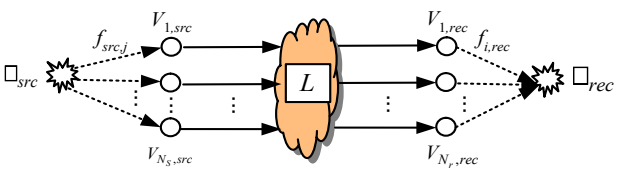


Figure 1: Graph's model with multiple sources (N_s) - multiple recipients (N_r)

In real systems depending on their parameters the following 5 types of maximum flow problems can be arised. In the first four of them the flow through the arcs f_{ij} is constant, and in the last one – it varies.

Problem 1: The arcs' capacities q_{ij} are known and the capacities of all graph's vertexes are infinity, i.e. $P_s \to \infty$.

Problem 2: The arcs' capacities q_{ij} are known and limited in a given range $q_{ij}^{\min} \le q_{ij} \le q_{ij}^{\max}$.

Problem 3: The arcs' capacities q_{ij} and the capacities P_s of the graph's vertexes $V_s \in V$ are known.

The Problems 1, 2 and 3 discuss the directed graph's model with one source and one recipient.

Problem 4: It consideres the directed graph with multiple sources $\{V_{src,l}\}$, $l = \overline{1, N_s}$ and multiple recipients $\{V_{l,rec}\}$, $l = \overline{1, N_r}$ (Figure 1). The arcs' capacities q_{ij} are known.

Problem 5: The arc's capacities q_{ij} of all the graph's arcs $L_{ij} \in L$ are known, but the balance condition (1a) is violated.

$$f_{in}\left(V_i, V_j\right) > f_{out}\left(V_i, V_j\right) \tag{3a}$$

or

$$f_{in}\left(V_i, V_j\right) < f_{out}\left(V_i, V_j\right), \tag{3b}$$

where $f_{in}\left(V_i, V_j\right)$ and $f_{out}\left(V_i, V_j\right)$ are the input and the output flows, associated to the arc $L_{ij} = \left\langle V_i, \quad V_j \right\rangle$, respectively.

For each of the above defined problem it obtains the maximum flow passing through the directed graph from the source to the recipient vertexes.

The basic aim and the contribution of the author of this paper consists of the suggesting the universal general algorithm for obtaining the maximum flow started from the source vertex V_{src} and reached to the final graph's vertex – either real V_{rec} or fictive $\mathbf{V}_{rec}$.

There are always losses when transmitting information or energy through the graph's model from the source to the recipient. They decrease the flow $f_{in}\left(V_i, V_j\right)$, i.e. the condition (3a). The condition (3b) is satisfied in some types of manufacturing systems.

For all 5 maximum flow problems defined above, *the condition* (1b).

The input flow for the system is generated by the source vertex V_{src} and it is named as $f_{src} = f(V_{src})$. The output flow $f_{rec} = f(V_{rec})$ is the flow, consumed by the recipient V_{rec}. It s assumed that the flow passing through the intermediate graph's vertexes is zero.

The different types of flows are defined in [1]. For the future analysis, we will describe them again in briefly.

Definition 1: The **allowed** flow is:

$$f^{allowed} = \sum_{M(V_i)} f_{out}\left(V_i, \quad V_j\right) - \sum_{M(V_j)} f_{in}\left(V_i, \quad V_j\right) = \begin{cases} f(V_{src}), & V_i = V_{src} \\ -f(V_{src}), & V_i = V_{rec} \\ 0, & \forall V \ne V_{src}, V_{rec} \end{cases}. \tag{4}$$

Definition 2: The **maximum flow** is:

$$f^{\max} = \max_{i=1, n_{\max}} \left[f_i^{allowed}\left(V_{src}, \quad V_{rec}\right) \right] \tag{5}$$

where $f_i^{allowed} = \left(V_{src}, \quad V_{rec}\right)$ are all allowed flows with number i passing through the directed graph from the source to the recipient vertexes.

Definition 3: The **conventionally optimal flow** $f^{opt} = \left(f^{opt}\left(V_{src}\right), \quad f^{opt}\left(V_{rec}\right)\right)$ is, if arbitrary allowed flow $f_i^{allowed} = \left(f_{src}, \quad f_{rec}\right)$ satisfies at least one of the conditions (6)

$$\begin{aligned} &f^{+}\left(V_{src}\right) \ge f^{-}\left(V_{src}\right), \ if \ \ f^{+}\left(V_{rec}\right) \ge f^{-}\left(V_{rec}\right) \\ &f^{-}\left(V_{rec}\right) \ge f^{+}\left(V_{rec}\right), \ if \ \ f^{+}\left(V_{src}\right) \ge f^{-}\left(V_{src}\right) \end{aligned}, \tag{6}$$

where: $f^{+}(V_{src})$ - the flow, generated from the vertex V_{src} passing through the graph's model from V_{src} to V_{rec};

$f^{-}(V_{src})$ - the flow, consumed of the vertex V_{src} passing through the graph's model from V_{rec} to V_{src};

$f^{+}(V_{rec})$ - the flow, consumed of the vertex V_{rec} passing through the graph's model from V_{src} to V_{rec};

$f^{-}(V_{rec})$ - the flow, generated from the vertex V_{rec} passing through the graph's model from V_{rec} to V_{src}.

Definition 4: The **maximum allowed flow** is $f^{*} = \left(f^{*}\left(V_{src}\right), \quad f^{*}\left(V_{rec}\right)\right)$, if it satisfies all the conditions (4), (5) and (6) simultaneously.

3. Algorithm of the universal approach for determining the maximum flow problems

Each problem defined in the previous paragraph can be reduced to the optimization maximum flow problem, where the the goal function is:

$$f_{\Sigma}^{\max} = \max \sum_{V_j} f\left(V_k, V_j\right) = \max \sum_{V_i} f\left(V_i, V_r\right) \tag{7}$$

and the constraints:

$$\begin{cases} \sum_{V_j} f(V_i, V_j) - \sum_{V_s} f(V_s, V_i) = 0 \ \ if \ \ V_i \neq V_{src}, \ V_i \neq V_{rec} \\ \sum_{V_j} f(V_i, V_j) - \sum_{V_s} f(V_s, V_i) = f_\Sigma; \ V_i = V_{src} \\ \sum_{V_j} f(V_i, V_j) - \sum_{V_s} f(V_s, V_i) = -f_\Sigma; \ V_i = V_{rec} \end{cases}, \quad (8)$$

where $f_\Sigma^{\max}$ is the maximum flow passing through the minimum intersection $L_{\min}^{C}$ of the considered directed graph.

The basic directed graph, which is divided into two separate ones G^A and G^B this function is:

$$f_\Sigma^{\max} = f_\Sigma^{\max}(L_{\min}^{C}) = \sum_{L_{\min}^{C}} q_{ij}, \quad (9)$$

where: $L_{\min}^{C} = \left[G^A \to G^B \right]$ $\quad V_{src} \in G^A$; $\quad V_{rec} \in G^B$;
$G^A \cap G^B = \varnothing$; $G^A \cup G^B = G$.

The approach, solving the optimization problem (7) with constraints (8) is proposed and described in details in [19]. Each iteration consists of two procedures – vertexes marking and flows correction. The iteration process stops when the minimal intersection $L_{\min}^{C}$ in the directed graph's model is determinated.

Before starting the algorithm, it is necessary to know the structure of the directed graph modeling the considered system as well as the capacities q_{ij} of each graph's arc. It is assumed that the initial values of the flows f_{ij} passing through the graph's arcs $L_{ij} \in L$ are zero.

In **Procedure 1** a vertexes marking is made. Each node is assigned the appropriate pair of marks, moving from source to the recipient vertexes. This marking indicates where the respective flow comes from and its maximum allowable value that can be passed through the respective arc. When all vertexes are marked then the algorithm goes to the **Procedure 2**. This procedure moves from the recipient to the source vertexes correcting the flows passing through each of the arcs in the respective sequences of arcs in the considered directed graph. If all the arcs entering the vertex V_{rec} are saturated, then the maximum flow $f_\Sigma^{\max}$ is determined and the procedure stops. Otherwise it returns to the beginning of Procedure 1.

Note: An arc is saturated if the flow that passes through it f_{ij} is exactly equal to its capacity q_{ij} .

4. Solutions to the defined 5 types of the problems

The maximum flow determination approach, proposed in [19], can be applied to solve each of the typical problems, defined in Section 2.

Problem 1

Before starting the algorithm the directed graph's model is determined. In this model the values of the arcs capacities $\{q_{ij}\}$ and vertexes infinity capacities $P_s \to \infty, \ s = \overline{1,n}$ are known.

To simplify the future exposition and without loss of the generality the maximum flow $(f_\Sigma^{\max})$ is obtained for the intersection of the directed graph's model shown in Figure 2. For simplicity the elements (vertexes and arcs) included in the beginning of the graph are aggregated in the set M^- .

In this problem the condition (1b) is satisfied, i.e. the flow f_{ij} passing through the arc L_{ij} , have to be smaller or equal its capacity q_{ij} .

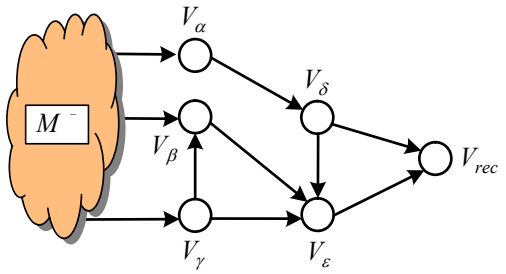


Figure 2: Idealized system described by the graph's model

Based on the minimum intersection $L_{\min}^{C}$ (see Figure 2), which satisfy the condition (4), the respective maximum flows are:

$$\begin{cases} f_{\Sigma_A}^{\max} = q_{\beta\varepsilon} + q_{\gamma\varepsilon} + q_{\delta\varepsilon} \\ f_{\Sigma_B}^{\max} = q_{\delta,rec} \end{cases}. \quad (10)$$

Conclusions:

- The maximum allowed flow $f_{\Sigma_A}^{\max}$ is equal to the sum of the capacities of the arcs $L_{\beta\varepsilon}$, $L_{\gamma\varepsilon}$, $L_{\delta\varepsilon}$, because they beyond to the minimum intersection $L_{\min}^{C}$ and they are saturated (their capacity is exhausted).
- The maximum allowed flow $f_{\Sigma_B}^{\max}$ for the arc $L_{\alpha\delta}$ is equal of the capacity of the arc $L_{\delta\varepsilon}$, because it is assumed that the condition $q_{\delta\varepsilon} < q_{\alpha\delta} + q_{\delta\varepsilon}$ is satisfied. Therefore, the arc $L_{\delta\varepsilon}$ is saturated and it is added to the minimum intersection $L_{\min}^{C}$.
- The final maximum flow, determined by Eq. (4) and Eq. (8) is $f_\Sigma^{\max} = f_\Sigma^{\max}(L_{\min}^{C}) = f_{\Sigma_A}^{\max} + f_{\Sigma_B}^{\max}$.

Problem 2

Before starting the algorithm the directed graph's model is determined. In this model the values of the arcs capacities $\{q_{ij}\}$ are known and each of them must belong to an initially defined range, i.e. $\leq q_{ij} \in \left[q_{ij}^{\min}, \ q_{ij}^{\max} \right]$.

The solution of the Problem 2 requires a transformation the basic directed graph's model to the new one, where:

- Each arc is replaced by two new arcs - L_{rg} and L_{gs}.
- The capacity of each of these new arcs determines as follows:

- about the arc $L_{rg} \to q_{rg} = q_{rs}^{\max} - q_{rs}^{\min}$;
- about the arc $L_{gs} \to q_{gs} = q_{rs}^{\min}$.
- It introduces the fictive arc L_{fict} with infinity capacity, i.e. $q_{fict} \to \infty$.

Then the Problem 2 has a solution only when the condition $f_{\Sigma}^{\max} = \sum_{q_{ij}^{\min} \neq 0} q_{ij}^{\min}$ is satisfied.

The transformed graph's model is the same as in Problem 1. Hence, the maximum flow $f_{\Sigma}^{\max}$ is determined in the same way.

Problem 3

Before starting the algorithm, the directed graph's model is determined. In this model the values of the arcs capacities $\{q_{ij}\}$ and vertexes capacities P_s, $s = \overline{1,n}$ are known.

The specificity of this problem requires that the additional condition (8) to be included:

$$f[L^{\oplus}(V_i)] \le P_s; \ \forall V_s \in V, \tag{11}$$

where: $f[L^{\oplus}(V_i)]$ is the flow entering to the vertex $V_s \in V$;

$L^{\oplus}(V_i)$ is the set of arcs, which are directly connected with the vertex $V_s \in V$ only by one arc.

To get the solution of this problem it is necessary to transform the basic directed graph's model to the new one with $(n+m)$ vertexes where *n* and *m* are the number of vertexes and arcs in the initial directed graph, respectively.

- Each vertex $V_s \in V$ is replaced by two vertexes V_{s_1} and V_{s_s}, which are connected by the arc $L_{s_1 \to s_2} = < V_{s_1}, V_{s_2} >$;
- The capacity of the arc $L_{s_1 \to s_2} = < V_{s_1}, V_{s_2} >$ is obtained as follows:

$$q_{s_1 \to s_2} = P_s; \ \forall V_s \in V. \tag{12}$$

The transformed graph's model is the same as in Problem 1. Hence, the maximum flow $f_{\Sigma}^{\max}$ is determined in the same way.

Problem 4

The initial conditions for this problem are the same as these in Problem 3 with the following additions. Fictive source $\mathbf{V}_{src} = \{V_{l,src}\}$, $l = \overline{1, N_s}$ and receiver $\mathbf{V}_{rec} = \{V_{l,rec}\}$, $l = \overline{1, N_r}$ are introduced, where each of them is associated with more than one real one (Figure 1). It is assumed that the graph's vertexes have infinite capacities, i.e. $q_{ij} \to \infty$.

The graph's model from the Problem 4 has $(n + N_s + N_r)$ vertexes and $(m+2)$ arcs compared to the models in Problems 1, 2 and 3, where *n* and *m* are the number of vertexes and arcs in the initial directed graph from Problem 1, respectively.

The resulting graph's model is the same as in Problem 1. Hence, the maximum flow studied $f_{\Sigma}^{\max}$ is determined in the same way.

Problem 5

Before starting the algorithm the directed graph's model is determined. In this model the values of the arcs capacities $\{q_{ij}\}$ and the values of the coefficients $\{\Psi_{ij}\}$, correcting the flow, passed through all of the graph's arcs are known. The latter are always positive and they reflect to the increasing ($\delta_{ij} > 0$), resp. decreasing ($\delta_{ij} > 0$) of the flow in the arc respective L_{ij}.

The next three steps describe the application of the considered heuristic algorithm [19] for solving Problems 1, 2 and 3, resp. 4 and 5, which are reduced to Problem 1.

1. Flows' ccorrection passing through the graph's arcs

The flows' correction is performed in the following way:

$$\Psi_{ij} . f_{in}(V_i, \ V_j) = f_{out}(V_i, \ V_j). \tag{13}$$

Note: If $\forall \ \Psi_{ij} = 1$ then the condition (1a) is satisfied and the directed graph's model is associated to the Problems 1, 2 or 3.

2. Determining the graph's paths from the source vertex V_{src} to the recipient vertex V_{rec}, where the resulting flow increases

If the conditions

$$\begin{aligned} &f_{in}(V_\alpha, \ V_\omega) < q_{\alpha\omega} \\ &\forall \text{"stright" arc } L_{\alpha\omega} = \langle V_\alpha, \ V_\omega \rangle \end{aligned}; \tag{14a}$$

and

$$\begin{aligned} &f_{in}(V_\alpha, \ V_\omega) > q_{\alpha\omega}, \\ &\forall \text{"opposite" arc } L_{\omega\alpha} = \langle V_\omega, \ V_\alpha \rangle \end{aligned}. \tag{14b}$$

are satisfied, then there is a path from the vertex $V_\alpha \in V$ to the vertex $V_\omega \in V$ in the analyzed directed graph, where the resulting flow increases.

The example of this path including two arcs with straight direction and two ones with opposite direction and their capacities q_{ij} and correcting coefficients Ψ_{ij}, is shown in Figure 3.

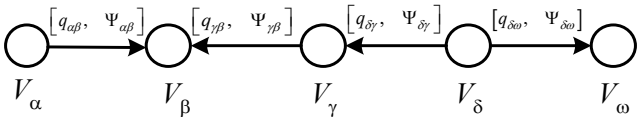


Figure 3: Path with initial vertex V_α and final vertex V_ω, including two straight and two opposite arcs

3. Obtaining the active cycles in directed graph's model

In directed graph's model $G = < V, L >$ there is an active cycle $C(G)$, connected by the recipient vertex V_{rec}, if the following conditions are satisfied:

- The value of the summary correcting cycle's coefficient is bigger than 1, i.e. $\Psi\left[C(G)\right] > 1$.
- The cycle's capacity is not zero, i.e. $q\left[C(G)\right] \neq 0$.
- The cycle includes a vertex V_λ, which is directly connected to the recipient vertex V_{rec}.

The example of this type cycle is shown in Figure 4.

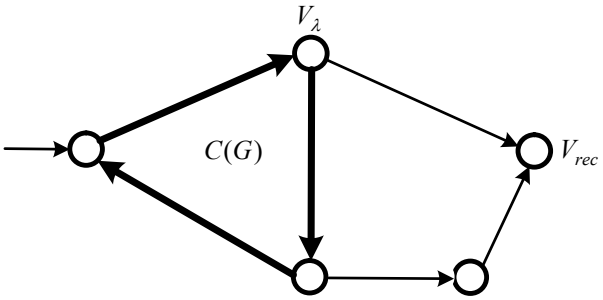


Figure 4: A cycle in directed graph's model

Note: It is allowed in directed graph's model to have cycles, connected to the source vertex V_{src}.

5. Application of the maximum flow suggested approach to an illustrative example about Problem 1

The algorithm for finding the maximum flow [19] will be illustrated for solving the Problem 1. Let's the following example of the directed graph's model is analyzed (Figure 5a). On the directed graph $G = < V, L >$ the following quantities are indicated:

1 The capacities q_{ij} of the respective arc $L_{ij} = \langle V_i, \ V_j \rangle \in L$ are shown below it.

2 The values of the flows $f_{ij} = f(V_i, V_j)$ passing through the arc L_{ij} are underlined and are written above to the arc. It is assumed that the initial values of these flows for all arcs are zeroes.

3 The vertexes marks $\left[M_1(V_i), \ M_2(V_i)\right]$ are noted in the following way:

- $M_1(V_i) = \pm V_j$, with respect to an increasing (+) or decreasing (–) flow passing through the arc L_{ij};
- $M_2(V_i) = f_{add}(V_i)$ - the maximum value of the additional flow $f_{add}(V_i)$ from the source vertex V_{src} to the vertex V_i passing through the sequence of the arcs, where the flow increases. This flow is determined as follows:

$$\begin{aligned} & f_{add}(V_i) = \min\left[f_{add}(V_j), \ q_{ij} - f\left(V_j, \ V_i\right) \right], \\ & \text{if } \ f\left(V_j, \ V_i\right) < q_{ij} \ \wedge \ V_i \in M^{+}(V_j) \\ & f_{add}(V_i) = \min\left[f_{add}(V_j), \ f\left(V_i, \ V_j\right) \right], \\ & \text{if } \ f\left(V_j, \ V_i\right) > 0 \ \wedge \ V_i \in M^{-}(V_j) \end{aligned} \qquad (15)$$

It is assumed that the initial marks of all vertexes are zeroes. And the source vertex V_{src} gets the mark $M(V_{src}) = \left[V_{src}, \ \infty\right]$, i.e. the generated flow from the source vertex is infinite.

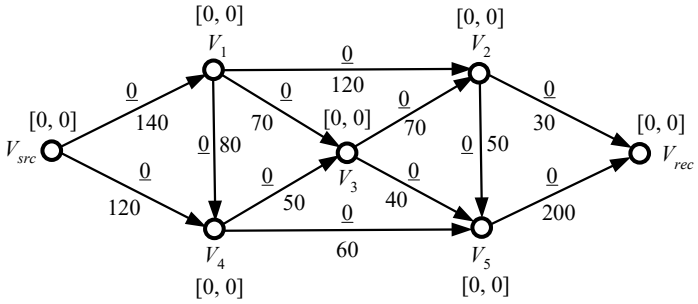


Figure 5a: Illustrative example for solving a maximum flow problem

When applying the heuristic algorithm [19], each iteration consists of 2 steps: first the vertices in the considered directed graph are marked and checked, and second the flows passed through the arcs in the graph's model are corrected. When marking the vertices, the algorithm moves from the beginning to the end of the graph, and when correcting the flows – vice versa, taking into account the capacity of each arc.

The required maximum flow is determined when all incoming arcs in the end graph's vertex are saturated. For the present example, this happens to the third iteration. The iterations are following.

Iteration 1

Vertexes marking (Figure 5b)

The source vertex V_{src} is marked with:

$$M\left(V_{src}\right) = \left[+V_{src}, \ \infty\right].$$

The vertexes V_1 and V_4, that are directly connected with V_{src} get the following marks:

$$\left| \begin{array}{l} M\left(V_1\right) = \left[+V_{src}, \ \ 140\right], \ M_2 = \min\left[\infty, \ 140 - 0\right] = 140 \\ M\left(V_4\right) = \left[+V_{src}, \ \ 120\right], \ M_2 = \min\left[\infty, \ 120 - 0\right] = 120 \end{array} \right. .$$

Therefore, the vertex V_{src} is marked and checked.

Let's move to the next vertex - V_1, which is marked but not checked. The latter is connected to the vertexes V_2 and V_3. So their marks are determined as follows:

$$\left| \begin{array}{l} M\left(V_2\right) = \left[+V_1, \ \ 120\right], \ M_2 = \min\left[140, \ 120 - 0\right] = 120 \\ M\left(V_3\right) = \max\left\{ \begin{array}{l} \left[+V_1, \ \ 70\right], \ M_2 = \min\left[140, \ 70 - 0\right] = 70 \\ \left[+V_4, \ \ 50\right], \ M_2 = \min\left[120, \ 50 - 0\right] = 50 \end{array} \right\} = \left[+V_1, \ \ 70\right] \end{array} \right.$$

Therefore, the vertex V_1 is marked and checked.

The marks of the other vertexes are determined by analogy as follows:

$$\left| \begin{array}{l} M\left(V_5\right) = \max\left\{ \begin{array}{l} \left[+V_2, \ \ 50\right], \ M_2 = \min\left[120, \ 50 - 0\right] = 50 \\ \left[+V_3, \ \ 40\right], \ M_2 = \min\left[70, \ 40 - 0\right] = 40 \\ \left[+V_4, \ \ 60\right], \ M_2 = \min\left[120, \ 60 - 0\right] = 60 \end{array} \right\} = \left[+V_4, \ \ 60\right] \\ M\left(V_{rec}\right) = \left[+V_2, \ \ 120\right], \ M_2 = \min\left[120, \ 30 - 0\right] = 30 \end{array} \right.$$

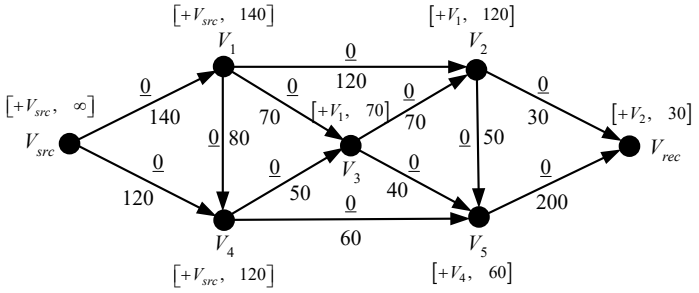


Figure 5b: Results from vertexes marking in an Iteration 1

So, all vertexes in the directed graph are marked and checked.

Flows correction (Figure 5c)

The vertex V_{rec} is chozen as a start vertex and we move from the end V_{rec} to the beginning V_{src} of the directed graph. Using Eq. (11) it determines the flow

$$f_{add}\left(V_{rec}\right) = M_2\left(V_{rec}\right) = 30\,.$$

Thus, the consequence of the arcs

$$\langle V_{rec},\ \ V_2\rangle \rightarrow \langle V_2,\ \ V_1\rangle \rightarrow \langle V_1,\ \ V_{src}\rangle$$

is formed basing on the first mark coefficients $M_1\left(V_i\right)$ of the respective vertexes. Then, using Eq. (12) the corrected values of the flows only in this arcs are calculated as follows:

$$f(V_2,V_{rec}) = f(V_1,V_2) = f(V_{src},V_1) = 0 + f_{add}\left(V_{rec}\right) = 30\,.$$

Then the condition

$$f(V_2,V_{rec}) = q_{2rec} = 30$$

is satisfied for the arc $\langle V_2,\ \ V_{rec}\rangle$. Then the arc is saturated and it is marked with bold.

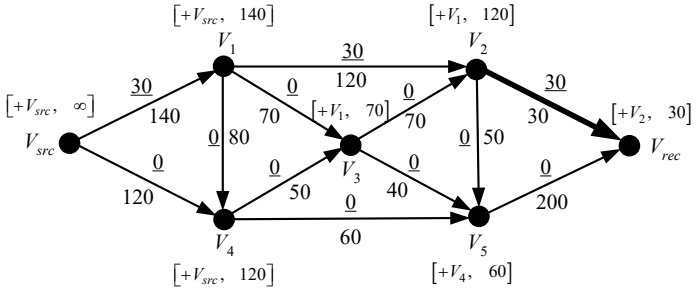


Figure 5c: Results from the flows corrections in the end of an Iteration 1

Since the arcs $\langle V_5,\ \ V_{rec}\rangle, \langle V_2,\ \ V_5\rangle, \langle V_3,\ \ V_5\rangle$ and $\langle V_4,\ \ V_5\rangle$ are not saturated the algorithm passes to next iteration – Iteration 2.

Moving from V_{rec} to the V_{src} as a result of this iteration a flow equal to 30 units is passed through each of the arcs $\langle V_{src},\ \ V_1\rangle, \langle V_1,\ \ V_2\rangle$ and $\langle V_2,\ \ V_{rec}\rangle$.

Iteration 2

Vertexes marking (Figure 5d)

Based on the information from Figure 5c, the new mark of the vertexes is set as follows:

$$\begin{aligned} M\left(V_{src}\right) &= \left[+V_{src},\ \ \infty\right] & M\left(V_2\right) &= \left[+V_1,\ \ 90\right] \\ M\left(V_1\right) &= \left[+V_{src},\ \ 110\right] & M\left(V_3\right) &= \left[+V_1,\ \ 70\right] \\ M\left(V_4\right) &= \left[+V_{src},\ \ 120\right] & M\left(V_5\right) &= \left[+V_4,\ \ 60\right] \\ & & M\left(V_{rec}\right) &= \left[+V_5,\ \ 60\right] \end{aligned}\,.$$

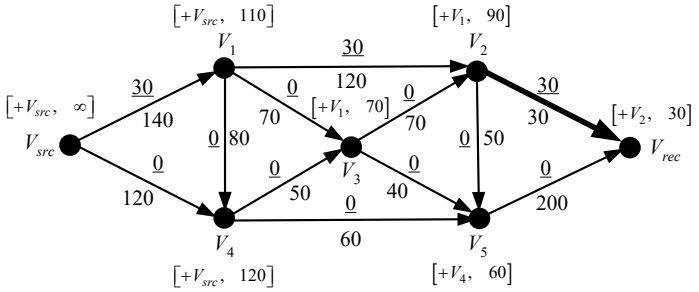


Figure 5d: Results from vertexes marking in an Iteration 2

Flows correction (Figure 5e)

Using Eq. (11) for the arc $\langle V_4,\ \ V_5\rangle$ the additional flow, is determined:

$$f_{add}\left(V_5\right) = M_2\left(V_5\right) = 60\,.$$

Thus, the consequence of the arcs

$$\langle V_{src},\ \ V_4\rangle \rightarrow \langle V_4,\ \ V_5\rangle$$

and

$$\langle V_{src},\ \ V_1\rangle \rightarrow \langle V_1,\ \ V_4\rangle \rightarrow \langle V_4,\ \ V_5\rangle$$

is formed basing on the first mark coefficients $M_1\left(V_i\right)$ of the respective vertexes. Using Eq. (12) the corrected values of the flows only in these arcs are corrected as follows:

$$\begin{aligned} f(V_{src},V_4) &= f(V_4,V_5) = 0 + f_{add}\left(V_5\right) = 60 \\ f(V_1,V_4) &= f(V_4,V_5) = 0 + f_{add}\left(V_5\right) = 60 \\ f(V_{src},V_1) &= 30 + f_{add}\left(V_5\right) = 30 + 60 = 90 \end{aligned}\,.$$

The arc $\langle V_4,\ \ V_5\rangle$ is saturated because the condition

$$f(V_4,V_5) = q_{45} = 60$$

is satisfied. Then the arc is marked bold.

Based on the additional flow $f_{add}\left(V_{rec}\right) = 60$, the corrected flows are:

$$f(V_5,V_{rec}) = f(V_4,V_5) = 0 + f_{add}\left(V_{rec}\right) = 60\,.$$

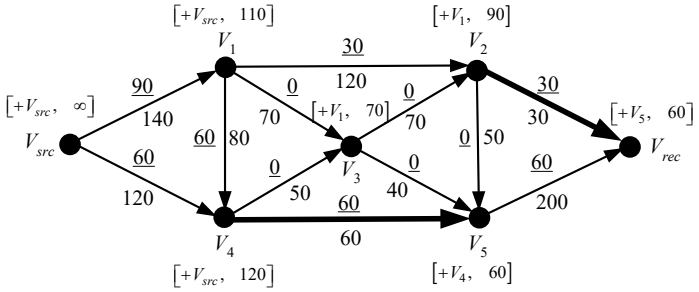


Figure 5e: Results from the flows corrections in the end of the Iteration 2

Since the arcs $\langle V_2, \ V_5 \rangle$ and $\langle V_3, \ V_5 \rangle$ are not saturated the algorithm passes to next iteration – Iteration 3.

Moving from V_{rec} to the V_{src} as a result of this iteration a flow equal to 60 units goes through each of the arcs $\langle V_{src}, \ V_4 \rangle, \langle V_4, \ V_5 \rangle$ and $\langle V_5, \ V_{rec} \rangle$. This additional flow of 60 units is then added to the current flows through the arcs $\langle V_1, \ V_4 \rangle$ and $\langle V_{src}, \ V_1 \rangle$, and finally, the flows of 60 and 90 units pass through the arcs, respectively.

Iteration 3

Vertexes marking (Figure 5f)

Based on the information from Figure 5e, the new mark of the vertex V_4 is set as follows (Figure 5f):

$$M(V_4) = [+V_{src}, \ \ 60].$$

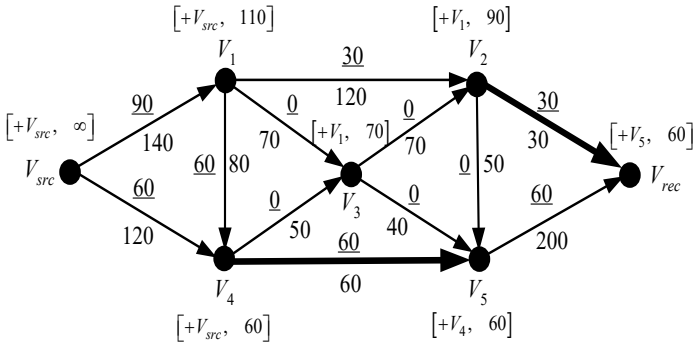


Figure 5f: Results from vertexes marking in an Iteration 3

3.2. Flows' correction (Figure 5g)

Using Eq. (11) the arcs $\langle V_2, \ V_5 \rangle$ and $\langle V_3, \ V_5 \rangle$ the following additional flows, are determined:

$$f_{add}^{V_2}(V_5) = \min\left[q_{25} = 50, \ M_2(V_5)\right] = \min[50, \ 60] = 50$$

and

$$f_{add}^{V_3}(V_5) = \min\left[q_{35} = 50, \ M_2(V_5)\right] = \min[40, \ 60] = 40 .$$

Thus, the other corrected flows are:

$$f(V_2, V_5) = f_{add}^{V_2}(V_5) = 50.$$
$$f(V_3, V_5) = f_{add}^{V_3}(V_5) = 40.$$

here

$$f(V_2, V_5) = 50, \ \text{and} \ \ q_{25} = 50 \Rightarrow \langle V_2, \ V_5 \rangle \ \text{is saturated.}$$
$$f(V_3, V_5) = 40, \ \text{and} \ \ q_{35} = 40 \Rightarrow \langle V_3, \ V_5 \rangle \ \text{is saturated.}$$

Therefore, both arcs $\langle V_2, \ V_5 \rangle$ and $\langle V_3, \ V_5 \rangle$ are marked bold.

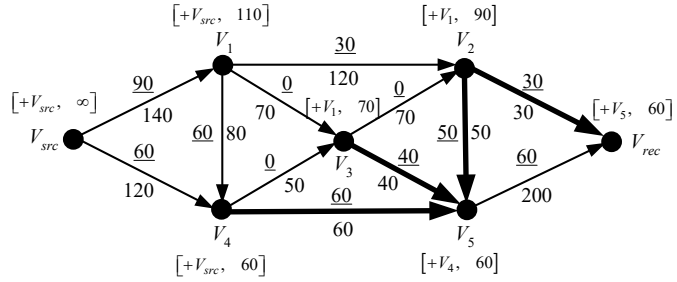


Figure 5g: Results from the flows corrections in the Iteration 3

The solution of the maximum flow problem is shown in Figure 5h, where the final vertexes are marked and flows corrections are made:

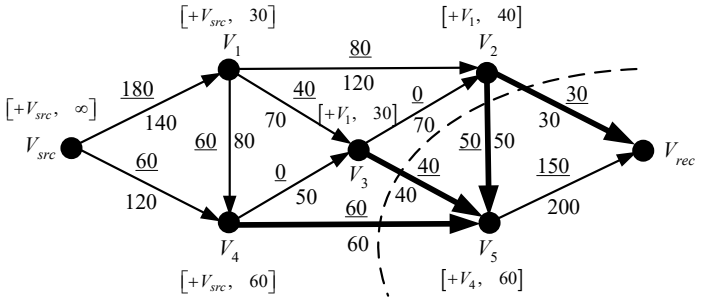


Figure 5h: Final maximum flow

Moving from V_{rec} to the V_{src}, as a result of this iteration, the following will happen. First, a flow equal to 40 units is passed through each of the arcs $\langle V_{src}, \ V_1 \rangle, \langle V_1, \ V_3 \rangle$ and $\langle V_3, \ V_5 \rangle$. This additional flow of 40 units is then added to the current flows through the arcs $\langle V_{src}, \ V_1 \rangle, \langle V_1, \ V_3 \rangle$ and $\langle V_3, \ V_5 \rangle$, and finally flows of 130, 40 and 40 units will pass through the arcs, respectively. Second, a flow equal to 50 units is passed through each of the arcs $\langle V_{src}, \ V_1 \rangle, \langle V_1, \ V_2 \rangle$ and $\langle V_2, \ V_5 \rangle$. This additional flow of 30 units is then added to the current flows through the arcs $\langle V_{src}, \ V_1 \rangle, \langle V_1, \ V_2 \rangle$ and $\langle V_2, \ V_5 \rangle$, and finally flows of 180, 50 and 50 units will pass through the arcs. Thus, the resulting flow through the arc will be 180 units.

After the last iteration, it is seen that all arcs, where a flow passing through, are saturated. Therefore, an end solution is reached and the final maximum flow is:

$$f_{\Sigma}^{\max} = f(V_2, \ V_{rec}) + f(V_2, \ V_5) + f(V_3, \ V_5) + f(V_4, \ V_5) =$$
$$= 30 + 50 + 40 + 60 = 180$$

In the analyzed illustrative example, the final maximum flow has been found after the third iteration, because all arcs of the cross section shown in Fig. 5h passing flow to the end vertex V_{rec} are saturated, i.e. the flow passing through them f_{ij} is exactly equal to the capacity of the corresponding arc q_{ij}.

A comparative study between the discussed heuristic algorithm and a standard min-cut algorithm for solving the maximum flow problem

The obtained solution with a heuristic approach of [19] for an illustrative problem shown in Figure 5a is compared with

analogous ones, determined by traditional methods based on min-cut theorem [20, 21, 22]. The latter is the following: "For any directed weight graph, the maximum flow passed through from the source to the recipient vertex is equal to the capacity of the minimum cut.". Then it is necessary to obtain the set of arcs with minimum weight that after removing from the graph there is no path from the source to the recipient vertexes. The following methods were used: Ford Fulkerson [23], Edmonds Karp [24] and Dinic algorithms [25, 26]. The calculations are made with free software Visualgo. The final the full solution of the considered problem about the maximum flow is obtained.

Note: In contrast to the method applied in this article, in the final solution Visualgo visualizes the flow that passes through each of the arcs and their remaining capacity, depending on the full arc's capacity, as shown in Figures 5a-5h.

The final solution obtained through the used methods are shown in Figures 6a, 6b and 6c.

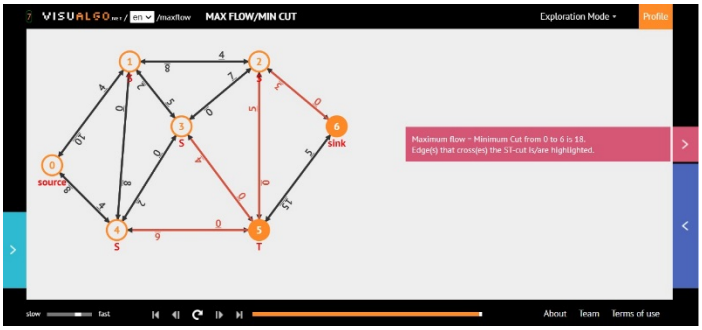


Figure 6a: Final maximum flow obtained by the Ford-Fulkerson's algorithm

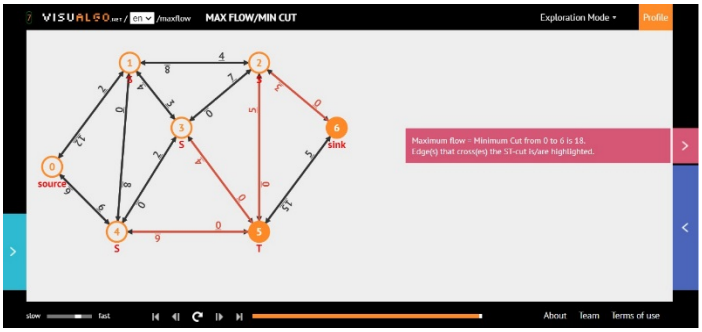


Figure 6b: Final maximum flow obtained by the Edmonds-Karp's algorithm

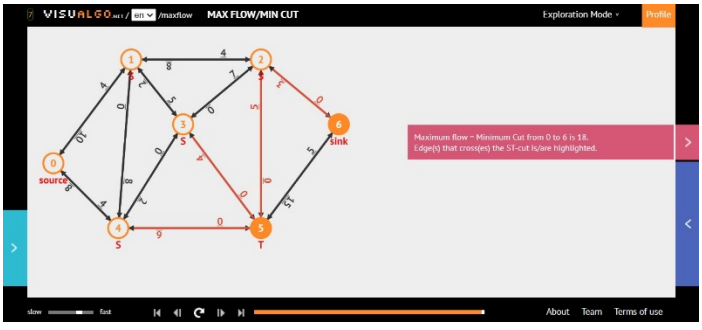


Figure 6c: Final maximum flow problem obtained by the Dinic's algorithm

The final maximum flow, determined by these three methods, based on min-cut theorem (Figures 6a, 6b and 6c), are the same as those of the heuristic approach applied in this article:

$$f_{\Sigma}^{\max} = f(V_2,\ V_{rec}) + f(V_2,\ V_5) + f(V_3,\ V_5) + f(V_4,\ V_5) = .$$
$$= 30 + 50 + 40 + 60 = 180$$

The advantage of the proposed approach is that the final solution is obtained on the 3rd iteration, and each iteration consists of two procedures.

In the method of Ford Fulkerson, the final solution is obtained on the 6th iteration. The Edmonds Karp algorithm consists of two built-in cycles. For the considered example the final solution is obtained on the 4th iteration. As the external cycle is executed 4 times and each time from its execution the internal is executed only once. The Dinic algorithm consists of two build-in cycles. For the considered example the final solution is determined after applying the external cycle 2 times. Initially, the internal cycle is rotated 2 times after that – it rotates 3 times. All this shows the better efficiency of the proposed algorithm

6. Conclusion

Determining the maximum flow and the distribution of the flows in the analyzed directed graph based on the interconnections in its topology is an important prerequisite for the efficient use of resources in the considered system. The solution of this optimization problem depends on the structure of the comsidered graph's model and the capacities of its arcs. The heuristic approach used in this article is universal not only in terms of the physical nature of the studied system, but also in terms of different types of problems considered - the input data of each of analyzed problem are different.

The solutions to the Problems 1, 2, 3 and 4 show that the flows passing through the arcs of the graph's model are constant. The solution to the Problem 5, however, shows that the flow is variable, so one can make a general conclusion when the new approach has been successfully applied in the case of violation of the balance condition.

In the present paper, the results of the solution for Problem 1 with heuristic algorithm are compared with these obtained from other analogous methods: Ford-Fulkerson's, Edmonds-Karp's and Dinic's algorithms. The time complexity of this algorithm is equal to that of the Ford-Fulkerson's and Edmonds-Karp's algorithms - $O\left[|V|.|E|^2\right]$, in contrast to that of the Dinic's algorithm, where it is $O\left[|V|^2.|E|\right]$. The illustrative example shows that it is faster to converge to the final solution comparing to the other considered algorithms.

The used universal algorithm is inapplicable to solving problems in which at least one of the three initial assumptions (defined in Problem statement) is not fulfilled, i.e. Balance condition, second condition for sequencing the flow passing through the arbitrary arc with its capacity and third condition for the invariability of the flow passing through the considered graph moving from the source to the recipient vertexes and vice versa.

One of the possible future development of the work is related to the program realization of the proposed approach for determining the maximum flow passing through a directed graph.

Other trend for the future development of the research on solving this problem is using an algorithm based on determining all possible paths in the considered directed graph, proposed in [27]. It uses two adjacent matrices corresponding to each vertex of the graph - these of the input and output arcs in the vertex, respectively. As a result of the algorithm, one will only be interested in the maximum flow that has passed the shortest path from the start to the end vertex in the directed graph.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

The publication of this manuscript was realized by financial support from the Scientific Research Sector of the Technical University of Sofia.

References

[1] H. A. Taha, Operations research: An introduction, Prentice Hall Price, 2017.

[2] J. A. Bondy, Murty, U. S. R., Graph Theory, Springer, 2008.

[3] N. Sayeekumar, Sh. Ahmed, S. Karthikeyan, S. Sahoo, I. Jacob, "Graph theory and its applications in power systems - a review", in 2015 IEEE International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 18-19 Dec. 2015, doi: 10.1109/ICCICCT.2015.7475267.

[4] J. Burch, Grudnitski, Information system. Theory and Practice, John Wiley & Sohn, 1997.

[5] S. L. Hakimi, "Application of graphs theory to problems in communication systems and networks", AD-A030 888. U. S. department of Commerce, National Technical Information Service, Northwestern Univ. Evenston III, Dept. of Electrical engineering, Air Force Office of Scientific Research, Boiling AFB, D C. 1976.

[6] G. Cherneva, H. Spiridonova, "Matching Filters Design for Information Transmit in Power Line Communication Systems", Mechanics, Transport, Communications Journal, **3/2**, XI-21- XI-27, 2018, (in Bulgarian).

[7] S. Guze, "Graph Theory Approach to Transportation Systems Design and Optimization", International Journal on Marine Navigation and Safety of Sea Transportation, **8**(3), Dec. 2014, doi: 10.12716/1001.08.04.12

[8] V. Radziukynas, I. Radziukyniene, Optimization Methods Application to Optimal Power Flow in Electric Power Systems, Optimization in the Energy Industry, Chapter 18, 409-436, Jan. 2009, doi: 10.1007/978-3-540-88965-6_18

[9] P. H. Nguyen, W. L. Kling, G. Georgiadis, M. Papatriantafilou, L. A. Tuan, L. Bertling, Application of the Graph Theory in Managing Power Flows in Future Electric Networks, New Frontiers in Graph Theory, Chapter 12, March 2012, doi: 10.5772/35527

[10] Schade, M. Begley. "Information flow diagram guidelines", 23 Sept. 2012, retrieved 20 March 2016.

[11] D. Christopher, T. Ashutosh, A. R. Jeffrey, "Modelling information flow for organisations: a review of approaches and future challenges", International Journal of Information Management, **33**(3), 597–610, June 2013, doi:10.1016/j.ijinfomgt.2013.01.009

[12] S. Riis, "Information flows, graphs and their guessing numbers", in 2006 IEEE 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 26 Feb. - 2 March 2006, doi: 10.1109/WIOPT.2006.1666516.

[13] Ch. Frey, A. Zufle, T. Emrich, M. Renz, "Efficient Information Flow Maximization in Probabilistic Graphs", IEEE Transactions on Knowledge & Data Engineering, **30**(5), 880-894, 2018, doi:10.1109/TKDE.2017.2780123

[14] K. A. Prasath, R. D. J. Johnson, "Fundamental Approach on Layout Optimization Using Graph Theory in Cellular Manufacturing Systems", Australian Journal of Basic and Applied Sciences, 9(33):180, October 2015.

[15] J. Dou, X.-Z. Dai, Z.-D. Meng, J. Li, "Optimization for single-part flow-line configurations of reconfigurable manufacturing system based on graph theory", Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing Systems, CIMS, **16**(1), 81-89, 2010.

[16] J. Dou, X. Dai, Zh. Meng "Graph theory-based approach to optimize single-product flow-line configurations of RMS", International Journal of Advanced Manufacturing Technologies, **41**, 916–931, 2009, doi: 10.1007/s00170-008-1541-2

[17] J. Dou, X. Dai, Zh. Meng, "Optimization for Flow-Line Configurations of RMS Based on Graph Theory", in 2007 International Conference on Mechatronics and Automation, 1261-1266, doi: 10.1109/ICMA.2007.4303730

[18] M. Mihail, N. Vishnoi, "On Generating Graphs with Prescribed Vertex Degrees for Complex Network Modeling", College of Computing Georgia Institute of Technology, 2002.

[19] S. Filipova-Petrakieva, "A New Universal Optimization Approach for Maximum Flows in Power Systems", in 2019 IEEE XI International Conference BulEF'2019, 11-14.09.2019, doi: 10.1109/BulEF48056.2019.9030731

[20] G. B. Dantzig, D. R. Fulkerson, "On the Max-Flow MinCut Theorem of Networks", in Linear Inequalities, Ann. Math. Studies, no. 38, Princeton, New Jersey, 1956.

[21] G. B. Dantzig, D.R. Fulkerson, "On the max-flow min-cut theorem of networks", RAND corporation: 13, 9 September 1964, retrieved 10 January 2018.

[22] G. Eusden, "Max-Flow, Min-Cut - History and Concepts Behind the Max-Flow, Min-Cut Theorem in Graph Theory", 2013.

[23] L. R. Ford, D. R. Fulkerson, "Maximal flow through a network", Canadian Journal of Mathematics, **8**, 399–404, 1956, doi: 10.4153/CJM-1956-045-5.

[24] J. Edmonds, R. M. Karp, "Theoretical improvements in algorithmic efficiency for network flow problems", Journal of the ACM, **19**(2), 248–264, 1972, doi:10.1145/321694.321699.

[25] E. A. Dinic, "Algorithm for solution of a problem of maximum flow in a network with power estimation", Soviet Mathematics – Doklady, Doklady, **11**, 1277–1280, 1970.

[26] Y. Dinitz, "Dinitz' Algorithm: The Original Version and Even's Version", In Oded Goldreich; Arnold L. Rosenberg; Alan L. Selman (eds.). Theoretical Computer Science: Essays in Memory of Shimon Even, Springer, 218–240, 2006.

[27] V. M. Mladenov, Kr. N. Filipova, S. K. Petrakieva, B. Dimov, H. Uhlmann, "Analysis of Signal Competition in Asynchronous Ultra High-Speed Digital Circuits", Przeglad Elektrotechniczny (Electrical Review), **83**(11), 197-200, 2007.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 4, 176-183 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Multi-Robot System Architecture Design in SysML and BPMN

Ahmed R. Sadik*, Christian Goerick

Honda Research Institute Europe, 63073, Germany

ARTICLE INFO

Article history:
Received: 30 November, 2020
Accepted: 08 July, 2021
Online: 20 July, 2021

Keywords:
Multi-robot System Model
Model-based System Engineering
Multi-agent Simulation
Systems Modeling Language
Process Model and Notation
Java Agent Development

ABSTRACT

Multi-Robot System (MRS) is a complex system that contains many different software and hardware components. This main problem addressed in this article is the MRS design complexity. The proposed solution provides a modular modeling and simulation technique that is based on formal system engineering method, therefore the MRS design complexity is decomposed and reduced. Modeling the MRS has been achieved via two formal Architecture Description Languages (ADLs), which are Systems Modeling Language (SysML) and Business Process Model and Notation (BPMN), to design the system blueprints. By using those abstract design ADLs, the implementation of the project becomes technology agnostic. This allows to transfer the design concept from on programming language to another. During the simulation phase, a multi-agent environment is used to simulate the MRS blueprints. The simulation has been implemented in Java Agent Development (JADE) middleware. Therefore, its results can be used to analysis and verify the proposed MRS model in form of performance evaluation matrix.

1. Introduction

This paper extends the work presented at the 2019 International Conference on Mechatronics, Robotics, and System Engineering (MoRSE) [1]. Related work can be also seen in [2].

Multi-Robot System (MRS) is a cyber-physical system that contains more than one robot, each of them owns a unique set of capabilities. The idea of an MRS is to solve a complex problem by collectively using the current capabilities of existing robots [3]. Therefore, the MRS must match the given problem with the existing robots' capabilities, to plan the solution steps. Many MRS applications can be seen in swarm robotics, cooperative automated transportation, unmanned aerial vehicles, and cooperative manufacturing [4]. The advantages of an MRS is increasing the performance by saving the time and the effort to solve the problem. Moreover, distributing the solution among different robots provides more computational processing power, this means faster and higher capacity to solve many problems simultaneously [5].

Implementing an MRS without a proper system architecture design is a crucial mistake that is often done by the system developers. Because the system requirements and functionalities are lost in a non-human readable machine code. Therefore, in this article we purpose a model driven development approach that uses the system model as the main software artifacts [6]. The proposed design approach in this article is based on the V-Model, which is a de facto solution for complex systems such as MRS.

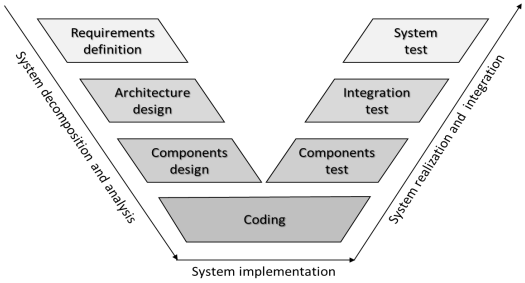


Figure 1: The V-Model simplified version – adapted from [6]

The V-Model shown in Figure 1 describes the required stages to build an MRS. In the first stage of the V-Model, the system is decomposed. In this stage, the system components and architecture are designed based on the system requirements. In the second stage, the implementation of the MRS is carried out. The implementation of an MRS often involves the coding the individual components. In the final stage, the MRS individual components are tested through unit tests, then integration tests are carried out over sub-systems and eventually the overall integrated system. This article focusses on the first stage of the V-Model to build an MRS. As the design stage is the most curtail stage of an

*Corresponding Author: Ahmed Rabee Sadik, ahmed.sadik@honda-ri.de

https://dx.doi.org/10.25046/aj060421

MRS system building, because all the following stages are depending on this design.

Section 2 of the article describes the problem and the use case of concern. Section 3 introduces the background that that is needed to model and simulate the use case. Section 4 discusses the system requirements that are used to build and evaluate the system performance Modeling the use case is explained in detail in section 5, while its simulation is shown in section 6. Therefore, the performance analysis is explained in section 7. Ultimately, the last section concludes the work and the future research.

2. Problem and use case

The main article objective is to design an MRS architecture model that can be simulated and evaluated due to a predefined evaluation criterion. An MRS architecture is an overall system description that abstracts its functionalities, logic, and constrains [7]. Accordingly, it provides an analysis tool to grasp and improve system characteristics, and a conceptual model that can be used as the system blueprints [8]. In this work, SysML block definition diagram is used to describe the proposed MRS architecture and components as shown in Figure 2 and Figure 3. SysML diagrams will be explained in the next section as many of them are used in constructing the proposed MRS design.

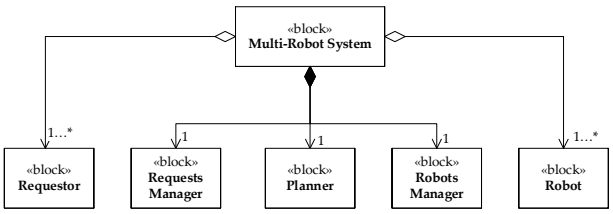


Figure 2: SysML block definition diagram for the proposed architecture

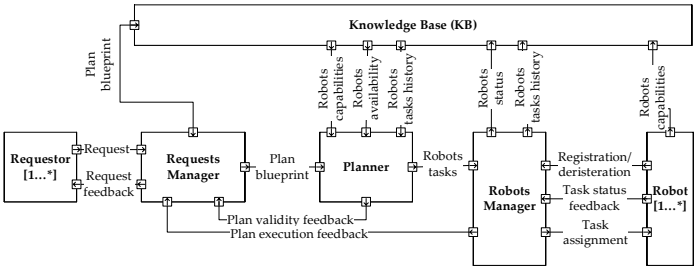


Figure 3: SysML internal block diagram for the proposed MRS architecture

Figure 2 is the proposed MRS block definition diagram. The block definition diagram defines the main components of the architecture, which are the Requests Manager (RqM), the planner (PLN), and the Robots Manager (RbM). Figure 3 shows the proposed MRS internal block diagram that describes the connections among the components as illustrated in Figure 2. When the RqM receives a request (Rq), it checks if there is a plan-blueprint (Pb) in the Knowledge Base (KB) to fulfill this Rq. A Pb is a sequence of tasks (T), i.e., $Pb_i = \{T_1,...,T_n\}$, where n is the number of tasks and could be different from one blueprint to another. A task is a function of the capabilities (C) of the robot (R), i.e., $T_i = f(C_x, C_y, \ldots)$, where each robot owns different capabilities set. If the RqM finds a match between Pb to and a Rq, it forwards the Pb to the PLN. The PLN checks the robots' availability, and their capabilities to achieve the tasks in the Pb. If more than a robot owns the capabilities to fulfill the task, the PLN compares the number of tasks that have been achieved by these robots in the past. Based on this comparison, the PLN selects a robot to assign for the task. If the PLN complete the matching of all the tasks with the robots, it sends a verified plan (Pv) to the RbM. The RbM sends the tasks to the robots and waits their response.

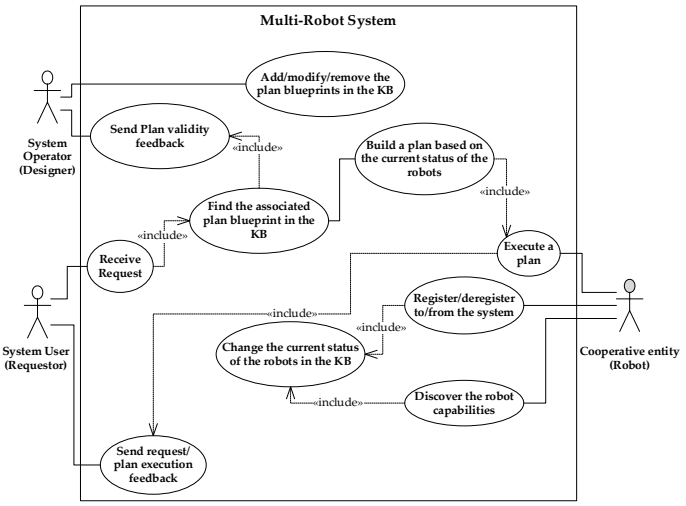


Figure 4: UML/SysML use case diagram for the proposed architecture

The use case diagram in Figure 4 shows three variation types that are considered during the simulation. First is the Pbs variation, by adding, editing, or omitting a Pb. Second is variation in the number of the available robots. The maximum number of robots that can exist is constrained to three. The robots are constrained to register or deregister through the RbM. Third is the variation in the robots' capabilities, by updating or editing the capabilities of a robot. the robot is constrained to deregister to be able to update its capabilities, then register again through the RbM, which automatically updates the robot new capabilities in the KB.

3. Solution preliminaries

3.1. Systems Modeling Language

SysML is a general-purpose modeling language that is derived from Unified Modeling Language (UML) [9]. SysML and UML belong are both developed by Object Management Group (OMG). UML is is a visual modeling language that is particularly used to construct, design, and document the software systems in fields such as web-development, telecommunication, banking, and enterprise services [10]. While SysML is extending and modifying UML diagrams to fit complex industrial systems that involve variety of hardware, software, information, and processes (e.g., Aviation, Space, Automotive) [11].

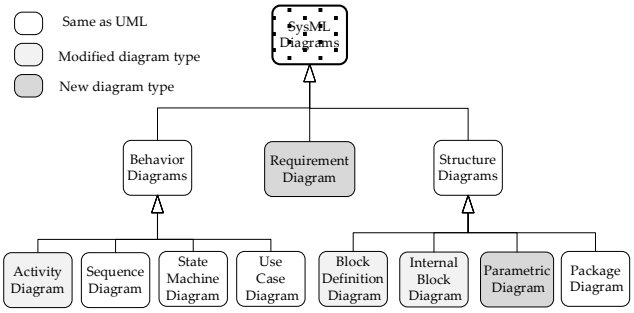


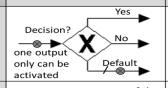
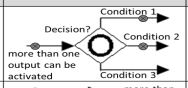
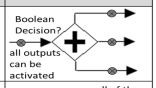
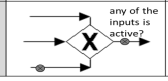
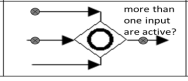
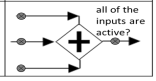
Figure 5: SysML Taxonomy and comparison to UML

Figure 5 shows the relation between SysML and UML graphs [12]. Requirement and parametric diagram are two new diagrams that distinguish SysML [13]. Section 4 of this article used the requirement diagram to define the system performance criteria requirements and their relations. Block definition diagram and internal block diagram are used to describe the main components of the system architecture and the connection among the components as illustrated in section 2, while the use case diagrams describe the interaction of the system as a black box with the external world or actors. The activity diagram is used in section 5 to represent the MRS components logic. The more detailed logic is represented in the activity model, the easier to automatically generate a low-level code from this activity model. For this reason, BPMN is used to build the MRS logic, as it extends the notations, semantics, and syntax of SysML and UML activity diagram. The state machine diagram is used in section 4 to model the internal states of the robot. While the sequence diagram is used in section 6 to represent the interaction and communication among the components during a simulation scenario.

3.2. Business Process Model and Notation

Since UML activity diagram provides an abstract high-level process description, BPMN extends the UML activity diagram to fulfill the following two drawbacks. First, UML activity diagram lakes the syntax and the logical execution among the actions. Second, the poverty in UML notations and semantics in comparison with BPMN [14].

Table 1: BPMN control gateways

Rule	Gateways		
	Exclusive OR (XOR)	Inclusive OR	Parallel AND
Split			
Merge			

Flow control gateways is the best example to demonstrate how BPMN is improving UML activity diagram. Flow control gateways are all equivalent to only one notation in UML, which is the decision notation. Table 1 shows the notations, semantics, and syntax of the basic gateways of BPMN. Three different notations are demonstrated in Table 1, which are exclusive-OR, inclusive-OR, and parallel-AND. The three mentioned gates operate either as spilt or merge context. In spilt context, exclusive-OR splits one input to only one output based on the conditions on the output branches. Inclusive-OR splits one input to more than one output simultaneously based on the conditions on the output branches. Parallel-AND splits one input to all the output simultaneously when the input branch is triggered. In merge context, exclusive-OR merges any of the input branches to only one output, when any of the input branches is triggered. Inclusive-OR merges more than one input branches to only one output, when these inputs are simultaneously triggered. Parallel-AND merges all the input branches to only one output, when all the inputs are simultaneously triggered [15].

3.3. Java Agent Development

JADE is a Multi-Agent System (MAS) middleware [16] that has been used in this research to deploy the proposed solution as shown in Figure 6-a. Each entity in the proposed SysML internal block diagram is implemented as a JADE agent. JADE Agent Management System (AMS) address each agent with a unique Identifier (AID) to facilitate the communication among the agents. While JADE directory Facilitator (DF) announces the services that every agent afford. JADE applies the Foundation for Intelligent Physical Agent (FIPA) specifications, to enable agent communication through FIPA-Agent Communication Language (FIPA-ACL) [17].

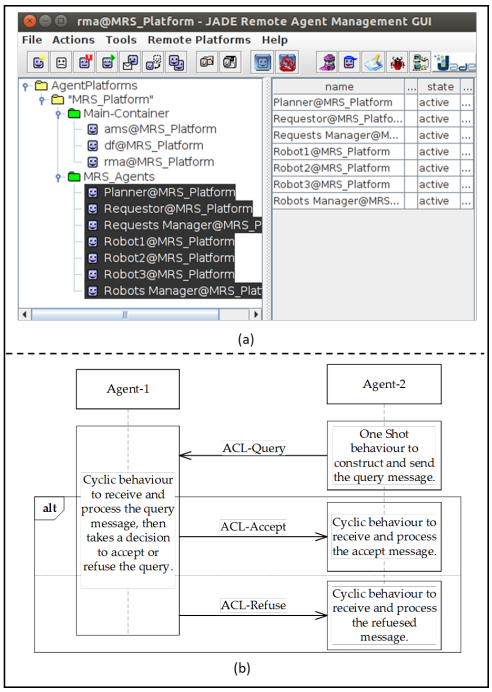


Figure 6 (a) JADE framework – (b) JADE sequence diagram example

Each JADE agent has a complex individual behaviour that can be seen as a composite of two simple behaviours. First is one-shot behaviour that is executed only once when it is triggered. Second is a cyclic behaviour that continuously executed when it is triggered. An example of JADE agent communication and decision making based on their behaviours can be seen in Figure 6-b. JADE is a suitable tool to build an agent simulation based on the MRS SysML/BPMN model. As the MRS logic and architecture can be easily translated to JADE implementation concepts [18].

4. Performance requirements

To evaluate the MRS design, it is necessary to measure the system performance during the simulation. Qualitative criteria such as reusability, scalability, extensibility, and interoperability have been proposed in [19]. However, these criteria are often relatively vague without quantitative performance measurements. Therefore, this research defines the quantitative indicators that are shown in Figure 7. The research assumes that the MRS is a black box that receives different Rq, that can either success or fail during the execution. The following measurements can be used to express the system performance:

- Throughput: the number of requests that are processed.
- Latency: the time needed from the request arrival till the request execution.

- Success rate: the number of request that success to be executed per the overall received requests number.
- Failure rate: the number of request that fail to be executed per the overall received requests number.
- Efficiency: the ration between the success rate and the failure rate.

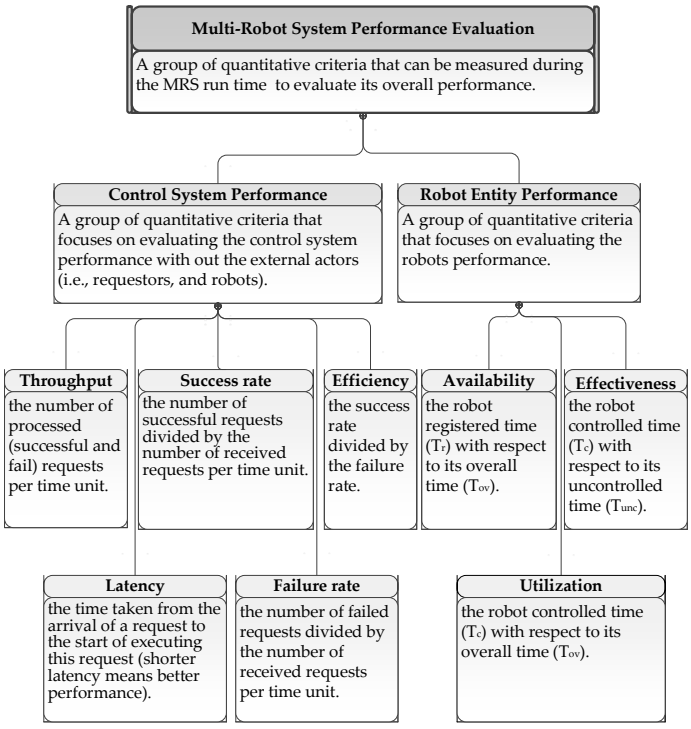


Figure 7: Requirement diagram of the MRS performance evaluation criteria

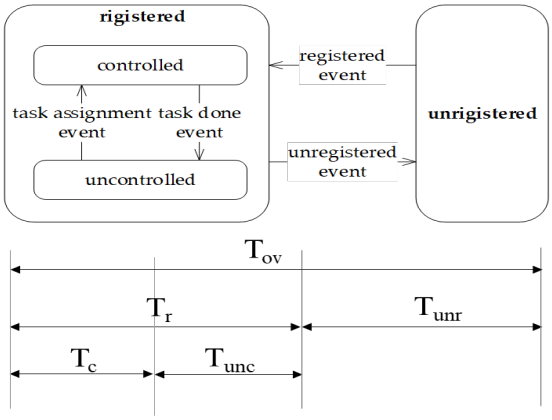


Figure 8: State machine diagram of the robot entity

The robot performance is also considered in this research as another measurement of the MRS performance [20]. The robot performance is fundamentally derived from its state machine diagram that is shown in Figure 8. The the robot state machine is built upon measuring the following times:

- Controlled time (T_c): the time that the robot needs to perform an assigned task.
- Uncontrolled time (T_{unc}): the robot waiting time to be assigned to a task after registration.
- Registered time (T_r): the sum of the controlled and the uncontrolled time of the robot.
- Unregistered time (T_{unr}): the accumulation of the robot unregistered time.
- Overall time (T_{ov}): the sum of the registered and the unregistered time of the robot.

Accordingly, the robot performance criteria are calculated as follows:

- Availability: the ration between the robot registered time (T_r) and the overall time (T_{ov}).
- Utilization: the ratio between the robot controlled time (T_c) and the overall time (T_{ov}).
- Effectiveness: the ration between the robot controlled time (T_c) and the uncontrolled time (T_{unc}).

5. System model

5.1. Requests manager

The RqM receives requests from various requestors, then it looks for an associated Pb within the KB. If the RqM finds the associated Pb, it forwards it to the PLN. The RqM decision making model is shown in Figure 9 via the BPMN activity diagram.

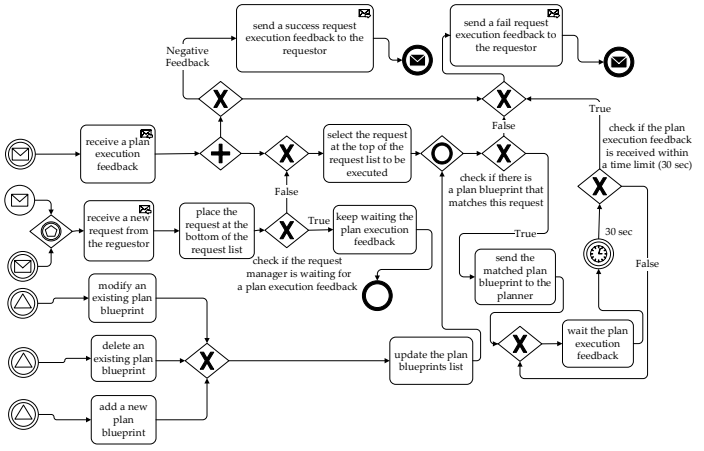


Figure 9: Requests-manager BPMN activity diagram

The RqM uses First Come First Serve (FCFS) technique to schedule the received requests. The RqM checks in the associated Pb for every received request. If there is no associated Pb with the request, the RqM directly sends a negative feedback to the requestor. If the RqM finds an associated Pb to the request, it forwards this Pb to the PLN, and waits for the feedback. If this feedback exceeds predefined limits, the RqM considers this request as a failure one. If not, it waits the execution feedback to forward it to the requestor.

5.2. Planner

The PLN receives the Pb and makes sure that it is visible to build a Pv instance according to the current system status. The PLN decision making model is shown in Figure 10 via the BPMN activity diagram.

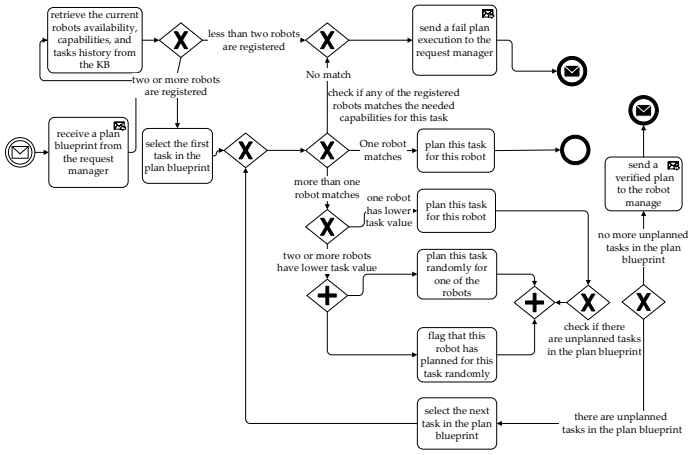


Figure 10: Planner BPMN activity diagram

To construct a Pv instance from a Pb, The PLN checks the available registered robots, the robots' capabilities, and the robots' tasks history. In case that there is only one available robot, the PLN directly considers a plan failure, as it is known in advance that a plan requires at least two robots to get executed. If at least two robots are available, the PLN compares the tasks in the Pb to the available robots' capabilities. If the robots' capabilities do not match the required tasks in the Pb, the PLN considers a plan failure. If there are two robots or more that can perform the same task, the PLN checks their tasks history, and assign the task to the robot that performed less tasks. This is to balance the task assignment among the available robots within the MRS. If all the tasks in the Pb could be assigned to robots, the PLN creates a Pv instance and sends it is the RbM to be executed.

5.3. Robots Manager

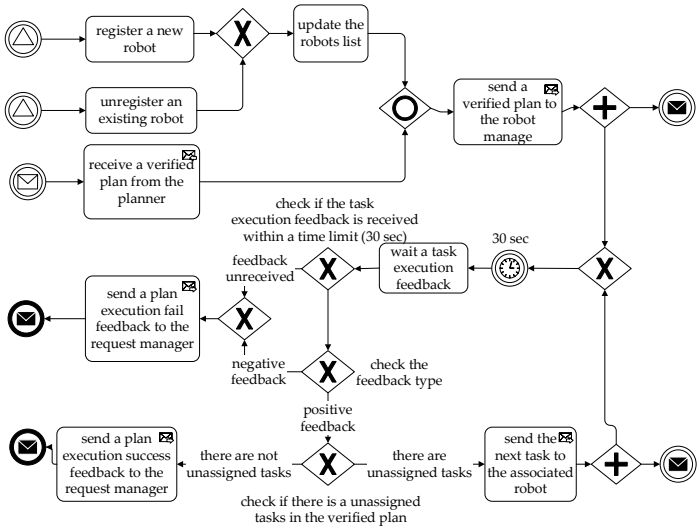


Figure 11: Robots-manager BPMN activity diagram

The RbM receives the Pv, then it assigns the tasks in this Pv to the available robot. Additionally, the RbM is also responsible for registering/unregister the robots from the MRS. this way it monitors the robots' availability. The RbM decision making model is shown in Figure 11 via the BPMN activity diagram.

When the RbM assigns a task to a robot, it waits the robot feedback within a time limit. If the robot feedback did not arrive within the predefined limits, the RbM sends a negative feedback to the RqM. This feedback means that the whole plan is failed to be executed. If the RbM received a positive feedback from the robot within the predefined time limits, it assigns the next task due to the Pv. If all the tasks in the Pv are executed, the RqM sends a positive feedback to the RqM, otherwise it sends a negative feedback.

6. Simulation

The activity diagrams that have been illustrated in the previous section are used as the MRS blueprints. JADE has been used in this research to deploy these blueprints, and hence enables the MRS simulation during the design phase. The Graphical User Interface (GUI) shown in Figure 12 has been created to achieve interact with every entity in the proposed architecture. The RqM GUI in Figure 12-a can be used to add/edit/remove the Pb. The PLN GUI in Figure 12-b is used to monitor the Pv execution, the robots' availability, the robots' status, the robots' capabilities, and the robots' tasks history. The RbM GUI in Figure 12-b is used to show the assigned tasks status.

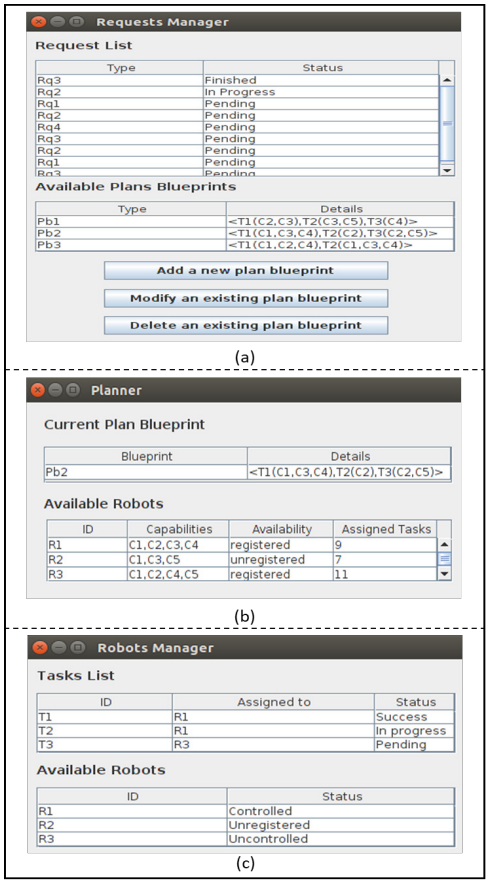


Figure 12: (a) Requests Manager GUI – (b) Planner GUI – (c) Robots Manager GUI

To illustrate the simulation scenario, an interaction example among the MRS entities is show in Figure 13. In this example, The RqM receives Rq_2. Therefore, the RqM sends the Pb in a form of the ACL-message shown in Figure 14-a to the PLN. Accordingly, the PLN constructs a Pv by matching the available robots' capabilities and tasks history with the received Pb. In this case, R_1 and R_3 were registered into the MRS as shown in Figure 14-b. As

T_1 needs (C_1, C_3, C_4) to be executed, T_1 was assigned to R_1, because (C_1, C_3, C_4) are unique capabilities of R_1. Similarly, T_3 was assigned to R_3, as T_3 needs (C_2, C_5) which is unique capability of R_3. However, in case of T_2, both R_1 and R_3 own the capability C_2 which is needed to execute this task. Therefore, the PLN checks both robots' task history to be able to assign T_2. The PLN finds out that R_1 task history is 9 while R_3 task history is 11. Accordingly, the PLN assigns T_2 to R_1, to balance the robots' tasks distribution.

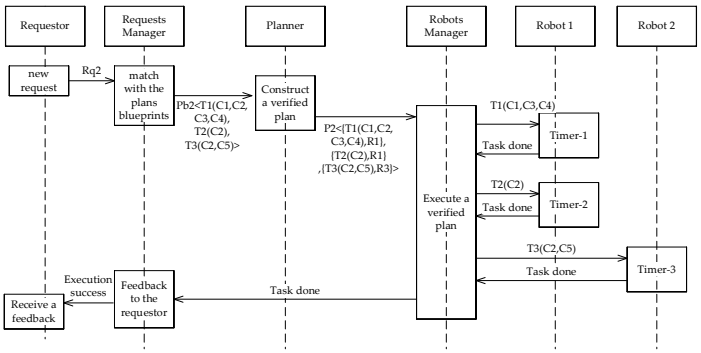


Figure 13: Simulation scenario plan execution sequence diagram – an example

Ultimately, the PLN sends the Pv in form of the ACL-message shown in Figure 14-b to the RbM. The RbM assigns the tasks to the associated robots according to the Pv. The task assignment is sent as an ACL-message as shown in Figure 14-c. The RbM waits the robots' feedback within a timeframe window. If all the RbM received success feedbacks for all the assigned tasks, it sends a plan success feedback to the RqM.

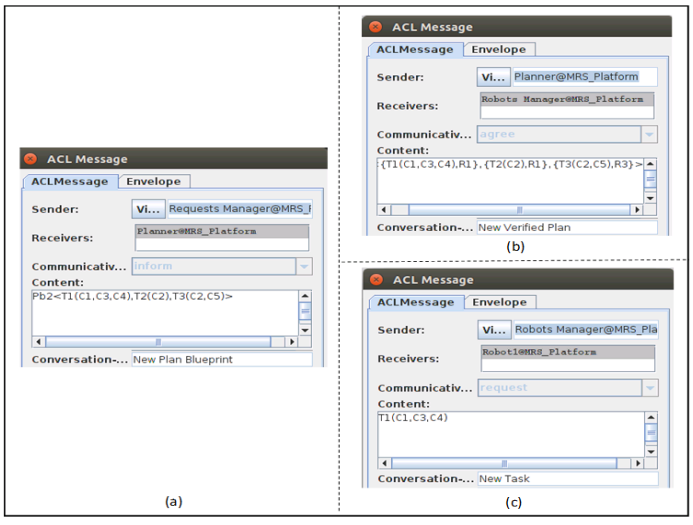


Figure 14: (a) Plan blueprint message – (b) Verified Plan message – (c) Assigned task message

7. Simulation results analysis

As it has been demonstrated in the previous section, the robots' availability, the robot's capabilities, and the the plan blueprints are the variables that can be used to build different simulation scenarios. Accordingly, to measure the system performance, the robots' availability was randomly altered during the run time. Thus, analyzing the simulation results has been done by running JADE MAS for 30 minutes as shown in Figure 15, then measuring the system performance indicators that are concluded in section 4. Each one minute, a new request is generated, one robot randomly unregister from JADE MAS, and one random robot register to JADE MAS. the robot's capabilities and the the plan blueprints do not change during the simulation scenario.

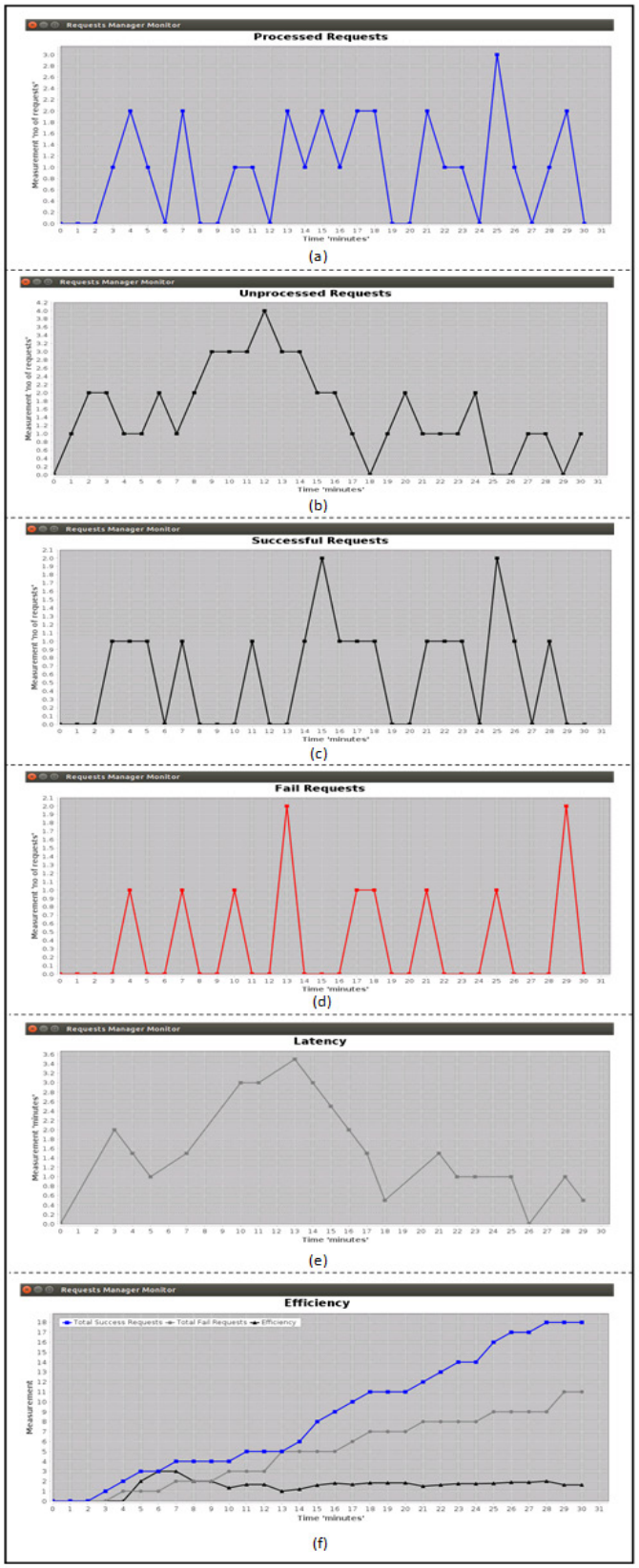


Figure 15: Simulation graphs (a) Processed requests – (b) Unprocessed request – (c) Successful requests – (d) Failed requests – (e) Latency – (d) Efficiency

One of the RqM responsibilities is to monitor the requests status. The number of processed requests by the RqM is shown in the graph in Figure 15-a. Accordingly, the MRS throughput can be directly calculated from this chart. On the one hand, MRS throughput expresses how fast the system, therefore it is a relative value. Thus, to understand the MRS throughput, Figure 15-c and Figure 15-d should be considered as well. For instance, the number of requests at minute 4 is two requests as can be seen in Figure 15-a. But, if we look closely into Figure 15-c and Figure 15-d, we will find out that one request is success and another fail. This means that, it is not important if the system is so fast, but most of the requests are failed to be executed. On the other hand, MRS latency expresses how much delay in the system as it can be seen in Figure 15-e. If the system delay value is equal to zero as can be seen in the 26th minutes of Figure 15-e, this means that the number of unprocessed requests is equal to zero as well, as can be seen in the 26th minutes of Figure 15-b.

The MRS efficiency graph shown in Figure 15-f is derived from dividing the data in Figure 15-c (successful requests) by the data Figure 15-d (fail requests). The MRS efficiency value is absolute. When the MRS efficiency is higher than one, this means that the number of success requests is higher than the number of fail request. Figure 15-f shows that the simulated MRS efficiency is higher than or equal to one during the simulation runtime.

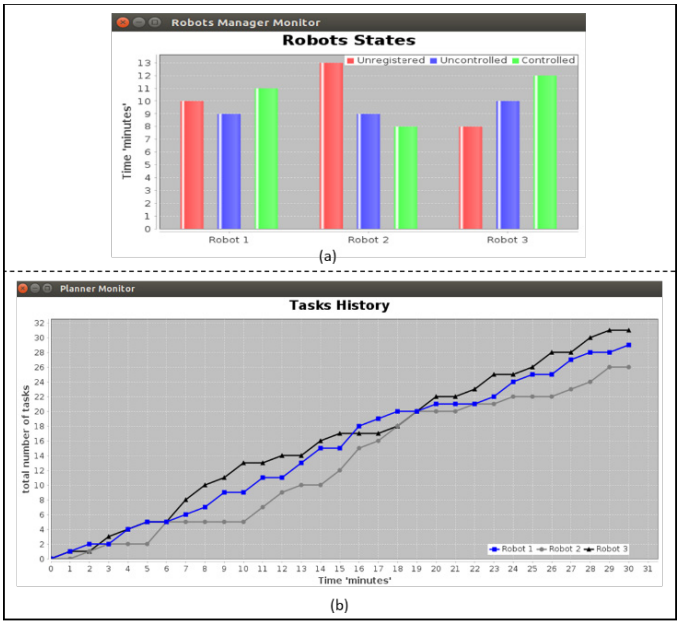


Figure 16: (a) Robots States – (b) Robots tasks history

Table 2: Robots availability, utilization, and effectiveness

	Robot 1 (R1)	Robot 2 (R2)	Robot 3 (R3)
Availability $\frac{T_r}{T_{ov}}$	$\frac{20\ min}{30\ min} = 0.67$	$\frac{17\ min}{30\ min} = 0.57$	$\frac{22\ min}{30\ min} = 0.73$
Utilization $\frac{T_c}{T_{ov}}$	$\frac{11\ min}{30\ min} = 0.37$	$\frac{8\ min}{30\ min} = 0.27$	$\frac{12\ min}{30\ min} = 0.4$
Effectiveness $\frac{T_c}{T_{unc}}$	$\frac{11\ min}{10\ min} = 1.1$	$\frac{8\ min}{9\ min} = 0.89$	$\frac{12\ min}{10\ min} = 1.2$

One of the PLN responsibilities is to monitor the balance the tasks among the available robots. Figure 16-a shows that the robots' available is changing over the simulation runtime. Simultaneously, Figure 16-b shows that the PLN compensates this variation by balancing the MRS. For instance, the task distribution among the available robot is converging to be 6 tasks per robot at the 6th minute of the simulation. Then, the robots' tasks distribution is diverging till it balanced again to be 20 tasks per robot at the 19th minute of the simulation. Table 2 can be also concluded from Figure 16-a. In this table, R_3 is the most utilized and available robot during the simulation runtime, and hence R_3 is the most effective in comparison to R_1 and R_2. Accordingly, the PLN compensates this variation by maximizing R_1 and R_2 task assignment, to balance them with R_3.

8. Summary and Discussion

This article has highlighted new dimensions of the MRS design problem, which are the formalization, simulation, and evaluation of the solution architecture. The proposed modeling approach is based on a formal generic ADLs, that can be used to transfer the solution concept over different system case studies, regardless the implementation technology. Furthermore, the illustrated simulation method can be used to verify different architecture design patterns, based on the concluded system performance measurements.

The fundamental SysML diagrams have been implemented to design the proposed MRS system model. Moreover, BPMN language has been used to implement the activity diagram as it extends UML/SysML notations, semantics, and syntax. The collection of these standard models is used as the MRS blueprints. Those blueprints can be easily coded in any programming environment that supports distributed system implementation. For instance, JADE has been used in this research to implement these blueprints, however Robot Operation System (ROS) or Web Service (WS) are very suitable candidates to deploy the system.

A group of MRS performance requirements have been defined during this article, to quantify the system performance during the simulation runtime. Those criteria can are technology agonistic as well, which means that they can be used to compare between the system performance when it is implemented with different technologies. Furthermore, the system simulation is not only used during the design phased, but it can be reused in a form of a real time digital twin during the implementation phase. For instance, to check in advance different planning and scheduling algorithms before executing them on the real system.

Using a formal description language such as SysML or BPMN enables separating the model from the code, which is a common domain specific programming method. Therefore, in the future work, we will write a code generator that can be used to automatically generate the implementation code. Therefore, the model that has been developed in this article will turn to be executable and will be used as the main software artifact of the project. This can dramatically reduce the coding time and effort and improve the system readability and maintainability. Additionally, in the future work, the same performance measurements that have been used in this article can be used in the implementation phase, as a part of the system visualization.

References

[1] A.R. Sadik, C. Goerick, M. Muehlig, "Modeling and Simulation of a Multi-Robot System Architecture," in Proceedings of the 2019 International

Conference on Mechatronics, Robotics and Systems Engineering, MoRSE 2019, 2019, doi:10.1109/MoRSE48060.2019.8998662.
[2] B. Sendhoff, H. Wersing, "Cooperative Intelligence-A Humane Perspective," in 2020 IEEE International Conference on Human-Machine Systems (ICHMS), 1–6, 2020.
[3] L.E. Parker, "Current research in multirobot systems," Artificial Life and Robotics, **7**(1), 1–5, 2003.
[4] R. Alami, "Multi-robot Cooperation : Architectures and Paradigms," in Journées nationales de la recherche en robotique, 2005.
[5] I. Jawhar, N. Mohamed, J. Wu, J. Al-Jaroodi, "Networking of multi-robot systems: Architectures and requirements," Journal of Sensor and Actuator Networks, **7**(4), 2018, doi:10.3390/jsan7040052.
[6] D.D. Walden, G.J. Roedler, K. Forsberg, "INCOSE Systems Engineering Handbook Version 4: Updating the Reference for Practitioners," INCOSE International Symposium, **25**(1), 2015, doi:10.1002/j.2334-5837.2015.00089.x.
[7] L. Bass, P. Clements, R. Kazman, "Software Architecture in Practice (3rd Edition)," Architecture, 2012.
[8] R. Tesoriero Tvedt, P. Costa, M. Lindvall, Evaluating Software Architectures, Advances in Computers, **61**(C), 2004, doi:10.1016/S0065-2458(03)61001-6.
[9] J. Holt, S. Perry, SysML for systems engineering: 2nd edition: A model-based approach, 2013, doi:10.1049/PBPC010E.
[10] J.E. Perez-Martinez, A. Sierra-Alonso, "UML 1.4 versus UML 2.0 as languages to describe software architectures," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 3047, 2004, doi:10.1007/978-3-540-24769-2_7.
[11] S. Friedenthal, A. Moore, R. Steiner, A Practical Guide to SysML, 2012, doi:10.1016/C2010-0-66331-0.
[12] I. Malavolta, P. Lago, H. Muccini, P. Pelliccione, A. Tang, "What industry needs from architectural languages: A survey," IEEE Transactions on Software Engineering, 39(6), 2013, doi:10.1109/TSE.2012.74.
[13] R. Eshuis, "Symbolic model checking of UML activity diagrams," ACM Transactions on Software Engineering and Methodology, 15(1), 2006, doi:10.1145/1125808.1125809.
[14] R. Petrasch, R. Hentschke, "Process modeling for industry 4.0 applications: Towards an industry 4.0 process modeling language and method," in 2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016, 2016, doi:10.1109/JCSSE.2016.7748885.
[15] S. Zor, D. Schumm, F. Leymann, "A Proposal of BPMN Extensions for the Manufacturing Domain," Proceedings of the 44th CIRP Conference on Manufacturing Systems (ICMS 2011); Madison, Wisconsin, June 1-3, 2011., 2011.
[16] N.R. Jennings, M.J. Wooldridge, Agent Technology: Foundations, Applications and Markets, 1998.
[17] S. Kumar, U. Kumar, Java Agent Development Framework, International Journal of Research, **1**(9), 2014.
[18] A.R. Sadik, A. Taramov, B. Urban, "Optimization of tasks scheduling in cooperative robotics manufacturing via Johnson's algorithm case-study: One collaborative robot in cooperation with two workers," in Proceedings - 2017 IEEE Conference on Systems, Process and Control, ICSPC 2017, 2017, doi:10.1109/SPC.2017.8313018.
[19] M. Hoffmann, "Analysis of the current state of enterprise architecture evaluation methods and practices," in ECIME 2007: European Conference on Information Management and Evaluation, 2007.
[20] I. Sommerville, Software engineering (10th edition), 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 944-950 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Empathy in Nursing Students that do the Non-Medical Internship in Three Universities in Lima, 2019

Liseth Acuña-Medina[1], Yumira Arias-Quispe[1], Yackeline Espeza-Veláquez[1], Brian Meneses-Claudio[2], Hernan Matta-Solis[1], Eduardo Matta-Solis[3]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, Lima, 15314, Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima, 15314, Perú*

[3]*Research and Intellectual Creativity Direction, Universidad María Auxiliadora, Lima, 15408, Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 19 November, 2020
Online: 08 December, 2020

Keywords:
Empathy
Students
Nursing internship
Non-medical internship

ABSTRACT

Empathy is an attribute of importance between the nursing professional and the patient, where it influences the recovery of their health. Several studies carried out have shown worrying results where they reflect a great decrease in empathy during the formative academic years. The objective of the study is to determine empathy in nursing interns from three university institutions in Lima, 2019. The study is a quantitative approach, with a non-experimental, prospective, descriptive and cross-sectional design. The application of the Jefferson Medical Empathic Scale instrument for students in Spanish version S is proposed, where the data was entered into a matrix in the Microsoft Excel 2013 program, then the IBM SPSS Statistics version 24.0 is subsequently analyzed obtaining the percentage of empathy of the internal, also take actions to improve these variables. 216 nursing interns from three universities in Lima participated. 51.4% are from Universidad Norbert Wiener, 89.4% (n = 193) were women, 56.4% (n = 123) are from the ninth cycle, 42.4% (n = 91) they study and work, 58 nursing interns from Universidad Norbert Wiener representing 52.3% have a medium level, 34 nursing interns from the Universidad de Ciencias y Humanidades, representing 30.6% have a medium level and finally 19 nursing interns from the Universidad Privada del Norte representing 17.1% have a medium level. Universidad Norbert Wiener obtained a higher score prevailing the medium level.

1. Introduction

The definition of empathy over time has been modified and expanded, coming to recognize it more than a static trait, also a state or ability that flows, being susceptible to change according to personal experiences and cultural interventions [1]. Empathy in the nurse-patient relationship should be used on a day-to-day basis, since it may affect their work of caring for patients, the decrease of it may lead to the loss of ability that allows better communication in the opinion of the patient since this is considered a relevant factor in the recovery process [2].

The comparison of the three universities is very important since it is going to differ in the type of teaching that each university offers to nursing interns, in turn the ability of nursing interns to function in the clinical field, where His training will stand out depending on each intern willing to improve and be a successful professional.

These circumstances may be due to the educational system because it must be oriented to the service of society and the individual, which should be based on the needs and demands that are important for development and competent from the professional point of view [3].

To carry out a relationship between nurse and patient, health personnel provide empathetic treatment to the patient to not worse the situation of the patient's disease, thereby also giving the staff the development of attitudes, capacities and skills, giving a care of quality, that's why, different causes of the deficit are visualized in its application.

The health personnel with the longest service time and high hierarchy level are those that present the greatest decrease in the level of empathy [4].

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & bmeneses@uch.edu.pe

In a study on health students, "Evaluation of empathy in human medicine students from the private university of Peru" [5], they found that those students who think that their specialty is linked to Patient Care presented greater empathy with 66.9%, the same as those who reported having a professional model in patient contact with 72%.

In a research work [6] studies carried out in nursing and medical students, indicate that students have found variations in empathy in terms of specialty and clinical competence, being worrying because some of them reflect a decrease in empathy over the years of training, the score is from 0.0 to 140.0; the women with 118.7 obtained a higher level of empathy score than the men with 110.4 points, differences that were statistically significant.

In the research work "Concepts on empathy: development of an instrument for formative evaluation based on a teaching experience with pediatric residents of the Universidad de Antioquía" [7], indicates that a deficit in student training on the basis and practice of empathy, because the participants had misconceptions about empathy, where they adopted universal criteria for all their patients, believing that they had to make the patient's wishes come true, make the patient or family feel that are important in the decisions of health personnel, among others.

Currently, professionals who can work normally in the socio-labor context that allows them to function interpersonally are required, the variable empathy is essential in the deployment of a professional in different medicine and social fields that its performs, its importance is appreciated in the development of professional skills [8].

The objective of the research work is to determine empathy in nursing interns from three University Institutions in Lima, 2019 according to their three dimensions of compassionate care, perspective taking and putting themselves in the other's place. The implementation of empathy on a daily basis will help, also the mental state of the patient and the health personnel, to recognize and identify mental problems that could be going unnoticed, being a national research priority 2015-2021 established by the Instituto Nacional de Salud [9].

The technique to be used is the survey that will be through the questionnaire and data collection instrument the Jefferson Scale of Empathy (JSE) for students in Spanish version S, created by the Jefferson group, led by Mohammadreza Hojat, consisting of 20 evaluated questions on a 7-point Likert-type scale, the score ranges from a minimum of 20 to a maximum of 140 points, there are 10 items written in a positive sense and 10 in a negative sense, whose items explore the dimensions of perspective taking, compassionate care and ability to put yourself in the other's place. Permission has been offered to use the Jefferson Empathy Scale in 85 countries. The JSE has been translated into 56 languages[10]. After this, the data will be entered into a matrix in Microsoft Excel 2010, and then exported to the IBM SPSS Statistics Base 22 statistical package, in which it will be analyzed.

The following research work is structured as follows: In section II, the development of the used data processing will be presented. In section III, the results of the tables and figures of sociodemographic agreement, institution and dimensions will be shown. In section IV, it will present the discussions of the research work and finally in section V, the conclusions, as well as the future work that is intended to be wanted with the research work.

2. Methodology

2.1. Research type and design

The present research is of a quantitative approach because the variables were measured, statistical data was used generalizing the hypothesis, the methodological design is a non-experimental, prospective research because according to the planning of data collection, they are primary, cross-sectional because the measure is non-experimental, it was for a single occasion, descriptive because it consists of a single variable.

2.2. Population and sample

The study population is made up of 216 nursing interns from three private universities in Lima, enrolled in 9th and 10th semesters, who carry out their pre-professional practices and are in contact with the patient, during the period of 2019, according to the Population type are non-probabilistic type, selected according to the inclusion and exclusion criteria.

Inclusion criteria

University students of both sexes who are studying nursing at the time of data collection, nursing interns enrolled during the academic semester 2019 who are taking the 9th and 10th semesters who voluntarily participate in the interview or fill out the data collection questionnaire, nursing students who agree to sign the informed consent.

Exclusion criteria

Nursing students who are from the 1st semester to the 8th semester and nursing interns who have not signed the informed consent to carry out the study.

2.3. Technique and instruments

The technique used was the survey, in the form of the Jefferson Scale of Empathy in Spanish for students, containing general student data such as: sex, age and year of training that will be used to obtain information of the students who participate in the survey, specially designed to measure empathy in the context of patient care, whose items consist of 3 dimensions of perspective taking, compassionate care and ability to put yourself in the patient's place. It consists of 20 items which are evaluated on a 7-point Likert-type scale ranging from: Strongly Agree (7 points) to Strongly Disagree (1 point). The score is from a minimum of 20 points to a maximum of 140 points, the highest scores correspond to a high level of empathy. There are 10 items written in a positive sense and 10 in a negative sense (inverted score) [10].

"Jefferson Scale of Empathy" has been validated in many countries around the world, including Latin American countries and previously applied in Peru, it has construct validity; convergent, parts that comprise are related to each other; and discriminating, its parts are not related to topics different to empathy. In addition, it has the Chi X^2 that is 0.5 in its various versions and translations, which is a reliable measurement scale and each of the questions are correlated with each other and with the entire instrument [11].

2.4. Data collection procedure

- An authorization was presented to the Nursing faculty for the execution of the project in order to obtain the respective permission.
- Subsequently, the nursing faculty provided facilities for the execution of this research, also will provide a cover letter.
- Once obtained, it was presented to teachers and delegates of the ninth and tenth semesters informing the objectives of the project and the informed consent, in order to obtain their acceptance, support and collaboration in the application of the instrument.
- The evaluators met teachers and delegates of ninth and tenth semester to discuss the date and time of application of the instrument.

2.5. Application of data collection instruments

- The evaluators attended the classrooms of each semester, ninth and tenth, on coordinated dates and times.
- The students were informed of the importance of their participation, a brief explanation about the instrument.
- The informed consent form was distributed to be signed for the student and then the questionnaire.
- The questionnaire development lasted approximately 10 to 15 minutes.
- Then the instrument was collected, verifying that all questions are answered.
- In closing, the students were thanked for their participation and collaboration in the research.

As shown in Figure 1, nursing interns were presented and explained how to carry out the survey. The protection aspects of the research participants and bioethical principles were taken into account, likewise informed consent was applied after clear information given to the participants, the application of the instrument is important, it will help especially in the mental state of the patient and the health personnel themselves, to recognize and identify mental problems that could be going unnoticed, being a national priority, which asks them to collaborate in providing with certain data that will allow to get the study objective .



Figure 1: Teacher permission and approval to apply the survey to students.

After obtaining the approval of the teacher of the classroom, the survey was applied, where advice was given before and during the application because questions were raised by the students, most of them were just to clarify questions, as shown in Figure 2.



Figure 2: Advice before and during the survey application

2.6. Data analysis and processing techniques

The data collection process was carried out during May 2020, the data was initially entered into a data matrix designed in the Microsoft Excel 2010 program, then exported to the IBM SPSS Statistics version 22.0 program. There, they were analyzed, obtaining tables and graphs that were later described and interpreted in the results and discussion sections, respectively.

For the statistical analysis of each main variable involved in the study, descriptive statistical tests such as absolute frequencies, relative frequencies, and measures of central tendency were applied, in addition to some correlational tests to verify the study hypothesis.

3. Results

A summary table of the surveys carried out following the guidelines corresponding to the research work will be shown, in the study participants were 216 nursing interns of 9th and 10th semester from the three University Institutions, in the following table, it will present results of the sociodemographic data of the study participants.

Table 1: Sociodemographic data of empathy in nursing interns from three University Institutions in Lima, 2019 (N = 216)

Participant Information	Total	
	N	%
Total	**216**	**100**
University		
Universidad de Ciencias y Humanidades	64	29,6
Universidad Norbert Wiener	111	51,4
Universidad Privada del Norte	41	19,0
Gender		
Female	193	89,4
Male	23	10,6
Study semester		
9th Semester	123	56,9
10th Semester	93	43,1
Occupation Status		
Study	61	28,2
Study and eventually work	64	29,6
Study and work	91	42,1

Regarding the university, 111 nursing interns representing 51.4% are from Universidad Norbert Wiener, 64 nursing interns representing 29.6% are from the Universidad de Ciencias y Humanidades, and finally 41 nursing interns representing 19% are from the Universidad Privada del Norte.

Regarding gender, the female sex predominated with 193 students representing 89.4%, followed by the male with 23 students representing 10.6%.

Regarding the study semester of nursing interns, the ninth semester is 123 representing 56.9%, and the tenth semester, 93, representing 43.1%.

Regarding the occupation status of nursing students, 91 with (42.1%) cases study and work, followed by 64 with (29.6%) cases study and eventually work and 61 with (28.2%) cases only study.

Table 2: Empathy in nursing interns from three University Institutions in Lima, 2019 (N = 216)

			Universidad de Ciencias y Humanidades	Universidad Norbert Wiener	Universidad Privada del Norte
Empathy	**Low**	N	0	0	1
		%	0,0	0,0	100,0
	Medium	N	34	58	19
		%	30,6	52,3	17,1
	High	N	30	53	21
		%	28,8	51,0	21,2

In Table 2, it can see Empathy in nursing interns from three University Institutions in Lima, 58 nursing interns from Universidad Norbert Wiener representing 52.3% have a medium level, 34 nursing interns from the Universidad de Ciencias y Humanidades, representing 30.6% have a medium level and finally 19 nursing interns from the Universidad Privada del Norte representing 17.1% have a medium level. Universidad Norbert Wiener obtained a higher score prevailing the medium level.

Table 3: Empathy in relation to gender of nursing interns from three universities in Lima, 2019 (N = 216)

		Empathy			
		Low	Medium	High	Total
Student gender	Female Count	0	98	95	193
	% within Student's gender	0,0%	50,8%	49,2%	100,0%
	Male Count	1	13	9	23
	% within Student's gender	4,3%	56,5%	39,1%	100,0%
Total	Count	1	111	104	216
	% within Student's gender	0,5%	51,4%	48,1%	100,0%

Chi-square tests

	Value	df	Asymptotic significance (bilateral)
Pearson's Chi-square	8,958[a]	2	,011
Likelihood ratio	5,069	2	,079
Linear by linear association	1,649	1	,199
N of valid cases	216		

a. 2 cells (33.3%) have expected a count less than 5. The minimum expected count is ,11.

In Table 3, empathy is related according to sex, where women present a higher level of empathy (49.2%) than in men (39.1%), in which it was determined with the chi-square test of Pearson (X^2). The level of significance of the test obtained a value of 0.11 (p> 0.05) (X^2 = 8.958; d.f = 2). Therefore, an association hypothesis is not rejected, for which there are statistical data that verify the relationship between empathy and the sex of nursing interns from three universities in Lima.

In Table 4, we can observe the relationship between the dimension of perspective taking with the Universities of the nursing interns, in which it was determined with Pearson's chi-square test (X^2). The level of significance of the test obtained a value of 0.95 (p> 0.05) (X^2 = 3.512; d.f = 4). Therefore, an association hypothesis is not rejected, for which there are statistical data that verify the relationship between the perspective taking dimension and the universities of the nursing interns in the research. By which, we can interpret that the interns of the Universidad Nolbert Wiener have a high perspective taking (79.3%).

Table 4: Empathy in its perspective-taking dimension in nursing interns from three universities in Lima, 2019 (N = 216)

			Institution			
			Universidad de Ciencias y Humanidades	Universidad Nolbert Wiener	Universidad Privada del Norte	Total
Perspective taking	Low	Count	1	3	1	5
		% within Institution	1,5%	2,7%	2,5%	2,3%
	Medium	Count	19	20	8	47
		% within Institution	29,7%	18,0%	19,5%	21,8%
	High	Count	44	88	32	164
		% within Institution	68,8%	79,3%	78,0%	75,9%
Total		Count	64	111	41	216
		% within Institution	100,0%	100,0%	100,0%	100,0%

Chi-square tests

	Value	df	Asymptotic significance (bilateral)
Pearson's Chi-square	3,512[a]	4	,476
Likelihood ratio	3,398	4	,494
Linear by linear association	,981	1	,322
N of valid cases	216		

a. 3 cells (33,3%) have expected a count less than 5. The minimum expected count is ,95.

Table 5: Empathy in its compassionate care dimension in nursing interns from three universities in Lima, 2019 (N = 216)

			Institution			
			Universidad de Ciencias y Humanidades	Universidad Nolbert Wiener	Universidad Privada del Norte	Total
Beware with compassion	Low	Count	8	16	6	30
		% within Institution	12,5%	14,5%	14,6%	13,9%
	Medium	Count	26	54	16	96
		% within Institution	40,6%	48,6%	39,0%	44,4%
	High	Count	30	41	19	90
		% within Institution	46,9%	36,9%	46,4%	41,7%
Total		Count	64	111	41	216
		% within Institution	100,0%	100,0%	100,0%	100,0%

Chi-square tests

	Value	df	Asymptotic significance (bilateral)
Pearson's Chi-square	2,278[a]	4	,685
Likelihood ratio	2,285	4	,684
Linear by linear association	,138	1	,710
N of valid cases	216		

a. 0 cells (0,0%) have expected a count less than 5. The minimum expected count is 5,69.

In Table 5, we can observe the relationship between the dimension of compassionate care with the Universities of nursing interns, in which it was determined with Pearson's chi-square test (X^2). The level of significance of the test obtained a value of 5.69 ($p > 0.05$) ($X^2 = 2.278$; d.f = 4). Therefore, it is not possible to reject an association hypothesis, in which there is no statistical evidence in a significant relationship in the compassionate care dimension with the universities of nursing interns. By which we can interpret that the interns of the Universidad de Ciencias y Humanidades have a high compassionate care (46.9%).

Table 6: Empathy in its dimension, ability to put oneself in the place of others in nursing interns from three universities in Lima, 2019 (N = 216)

			Institution			
			Universidad de Ciencias y Humanidades	Universidad Nolbert Wiener	Universidad Privada del Norte	Total
Ability to put oneself in the place of others.	Low	Count	17	41	12	70
		% within Institution	26,5%	36,9%	29,2%	32,4%
	Medium	Count	41	54	25	120
		% within Institution	64,1%	48,6%	61,0%	55,6%
	High	Count	6	16	4	26
		% within Institution	9,4%	14,5%	9,8%	12,0%
Total		Count	64	111	41	216
		% within Institution	100,0%	100,0%	100,0%	100,0%

Chi-square tests

	Value	df	Asymptotic significance (bilateral)

Pearson's Chi-square	4,581[a]	4	,333
Likelihood ratio	4,607	4	,330
Linear by linear association	,068	1	,794
N of valid cases	216		

a. 1 cells (11,1%) have expected a count less than 5. The minimum expected count is 4,94.

In Table 6, we can observe the relationship between the dimension ability to put oneself in the place of others with the Universities of nursing interns, in which it was determined with Pearson's chi-square test (X^2). The level of significance of the test obtained a value of 4.94 ($p > 0.05$) ($X^2 = 4.581$; $d.f = 4$). Therefore, an association hypothesis cannot be rejected. In which, there is no statistical evidence in a significant relationship in the compassionate care dimension with the universities of nursing interns. By which it can be interpreted that the interns of the Universidad de Ciencias y Humanidades have a medium level of empathy (64.1%).

This dimension was the most affected, it makes understand that being nursing interns where the work of a professional is being carried out having a great responsibility, the first thing the hospital prioritize is the technological or the material and forget the emotional, the feelings; this makes understand that during academic training, it does not have a good empathetic training where it must be emotionally prepared and not as professional competitiveness among University institutions.

These results will be used as antecedents for future researchers to give greater emphasis to humanized treatment and improve the level of empathy at first base that are the students.

4. Dicussion

In the present study, an increase in medium empathy was found in nursing students where the level of empathy occurred more in women than in men, this is because women, because they are more emotionally sensitive, feel a more empathic concern, personal discomfort where concern for the care of the child, adult and older adults perceive themselves based on their understandings and at the same time on non-verbal language. Our findings are similar to the results of the authors M. Montilva and her collaborators in their study "Empathy according to the Jefferson scale in students of Medicine and Nursing in Venezuela", the average score of the EMed was 120.3, and the EEnf was 117.9; found a non-significant difference in the score with a higher level of empathy in women than in men [6]. The presence of empathy in the Peruvian population of three universities was 111, also have a medium empathy.

According to the study, empathy in nursing students according to their compassionate care dimension, it found with a high frequency, regarding the author M. Mejía and her collaborators in their study "Empathic behavior in science students of the health from Universidad de los Andes ", it is observed that there is not as much relationship since the students of the Universidad de los Andes are less compassionate [12].

In the found findings, it can observe that empathy in nursing students who carry out the non-medical internship, in its dimension perspective taking is high, which is related to the study of the author R. Lorch and his collaborators "a phenomenological study of nature of empathy ", which leads to the importance of developing empathy as an aspect of emotional competence, this study is related to the understanding that nursing students perceive and experience [13].

Another findings of empathy in health students, elaborated by J. Huancaya and his collaborators in their study "Evaluation of empathy in human medicine students of the private university of Peru", they found that those students who think their specialty was linked to patient care, they showed greater empathy, like those who reported having a professional role in patient contact, a higher percentage of the level of empathy in women was also found [5].

In the study by J. Blanco and collaborators in their study "Medical empathy in residents and medicine tutors of family and community. The vision of the professional and the patient "found that the health personnel, who have the longest service time, presented low levels of empathy compared to other professionals [14]. Regarding nursing students, it is observed that the majority have medium empathy, comparing with the mentioned study, the difference between them was that in the non-medical internship, there is a decrease in empathy.

It can be seen that there is a similar relationship in terms of gender, empathic levels in women are high, which indicates that women are superior to men.

5. Conclusions

It is concluded that empathy in nursing interns from three university institutions in Lima, 2019, medium empathy predominates. Regarding the dimensions of the three universities, the most affected dimension was putting themselves in the place of others the low level predominates, according to their compassionate care dimension, the medium level empathy predominates and, in the perspective, taking dimension the high level predominates.

In the bibliographic reviews of the national and international scientific databases, it was determined that there is not an adequate number of national scientific articles that involve this type of subject: Empathy in nursing interns and the comparison between different universities. Knowing that these students of the new generation must be prepared not only with scientific aspects but also with the human part. For this reason, the present research wants to complete this research gap.

The study contributes to the development and consolidation of the topic of research in empathy, and also introduces the concept of empathy in nursing interns through an instrument, which its new version is adapted to the field of nursing-patient. The results will contribute to the improvement of the humanized treatment of the patient.

Conflicts of Interest

The authors declare no conflict of interest

References

[1] J. McKinnon, "In their shoes: An ontological perspective on empathy in

nursing practice.," Journal of Clinical Nursing, **27**(21), 3882–3893, 2018, doi:10.1111/jocn.14610.

[2] J. Álvarez, J. Sáchica, J. Villalba, "Percepción de los pacientes acerca de la empatía de las enfermeras en Monterrey (México).," Revista Española De Comunicación En Salud, **9**(1), 46, 2018, doi:10.20318/recs.2018.4252.

[3] Navarro G., Vaccarezza G., Gracia M., Catalán R., Construcción de conocimiento en educación superior., 2015.

[4] M. Neumann, F. Edelhäuser, D. Tauschel, M. Fischer, M. Wirtz, C. Woopen, A. Haramati, C. Scheffer, "Empathy decline and its reasons: A systematic review of studies with medical students and residents.," Academic Medicine, **86**(8), 996–1009, 2011, doi:10.1097/ACM.0b013e318221e615.

[5] J. Huarcaya, M. Rivera, "Evaluación de la empatía en alumnos de medicina humana de una universidad privada del Perú.," Educacion Medica, **20**, 100–107, 2019, doi:10.1016/j.edumed.2018.05.008.

[6] M. Montilva, M. García, A. Torres, M. Puertas, E. Zapata, "Empatía según la escala de Jefferson en estudiantes de Medicina y Enfermería en Venezuela.," Investigación En Educación Médica, **4**(16), 223–228, 2015, doi:10.1016/j.riem.2015.04.006.

[7] O. Salazar, L. Dupuy, "Concepciones sobre empatía: desarrollo de un instrumento para la evaluación formativa a partir de una experiencia docente con residentes de pediatría de la Universidad de Antioquia.," Revista de La Fundación Educación Médica, **20**(5), 227, 2017, doi:10.33588/fem.205.910.

[8] L. Loyola, S. Landgrave, E. Ponce, A. González, "Nivel de empatía en médicos del servicio de urgencias de un hospital público de la ciudad de México.," Revista Atención Familiar, **22**(2), 43–45, 2015.

[9] Instituto Nacional de Salud, Prioridades de Investigación en Salud en el Perú 2019-2023., 2019.

[10] A. Callhan, J. Gonnella, V. PharmD, J. Veloski, M. Hojat, M. LaNove, Center for Research in Medical Education and Health Care., Jefferson Philadelphia University Thomas Jefferson University, 1–39, 2018.

[11] M. Hojat, Empatía en el cuidado del paciente: Antecedentes, desarrollo, medición y resultados., 1–318, 2012.

[12] M. Mejía, J. Poveda, V. Díaz, "Comportamiento empático en los estudiantes de las ciencias de la salud. Universidad de Los Andes.," Fermentum, **23**(67), 203–220, 2013.

[13] R. Lorch, P. Maras, R. Ruiz, L. Verschaffel, "Descenso de empatía en estudiantes de enfermería y análisis de posibles factores implicados.," Psicología Educativa, **20**, 53–60, 2014.

[14] J. Blanco, C. Valcárcel, M. Guerra, E. Ortigosa, F. García, F. Caballero, L. Cabello, D. Monge, "Empatía médica en residentes y tutores de medicina familiar y comunitaria. La visión del profesional y del paciente.," Atención Primaria, **52**(3), 185–192, 2020, doi:10.1016/j.aprim.2018.04.009.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 178-183 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Level of Empathy in Nursing Students Attending Clinical Practices of the Universidad de Ciencias y Humanidades

Walter Cervera-Flores[1], Yenifer Choque-Garibay[1], Nahuel Gonzalez-Cordero[1], Brian Meneses-Claudio[2, *], Hernan Solis-Matta[1], Lourdes Matta-Zamudio[1]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 25 December, 2020
Online: 15 January, 2021

Keywords:
Empathy
Nursing students
Clinical practices
Compassionate care
Professional training

ABSTRACT

Empathy in the care of the patient by the nursing students is important because it allows having the capacity of response towards the patient, this study aimed to determine the level of empathy in nursing students who attend in clinical practices of the Universidad de Ciencias y Humanidades. This is a cross-sectional study, with a population of 289 nursing students who answered a questionnaire and the Jefferson medical empathy scale. 210 participants (72.7%) had a medium level of empathy, 74 participants (25.6%) with a high level of empathy and 5 participants (1.7%) with a low level of empathy. Regarding the empathy dimensions, the high level in the perspective taking dimension predominated with 81.7%, followed by the average level of the capacity dimension to put oneself in the patient's place with 59.9% and the compassionate care dimension predominated in the low level with a percentage of 45%. An improvement program is required in terms of professional training so that nursing students can carry out compassionate care when performing their clinical practices.

1. Introduction

Empathy in patient care for nursing students is a cognitive attribute that implies the ability to understand pain, suffering and the patient's perspective, it is the ability to stablish this understanding and the intention to help [1].

Nursing students must be empathetic and this is governed based on their behavior, this will be influenced in the well-being of themselves and also of the patient [2].

Empathy depends on the personality and functional characteristics of the nurse so that the nursing student has a good performance in their study plan [3]. Therefore, there are numerous exploration methods that have been designed in educational programs to improve empathy [4]. In addition, empathy is attributed to factors that influence the nursing student such as interpersonal relationships, self-esteem, communication, and self-efficacy [5].

In [6], the authors explain a study carried out in Spain, it was found that technological advance acts on empathy in nursing professionals, leading to a lack of interest in interaction with patients.

In [7], the authors presented a study carried out in Venezuela, it was observed in the results that 90.1% of the nursing students were women and those who obtained a higher score in level of empathy.

In [8], the authors described a study carried out in Canada, they stated in their results that the instructional and practice sessions increased students' self-awareness about prejudices and interest in learning empathy through video tag feedback.

This study is important since it will give relevant and real data about empathy as students go in their academic plan and, observe if they can have an empathetic attitude when they carry out their activities in the clinical practical field.

The following research work is structured as follows: In section II, the development of the data collection process of each nursing student will be presented, as well as the guidelines to consider so that they are within the research work. In section III, the results will show the level of empathy that nursing students

*Corresponding Author: Brian Meneses-Claudio, bmeneses@uch.edu.pe

will present according to the specified dimensions of the instrument in the measurement of the variable. In section IV, it presents the discussions of the research work, in section V, the conclusions and in section VI the recommendations as well as the future work that is intended to come with the research work.

2. Methodology

2.1. Materials and Methods

In this part, the type and design of the research will be developed, as well as the population and sample that will be carried out in the research work, the inclusion and exclusion criteria in detail and finally the technique and the instrument of collection of data.

2.2. Research type and Design

It is a study with a quantitative approach, a non-experimental, descriptive and cross-sectional design [9].

2.3. Population

The population was 289 nursing students from the Universidad de Ciencias y Humanidades.

2.3.1. Inclusion Criteria

Nursing students who are in the third semester until the eighth semester and who have signed the informed consent that was evaluated and approved by the ethics committee of the Universidad de Ciencias y Humanidades.

2.4. Technique and Instrument

In the study, the data collection instrument Escala de Empatía Médica de Jefferson (EEMJ) from Hojat Mohammandreza was applied, which has been demonstrated to be useful in assessing level of empathy in nursing students.

The JSE instrument is a shortened and derived tool of the Escala de Empatía Médica de Jefferson, which consists of 20 questions developed in parallel in more than 85 countries and translated into 56 languages.

To measure empathy, the instrument Escala de Empatía Médica de Jefferson (EEMJ) from Hojat Mohammandreza, which includes 20 items grouped in 3 dimensions, will be used: perspective taking, compassionate care and ability to put oneself in the patient's place where they are governed by the high, medium, and low levels according to the instrument manual. The answers will have two significant frames for each of the items, if it disagrees it is 0 and if there is agreement it is valued at 7, with a variation score of 20 to 140. Between 20 to 60 refers to low empathy, of 61 to 100 refers to medium empathy and 101 to 140 refers to high empathy [10].

2.5. Place and Application of the Instrument

The questionnaire was carried out in two consecutive schedules, morning, and night at the same time as the nursing students who practice at the university, the questionnaire was taken to the nursing students with an approximate time of 15 minutes each selected semester (III semester to VIII semester according to inclusion criteria) in the research work, concluding with good satisfaction when collecting the questionnaires as the students provided support to carry out the research work.

The data collection was processed through the survey of nursing students, the data was carried out in a data matrix that will be designed in the statistical program SPSS (Statistical Package for the Social Sciences) in its version 24.0, in which it will allow a better data processing to make statistical tables and graphs so that they can then be described and interpreted in results and discussions, respectively.

It is important to emphasize the presence of medical personnel at the time of completing the student questionnaire, since most of them presented signs of having a health problem.

3. Results

Regarding empathy in nursing students who carry out pre-professional practices at the Universidad de Ciencias y Humanidades it can see that 210 participants represent 72.7% with a medium level of empathy, followed by 74 participants who represent 25.6% with a high level of empathy and 5 participants representing 1.7% with a low level of empathy.

Table 1: Level of Empathy in Nursing Students attending Clinical Practices of the Universidad de Ciencias y Humanidades, 2019 (N = 289)

Empathy					
		Frequency	**Percentage**	**Valid percentage**	**Accumulated percentage**
Valid	**Low**	5	1,7	1,7	1,7
	Medium	210	72,7	72,7	74,4
	High	74	25,6	25,6	100,0
	Total	289	100,0	100,0	

These results indicate that the students present a medium empathy, because many of them possess skills that would demonstrate high empathy, however, they have the difficulty of expressing or manifesting what another person may experience, in which they cannot to be developed without having the adequate knowledge, for this, they seek solutions either at an academic or practical level in which they can complement that space, in order to be able to develop their skills and have the greatest capacity for evaluation in what another person may experience.

In the Figure 1 shown, the results regarding the dimensions of empathy are evidenced according to the data based on the research instrument:

In Figure 1, it can see the empathy according to its dimensions, in nursing students who attend clinical practices of the Universidad de Ciencias y Humanidades, where the dimension most affected was the compassionate care where 130 students representing 45% have low empathy. This dimension has been the most affected, although it is true it is related to patient care, however, it has been proven that the same patients

attribute that their care is not adequate, that they are not provided with compassionate care, for the Nursing students, this dimension assumes an important role due to the fact that the students are in full training, that compassionate care is very important, for this, education at an academic or practical level, it is necessary to distinguish a very important role on compassion, due to this fundamental role is being lost in nursing students due to the lack of interventions on understanding compassionate care for the patient.

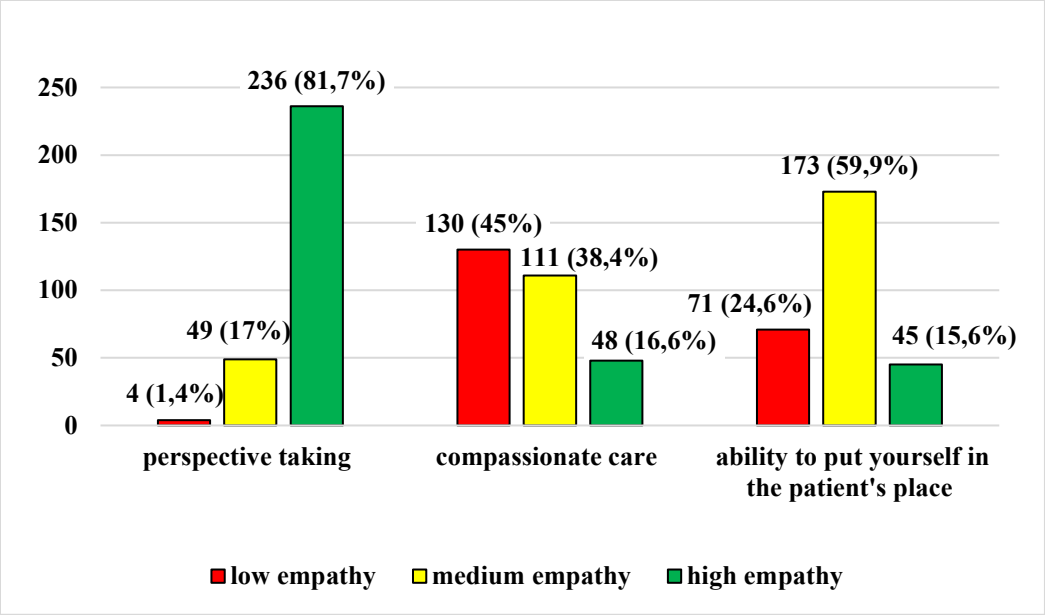


Figure 1: Empathy according to its dimensions, in nursing students attending clinical practices at the Universidad de Ciencias y Humanidades, 2019 (N = 289).

Table 2: Empathy in relation to the Semester of study of Nursing Students who attend in clinical practices of the Universidad de Ciencias y Humanidades, 2019 (N = 289)

			Empathy			Total
			Low	Medium	High	
STUDY SEMESTER	III SEMESTER	Count	0	61	16	77
		% within the Semester	0,0%	79,2%	20,8%	100,0%
	IV SEMESTER	Count	3	25	9	37
		% within the Semester	8,1%	67,6%	24,3%	100,0%
	V SEMESTER	Count	0	47	10	57
		% within the Semester	0,0%	82,5%	17,5%	100,0%
	VI SEMESTER	Count	0	34	16	50
		% within the Semester	0,0%	68,0%	32,0%	100,0%
	VII SEMESTER	Count	1	28	9	38
		% within the Semester	2,6%	73,7%	23,7%	100,0%
	VIII SEMESTER	Count	1	15	14	30
		% within the Semester	3,3%	50,0%	46,7%	100,0%
Total		Count	5	210	74	289
		% within the Semester	1,7%	72,7%	25,6%	100,0%
Chi-square tests						

	Value	gl	Asymptotic significance (bilateral)
Pearson's Chi-square	24,344[a]	10	,007
Likelihood ratio	22,625	10	,012
Linear by linear association	4,130	1	,042
N°. of valid cases	289		

a. 6 cells (33.3%) have expected a count less than 5. The minimum expected count is .52.

In Table II, empathy is related to the semester of study of nursing students, in which it was determined with the Pearson chi-square test (X^2), the level of significance of the test obtained a value of 0.52 ($p > 0.05$) ($X^2 = 24.344$; $df = 10$). Therefore, a hypothesis of association between variables is not rejected, thus proving that there is a relationship between empathy and the semester of study of nursing students.

4. Discussion

In this study, we raise the issue of empathy is raised from the point of view of professional nursing training, where the level of empathy in nursing students who attend clinical practices at the Universidad de Ciencias y Humanidades, 2019 is assessed.

Regarding the level of empathy in nursing students who attend clinical practices of the Universidad de Ciencias y Humanidades, 2019, the medium level predominated followed by the high and low. These results may be because students have difficulty being able to understand or feel what the other person is experiencing.

It should be noted that the previous research work on the Empathy study were of different sociodemographic characteristics, the study in [3], the author indicates that the results of empathy were given by the different characteristics such as sex, academic year, academic performance, where the female sex showed high levels of empathy. Where empathy can change as the student progresses in their study plan. In [7], the author mention that the decrease in empathy is influenced by the absence of good teachers, the increased pressure for activities, long hours of service and evaluations in students, can lead them away from feelings of human problems and have a lack of opportunities to manage their emotions. Given the aforementioned, the nursing student can be related to their emotions, although it is true that empathy is not only based on doing a good job but, the way that the emotions are going to be applicated in this situation, it is very important to know that emotions have an important role in the students since they are in full formation and need the emotional maturity that allows them to perform humanized care, this means that it is not only enough to perform academically, because it would only enable their cognitive abilities, and to develop their emotional skills in a successful way, it is necessary to have training for it, so that the nursing students themselves carry out their humanized care at the time of carrying out their clinical practices.

Regarding the level of empathy in its dimension perspective taking in nursing students who attend clinical practices of the Universidad de Ciencias y Humanidades, 2019, the high level predominated followed by the medium and low. The perspective taking dimension refers to cognitive flexibility, understanding mood states, having good communication and interpersonal relationships with the patient. In [8], the author mentioned that students were concerned with learning about the process of conscious or consciously driven perspective taking and practicing self-awareness to help identify and then put aside biases and beliefs about health risk behaviors allowing better patient care. In [11], the authors indicate that empathy in a quality with great cognitive predominance, where understanding is more important. Since in this way, the patient's pain can be understood to show adequate communication and support, where the nurse-patient relationship will be strengthened. That is why having the ability to understand the patient's pain will be related to motivational factors in academic nursing training.

Regarding the level of empathy in its dimension compassionate care in nursing students who attend clinical practices at the Universidad de Ciencias y Humanidades, the low level prevailed followed by the medium and high levels, where this dimension was the most affected. The compassionate care dimension refers to the emotional states of others, literally feeling the person's pain. It originates from primitive parts of the brain that allow to feel quickly without thinking deeply. Compassion has global implications for nursing care. However, patients report that nurses do not provide compassionate care. Nursing education has been examined on the impact training can have on student compassion as there is a shortage of interventions aimed at teaching compassion to nursing students [12]. In [6], the authors mention the need for high empathic results, since emotional understanding is needed, listening to the patient's concerns so that it can alleviate its anguish thanks to the emotional expression of the nurse. That is why the study in [8], the authors indicated that the students said that it is important to listen to the stories of the patients, it was revealing as the stressors related to care and their impact on self-care. Patient feedback made it easier for students to better understand and take ownership of their risky health behavior.

Regarding the level of empathy in its dimension, ability to put yourself in the patient's place in nursing students who attend clinical practices of the Universidad de Ciencias y Humanidades, 2019, the medium level prevailed followed by the low and high. The dimension of putting yourself in the patient's place refers to feel not only how other people feel, but also to know what they need from you. It involves putting yourself in the other's shoes, examining issues from their perspective, trying to think and feel as the other would. Lack of communication produces a loss of

empathy and understanding between people, which translates into bad humor, fights, and an atmosphere of tension. Despite the burdens it has on the journey, it must not forget that others also have their own problems. Listening, dialoguing, understanding and helping are just some of the actions laden with value that can improve coexistence between people and allow to live in a more peaceful society [13].

Regarding the level of empathy in nursing students who attend clinical practices at the Universidad de Ciencias y Humanidades, 2019, studies carried out in [14], they concludes that women score significantly more than men in fantasy, empathic concern and personal discomfort.

In [15], the authors showed that women are more empathetic than men. This was attributed to an increased ability for perspective taking and compassionate attention, a situation that can be explained by social and biological factors. Among social factors, stereotypes attribute to women greater emotional sensitivity, concern for social aspects, a tendency to care for children and the elderly people, and greater ability to perceive feelings and non-verbal language.

In [7], they mentioned that the female sex has greater empathy than the male sex. This is due to the brain structure, where the nerve fibers between the two hemispheres perform better information transfer in communication and empathy, which occurs in the case of women. This would also explain the synchronization they have with the emotions expressed in non-verbal language with the patient.

The implementation of empathy education in the nursing professional training process brings good results. A systematic review in [16], titled "A systematic review of the effectiveness of empathy education for undergraduate nursing students," they indicate that in 23 studies of empathy education in undergraduate education in nursing, practical improvements in empathy were demonstrated. In all of them, interventions based on interview and experimental simulations were carried out, focusing on vulnerable patients.

5. Conclusions

It is concluded that the level of empathy in nursing students have been high, it could be specified that as students' progress academically, they tend to increase their empathy, depending on how the students can express themselves regarding their academic performance either theory or practical.

It is concluded that to have a high level of empathy, it is not only requires having a good academic ability, but also having a good ability to show emotions during clinical practices to provide self-confidence and gain the confidence of the patient during their care.

It is concluded that within the established dimensions of the research variable, in terms of the compassionate care dimension, an improvement program is required in terms of their professional training so that students can carry out compassionate care when performing their clinical practices.

As a future work, it is important to keep in mind that in universities there must be student nursing programs that allow them to benefit from their own empathic skills, so students could choose to have more knowledge and, at the same time, express their emotions through methods that allow to feel safe and have confidence in themselves when performing their clinical practices so that they are able to resolve any situation in patient care.

6. Recommendations

Regarding empathy in nursing students who attend clinical practices of the Universidad de Ciencias y Humanidades, 2019, the medium level predominated, followed by the high and low. It is suggested to the university authorities that activities be established to encourage students to have a good comprehensive education, in which the student can be involved in various topics in which the student can develop socially.

Regarding the empathy dimensions in all of them, the high level prevailed in the perspective taking dimension, it is suggested to the university authorities that students be fostered capacities in which they can distinguish when carrying out their clinical practices.

Training with the knowledge that the students could get a positive interaction with the people who will provide their care, involving both the faculty and the health institutions in which they will carry out their clinical practices.

It is suggested that, in future studies, a relationship between other variables linked to empathy be involved or exist, so that this problem can be better understood.

It is suggested to develop studies with qualitative or mixed approaches to complement the understanding of empathy in nursing students.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] U. Philadelphia, U. Jefferson, Jefferson Scale of Empathy. Center for Research in Medical Education and Health Care, Sidney Kimmel Medical College, 2019.

[2] H. Rozengway, A. García, I. Vallecillo, "Niveles de empatía según la escala de Jefferson en estu - diantes de Medicina, Enfermería y Odontología de Honduras.," Rev. Cient. Cien. Med., **19**(2), 14–19, 2016.

[3] M. Madera, L. Tirado, F. González, "Factores relacionados con la empatía en estudiantes de Enfermería de la Universidad de Cartagena.," Enfermeria Clinica, **5**(26), 282–289, 2016, doi:10.1016/j.enfcli.2016.06.004.

[4] P. Montanari, C. Petrucci, S. Russo, I. Murray, V. Dimonte, L. Lancia, "Psychometric properties of the Jefferson Scale of Empathy-Health Professional Student's version: An Italian validation study with nursing students.," Nursing and Health Sciences, **17**(4), 483–491, 2015, doi:10.1111/nhs.12221.

[5] J. Kim, "Factors influencing nursing students' empathy.," Korean Journal Of Medical Education, **30**(3), 229–2236, 2018, doi:10.3946/kjme.2018.97.

[6] M. Giménez, S. Avivar, V. Prado, "Niveles de empatía en una muestra de enfermeras españolas.," Calidad de Vida y Salud, **9**(2), 120–130, 2016.

[7] M. Montilva, M. García, A. Torres, M. Puertas, E. Zapata, "Empatía según la escala de Jefferson en estudiantes de Medicina y Enfermería en Venezuela.," Investigación En Educación Médica, **4**(16), 223–228, 2015, doi:10.1016/j.riem.2015.04.006.

[8] M. Lobchuk, G. Halas, C. West, N. Harder, Z. Tursunova, C. Ramraj, "Development of a novel empathy-related video-feedback intervention to improve empathic accuracy of nursing students: A pilot study.," Nurse Education Today, **46**, 86–93, 2016, doi:10.1016/j.nedt.2016.08.034.

[9] C. Fernández, P. Baptista, Metodología de la Investigación. 6ta ed. México: Mc Graw-Hill/Interamericana., 2015.

[10] A. Garza, J. González, S. Tavita, F. Rodriguez, M. Hojat, "Validación de la Escala de Empatia Médica de Jefferson en Estudiantes de Medicina Mexicanos.," Salud Mental, **28**(5), 57–63, 2005.

[11] J. Yu, M. Kirk, "Measurement of empathy in nursing research: Systematic review.," Journal of Advanced Nursing, **64**(5), 440–454, 2008, doi:10.1111/j.1365-2648.2008.04831.x.
[12] M. Durkin, R. Gurbutt, J. Carson, "Qualities, teaching, and measurement of compassion in nursing: A systematic review.," Nurse Education Today, **63**, 50-58., 2018, doi:10.1016/j.nedt.2018.01.025.
[13] G. Karayiannis, E. Papastavrou, A. Farmakas, H. Tsangari, M. Noula, Z. Roupa, "Exploration of empathy in Cyprus nursing and health care students: A mixed method study.," Nurse Education in Practice, **42**, 102686, 2020, doi:10.1016/j.nepr.2019.102686.
[14] V. Díaz, G. Muñoz, N. Duarte, M. Reyes, S. Caro, A. Calzadilla, L. Alonso, "Empatía en estudiantes de enfermería de la Universidad Mayor, Sede Temuco, IX región, Chile.," Aquichan, **14**(3), 388–402, 2014, doi:10.5294/aqui.2014.14.3.9.
[15] G. Parra, R. Cámara, "Nivel de empatía médica y factores asociados en estudiantes de medicina.," Investigación En Educación Médica, **6**(24), 221–227, 2017, doi:10.1016/j.riem.2016.11.001.
[16] T. Levett, R. Cant, S. Lapkin, "A systematic review of the effectiveness of empathy education for undergraduate nursing students.," Nurse Education Today, **75**, 80–94, 2019, doi:10.1016/j.nedt.2019.01.006.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 184-199 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

QoE-aware Bandwidth Allocation for Multiple Video Streaming Players over HTTP and SDN

Pham Hong Thinh[1,2], Tran Thi Thanh Huyen[3], Nguyen Ngoc Quang[1], Pham Ngoc Nam[4], Truong Cong Thang[2], Nguyen Viet Hung[5], Truong Thu Huong[*,1]

[1]*School of Electronics and Telecommunications, Hanoi University of Science and Technology, 100000, Vietnam*

[2]*Faculty of Engineering and Technology, Quy Nhon University, 820000, Vietnam*

[3]*The University of Aizu, Aizu-Wakamatsu, 965-0006, Japan*

[4]*College of Engineering & Computer Science, VinUniversity, Hanoi, 100000, Vietnam*

[5]*Faculty of Information Technology, East Asia University of Technology, Hanoi, 100000, Vietnam*

ARTICLE INFO

Article history:
Received: 30 October, 2020
Accepted: 06 January, 2021
Online: 15 January, 2021

Keywords:
HTTP Video streaming
Resource allocation
QoE
SDN
Traffic shaping
Adaptive Bitrate Streaming - ABR

ABSTRACT

For many years, the most popular technique for Internet video streaming is hypertext transfer protocol-based adaptive streaming, known as HAS (HTTP Adaptive Streaming). However, a seamless viewing experience can not be just simply guaranteed by HAS only. In the management network, the adaptation of HAS copes with a huge challenge since client-driven schemes lead to unfair share of available bandwidth when multiple players request adaptive bitrates (i.e bandwidth) through a bottleneck network link. Each client's requesting to maximize its needed bandwidth leads to the competition of network resources. This causes great QoE (Quality of Experience) reduction in terms of main metrics for each player: fairness, efficiency, and stability. In this paper, we propose an integration scheme of bitrate adaptation and Software Defined Networking-based resource allocation that can improve the QoE of competing clients. Our experiments show that the proposed scheme increases at least 20% up to 124% in terms of QoE scores compared with some existing methods as well as gains smoother viewing experience than the solutions of the traditional Internet.

1 Introduction

Recently we have observed a huge volume of media content, especially high-definition videos over the Internet. According to the report from [1], video traffic is estimated to occupy around 82% of the overall Internet traffic by 2021. Although the quality of the Internet has been improved significantly thanks to the advanced technologies such as 4G, 5G and Fiber to the Home (FTTH), current technologies are not yet ready to provide sufficient Internet bandwidth for new multimedia types such as 360-degree Videos and 12K Videos. Those new types of multimedia cause an enormous portion of Internet traffic.

The quality of video streaming over the Internet have been enhanced by several approaches such as: Microsoft's Smooth Streaming, Real-time Transmission Protocol, Adobe's Adaptive Streaming, Apple's HTTP Live Streaming, and MPEG-DASH. Especially, with MPEG-DASH, the vendor dependence on previously invented protocols were removed so as content providers can extend their services easily.

Based on this DASH standard emerged a number of solutions in which a source video is segmented into short segments of 2-10 seconds, each of which is encoded at a different bitrate resolution and level. During the playback, each client uses its bitrate adaptation logic to dynamically request and fetch desired encoded segments based on metrics such as average throughput and buffer occupancy. In the DASH technology, most of the current adaptive algorithms that decide the bitrate for the next downloaded segment are based on two parameters at the client side: throughput variation and buffer level. The buffer-based bitrate adaptation methods determine the bitrate mainly based on the characteristics of the buffer. These methods can also consider throughput, but the characteristics of the buffer are the preferred factors. Typically, buffers are divided into

[*]Corresponding Truong Thu Huong. Email: huong.truongthu@hust.edu.vn

multiple ranges, and for each of these ranges, different actions will be applied to determine the bitrate. Note that the specific buffer thresholds depend on the specific adaptation method. Typically, if the buffer level is high, the next segment will be selected with a higher bitrate than the current segment. If the buffer level is average, the selected bitrate will not be too different from the current bitrate. On the contrary, if the buffer level is low, the next segment will be selected with a low bitrate to avoid rebuffering and video freezes.

In addition, throughput-based methods are only based on estimated throughput to define a bitrate for the next segment. The main difference between these types of algorithms is the way for estimating throughput and how to utilize throughput. The simplest way to determine the throughput is based on the ratio of the given segment duration over the download time. The segment delivery time is determined from the time of sending a HTTP request to the time of receiving a HTTP response message that contains that segment.

On the other hand, Software-Defined Networking (SDN) [2] has emerged as a new networking paradigm in which a SDN Controller can make network management more flexibly by eliminating proprietary protocols from hardware vendors. With the network control plane implemented in software, rather than in firmware, network traffic can be managed more dynamically and at a more granular level. SDN switches/routers send and receive information to and from the SDN controller via southbound APIs; and the SDN controller relays information to the applications and business logic via northbound APIs. Within the scope of SDN, the OpenFlow Controller [3] is a type of SDN Controller that uses the OpenFlow protocol to configure SDN switches.

In any network exists always a popular issue called bottleneck link that could be congested when the amount of bandwidth demands go beyond the bottleneck link capacity. In case of bottleneck, video streams for each resource-sharing users could experience a remarkable decrease in perception of the video service, called Quality of Experience (QoE). Therefore, a method to solve the bottleneck problem in video streaming is neccesary.

To extend our previous work [4], in this work, we deploy a streaming architecture which make use of the MPEG-DASH technique and the SDN technology to solve the bottleneck problem. The scheme can increase the overall video quality for all users sharing the same bottleneck point. At the client's side, we develop a throughput-buffer based bitrate adaptation algorithm to utilize most of the available bandwidth in order to increase users' QoE. On the SDN controller, we establish a flow management mechanism which monitors the network to ensure fairness and avoid congestion when the network does not have sufficient bandwidth to serve new incoming stream requests. In the SDN controller, we develop the so-called Media Streaming Multiple Access (MSMA) module to monitor the network, detecting the number of ongoing streams, and deciding a fair share of bandwidth to all clients. Additionally, MSMA can also reallocate bandwidth or reject a new incoming request if no sufficient bandwidth is found; and thereby preventing congestion at the bottleneck point. The proposed adaptation and bandwidth allocation methods aim at the three following goals:

• $Efficiency$: all players joining the system can choose for themselves a level of video quality as high as possible.

• $Fairness$: When there are multiple clients accessing the bottleneck link at the same time, all of the available network resources must be allocated as fairly as possible.

• $Stability$: Each player should avoid video stalling, unnecessary quality switches, which will adversely affect users' QoE.

The proposed solution not only allocates bandwidth fairly and adaptably on the network side, but also adaptive bitrate at the client side to improve video streaming quality. When there are many clients streaming on bottleneck link, the problem of sharing bandwidth for optimization is very important. There are three main approach groups: dividing bandwidth equally, dividing it by groups or choosing which user should be the one to share bandwidth with new coming users or new request of an existing user. Each solution has its own advantages and disadvantages. So, our proposed technique's limitation is to trade off a small QoE degradation of one or a small number of users sharing the bandwidth. However, the advantage of the proposal is to maximize the number of access clients while keeping QoE in a good shape.

The rest of this paper is organized as follows. Section 2 presents related work of typical bitrate adaptation algorithms and bandwidth allocation schemes. Section 3 provides information on throughput estimation and the proposed adaptation algorithm. Our proposed SDN-based dynamic resource allocation is described in Section 4 . A performance evaluation is presented in Section 5. Finally, conclusions and future directions are discussed in Section 6.

2 Related Work

HTTP Adaptive Streaming (HAS) has become the key solution for video streaming. There are a number of recent researches conducted to improve users' QoE in the context of multiple clients requesting for DASH service. As discussed in [5] and [6], the authors proposed a so-called Aggressive method in which a rate adaptation algorithm based on the estimated throughput is implemented where the last segment throughput is simply used as the estimated throughput. The quality level will increase or decrease depending entirely on estimated bandwidth. The Aggressive method is sensitive to cases where the bandwidth increases sharply in a short period of time and then abruptly decreases (short-term bitrate peaks). The client then requires a high-quality version, but the high bandwidth level does not last long enough to load the required segment. This may cause video freezing during playback. The Aggressive algorithm aims to use most of the available bandwidth for the streaming client by always adapting to the highest quality version of video segment that does not surpass the available bandwidth. By doing this, the expectation of streaming users is also met. In [7], [8] the authors proposed a buffer-based bitrate adaptation algorithm for HTTP video streaming, called BBA. BBA is based on the current playback buffer occupancy only by selecting bitrate heuristically without throughput estimation. The goal of BBA is avoiding stalling events and maximizing the average video quality by choosing the highest available bitrate level. However, this scheme shows many limitations such as QoE degradation in case of long-term bandwidth variations and slow bandwidth fluctuation, especially in case of bandwidth competition in a shared network. In the literature, there are some ABR (Adaptive Bitrate) algorithms that take into account both of the estimated throughput and current buffer occupancy to define the most suitable video version

for the next segment. In [9], [10], the authors propose a buffer-based bitrate adaptation algorithm for VBR (variable bitrate) videos. The study in [11] proposes a throughput-buffer based algorithm relying on the buffer target, the playback buffer filling level, and the current bandwidth to determine the next segment's version. Work [12] proposes a segment-aware rate adaptation algorithm (SARA), in which segment size variation, the estimated throughput and the current buffer occupancy are considered to predict precisely the time required to download the next segment. In SARA, the MPD file not only contains the average bit rate information of the segment but also the size of the segment. SARA has expanded and changed the structure of the MPD file to accomplish this. Estimated throughput is based on the measured throughputs in the past and divides the buffer into different levels to determine the bitrate. This assures to download the best possible representation of the video while avoiding video buffer starvation. However, SARA shows low performance and poor QoE in the context of bandwidth sharing competition.

In a shared network environment, with so many Internet services available, bandwidth resource always gets risk of becoming insufficient to the increasing number of users, despite the maturity of cutting edge technologies nowadays. Therefore, it is always a huge concern for researchers to deal with limited resource allocation in any networking circumstances. Most of the existing researches tend to increase the Quality of Experience of clients while optimizing the transmission capability. Now exist three solution categories to deal with the problem where many clients compete through a bottleneck point. The first category is to solve the issue of sharing video streaming services at the client side as mentioned in [13]–[15]. In [13], an algorithm is proposed to improve efficiency, fairness, and stability named FESTIVE by providing schemes to ensure a trade-off of these factors. FESTIVE is the first known client-side adaptive bitrate algorithm that is designed for the multiple client context specifically. FESTIVE contains three main components: a scheduler that determines the time of starting to download a segment, an algorithm that estimates the throughput based on all previous downloaded segments and a mechanism to select the bitrate of the next segment. However, with FESTIVE, when the number of users grows sharply in a competition, the system's stability could be reduced remarkably since the estimation of the bandwidth is too high, perhaps. In [14], a rate adaptation algorithm for HAS service is proposed based on a "probe-and-adapt" principle at the client side so-called PANDA. "Probe-and-adapt" is a kind of the additive increase and multiplicative decrease principle (AIMD) that is used in TCP. However Probe-and-adapt operates at the application layer and in a longer time scale. PANDA is a HAS client rate adaptation algorithm designed to provide high stability and fast responsiveness to bandwidth variations in the context of multiple HAS clients sharing the same bottleneck links. For detecting the available bandwidth, the network is probed by PANDA which additively increments its sending rate at each adaptation step and multiplicatively decreases its rate if a congestion is detected, in order to adapt its video bitrate accordingly. During the estimation of the available bandwidth, the PANDA probing mechanism tries to increase its sending rate, instead of transmitting auxiliary piggybacking traffic. PANDA also allows HAS clients to respond to bandwidth fluctuation quickly and to achieve stability. But PANDA does not consider the resource allocation strategy in the context of multiple user negotiation. In [15], the streaming framework in VHS (VideoHomeShaper) is designed and implemented to evaluate performance in home's last access hop. In this study, the evaluation of QoE for competing clients takes into account some metrics such as Stability, utilization, and fairness. Overall, these researches mainly focus on adapting bit rate at the client side to improve the parameters that can affect user's QoE.

The second category focuses on the sever side. In this area, servers are normally considered as the only control point responsible for managing video transmission for all clients. The authors of [16] introduce a server-side adaptation scheme, which uses TCP to limit the video bit rate. A server-based traffic shaping solution is proposed in [17] to notably decrease such oscillations, at the expense of a small loss in bandwidth utilization. However, we can only be observe the actual bandwidth congestion problem at the network's side. Consequently, not too many clients might be able to handled by the server.

Finally, the category of in-network methods uses active bandwidth allocation to gain QoE fairness for multiple clients. Fair bandwidth allocation for players at a bottleneck link has been recently suggested by several work. Based on the SDN technology, [18] proposes traffic shaping techniques and do an analysis of ABR video streams when sharing resources through a congested link. In this method, all shared connections at the bottleneck point have the same throughput; and the total bandwidths assigned to requiring clients must not exceed the available bandwidth at the bottleneck. SDN-based traffic shaping in [18] can improve QoE for video streaming in terms of efficiency, fairness, and quality. The authors in [19] also consider the bandwidth sharing within a bottleneck link. However, in the network sharing model, the network sharing is calculated periodically, not in the per-request basis. In adaptive streaming, clients can require changing the bit rate at any time and the SDN controller must calculate the network sharing based on all of requests at the certain time. Authors in [20] proposed a system for increasing QoE of SVC-DASH clients based on SDN. The SDN architecture enables to select different streaming path to transfer different video layer flows between the server and the clients. But this paper does not consider the resource allocation strategy in the context of multiple user negotiation. In [21], the author proposed an OpenFlow-based QoE fairness framework (QFF). QFF manages the player statuses and allocates network resources dynamically to achieve the maximum user-level fairness for all clients that compete in a shared network. A proof-of-concept implementation is provided considering a small number of concurrent flows. However, only the fairness parameter of QoE is considered in this research. In [22], network-assisted strategies are considered to provide service differentiation to concurrent video flows such as the Bitrate Guidance approach (BG), the Bandwidth Reservation approach (BR), and the hybrid approach (BG+BR). A centralized Video Control Plane (VCP) is built to provide Video Quality Fairness (VQF) to concurrent video sessions that are sharing a common bottleneck. Work [23] argue that the network and clients shall collaborate to gain QoE fairness while video streams are encrypted. Placing cDVD in the greater context of QoE delivery methods is instructive. The resource-sharing clients only focus on fairness but not efficiency and stability. In [24], centralized and distributed architectures are proposed for collaboration between video service provider (VSP), network service provider

(NSP), and DASH clients to provide SDN-based NSP-managed or VSP-managed DASH services with quality-of-service (QoS) reserved network slices. The SDN controller computes the shortest-path route and bitrate for each client. This approach is to gain QoE fairness among heterogeneous DASH clients. However, again the resource-sharing clients only focus on fairness but not efficiency and stability. In [25], the author proposed an optimization model aiming at fairly sharing the bottleneck link while maximizing the client's perceived video quality. However, the work focuses only in bandwidth allocation and maximal link capacity of each client without considering how bit rate should be adapted to receive desired quality experience of the clients based on the real condition of the clients. Research [26] proposes a SDN-assisted adaptive streaming solution for tile-based contents using DASH to improve the user's quality of experience. The paper also summarizes the difference in immersive content streaming between SDN-based and traditional networks and introduce a SDN-based framework to support tile-based immersive content streaming. However, this paper does not consider the bandwidth allocation and bitrate adaptation problem. In [27] and [28], the author proposed a SDN based management architecture and dynamic resource allocation for HAS systems aiming to alleviate scalability and improve the QoE of each client. This architecture is in charge of allocating the network resources dynamically for each client depending on the client's QoE expectation. The same authors in [29] propose a SDN-based QoE-aware bandwidth broker for HTTP adaptive streams in an hybrid fiber coax (HFC) network, so-called Bandwidth Management Solution - BMS that dynamically chooses the respective bandwidth allocation decisions and the optimal joint representation in oder to meet the per-session and per-group QoE objectives. In [30], there is a network-driven video streaming architecture is built to design robust mechanisms for multiple players so as to achieve end-user QoE fairness. The SDN controller observes the environment state, computing the reward, and updating Q-value and taking the bandwidth allocation action by forwarding flow tables. However, the author considered how bandwidth should be shared to achieve fairness while QoE is taken into account from the perspective of receiving bit rate that is suitable to the link bandwidth. The work does not consider how QoE of each client can be acquired based on the client's real need in terms of bit rate at a certain moment QoE. In [31], a SDN-based model is built over the DASH protocol to improve the QoE whilst taking into account the variety of video parameters, devices and the network requirements. The proposed scheme comprises two phases: Machine Learning-based estimation phase of available resources, and adaptation and selection phase based on the results of the first phase. The goal of the approach is to have the best quality perceived video. The authors in [32] proposed a SDN-assisted solution to use the network management plan to reinforce policies for improving QoE of clients. However, again, this work does not consider the case of lack of network resource so multiple users need to negotiate for the network share in trade off with QoE. In [33], the authors present SAND/3 that is an SDN-Assisted QoE control for DASH over HTTP/3 for improving the QoE at the client side based on criteria such as buffer size, display size, type of device and subscription plan, and available memory. From the SDN network perspective, SDN tries to find a route for each requested connection from each client based on available resource. But this paper does not consider the resource allocation strategy in the context of multiple user negotiation.

Generally, SDN-based dynamic bandwidth allocation shows a good performance in improving video quality. However, the interaction between the network control and the client-side control in the context of multiple users sharing one bottleneck link has not been well investigated. Therefore, deploying solutions in a large-scale system is still a pending question.

3 Bitrate Adaptation Algorithm

Our proposed framework to the bottleneck problem is a integration of the bitrate adaptation algorithm at the client's side and the centralized SDN-based resource allocation, i.e. a cooperation between the network and video streaming applications. In the streaming network, a flow management mechanism called MSMA (Media Streaming Multiple Access) is proposed (discussed in Section 4.2).

In fact recently, there have been two widely-adopted types of algorithms at a MPEG-DASH Client, which belong to either the throughput-based group or buffer-based group. Therefore, we consider the context of multiple video-on-demand players, provided that the throughput-based adaptation algorithm called Aggressive [6] is deployed at each client.

Besides, we have also proposed a buffer based bitrate adaptation algorithm to determine a video quality level for the next segment. Compared to the throughput-based algorithms, buffer-based algorithms provide smoother bitrate curves in video-on-demand streaming. In the next subsections, the description of the throughput estimation and the proposed adaptation algorithm are presented.

Table 1: Symbols use in this paper

Symbol	Description
B_i	Current buffer level at segment i
B_e^{i+1}	Estimated buffer for segment $i+1$
B_{max}	Maximum buffer size
N	Number of clients
I_{i+1}	The chosen version for the next segment
T_i	The throughput for the current segment
T_{i+1}^e	Estimated throughput for the next segment
r_i	Bitrate of client i
$r_i^{'}$	Bitrate of r_i after the SDN controller is recalculated
$R_{0,\ldots,V-1}$	Bitrate of version 0, 1, ..., V-1
R_{max}	Bitrate corresponds to the highest level of quality
RTT	Round Trip Time
SD	Segment Duration
V	Total number of video versions

3.1 Throughput Estimation

In simple terms, the bitrate adaptation mechanism adapts its requested bitrate based on the estimated throughput. Video segments are delivered from the server to the clients in sequence of consecutive HTTP request-response transactions. In general, throughput is obtained by dividing the data size by the delivery time. Knowing the actual amount of data delivered by a request-response transaction is

important. The segment throughput T_i^e for each delivered segment i is computed using the request-response duration (i.e. from the instant of sending the request to the instant of receiving the last byte of the response). Table 1 lists the symbols used in this paper.

A video version is chosen based on Equation (1). Bitrate R_j is selected as the maximum value not exceeding the product of the current estimated throughput T_{i+1}^e and safety margin μ (μ varies from 0.0 to 0.5). μ is defined to solve the instability of the Internet connection and delay caused by other streaming processes such as bitrate adaptation calculation and file decoding.

$$R_j \leq (1-\mu) \times T_{i+1}^e \tag{1}$$

Next, we propose to estimate the next throughput based on two previous measured throughputs at the client side (as described in Equation (2)). In the Equation, γ is a dynamic constant range from [0, 1] that allows a client to adapt well with highly bandwidth fluctuations, especially with highly variable bandwidth since it is the main cause of video freezing. γ is usually set to be 0.5. When γ = 1.0, the estimated result will be exactly the same as the one of the aforementioned aggressive method.

$$T_{i+1}^e = \gamma \times T_i + (1-\gamma) \times T_{i-1} \tag{2}$$

Algorithm 1: Proposed Bitrate Adaptation Algorithm

```
Input: T_i, R_n, B_i, RTT, SD, I_i, Δ_B ;
Output: I_{i+1} ;
T^e_{i+1} ← γ × T_i + (1 − γ) × T_{i−1};
I_{i+1} = 0;
for j ← V − 1 to 0 do
    B^e_{i+1} ← B_i + SD − RTT − SD×R_i / T^e_{i+1};
    if j > I_i and B^e_{i+1} ≥ B_i + Δ_B then
        I_{i+1} = j;
        Break;
    end
    if j ≤ I_i and B^e_{i+1} ≥ B_i − Δ_B then
        New request to controller;
        I_{i+1} = j;
        Break;
    end
end
```

3.2 Proposed adaptation algorithm

In this paper, we design a new quality-selecting scheme based on the estimate buffer level so as to avoid video freezing. According to our previous work [34], we use Equation (3) to estimate the next buffer level, provided the client selects version j for segment i where SD is segment duration. RTT is the duration defined from the instant a client sends its request to the instant it receives a response and calculated by timestamps of the westward and eastward packets captured at the network card of the SDN controller. The westward packet is the request packet sent from a client to the controller and the eastward packet is the packet carrying information of calculated BW from the controller to that client.

The main goal of our solution is to choose a suitable quality level that can keep the buffer greater than the current buffer B_i to prevent stalling events. Our method details are shown in Algorithm 1. This algorithm is based on an estimated buffer to determine a new video version for the next segment.

In fact, increasing the quality version has a lower priority than keeping the quality version stable, and the quality should be reduced as little as possible through threshold Δ_B. The larger Δ_B , the less varied the video quality version, but the worse the connection adaptation. On the contrary, the smaller Δ_B, the better the connection adaptation, but more the stability is lacked. This parameter usually ranges from 0.25s to 2s. In this paper, we choose $\Delta_B = 1s$, since this figure is appropriate to trade off the issues above. When there is network fluctuation, a client shall reduce its quality version accordingly. In that case, the client will request the SDN controller to recalculate the bitrate, and sending this bitrate information to all clients that are having ongoing connections so that the quality version of all clients are kept as stably as possible.

$$B_{i+1}^e = B_i + SD - RTT - \frac{SD \times R_j}{T_{i+1}^e} \tag{3}$$

4 Proposed SDN-based Solution

4.1 Bandwidth Allocation Policy

Provided that N clients go through the same bottleneck with bandwidth capacity C. $BW_{i,t}$ is the amount of bandwidth allocated for client i at time t. The bandwidth allocation policy is to map the bitrate requested by each client to the actual bandwidth that could be granted to that client:

$$BW_{i,t} = f(r_{i,t}) \tag{4}$$

where $r_{i,t}$ is the bitrate that client i requests at time t. r_i is the bit rate the clients desire to have at the beginning. But BW_i is the actual bit rate that the clients will adapt to after being reconsidered by the SDN controller. In case the bandwidth resource is sufficient, all clients will have BW_i which are the same to r_i. In case network resource lacks, one or some r_i may be different from BW_i.

If the network resource (bottleneck link capacity) is sufficient for all requested bandwidths of each client, the clients will be allocated the same bandwidth as they request, i.e.:

$$BW_{i,t} = \frac{C}{N}, \forall i \in N, \forall t \tag{5}$$

However, in reality, HTTP networks are not ideal because clients frequently require bitrates that can be higher than the then-actually allocated bandwidth. As described in [35], [36], the bandwidth allocation policy for this non-ideal HTTP network condition is as follows:

. if $r_i = r_j$ then $f_i(r) = f_j(r)$;
. if $r_i > r_j$ then $f_i(r) > f_j(r)$;
. $\frac{\partial f_i(r)}{\partial r_i} > 0$, $\frac{\partial f_i(r)}{\partial r_j} < 0$, where $i \neq j$;
. $lim_{r_i \to \infty} f_i(r) < C, lim_{r_i \to 0} f_i(r) > 0$;
. $f(.)$ is a function of r and value of f is independent of the client order.

In fact, requests do not come to the edge switch at the same time, therefore naturally, a solution to each request is thought appropriate.

When a request comes, the system calcualte to determine if it should accept the request, so as to ensure no congestion occured at the bottleneck point, causing the drop of quality of experience of all N clients.

In DASH, content provider encode videos at different quantity levels of different bit rates and then brake the videos into segments. That is the base to determine the minimum bandwidth needed by each client in order to obtain an uninterrupted video streaming. That minimum bandwidth is equal to the average bitrate of the video with the lowest quality level. This value is denoted as R_0.

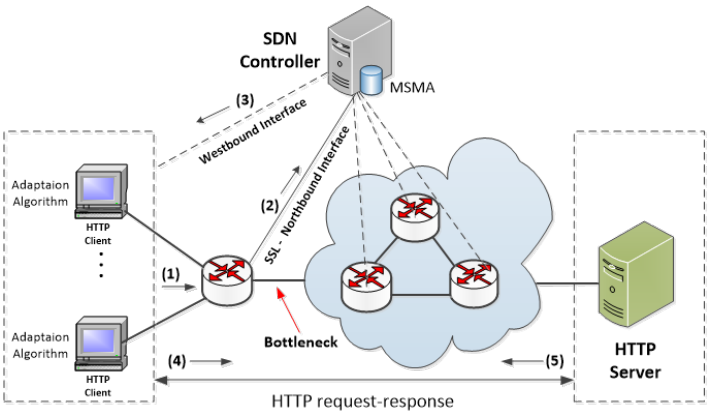


Figure 1: Proposed architecture and operation for the bottleneck problem

4.2 Proposed fair bandwidth allocation solution (MSMA-0)

The proposed architecture is illustrated in Figure 1. As it can be seen, the Aggressive adaptation algorithm and the MSMA module are implemented in each client and the SDN controller respectively. The whole solution is called MSMA-0. In our proposed architecture, the edge switch (OFS - OpenFlow Switch) communicates to the SDN controller through 2 interfaces: a standard Northbound interface of Openflow that connects OFS and the SDN controller via TSL protocol; and a legacy network interface that connect the OFS and controller as 2 legacy network devices called Westbound interface. The operation of the proposed architecture is as follows:

• Phase 1: Negotiation of appropriate bit rates requested by each client

(1) Each client calculates its desired bit rates according to the bit rate adaptation algorithm implemented at client.

(2) Client sends that streaming request (i.e. desired bit rate) to the edge OFS switch. Then, the OFS encodes the request information in a *Packet − In* message and sending it to the SDN controller via the Northbound interface (which is defined in OpenFlow protocol).

(3) The Controller, via the Northbound interface with the OFS edge switch, having already knowledge of link capacity of the network cards of the switch can calculate new BW values, then sending these new information to the clients via Westbound interface (as a legacy network device).

• Phase 2: Requesting the finally negotiated bit rate to the application server

After receiving information on BW from the SDN controller, the clients operate as follows:

(4) N ongoing clients send new requested BW_i to the application server via an open connection.The new client $(N+1)$ sends a request to connect to the application server, the SDN controller will order the edge OFS to install a new flow entry to establish the connection between Client N+1 and the application server.

(5) Upon receiving the request on bit rate (i.e BW_{N+1}), the app server will push a video quality version segments accordingly to all streaming clients.

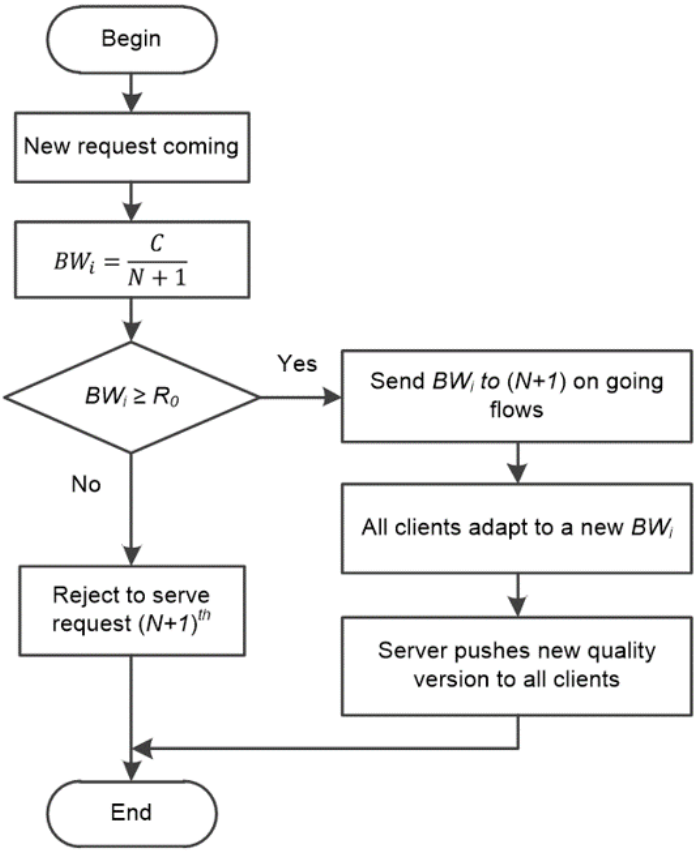


Figure 2: MSMA Flowchart

As the implication of its name, a network bottleneck causes a slow connection speed, limiting the experience of media streaming. Solving the aforementioned issue is the goal of our proposed MSMA algorithm. Figure 2 depicts the flowchart of MSMA deployment with the working principle described as follows:

1) Given that there are N ongoing clients requesting new quality versions for each segment, the average bandwidth BW_i distributed to $(N + 1)$ streams is recalculated by the MSMA module, according to the Equation (6) below:

$$BW_i = \frac{C}{N+1} \tag{6}$$

2) BW_i is then compared with the minimum acceptable bitrate R_0 by the SDN controller that then decides what action to take. Either one of these two actions depending on the result of the comparison will be carried out by the SDN controller.

3) A flow for the incoming stream is added by the SDN controller and BW_i is sent to all on going streams by the controller if $BW_i \geq R_0$ i.e. the bandwidth allocated to each client is higher than the one of the lowest quality version and is sufficient to stream that video version.

4) The incoming request is rejected to serve by the controller if $BW_i < R_0$ i.e. the bandwidth allocated to each client is insufficient to stream even the lowest quality version of the video.

Then after all, all streaming clients shall adapt to BW_i decided by the SDN controller. The clients then request the DASH server for the decided bitrate. The adaptation logic is described in the next section. TheDASH server pushes a new quality version to all streaming clients.

4.3 Proposed adaptive bandwidth allocation solution (MSMA-1)

In this work, we propose a new method (called MSMA-1) that combines a bit rate adaptation algorithm at the client side (i.e. Algorithm 1), with modul MSMA deployed in the SDN controller as Algorithm 2. The goal of MSMA-1 algorithm is trying to improve the Quality of Experience of each client on the streaming service by providing the bit rate that a client requires while trying to increase the number of clients that can access to the system within a limited resource.

At the network side, we allocate bandwidth dynamically (i.e. Algorithm 2) for all clients in order to maximize the number of clients that can be given an access to the network and minimize the number of clients who are requested to change their bitrate expectations (bitrate to satisfy the QoE criterion). Provided the system has N clients that have HTTP streaming over a bottleneck point. The capacity of the output link of the bottleneck point is assumed C (Mbps).

At the client side, each client deploys the bitrate adaptation algorithms proposed in Algorithm 1 to find out what is their bitrate expectation in order to achieve desired QoE. Actually, each bitrate is corresponding to a quality version defined by DASH and stored at the server.

In case, the sum of all requested bitrates r_i of N clients (i = 1 to N) is smaller or equal to capacity C, the SDN controller will request the DASH server to push the quality version desired by each client.

In case, the sum of all bitrates r_i (i = 1 to N) is larger than C, i.e:

$$\sum_{i=1}^{N} r_i > C \tag{7}$$

then the SDN controller recalculates bandwidth allocated for each client. The SDN controller operates in the following criteria:

1. The SDN controller tries to keep the bitrates r_i required by the clients as stably as possible with the goal of meeting these clients' QoE requirement.

2. The SDN controller selects a client holding the highest rank, and taking a fraction of its required bandwidth to give to another client. In this manner, the network take a fraction of the network resource of this client but in return it will give an access for an additional client (i.e. client (N + 1)). So, the goal of serving as many clients as possible is fulfilled.

In fact, as in the solution described in section 4.1, the link capacity C is divided equally to all (N + 1) clients, therefore, all (N + 1) clients suffer an undesirable QoE degradation since distributed resource may be lower than their requirements to achieve each client's satisfaction. Therefore, in this new proposed solution, the SDN controller, firstly, ranks all clients in a certain manner. Then the controller will set a policy to allocate resource to all clients by rank. Technically, this ranking function (or objective function) can be defined as a function of multiple objectives. Within the scope of this paper, we define the rank of client by Equation (8). Looking at the Equation (8) shows that the client that have a higher buffer level and a bitrate is likely to have the highest rank function. Therefore, this client will bring in some of the resources available for the new service. In addition, the choice of which client to share the resource will depend on for the two parameters which are bit rate r_i and buffer level B_i such that rank function is highest.

$$Rank(i) = f(B, r) = \alpha \times \frac{B_i}{B_{max}} + \beta \times \frac{r_i}{R_{max}} \tag{8}$$

where α, β: non-negative weighting factors. And r_i is bitrate required by the ongoing client, B_i is current buffer level of client.

The set of $Rank(i)$ is arranged in the decreasing order according to the rank function value.

Details of the proposed solution are described in Algorithm 2. The details of this algorithm are explained as follows:

Firstly, the requested bitrate of a client who has the highest rank is recalculated to share the resource to other client. Provided the client with the highest rank is Client N. Then the bandwidths of client j and client N are reallocated as Equation (9) and (10) accordingly:

$$r'_j = \frac{r_j}{r_j + r_N} \times \Delta C \tag{9}$$

$$r'_N = \frac{r_N}{r_j + r_N} \times \Delta C \tag{10}$$

where r_N is the requested bitrate of client N who has the highest rank value and r_j is the requested bitrate of client who requests a new bitrate. r'_j and r'_N are the varied bitrates of r_j and r_N after the SDN controller recalculate bandwidth distribution. ΔC is the remaining bandwidth after reserving the bandwidth portion for all left clients that have not been shared it's resource. (i.e. all clients that retain their bitrates r_i). ΔC is defined by Equation (11).

$$\Delta C = C - \sum_{i=m}^{N} r_i \tag{11}$$

where m is the number of clients who are sacrificed.

In case if the calculated bitrate of client j is smaller than the smallest acceptable quality version ($r_j < R_0$), continue to take the second highest bitrate requested by an ongoing client and redivide

proportionally for all 3 clients. The process continues until all N clients satisfy a bitrate $r_i \geq R_0$ ($i = 1 \div N$).

Algorithm 2: . MSMA-1 algorithm

Input: r, C, Δ_C
Output: r
for $i \leftarrow$ *to* N **do**
  $r'_{N+1} = \frac{r_{N+1}}{\sum_{k=1}^{i} r_k + r_{N+1}} \times \Delta_C$;
  for $j \leftarrow 1$ *to* i **do**
    $r'_j = \frac{r_j}{\sum_{k=1}^{i} r_k + r_{N+1}} \times \Delta_C$;
  end
  if $min(r') \geq R_0$ **then**
    $r = r'$;
    Break ;
  end
  else if $i = N$ *and* $min(r') < R_0$ **then**
    Reject client $(N+1)^{th}$;
  end
end

5 Performance Evaluation

5.1 Experiment Setup

Bitmovin, a multimedia technology company providing services, which transcode digital video and audio to streaming formats using cloud computing, and streaming media players has developed and maintained a library named libdash [37]. In this paper, QtSamplePlayer is used as the media player of the streaming architecture. QtSamplePlayer is a Qt-based player recommended for the libdash library by Bitmovin. Videos pushed from the HTTP server to the clients are 2 types of videos: constant bit rate CBR and variable bit rate VBR. Besides, traffic from the clients to other devices (i.e. OFS switch and the SDN controller) is only control information on bit rate requests to adapt to required quality version.

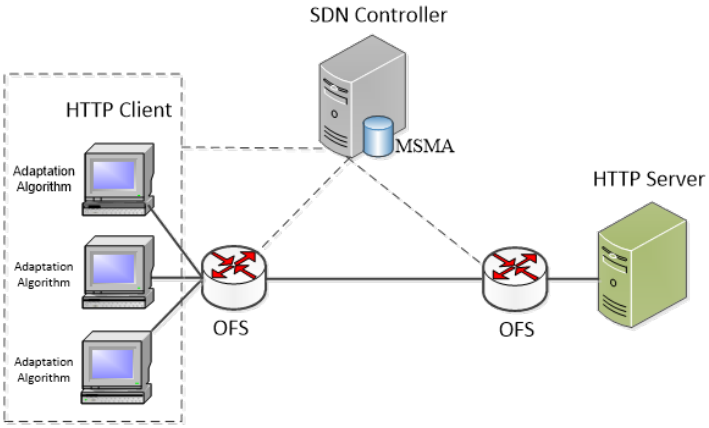


Figure 3: Network topology used for experiment.

Figure 3 shows the network setup for performance evaluation. In our experiment, five machines are used to conduct the experiment: the media server, one machine for emulation of the streaming network including the SDN switches and the SDN controller - Floodlight [38], and 3 video streaming clients. One clients is implemented on a physical PC and all other machines are VMWare machines. Due to limited testbed resource, we just make a proof of concept with 3 clients. In the future, we will improve the testbed to scale up the experimental scenario. The network consists of two OpenFlow switches and a bottleneck link between them. An OFS is connected to the three client machines and the other OFS is connected to the media server ,in which MPD (Media Presentation Description) and video files are stored. The MPD file is an XML (Extensible Markup Language) file that represents the different quality levels of the media content and the individual segments of each quality level with Uniform Resource Locators (URLs). For this evaluation, Mininet [39] is used to create a realistic virtual network.

There are many different types of videos that can be used in evaluation simulations, but all of them focus on two types of video: variable bit variable (VBR) and constant bitrate video (CBR). In this paper, the experiment videos are two short movies named Elephants Dream [40] and Destiny [41] to represent the 2 aforementioned video types. In the Elephants Dream movie, the video is divided into 327 segments with the total duration of 654 seconds (i.e. 2 seconds per segment). This video is encoded in the VBR mode of 12 different versions corresponding to 12 different average bitrates. The version index, Quantization Parameter (QP), and the average bitrate of each version are listed in Table 2. The Destiny short file is a constant bitrate video (CBR) and is divided into 340 2-seconds-long segments of 16 quality levels of different bitrate (50, 100, 250, 400, 600, 800, 1000, 1200, 1600, 2000, 3000, 4000, 6000, 8000, 12000, 16000 Kbps).

Table 2: Version information of the VBR video

Version	QP	Average Bitrate (Kbps)
0	46	354
1	43	472
2	40	638
3	37	882
4	34	1234
5	31	1779
6	28	2588
7	25	3823
8	22	5613
9	19	8028
10	16	11156
11	13	15227

5.2 Experimental Scenarios

Six experiments are deployed to evaluate our proposed system:

• Experiment 1: The link capacity of the bottleneck is 10 Mbps, 7 Mbps and 4 Mbps. The controller uses the aforementioned method MSMA-0. Client 1 starts streaming 110 seconds (55 segments). After the first 110 seconds of Client 1, Client 2 starts streaming in 110 seconds and in the same way with Client 3.

• Experiment 2: The link capacity of the bottleneck is set at 10 Mbps, 7 Mbps and 4 Mbps. The controller uses the aforementioned MSMA-1 algorithm. Each of Client 1, Client 2 and Client 3 starts

streaming in 110 seconds, one after another.

• Experiment 3: The link capacity of the bottleneck is set at 10 Mbps, 7 Mbps and 4 Mbps. The Aggressive, SARA, BBA, PANDA algorithms are used. MSMA is not available. Each of Client 1, Client 2, and Client 3 starts streaming in 110 seconds, one after another.

• Experiment 4: The link capacity of the bottleneck is set at 10 Mbps, 7 Mbps and 4 Mbps. The SARA, BBA algorithms are used. The MSMA module is activated. Each of Client 1, Client 2, and Client 3 starts streaming in 110 seconds, one after another. This experiment simulates methods "SDN-based SARA" and "SDN-based BBA".

• Experiment 5: The link capacity of the bottleneck is set at 500 Kbps. The controller uses the MSMA modul. The clients request for the VBR video an independent time.

• Experiment 6: The link capacity of the bottleneck is 500 Kbps. The controller does not use the MSMA modul. The clients request for the VBR video at independent time.

We use the Aggressive method as the bitrate adaptation algorithm [6] with a safety margin $\mu = 0.2$.

We perform each of the five experiments in three times and record the average results for later comparison. The purpose of Experiment 1, Experiment 2, and Experiment 3 is to determine stability, efficiency, and fairness; The purpose of Experiment 4 and Experiment 5 is for checking the possibility of rejecting new requests when a congestion is possible.

With Experiment 5, since $R_0 = 354Kbps$, only the connection request of the first client will be accepted by the controller; and video with the minimum average bitrate will be streamed by the client. The evaluation results illustrates that the number of video freezes is 5 in case the segments contain a high bitrate compared to the representative bitrate (i.e. average bitrate). Our proposed system does not accept to provide services to requests from client 2 and client 3. Experiment 6 shows the scenario of a traditional network where the controller does not operate with the MSMA modul. In that scenario, the 3 requests from 3 clients are accepted in turn.

So, because the available bandwidth is not sufficient for all 3 clients sharing bandwidth and asking an acceptable bandwidth, all 3 clients are frozen immediately after being connected and also experience buffering most of the time during the whole course of their streaming section. This proves that at the circumstance of poor bandwidth, MSMA-0 is still able to serve the modest number of players, rather than the other methods where no user will be served.

5.3 Simulation Results and Discussion

To evaluate the performance of the proposed solution, we use three metrics of fairness, efficiency and stability as presented in [42] and [13] and a QoE metric. In the following subsections, these metrics will be described in detail.

5.3.1 Fairness, efficiency and stability metrics

In this subsection, the 3 metrics of stability, fairness, and efficiency are presented. In particular, the fairness score is measured by Equation (12)

$$Fairness = \frac{|\sum_{i=1}^{N} r_{i,t}|^2}{N\sum_{i=1}^{N} {r_{i,t}}^2} \tag{12}$$

where $r_{i,t}$ is the video bitrate allocated to client i at time t. The fairness score should be in the range of [0 to 1]. As the matter of fact, the closer to 1 the fairness score is, the fairer the system is.

The efficiency score of the bandwidth distribution is computed using Equation (13), where W is the available bandwidth of the shared connection. The closer to 1 the Efficiency score is, the better the system efficiency is.

$$Efficiency = \frac{|\sum_i r_{i,t}|}{W} \tag{13}$$

Equation (14) defines the stability score for client i at time t. This score is calculated based on the absolute amplitudes of the quality fluctuations (i.e. bitrate switches).

$$Stability = 1 - \frac{\sum_{d=0}^{k-1} |r_{i,t-d} - r_{i,t-d-1}| \times \omega(d)}{\sum_{d=1}^{k} r_{i,t-d} \times \omega(d)} \tag{14}$$

where: $\omega(d) = k - d$: weight function; and k: 20 seconds (10 segments) in this experiment. The closer to 1 the Stability score is, the more stable the system will be.

5.3.2 QoE Metric of Video Streaming

In this section, we first present parameters that have influences on the QoE of video streaming. Then, a QoE metric taking into account the parameters is provided. According to [43], the most important parameters are as follows:

Initial Delay (D_t): Initial Delay is computed as the duration from the point the first segment of a video is requested until the point the video is played out. For an effective video streaming system, the buffer time required for this parameter must be much smaller than the maximum client-side buffer level. According to [27], the initial delay should be less than 6 seconds.

$$D_t \ll B_{max} \tag{15}$$

Average Version Quality (AVQ): In the DASH system, the quality of the segments received at the client is variable. This parameter is defined as the average version of all the downloaded segments. Assume a video stream consists of K segments and each segment is encoded with V bit rate corresponding to V quality level, where v represents a specific quality level .

$$AVQ = f(i, v) = \frac{1}{K}\sum_{i=1}^{K} q_{i,v} \tag{16}$$

where $q_{i,v}$ is the quality value of the i^{th} segment at the v quality session level, for example $q_{i,v_{max}}$ is the quality value of the i^{th} segment with the highest quality level.

Number of Version Switch-downs (NVS): is the number of times that a downloaded segment has a lower bitrate than the previous segment. We can use this paremeter along with Average Video Quality to provide quantitative inferences of QoE. If video flows have similar Average Video Quality, a flow with the lower number

of Version Switch-downs will be perceived,by users, to be the better one .

$$NVS = \frac{1}{K-1}\sum_{i=1}^{k-1}|q_{i+1,v_{i+1}} - q_{i,v_i}| \quad (17)$$

Number of Video Stalling (S_t): Video streaming will be stalled if the buffer at the client side is empty. In other words, if the playback buffer size is reduced to the minimal buffer level but has not downloaded the current segments, then a video stalling will occur. Number of Video Stalling can be calculated as the total number of stalls in a video playback time period.

$$S_t = \frac{1}{K}\sum_{i=1}^{K} S_i, \forall 0 \le B_i \le B_{min} \quad (18)$$

To evaluate the performance of our bitrate adaptation scheme, we use an QoE metric based on the above parameters and also from the reference model provided by A. Bentaleb et al. [27] as follows:

$$QoE_i = \alpha_1 \times AVQ - \alpha_2 \times NVS - \alpha_3 \times S_t - \alpha_4 \times D_t \quad (19)$$

Both S_t and D_t are in seconds. ($\sum_{n=1}^{4}\alpha_n = 1$) are non-negative tunable values and selected based on [13] to test in our experiments. We convert the QoE score to a normalized QoE (N-QoE) accordingly to Table 3 to represent a user's satisfaction MOS (mean opinion score) from 1 to 5.

Table 3: Normalized QoE

QoE_i	N-QoE	Quality
$QoE_i \le 1$	1	Bad
$1 < QoE_i \le 2$	2	Below Average
$2 < QoE_i \le 3$	3	Average
$3 < QoE_i \le 4$	4	Good
$QoE_i > 4$	5	Excellent

5.3.3 Performance Evaluation

In this section, we compare our proposals (MSMA-0 and MSMA-1) with the existing methods such as PANDA, BBA, SARA, Aggressive (called AGG), SDN-based SARA and SDN-based BBA. In which, the SDN-based BBA and SDN-based SARA methods are the BBA and SARA methods that are combined with the fair bandwidth allocation of the MSMA module. To understand more about the existing solutions, let us to highlight some main points of the existing methods:

1. AGG only based on the throughput of the previous segment to estimate the throughput for the next segment. This method has advantages in fast adaptability to the strongly fluctuating bandwidth conditions, but has disadvantages in the stable bandwidth conditions compared to other methods.

2. For SARA method, the bitrate adaptation algorithm of SARA is both throughput-based and buffer based to determine which quality level would be chosen. SARA has shown a number of advantages over the AGG algorithm as it is able to maintain a higher quality level for longer time under stable bandwidth conditions because of having the segment capacity information, thus it could predict download time.

3. BBA is the buffer-based method only to select a version quality without estimating throughput.

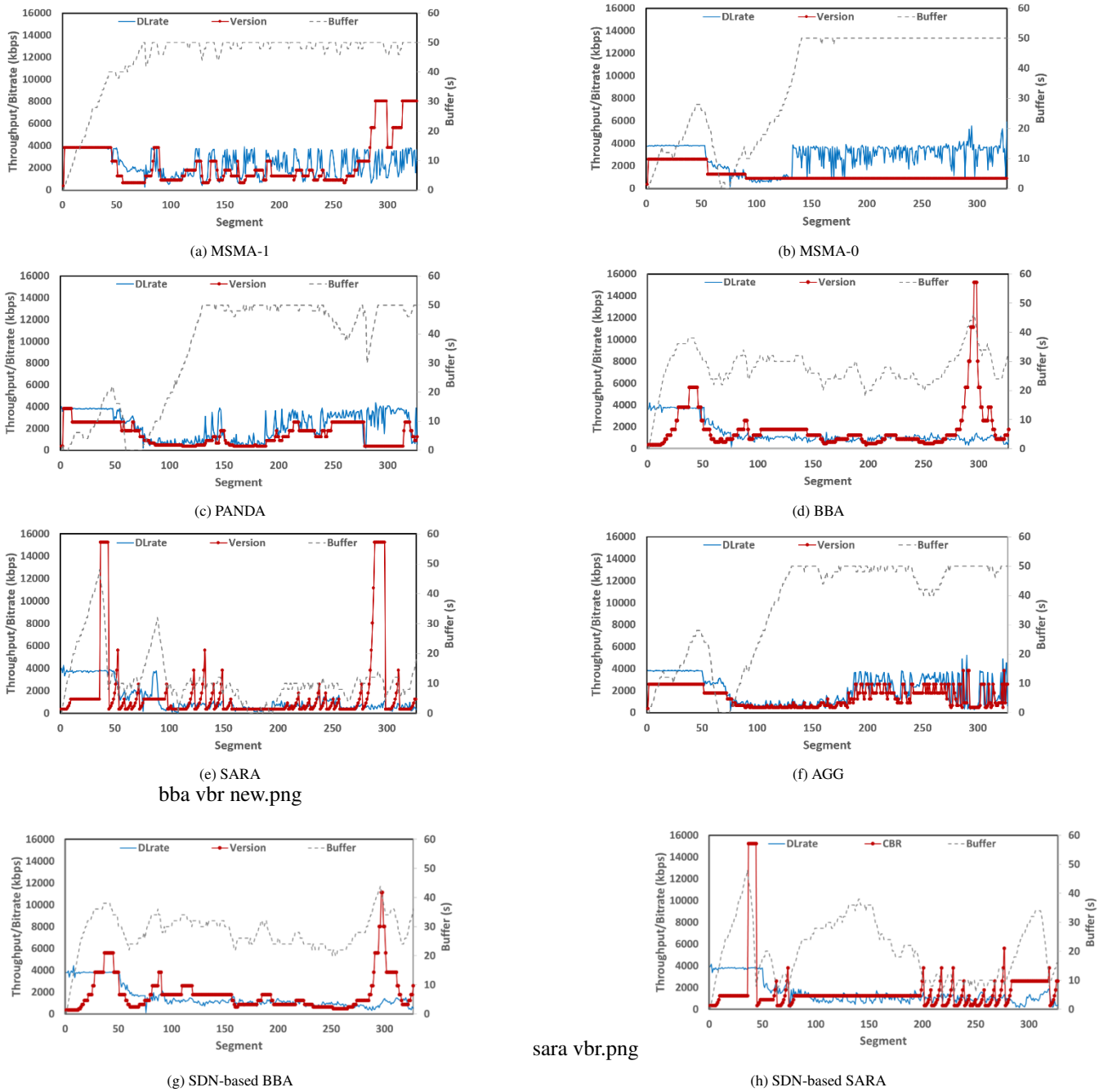
4. PANDA uses "probe-and-adapt" to build the rate adaptation algorithm and fair-share bandwidth for players. This solution achieves good stability but the averrage video quality is not high. On the other hand, PANDA's resource allocation is in a distributed rather than centralized manner like our proposed one so fairness is not high.

The download rate (DLrate), buffer level, and selected versions of all the methods are shown in Figure 4 (with VBR video) and Figure 5 (with CBR video). In experiments for the 10 Mbps, 7 Mbps, and 4 Mbps cases, the results show similar performances in terms of version and download rate. Therefore we only count in the 4 Mbps case as the representative for this type of evaluation. As Figure 4 demonstrate, in case of using the SDN controller with MSMA-0 and MSMA-1, video versions of the client are stable during the whole course without requesting any new connection.

Especially, MSMA-0 is very stable, however this method has average video bitrate much lower than of MSMA-1. The other methods have relatively the same average video bitrate. For MSMA-1 and BBA, due to buffer preestimation, the buffer does not go to zero. This means that the video will not be stalled. Other methods have almost video stalling for some times. For MSMA-0, the buffer is dropped low in segments of 68-78. In PANDA, the video is interrupted from segment 60 to segment 78. In SARA, the average buffer is very low, and there are 2 video interruptions, one at around segment 110 and one from segment 174 to 198. In the Aggressive method, video quality is fluctuated continuously which can be annoying to viewers; and video is also stalled from segment 66 to segment 78. With the support in bandwidth sharing of the SDN controller, the SDN-based BBA and SDN-based SARA methods have better performance in comparison with the BBA and SARA, as described in Figure 4g and Figure 4h. As we can see, the SDN-based SARA method has more stable quality levels compared to SARA in the duration from segment 80 to 200. This is because the fair bandwidth allocation of the MSMA module incorporates the segment-aware rate adaptive algorithm of SARA when multiple clients access the system. With good buffer-based bitrate adaptation, SDN-based BBA as well as BBA can avoid video freezing.

As Figure 5 illustrates, the status of clients in the case of the CBR video is more stable than the one of the VBR case, since the segment sizes of the CBR video are quite similar while segment sizes of the VBR video are varied and unpredictable. Average version of SDN-based BBA is increased by about 3.2% compared to BBA. SDN-based SARA had similar average quality levels to SARA, but the variation in quality levels was significantly reduced. This is quite an important parameter in QoE evaluation.

For the performance evaluation, the time when all clients begin to stream at the same time is selected. It take 20 seconds (10 segments) to to calculate results of the performance. Figure 6 compares the performance of our proposed solutions (MSMA-0 and MSMA-1) with PANDA, BBA, SARA and AGG using the three



(a) MSMA-1

(b) MSMA-0

(c) PANDA

(d) BBA

(e) SARA

(f) AGG

bba vbr new.png

(g) SDN-based BBA

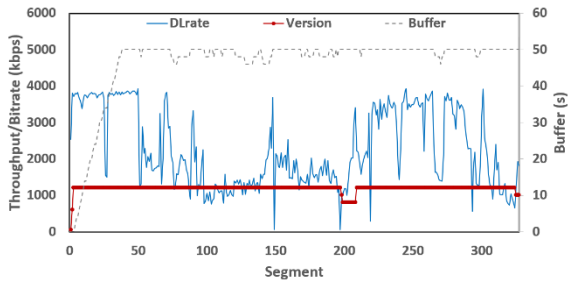
sara vbr.png

(h) SDN-based SARA

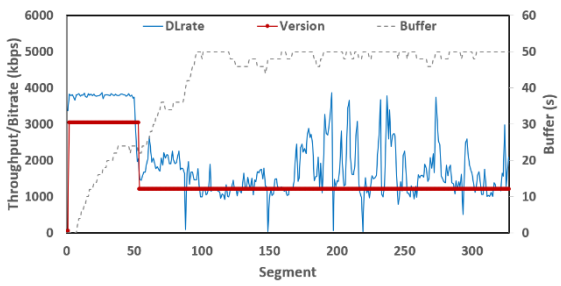
Figure 4: Download speed, buffer and version quality for all methods at a bottleneck link of 4 Mbps with VBR video

evaluation parameters: *Fairness*, *Efficiency* and *Stability*. For a more accurate evaluation, in this Figure 6, we get the average for all the three cases with the link capacities of 4 Mbps, 7 Mbps, and 10 Mbps. In real time, we set the mean value over three runs for 3 players. The results show that the fairness, efficiency, and stability obtained using the existing methods are much lower than those obtained using our proposed solution. However, due to the segment aware rate adaptation mechanism, in Figure 6a, the fairness between segments 4 and 7, SARA performs better than all of the others. This scheme uses all of the segment sizes, the estimated path bandwidth and current buffer occupancy to predict the amount of time needed to download the next segment. Therefore this method guarantees the fairness for all players. Using the existing streaming solutions as PANDA, BBA, SARA and AGG, the fairness in these cases varies from 0.634 to 0.860 and has an average of 0.699, 0.723, 0.77 and 0.687, respectively. The experiment also shows that the proposed architecture has improved the fairness score significantly, with scores ranging from 0.791 to 0.863 that results in an average of 0.816 which is very high. The efficiency of bandwidth allocation in the current solutions also has a considerable deterioration with values from 0.488 to 0.761 (i.e. average of 0.654, 0.655, 0.602 and 0.604, respectively). Whilst the efficiency of MSMA-1 varies from 0.716 to 0.970 (i.e average of 0.830).

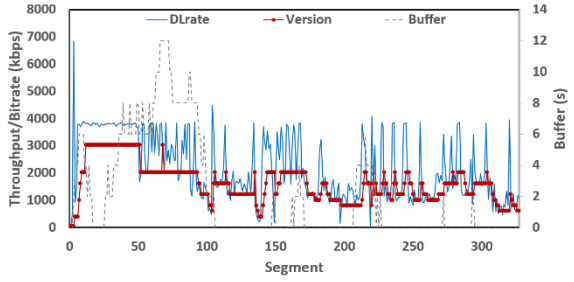
Figure6.c illustrates the comparison of the average stability of



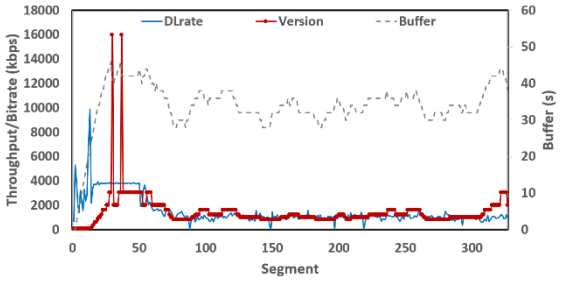
(a) MSMA-1



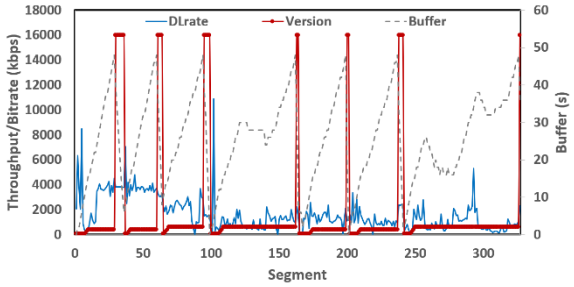
(b) MSMA-0



(c) PANDA

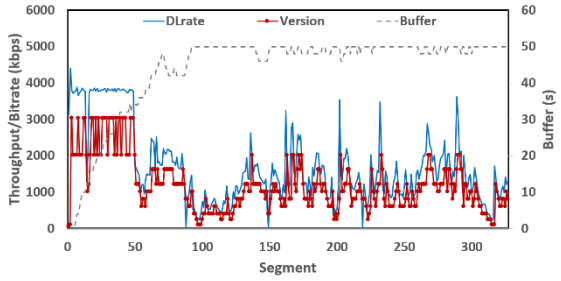


(d) BBA



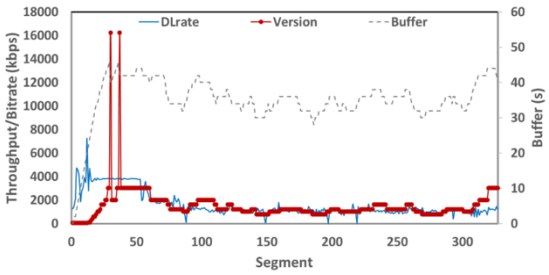
(e) SARA

bba cbr new.png

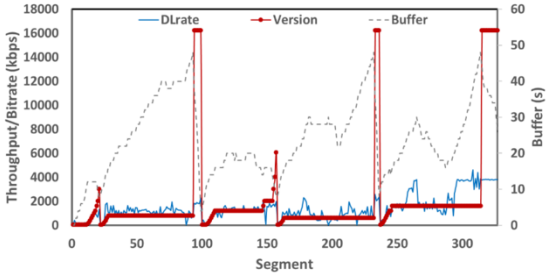


(f) AGG

sara cbr new.png



(g) SDN-based BBA



(h) SDN-based SARA

Figure 5: Download speed, buffer and version quality for all methods at a bottleneck link of 4 Mbps with CBR video

the 3 clients in all methods. It can be seen that the stability of our proposal is better than other schemes. Table 4 and Table 5 show the average fairness, efficiency and stability of the 3 clients in case of using the existing methods and our method. The shown results are taken from 20 seconds (i.e 10 segments) when all players are streaming at the same time.

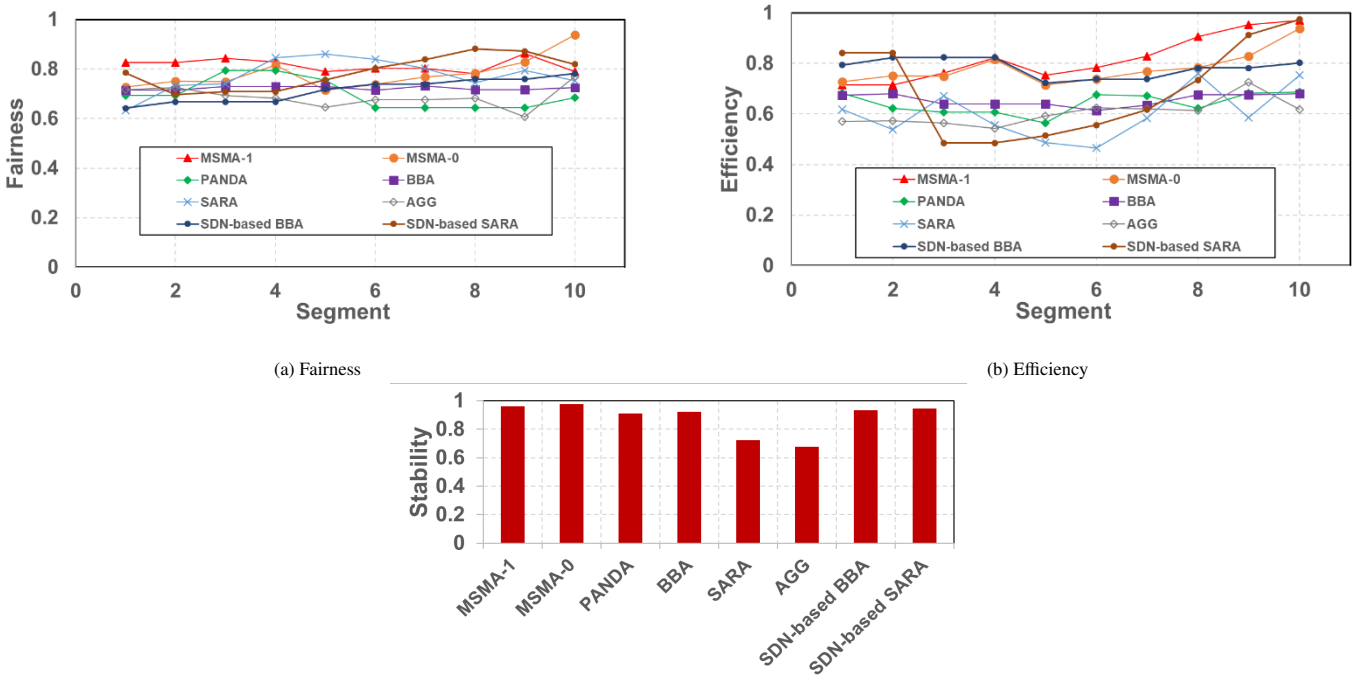
For VBR video Elephants Dream, in the following summary list, the consecutive numbers represent the results for MSMA-0, PANDA, BBA, SARA and AGG, respectively.

1) MSMA-1 outperforms the existing schemes in terms of fairness by an increase of 3.9%, 16.7%, 12.9%, 5.3% and 18.8%.

2) MSMA-1 improves bandwidth utilization in terms of efficiency by 3.3%, 26.9%, 26.7%, 37.8%, and 37.4%.

3) MSMA-1 increases the video stability by -1.5%, 5.6%, 4.6%, 33.1%, and 42.5%.

MSMA-1 improves two parameters: fairness and efficiency compared to MSMA-0. However, MSMA-0 has better stability than MSMA-1 because the clients in MSMA-1 are free to adapt their bitrates based on the throughput. And only when the throughput does not meet the requirements, the client will send the request to the controller. The SDN controller is then responsible for reallocating bandwidth to the clients accessing through that node, so that



(a) Fairness

(b) Efficiency

(c) Stability

Figure 6: Comparison average fairness, efficiency and stability of all methods

Table 4: Performance Comparison with VBR video

Metrics	MSMA-1	MSMA-0	PANDA	BBA	SARA	AGG	SDN-based BBA	SDN-based SARA
Fairness	0.816	0.785	0.699	0.723	0.775	0.687	0.714	0.807
Efficiency	0.830	0.781	0.654	0.655	0.602	0.604	0.783	0.697
Stability	0.962	0.977	0.911	0.920	0.723	0.675	0.933	0.945

the requesting clients receive the new bandwidth that can keep the most appropriate video version. As for MSMA-0, due to the use of the aggressive adaptive bitrate algorithm based on the bandwidth that has been divided evenly from the controller, it is still stable in choosing the bitrate or quality version of the next segment.

Similarly, with statistics in Table 5 (i.e. for CBR video "Destiny"), we can see that the two proposed methods (MSMA-1 and MSMA-0) outperform the rest. On another side, compared with traffic of the VBR video, traffic of the CBR video has higher Stability. The fairness of SDN-based BBA is almost similar to BBA, while the one of SDN-based SARA is slightly increased compared to SARA, by only 4%. SDN-based BBA effectively increases the bandwidth usage by 19.6% compared with BBA and SDN-based SARA increases by 15% compared to SARA. Thanks to the SDN-based fair bandwidth allocation, BBA and SARA have much better stability than the non-SDN solution.

Table 6 shows statistics from the experiments. As shown in the table, MSMA-1 outperforms the other methods. At first, with a reasonable throughput and buffer-based estimation mechanism, MSMA-1 is able to avoid stalling conditions whilst keeping the highest average value of buffer occupancy (i.e. 44.6s).

Secondly, MSMA-1, with the dynamic bandwidth allocation mechanism, can achieve the highest average bitrate of 2445.6 Kbps, which is an improvement of about 103.8%, 76%, 36.7%, 42.4% and 67.3% compared to MSMA-0, PANDA, BBA, SARA and AGG, respectively. Based on the QoE formula in Equation (19), the final QoE scores of MSMA-1 and MSMA-0 are calculated in comparison with 4 other solutions such as PANDA, BBA, SARA, AGG (Table 6). In fact, MSMA shows to achieve the highest QoE score compared to other competitors.

For the AGG algorithm, by requiring the highest bitrate according to the throughput without relying on the current buffer, the average video quality is low and the number of quality level switches-down is largest among all the methods. On the other hand, AGG also has the lowest QoE (i.e. 1.9).

SARA is the mixed throughput-buffer based algorithm, having the average video quality and QoE higher than AGG, but it is only medium compared to the other algorithms. BBA is the buffer-based algorithm, BBA successfully avoids stalling events which are very annoying to video viewers. The average bitrate of BBA is medium at only 1788.7 Kbps and its QoE stays at the good level (3.6).

With the estimation of the available bandwidth, a probing mech-

Table 5: Performance Comparison with CBR video

Metrics	MSMA-1	MSMA-0	PANDA	BBA	SARA	AGG	SDN-based BBA	SDN-based SARA
Fairness	0.971	0.965	0.764	0.713	0.716	0.873	0.932	0.917
Efficiency	0.879	0.796	0.904	0.665	0.726	0.708	0.812	0.803
Stability	0.975	0.957	0.931	0.949	0.826	0.703	0.941	0.955

Table 6: Quality Statistics of the solutions

Criteria	MSMA-1	MSMA-0	PANDA	BBA	SARA	AGG	SDN-based BBA	SDN-based SARA
Average Bitrate (kbps)	2445.6	1199.6	1389.5	1788.7	1717.4	1461.0	1812.3	1645.1
Average Buffer (s)	44.6	35.8	34.2	27.7	11.6	36.4	28.2	18.4
Number of Stalling	0	4	12	0	17	11	0	0
Average Version Index	6	4.5	5.1	5	5.2	5	5.4	5.1
Number of Version Switch-downs	21	3	7	17	20	40	16	15
N-QoE	4.25	3.6	3.54	3.55	2.8	1.9	3.96	3.74

anism is leveraged by PANDA to additively increase its sending rate, whilst decreasing sending rates multiplicatively, when congestion occurs. PANDA has a medium average bitrate at 1389.5 kbps and a relatively low number of switch-downs (7 times). The QoE score of PANDA also reaches 3.6.

For MSMA-0, the average bitrate is the lowest due to the use of the Aggressive bitrate adaptive algorithm on the client side, and MSMA makes its best to fairly allocate the bandwidth afor all clients. However, MSMA-0 has the lowest number of switch-downs (3 times). Besides QoE of MSMA-0 also reaches good quality (i.e. score of 3.6).

At the bottom lines, it's difficult to improve for all QoS parameters as well as the Quality of Experiment (QoE)- related parameters of video streaming. In fact, all of the proposed solutions need a tradeoff among these parameters to improve video streaming quality. However, our proposed solutions MSMA-0 and MSMA-1 have significantly improved average video quality, ensuring the highest average QoE compared to existing methods such as PANDA, BBA, SARA and AGG. Especially, MSMA-1 increases the QoE score by 18%, 20%, 20%, 51%, and 124% compare with MSMA-0, PANDA, BBA, SARA and Aggressive, respectively. As we can see, MSMA-0 does not SDN-based BBA and SDN-based SARA in terms of QoE, but our proposed MSMA-1 outperforms all compared methods.

Overall, our existing innovative algorithm is a combination of bitrate adaptation at the client side and resource allocation at the network side. The proposed rate adaptation method is to choose the best quality version on the current network condition and to avoid video stalling events. At the SDN controller, the resource allocation handles fair and adaptive bandwidth sharing in order to maximize the number of access clients, while trying to minimize the number of clients who got QoE degraded.

6 Conclusion and Future Work

In this work, we have presented an integrated solution of centralized SDN-based resource allocation at the network side and bitrate adaptation at the client side to resolve the problem of multi clients streaming video through a bottleneck connection. This causes several problems such as low quality of service, congestion, delay, and especially quality of experience of users. The experiments show that our proposed solution outperforms the predecessors in terms of Quality of Experience.The proposed scheme ensures bandwidth fairness, efficiency and stability. Another contribution of our work is to ensure QoE of as many ongoing clients as possible. Additionally, the method can prevent congestion at bottleneck point when there is insufficient bandwidth by rejecting a new incoming request

Our future work will be building the whole SDN-based architecture to allocate network resources and to maximize QoE for each client, while avoiding scalability issues among many competing clients in a shared environment. Multiple heterogeneous scenarios with multiple SDN controllers and various clients will also be taken into account.

Acknowledgment

This paper is an extension of work originally presented in the ICTC conference [4]. This work was funded by Vingroup and supported by Vingroup Innovation Foundation (VINIF) under project code VINIF.2020.DA03.

References

[1] "Cisco Visual Networking Index: Forecast and Methodology, 2016-2021"; San Jose, CA, USA, Cisco, White Paper, 2016.

[2] Open Networking Foundation (ONF), "Software Defined Networking: the new norm for network,"; White paper, 2012.

[3] N. McKeown et al., "OpenFlow: Enabling Innovation in Campus Networks," ACM SIGCOMM Comput. Commun. Rev., **38**(2), 69–74, 2008, doi: 10.1145/1355734.1355746.

[4] P. H. Thinh. P. N. Nam, N. H. Thanh, T. T. Huong, "SDN – based Dynamic Bandwidth Allocation for Multiple Video-on-Demand Players over HTTP," proceeding 2019 Int. Conf. Inf. Commun. Technol. Converg. (ICTC), 16-18 Oct. 2019, Jeju, Korea, 163–168, 2019, doi: 10.1109/ICTC46691.2019.8939834.

[5] L. R. Romero, "A Dynamic Adaptive HTTP Streaming Video Service for Google Android," M.S. Thesis, R. Inst. Technol. (KTH), Stock., 148, 2011.

[6] T. C. Truong, D. Q. Ho, J. W. Kang, A. T. Pham, and S. Member, "Adaptive Streaming of Audiovisual Content using MPEG DASH," IEEE Trans. Consum. Electron., **58**(1), 78–85, 2012, doi: 10.1109/TCE.2012.6170058.

[7] T.-Y. Huang, R. Johari, M. Watson, M. Trunnell, and N. McKeown, "A buffer-based approach to rate adaptation: Evidence from a large video streaming service," ACM SIGCOMM Comput. Commun. Rev., **44**(4), 187–198, 2015.

[8] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson, "Using the Buffer to Avoid Rebuffers: Evidence from a Large Video Streaming Service," arXiv Prepr. arXiv1401.2209, 2014, https://arxiv.org/abs/1401.2209.

[9] Y. Zhou, Y. Duan, J. Sun, and Z. Guo, "Towards simple and smooth rate adaption for VBR video in DASH," 2014 IEEE Vis. Commun. Image Process. Conf. VCIP 2014, 9–12, 2015, doi: 10.1109/VCIP.2014.7051491.

[10] H. N. Nguyen, T. Vu, H. T. Le, N. N. Pham, and T. C. Truong, "Smooth quality adaptation method for VBR video streaming over HTTP," 2015 Int. Conf. Comput. Manag. Telecommun. ComManTel 2015, 184–188, 2016, doi: 10.1109/ComManTel.2015.7394284.

[11] S. Latré, J. Famaey, F. De Turck, M. Claeys, and S. Petrangeli, "QoE-Driven Rate Adaptation Heuristic for Fair Adaptive Video Streaming," ACM Transaction Multimedia Computing Communication Application, **12**(2), 1–24, 2015, https://doi.org/10.1145/2818361.

[12] P. Juluri, V. Tamarapalli, and D. Medhi, "SARA: Segment aware rate adaptation algorithm for dynamic adaptive streaming over HTTP," 2015 IEEE Int. Conf. Commun. Work. ICCW 2015, 1765–1770, 2015, 10.1109/ICCW.2015.7247436.

[13] J. Jiang, "Improving Fairness , Efficiency , and Stability in HTTP-based Adaptive Video Streaming with FESTIVE," Proc. 8th Int. Conf. Emerg. Netw. Exp. Technol., 97–108, 2012, https://doi.org/10.1145/2413176.2413189.

[14] Z. Li et al., "Probe and Adapt: Rate Adaptation for HTTP Video Streaming At Scale," IEEE J. Sel. Areas Commun. **32**(4), 719–733, 2014, doi: 10.1109/JSAC.2014.140405.

[15] A. Mansy, "Network-layer Fairness for Adaptive Video Streams," IEEE, IFIP Netw. Conf. (IFIP Networking), pp. 1–9, 2015, doi: 10.1109/IFIPNetworking.2015.7145310.

[16] M. Ghobadi and M. Mathis, "Trickle: Rate Limiting YouTube Video Streaming," Present. as part 2012 USENIX Annu. Tech. Conf. (USENIX-ATC 12), 191–196, 2012, https://www.usenix.org/conference/atc12/technical-sessions/presentation/ghobadi.

[17] S. Akhshabi, L. Anantakrishnan, C. Dovrolis, and A. C. Begen, "Server-Based Traffic Shaping for Stabilizing Oscillating Adaptive Streaming Players," Proceeding 23rd ACM Work. Netw. Oper. Syst. Support Digit. audio video, 19–24, 2013, doi: 10.1145/2460782.2460786.

[18] J. J. Quinlan, A. H. Zahran, K. K. Ramakrishnan, and C. J. Sreenan, "Delivery of Adaptive Bit Rate Video: Balancing Fairness , Efficiency and Quality," 21st IEEE Int. Work. Local Metrop. Area Networks, 1–6, 2015, doi: 10.1109/LANMAN.2015.7114736.

[19] S. Ramakrishnan, X. Zhu, F. Chan, and K. Kambhatla, "SDN Based QoE Optimization for HTTP-Based Adaptive Video Streaming," 2015 IEEE Int. Symp. Multimed., 120–123, 2015, doi: 10.1109/ISM.2015.53.

[20] C. Cetinkaya, Y. Ozveren, and M. Sayit, "An SDN-assisted System Design for Improving performance of SVC-DASH," 2015 Fed. Conf. Comput. Sci. Inf. Syst. (FedCSIS). IEEE, **5**, 819–826, 2015, doi: 10.15439/2015F333.

[21] P. Georgopoulos, Y. Elkhatib, M. Broadbent, M. Mu, and N. Race, "Towards Network-wide QoE Fairness Using OpenFlow-assisted Adaptive Video Streaming," Proc. 2013 ACM SIGCOMM Work. Futur. human-centric Multimed. Netw., 15–20, 2013, doi: 10.1145/2491172.2491181.

[22] G. Cofano, L. De Cicco, T. Zinner, and S. Mascolo, "Design and Experimental Evaluation of Network-assisted Strategies for HTTP Adaptive Streaming," Proc. 7th ACM Multimed. Syst. Conf. (MMSys 2016), 1–12, 2016, doi: 10.1145/2910017.2910597.

[23] J. Chen, M. Ammar, M. Fayed, and R. Fonseca, "Client-Driven Network-level QoE fairness for Encrypted 'DASH-S'," Proc. 2016 Work. QoE-based Anal. Manag. Data Commun. Networks, 55–60, 2016, doi: 10.1145/2940136.2940144.

[24] K. T. Bagci, S. Member, K. E. Sahin, S. Member, and A. M. Tekalp, "Compete or Collaborate: Architectures for Collaborative DASH Video over Future Networks," IEEE Trans. Multimed., **19**(10), 2152–2165, 2017, doi: 10.1109/TMM.2017.2736638.

[25] O. El Marai and T. Taleb, "Online Server-side Optimization Approach for Improving QoE of DASH Clients," GLOBECOM 2017-2017 IEEE Glob. Commun. Conf. IEEE, 1–6, 2017, doi: 10.1109/GLOCOM.2017.8254128.

[26] S. Zhao and D. Medhi, "SDN-Assisted Adaptive Streaming Framework for Tile-Based Immersive Content Using MPEG-DASH," 2017 IEEE Conf. Netw. Funct. Virtualization Softw. Defin. Networks, 1–6, 2017, doi: 10.1109/NFV-SDN.2017.8169831.

[27] A. Bentaleb, A. C. Begen, and R. Zimmermann, "SDNDASH: Improving QoE of HTTP Adaptive Streaming Using Software Defined Networking," ACM Multimed., 1296–1305, 2016, doi: 10.1145/2964284.2964332.

[28] A. Bentaleb, A. C. Begen, R. Zimmermann, and S. Harous, "SDNHAS: An SDN-enabled architecture to optimize QoE in HTTP adaptive streaming," IEEE Trans. Multimed., **19**(10), 2136–2151, 2017, doi: 10.1109/TMM.2017.2733344.

[29] A. Bentaleb, A. C. Begen, S. Member, R. Zimmermann, and S. Member, "QoE-Aware Bandwidth Broker for HTTP Adaptive Streaming Flows in an SDN-Enabled HFC Network," IEEE Trans. Broadcast., **64**(2), 575–589, 2018, doi: 10.1109/TBC.2018.2816789.

[30] J. Jiang, L. Hu, P. Hao, and R. Sun, "Q-FDBA: improving QoE fairness for video streaming," Multimed. Tools Appl., **77**(9), 10787–10806, 2018, doi: 10.1007/s11042-017-4917-1.

[31] T. Abar, A. Ben, L. Sadok, and E. Asmi, "Enhancing QoE based on Machine Learning and DASH in SDN networks," 2018 32nd Int. Conf. Adv. Inf. Netw. Appl. Work., 258–263, 2018, doi: 10.1109/WAINA.2018.00095.

[32] A. Ahmad, A. Floris, and L. Atzori, "Towards Information-centric Collaborative QoE Management using SDN," 2019 IEEE Wirel. Commun. Netw. Conf., 1–6, 2019, doi: 10.1109/WCNC.2019.8885412.

[33] L. Guillen, S. Izumi, and T. Abe, "SAND/3: SDN-Assisted Novel QoE Control Method for Dynamic Adaptive Streaming over HTTP/3," Electron. Multidiscip. Digit. Publ. Inst., **8**(8), 1–17, 2019, doi: 10.3390/electronics8080864.

[34] Pham Hong Thinh, Nguyen Thanh Dat, Pham Ngoc Nam, Nguyen Huu Thanh, Nguyen Minh Hien & Truong Thu Huong. "An Efficient QoE-Aware HTTP Adaptive Streaming over Software Defined Networking,"Mobile Networks and Applications, 1-13,2020, https://doi.org/10.1007/s11036-020-01543-1.

[35] T. Huang, B. Heller, and N. Mckeown, "Confused , Timid , and Unstable: Picking a Video Streaming Rate is Hard," Proceedings of the 2012 internet measurement conference.ACM, 225–238, 2012, doi: 10.1145/2398776.2398800.

[36] X. Yin, M. Bartulovi, V. Sekar, and B. Sinopoli, "On the Efficiency and Fairness of Multiplayer HTTP-based Adaptive Video Streaming" 2017 Am. Control Conf. (ACC). IEEE, 4236–4241, 2017, doi: 10.23919/ACC.2017.7963606.

[37] Bitmovin, "Libdash - bitmovin." [Online]. Available: https://github.com/bitmovin/libdash. [Accessed: 25-Nov-2020].

[38] "Floodlight." [Online]. Available: http://www.projectfloodlight.org/floodlight/. [Accessed: 25-Nov-2020].

[39] "Mininet." [Online]. Available: http://mininet.github.io/. [Accessed: 25-Nov-2020].

[40] O. O. M. P. Studio, "Elephants dream," 2009. [Online]. Available: https://orange.blender.org/. [Accessed: 25-Nov-2020].

[41] V. D. Fabien Weibel, Manuel Alligne, Sandrine Wurster, "Destiny," 2012. [Online]. Available: http://fweibel.com/destiny . [Accessed: 05-Nov-2020].

[42] W. R. Jain, Rajendra K and Chiu, Dah-Ming W and Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," East. Res. Lab. Digit. Equip. Corp. Hudson, MA, 1984.

[43] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hossfeld, and P. Tran-gia, "A Survey on Quality of Experience of HTTP Adaptive Streaming," Ieee Commun. Surv. Tutorials, **17**(1), 469–492, 2015, doi: 10.1109/COMST.2014.2360940.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 200-207 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

PrOMor: A Proposed Prototype of V2V and V2I for Crash Prevention in the Moroccan Case

Zakaria Sabir*, Aouatif Amine

Ibn Tofail University, ILM Department, ENSA, Kenitra, 14000, Morocco

ARTICLE INFO

Article history:
Received: 03 October, 2020
Accepted: 02 January, 2021
Online: 15 January, 2021

Keywords:
Vehicular Communications
Road Safety
V2V
V2I
VANET
Connected Vehicles

ABSTRACT

Road safety has become an object of research and many research institutes have invested in this field because a lot of people die and many others are injured every year due to road accidents. The deployment of wireless communication technologies for vehicular networks can considerably improve road safety. It can enable new services such as traffic management, collision detection, and additional communication ease between moving vehicles. This paper presents a complete implementation of Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications. Raspberry Pi boards, ultrasonic sensor, infrared obstacle detector, and line follower sensors are used in order to implement the complete prototype. The results show the usefulness of this road safety prototype named PrOMor (Prevention of Obstacles in Morocco). Based on these results, it can be concluded that the presented scenarios can be applied to the field of road safety related to the Moroccan case. This should reduce the number of accidents and save more human lives.

1 Introduction

This paper is an extension of work originally presented in the International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) [1]. In the current work, the authors will deal with the whole vehicular communication system to address road safety issues, especially in the Moroccan case.

Since millions of people die every year and many more are injured because of car accidents, road safety has been gaining importance in recent years. Statistics [2] show that road traffic casualties in the world amount to 1.17 million per year. According to the World Health Organization (WHO), road accidents claimed the lives of 1.25 million people in 2013. In 2014, more than a million accidents occurred in Europe, killing 25 900 people and 1.4 million wounded. In 2017, 3,499 deaths were recorded on Moroccan roads, against 3,593 in 2016, a small decrease of 2.6% year-on-year. While statistics show a slight advance in road safety, they hide other, more disturbing realities[1]. These statistics are blatant, so the Ministry of Equipment, Transport, Logistics, and Water (METLE) thought to launch research projects in partnership with Moroccan universities and the National Center for Scientific and Technical Research (CNRST) in order to reduce the number of victims. This project based on the detection and the anticipation of the accidents aims to help to reduce the accident rate using the V2V and V2I communications to detect obstacles and prevent the crash in the Moroccan case. According to the traffic accident statistics carried out by the METLE in 2016 [3], there were 80680 accidents caused by obstacles.

Today the road network in Morocco [4], whose routes are traditionally hierarchical in the main, collector, and service according to the function they fulfill for the network itself, is not in adequacy with its urban environment and the needs of the inhabitants and users. Thus, for example, very wide paths designed for cars can pass through neighborhoods endangering pedestrians who go there to shop, go to school or work, or simply go to the neighbor. It must be recognized that, particularly in dense urban areas, the lanes fulfill a multiplicity of functions that go beyond the simple role of channeling private vehicles (social functions, economic, leisure, etc.).

To put a stop to this massacre, in addition to legislative measures, digital technologies will play a major role. The automation/semi-automation of vehicles will limit the human error, responsible for the majority of road accidents. Manufacturers are working on connected vehicles that will be equipped with several sensors. They will be able to interpret their environment in a dynamic way to alert drivers of the various hazards which could be found on the road.

In this context, the main purpose of Intelligent Transportation

*Corresponding Author: Zakaria Sabir, University Campus P. O. Box 241, Kenitra, Morocco, zakaria.sabir@uit.ac.ma
[1]Leconomiste, https://www.leconomiste.com (accessed Jul. 08, 2019)

Systems (ITS) is to improve transportation safety, increase productivity, reduce traffic jams and reduce the number of accidents on our future roads, thanks to the use of advanced information technology. Nonetheless, the big challenge is to constitute an interdependent environment that merges all the technologies and can solve a variety of transportation problems. If the electronic, computing, and networking technologies are well integrated, the mobility will be more efficient, the roads will be safer and the impact on the environment will be minimized [5].

These systems will allow vehicles to communicate their relative positions, for example, to signal the presence of another vehicle in a blind spot. Mainly, it will be connected to various sensors inside the vehicle: stability, hazard warning lights, airbags, and other diagnostic tools. Therefore, a vehicle can automatically communicate to others the presence of a slippery roadway, if it is down or in an accident. The other drivers will then adjust their approach in the concerned area. The ITS will also help to know the position of emergency vehicles (ambulances, firefighters, police) on approach. Their only sirens today give little information on their provenance and direction. Civilian vehicles will then clear the way more efficiently using the ITS.

One of the most important components of ITS, which has been proposed to address challenges in the transportation surface is the connected vehicle system. It is a hopeful technology that uses wireless communication. In order to treat transportation problems, connected vehicle system will use Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications [6] [7].

Different definitions have been given recently for a "connected" vehicle. A definition given at the 2013 Automotive News World Congress by IBM, identifies a connected vehicle as "a vehicle capable of seamless integration with multiple systems, connecting consumers to the digital world". More precise definitions include all the other connections of a modern vehicle: according to the US Department of Transportation (USDOT), "Connected vehicle applications provide connectivity between vehicles to prevent accidents, between vehicles and infrastructure, in order to generate safety, mobility and environmental benefits; between vehicles, infrastructure and wireless devices to provide continuous real-time connectivity to all users of the system" [8].

Different studies have already tried to develop prototypes of connected vehicles and systems that support V2V and V2I communications. But, as will be explained later, those systems were developed in foreign countries and cannot be applied in Morocco. Thus, the authors thought of a system that will fit the Moroccan case. This paper presents the implementation of a vehicular communication environment using both V2V and V2I technologies. This system is the first of its kind in Morocco. It requires a real test on real cars, but the authors carried out the tests on two robots only and it will be spread over a large number of cars in the real case, the used robots are based on Raspberry Pi-based mini cars kits, which were personalized by adding different features and other sensors in order to get to the target. The remainder of the paper is organized into four sections: Section 2 explain the proposed system. Experiments results will be presented and discussed in Section 3, while Section 4 concludes the paper.

2 Related works

This work is inspired by several prior studies that used various technologies on connected vehicle networks. A summary of related works is discussed in this section.

A recent work [9] explored the future beneficiaries of 5G and selected the conditions where 5G can make an effect. Particularly, three use-cases were considered: V2X (Vehicle-to-Everything Communication), Healthcare, and Drones. The authors investigated and emphasized some imperfections and failures of the current cellular topologies regarding the mentioned use-cases, then they determined how 5G will surmount those failures.

Kuramoto et al. [10] described the multi-class zone ITS communication scheme regarding short periodical communications and real-time for vehicle safety in intersections. In the proposed scheme, two types of areas around an intersection were designed, the nearest area and quasi-nearest area. In the first one, the communication of vehicles is done every 20 ms with a base station. While in the second one, vehicles make reservations for the communication for the first type and also communicate with the base station for safety. The required frequency bandwidth needed to realize the proposed design was shown in the simulation results.

A strategy of centralized and localized data congestion control is proposed in [11]. It uses RSUs (Road-Side Units) at intersections to control data congestion. It includes three entities for clustering messages, detecting congestion, and controlling data congestion. In order to detect data congestion in the channels, the strategy measures the channel usage. The messages are refined, assembled, and then grouped using algorithms of machine learning. Simulation results affirmed the efficacy of the proposed strategy in terms of improving the delay, packet loss ratio, and throughput.

Another system is presented in [12], it is an integrated vehicular system that uses an integrated V2I and V2V communication model improved with an information management system. The aim is the management, accumulation, and arrangement of context-aware information on driver location and traffic. The concept of smart roads is featured in the vehicular system. Significant information and all the risks related to the safety and detected by vehicles are managed by the infrastructure and adjusted to the driver's choices and the vehicle's context. The authors implemented the required software and used real hardware for infrastructure as well as a real vehicle to build the entire system.

Other authors [13] proposed to substitute the currently proposed standard (802.11p, DSRC/WAVE) to build an inter-vehicle communication system. They suggested using android smartphones along with Wi-Fi to replace the OBU (On-Board Unit) on consumer vehicles and to enable thereby V2I and V2V communications. The system is compatible with android devices with the 2.3 version and above. The users will then acquire an OBU that communicate with the nearby vehicles to broadcast an emergency alert or to warn about an accident. Experiment results showed the ease of use of the system, its efficiency, and its compatibility with Android devices.

Day et al. [14] checked out the performance of Het-Net (Heterogeneous Networks) consisting of LTE, DSRC, and Wi-Fi technologies for V2I and V2V communications. To enable Het-Net for vehicular communication applications, the authors developed an application layer handoff method. It was used to collect traffic

data and to warn about the collision. This research proved through the conducted field studies the usefulness of Het-Net in vehicular communications. Simulation experiments were run using a big number of connected vehicles including LTE and DSRC Het-Net scenarios. The results in terms of packet delivery error and latency from simulation were found to be identical to the field experiment results.

In this paper [15], the authors used data from the University of Michigan Safety Pilot Model Deployment project to analyze the performance of DSRC. They used selected road-side equipment and 1050 vehicles to evaluate the packet delivery ratio and the effective range of V2I communications. They have also studied the impact of some environmental factors like weather, trees, road elevation, buildings, and so on. The results are useful in the development of DSRC communication models and guiding future DSRC equipment placement.

Du et al. [16] used a distributed message delivery platform to develop a strategy. The aim is the creation of an effective and low-latency distributed message delivery system for connected vehicle applications. The strategy can transform the unstructured data into labeled topics for large consumers like data centers, mobile devices, and connected vehicles. The performance of the strategy was evaluated using a prototype and Apache Kafka, an open-source delivery system, and comparing its efficiency with the latency qualifications of connected vehicle applications. The experiments disclosed the correspondence of the measured latencies with USDOT (U.S. Department of Transportation) recommendation for connected vehicle applications.

In this paper, different from those studies, the proposed system considers employing vehicular communications in the Moroccan case. Authors have focused on the V2V and V2I communications since the conventional scheme can't be applied in Morocco. Taking into consideration the situation in Morocco and many other differences in comparison to other countries, a foreign system cannot work well for Morocco, the adjustment is required and many factors are to bear in mind. The architecture of Moroccan cities is different from other countries, in the big cities, for example, there are administrative regions, industrial regions, residential districts, the old medina, the new medina, etc. Other factors to take into consideration are culture and environment. The driver behavior and the usage of technology by the population need to be considered. Road conditions, alleys in one or two directions, signs boards, and narrow alleys also have to be considered.

The authors started with the realization using mini cars on a model and they have successfully achieved the goal. In the next step, they will move on to real scenarios with real cars in order to specify the necessary adjustments. The proposal represents also a contribution to the development of smart cities as there is already a trend for smart cities in Morocco. In June 2013, e-Madina was born in APEBI, the Moroccan federation of information technologies, telecommunications, and offshoring. e-Madina is a Smart Cities Cluster that works to make the city of Casablanca more attractive and more efficient using digital technologies and available material and immaterial resources. In March 2017, the government launched the "Cité Mohammed VI Tanger Tech" project, a smart city that will extend over 200 hectares within 10 years. Funding for the operation is estimated at $ 1 billion. Nearly 200 companies operating in the automotive, aeronautic, textile, electronics, and machine tool industries will set up there.

To the best of their knowledge, authors believe this is the first paper that deals with vehicular communications in the Moroccan case, and this is one of the reasons that made this work chosen, accepted, and financed by the METLE and the CNRST.

3 Proposed system

Connected vehicles are able to revolutionize the current traveling style by creating a protected and interoperable wireless communications network, a system that involves several elements like cars, traffic lights, buses, and trains. Technologies of connected vehicles aim to address some serious challenges in the transportation fields, especially in environment, mobility, and safety.

The purpose of this work belongs to the last field, which is road safety, especially in the Moroccan case to solve the problems cited above, since the conventional scheme can't be applied in Morocco as it is as explained previously. Authors propose three safety scenarios which will be used to alert Moroccan drivers in different situations, in order to decrease the probability of road accidents and to stay away from collision scenes that repeatedly happen between vehicles and pedestrians, animals, trees, and such other objects. These scenarios are based on reel time information and use V2V and V2I modes. Roadside units and vehicles exchange information that will be very useful to anticipate a hazardous situation. This information is delivered to Moroccan drivers in order to help them avoid the danger presented in unstructured roads.

Taking into consideration the statistics presented in [3], the first scenario can notify a driver of a hard-braking vehicle in the path ahead. For example, three vehicles are traveling in the same lane (see Fig. 1 (a)); the driver of the third vehicle can't see the first vehicle because it's being blocked by the second one, which is located directly in front of him. Unexpectedly, the first vehicle slams on its brakes. Thanks to V2V communication, the third vehicle is able to provide a warning of the hard-braking vehicle ahead. This warning will help the driver to drive safely and avoid a potential crash, as can be seen in Fig. 1 (a).

The second and the third scenarios aim to notify a driver of a special traffic condition in the path ahead. Whenever a traffic event is detected (accident, traffic jam, pedestrian crossing the road, animal, road work...), the driver is advised to slow down, change the lane, and take another route. Using V2V communication, the warning notification is transferred by another vehicle ahead, which has previously detected the problem of the related traffic event. V2I communication also was used for the same purpose (see Fig. 1 (b)). In this case, the warning notification is transferred by the RSU placed on the border of the road instead. The location of the RSU depends on the number of accidents recorded in this area and given by the METLE Ministry.

The diagrams in Fig. 2 summarize the logical flow of the methodology in each car. In this scenario, V1 and V2 which are equipped with a Raspberry Pi ran on a straight path using line follower sensors. V1 keeps using the infrared sensor along the way in order to detect any obstacle before the destination. If V1 detects an obstacle, it will stop and send the data as an alert to V2. Similarly,

the latter continues on the path to the destination unless it receives an alert from V1. If so, it will stop and change the road to avoid the obstacle and get to the destination.

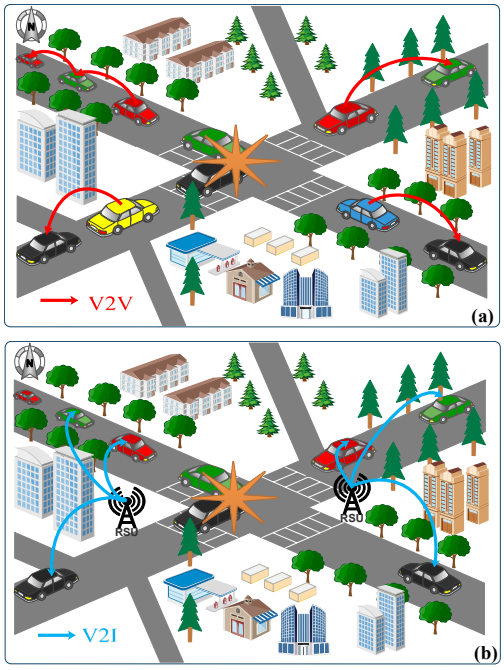


Figure 1: Operation of the system, (a) represents V2V mode and (b) represents V2I mode

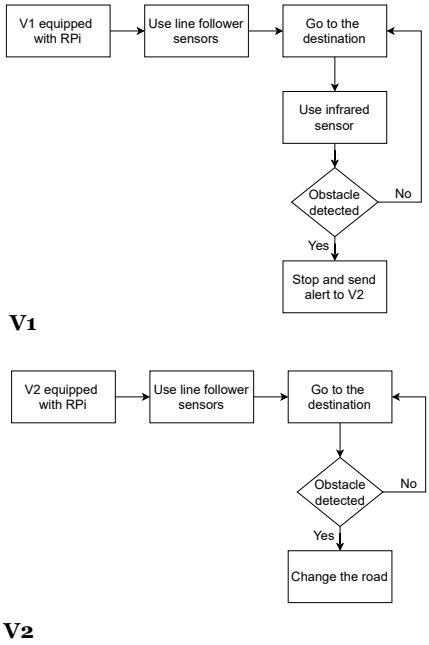


Figure 2: Flowchart for V2V communication.

Reducing possible points of conflict between vehicles and with other users is a common practice of road safety. For this, it is necessary to act in particular on the geometry of the path and the management of the priorities, especially for the zones of intersection and crossing with the presence of different users. According to the compendium of statistics of road traffic accidents [3], carried out in 2017 by the METLE, 14 998 accidents were recorded in the intersections. Thus, taking into account the statistics, the authors proposed the METLE concerning the locations of the RSUs concerning the areas likely to have accidents in the city of Kenitra (as can be seen in Fig. 3). The communication between these RSUs and also between these RSUs and the vehicles in the area can help to avoid the slowed traffic during a rush hour as shown in Fig.3.

This communication can also help to avoid accidents in hazardous regions especially unstructured ones. A real case is shown in Fig. 4 where the area with diagonal lines was recently transformed into a roundabout, but it was not well adjusted to represent the roundabout. Therefore, someone who is a stranger to the city can take the road straight while it is forbidden. This can cause accidents in this area, as shown in figure 3.

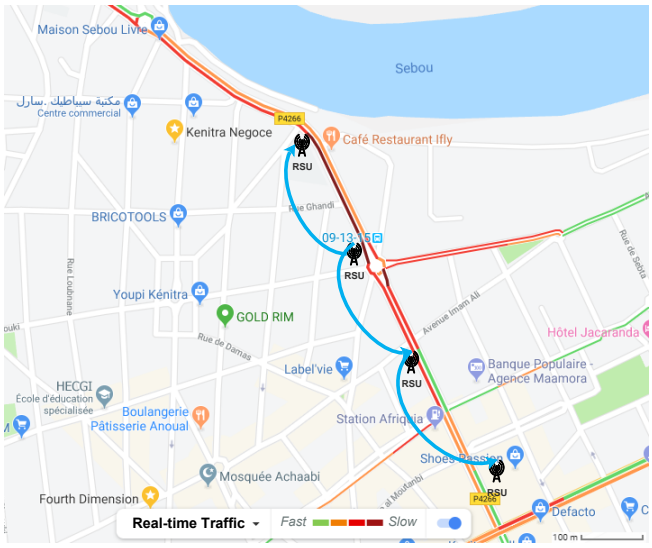


Figure 3: The locations of the RSUs to help mitigate the congestion in Kenitra city

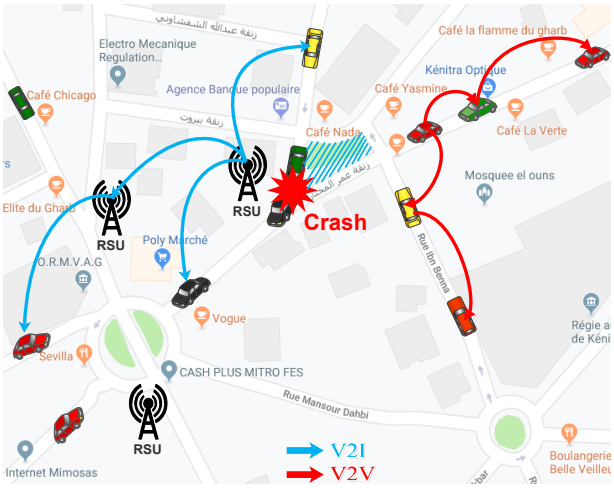


Figure 4: A Real case of an unstructured area in Kenitra city

In order to model the system, the authors proposed two algorithms. One on the client-side and the other on the server-side. Both algorithms start by importing the necessary modules and assigning

pins numbers. The next step is to create a socket so the two cars can communicate. After that, various methods used in the algorithms are defined :

- setup(): sets the setup of the system (variables, frequency, sensor modes ...).
- setSpeed(): manipulates the speed of the vehicle and handles the error related to the trajectory.
- stop(): stops the motors of the vehicle.
- read_sensors(): reads the line follower sensors and stores a variable related to the error of the trajectory.

Algorithm 1 : Client-Side

```
Import necessary modules
Assign pins numbers
Create a socket
Define the setup method
Define the setSpeed method
Define the stop method
Define the read_sensors method
setup()
while (true) do
    val ← IR_value
    if val = 1 then
        data ← "go"
        send data
        read_sensors()
        set_sensors()
    else
        stop()
        data ← "stop"
        send data
Close connection to the server
```

Algorithm 2 : Server-Side

```
Import necessary modules
Assign pins numbers
Create a socket
Bind address to socket
Start TCP listener
Accept TCP client connection
Define the setup method
Define the setSpeed method
Define the stop method
Define the read_sensors method
setup()
while (true) do
    get data
    if data = "go" then
        read_sensors()
        set_sensors()
    else if data = "stop" then
        stop()
End server
```

An infinite loop is necessary to keep reading data from sensors, sending it by the client, and receiving it by the server. If the infrared sensor (located on the client side) detected an obstacle, the IR_value receives the value "0". The client sends then the data to the server which will stop immediately. Else, the IR_value remains in "1". The client sends then the data to the server which will continue the path to get to the destination. The proposed algorithms are presented in Algorithm 1 and 2.

Table 1 presents a comparison of the proposed system with other studies.

4 Experimentation

4.1 Experimental setup

In the validation step, the authors preferred to use prototypes in the tests before applying the system to real cars. The experiments were conducted using two mini cars from Sunfounder[2] (Fig. 5) to develop the prototype system. Each car has a Raspberry Pi 3 B+ with basic necessary configuration to enable communication and control movements and sensors. A third Raspberry Pi 3 B board was used to play the role of the RSU. The used boards consist of ARM Cortex-A53 quad-core, which is operating at 1.4 GHz for the model B+ and 1.2 GHz for the model B. They both have a memory of 1GB (RAM), 40GPIO, Audio jack, Ethernet port, Camera interface, HDMI port, Display interface, 4 USB port 2.0, and micro SD slot. The two devices are powered by 5V / 2.5A via a micro USB port. Raspbian was chosen as an operating system. It is installed and boots from the micro SD card.

Both the cars integrate Wi-Fi, so they are able to communicate within an ad hoc network while moving in the test zone. The movement of the cars is powered by the two rear wheels using a DC motor driver. The cars came also with a USB camera, a step-down DC-DC converter module, and a PCA9685-based servo controller. The camera is not used in the current prototype, but it is used to prevent pre-crash by detecting pedestrians in another part of the same project SAFEROAD [21].

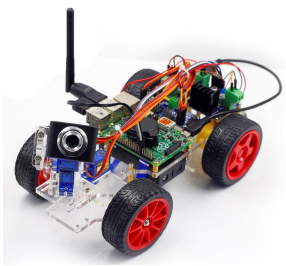


Figure 5: Sunfounder mini car

Since the original car kit didn't come with all the hardware needed to run the tests, the authors used some other sensors to complete the prototype, as can be seen in ((a) and (b)) Fig. 6. Five line follower sensors were added below each car, so they help the car following the line (road), in addition to detectors that have the ability to sense the existence of an obstacle in the path.

[2]Sunfounder, https://www.sunfounder.com/learn/category/Smart-Video-Car-for-Raspberry-Pi.html/ (accessed Jul. 15, 2019)

Table 1: Comparison of the current work with other studies

Study	Communication Mode	Implementation	Tools	Scenario	Purpose	Special Case
[17]	V2V	Real Time	RPi; Ultrasonic sensors; Audible Alarm; GPS Sensor; RF Module	Blind Spot Detection	Improving Traffic Congestion	Not specified
[18]	V2V	Real Time	OsmAnd; Android Devices; GRCBox	Displaying location of relevant vehicles	Warning about emergency vehicles	Not specified
[19]	V2V	Real Time	RPi; Controller; IR Sensor	Autonomous Driving (level 3)	Smooth Traffic Flow	Not specified
[20]	V2V and V2I	Real Time	ECU sensors, Arduino; GPS sensor; LM393 speed sensor module; CAN communication module; Wi-Fi modem	Detect sudden changes in status of the vehicles (speed, location, brake position)	Help in Advanced Driver Assistance System (ADAS)	Not specified
This paper	V2V and V2I	Real Time	RPi mini cars; IR sensor; Line Followers	Obstacle Detection	Preventing Accidents	Moroccan Case

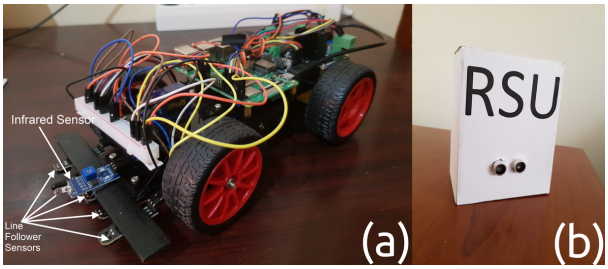


Figure 6: The prototypes, (a) Robot & (b) RSU

One infrared obstacle detector was used at the front of the first car, in addition to one ultrasonic sensor used in the RSU. The specifications of the sensors are given in Table 2.

Table 2: The specifications of the used sensors

Sensor name	Line Follower Sensor	Ultrasonic Sensor	Infrared Obstacle Detector
Reference	TCRT5000	HC-SR04	FC-51
Detection Distance	1 mm - 25 mm	2 cm - 4 m	2 cm - 30 cm, Angle: 35
Operating Voltage	5 V	5 V	3.3 - 5 V
Quantity	5	1	1

4.2 Results and discussion

The aim of the first and second scenarios was the dissemination of a warning message indicating the presence of an obstacle in the path ahead, using V2V communication. First, the authors tried to move the robots following a map (pre-programmed distances), but since the motors are not completely identical, they ran at different speeds. Consequently, the car couldn't move forward in a straight movement. For that reason, the authors added line follower sensors, to keep the car on the line (road).

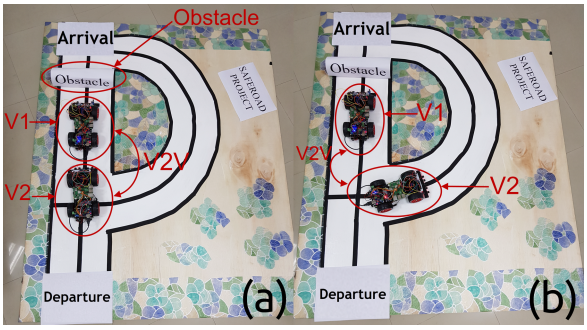


Figure 7: (a) The first scenario (V2V communication): V2 stops, once receiving the alert from V1 indicating the presence of an obstacle. (b) The second scenario (V2V communication): V2 changes the road, once receiving the alert from V1 indicating the presence of an obstacle.

In the first scenario, V1 and V2 ran at the same time on a straight path, and an obstacle was placed just before the destination. V1 is equipped with an infrared sensor (see Fig. 6 (a)), which detect the obstacle. Once the obstacle is detected, V1, which plays the role of the client, stops immediately, and sends the data to the server (V2) as a warning notification, using the Wi-Fi interface. As soon as it receives the warning, V2 stops in its turn (Fig. 7 (a)). Receiving the notification depends on the range of the Wi-Fi (the maximum distance between the first and the second car).

Similarly, in the second scenario, V1 and V2 ran at the same time on a straight path, and an obstacle was placed just before the destination. The infrared sensor in V1 (the client) detects the obstacle and stops immediately. It sends then an alert to V2 (the server) to change the road in order to avoid the obstacle, and get to the

destination. Therefore, once V2 reaches the intersection, it changes the road taken by the first vehicle and chooses the safe road (Fig. 7 (b)).

For the third scenario, the intention was the sending of the warning message using V2I communication. For this purpose, a third Raspberry Pi board was used with an ultrasonic sensor to play the role of the roadside unit (RSU). In this case, one vehicle was put in the path, and the obstacle was placed on the road. The RSU senses the presence of the obstacle using the ultrasonic sensor (see (b) in Fig. 6). It uses then the Wi-Fi interface to send a warning to the vehicle presented in the zone. As soon as the vehicle reaches the intersection, it takes the safe road (Fig. 8).

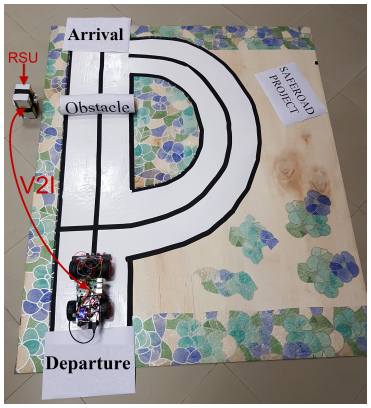


Figure 8: The third scenario (V2I communication): once getting to the intersection, the vehicle will take the other road, thanks to the alert received from the RSU.

The realization can be applied in the real world, as can be seen in Fig. 4 which depicts an example of Kenitra city where the area mentioned is likely to have accidents. The presence of the RSUs in this zone can help to avoid accidents. The RSU detects the crash and sends the alert to the vehicles in the area and also to the nearby RSUs. Another type of obstacles that can be detected using RSUs is the slowed traffic, as it is illustrated in Fig. 3. The communication between the RSUs in the area can help Moroccan drivers to avoid the traffic jam by taking another road where the traffic is relatively fast.

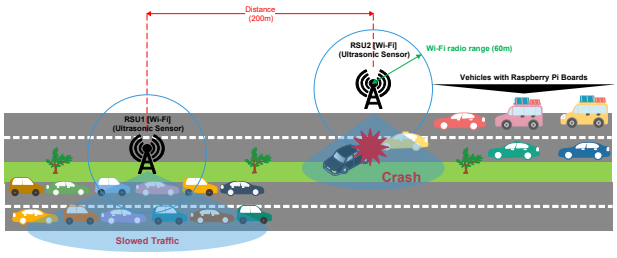


Figure 9: Summary of the contribution to Moroccan roads.

As explained above, the application of the system in real life in Morocco can play an important role in addressing the issues of road safety. It can be used to serve to avoid accidents, as it can also be useful to avoid slowed traffic. Fig. 9 summarizes the contribution of the system in Moroccan roads. In the first direction, the RSU2 detects the crash and sends an alert to nearby vehicles using the Wi-Fi interface. In the opposite way, the RSU1 detects the slowed traffic and sends a warning to the vehicles in the zone. The two RSUs can also communicate by sending information about the traffic in the two directions.

From Fig. 9 it can be seen that the crash was caused by a vehicle that took the lane in the opposite way. In order to resolve this kind of situation, the authors made another proposition to the METLE. They suggest building the new roads without a separation between the lanes of the opposite directions. Therefore, in the case of an accident or traffic jams, it will be possible to use a lane or more in the opposite direction to alleviate the traffic (see Fig. 10 and Fig. 11).



Figure 10: Road with separation between opposite lanes - Morocco



Figure 11: Road without separation between opposite lanes - Morocco

Due to V2I communication, the RSU will detect the obstacle (accident), send a warning to inform the nearby drivers and send also a notification to the traffic management center (direction of the roads or gendarmerie). Consequently, the center will change the light panels to indicate the use of additional lanes. The drivers will then keep the correct lane.

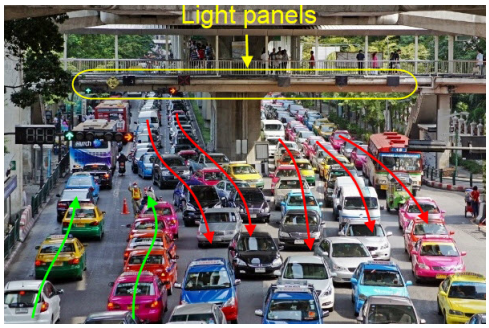


Figure 12: Avoiding Traffic Jams in Bangkok, Thailand by adding opposit lanes using the light panels

[3]https://www.shutterstock.com/video/clip-17081557-bangkok-thailand-june-2016–traffic-jam (accessed Oct. 01, 2019).

An example of Bangkok city[3] in Thailand is illustrated in Fig. 12. The road contains six lanes, but two of them were added in the opposite direction to mitigate the congestion. This is one of the propositions that can be made in the Moroccan case to solve rush hour problems, especially in big cities.

5 Conclusion

The main concern of the paper was to present an implementation of vehicular communications using both V2V and V2I technologies. The goal was to notify the driver at the right time so that he can take the navigation choice depending on the received alert. The tests were conducted using the Raspberry Pi 3 boards, model B and model B+, in addition to a specific number of sensors. Two mini cars were used to realize V2V communication by testing two different scenarios which worked well. Authors realized also the V2I communication which worked as well. In addition to that, they made proposals to the Ministry of transportation concerning the deployment of the RSUs and the management of traffic jams in Kenitra City. Based on the results, it can be concluded that these scenarios can be used in the field of road safety, related to the Moroccan case, to address the problems of transportation in kenitra city for example. This will decrease the number of accidents in Morocco and help save lives. As the current work may have some limitations that can be faced in real life, the authors intend in future research to conduct real experiments using two real vehicles. This will reveal all the limitations and lead to a complete system. Further research could bear from this work, such as the integration of cameras in this system in order to study different scenarios of accidents.

Conflict of Interest The authors declare no conflict of interest.

References

[1] Z. Sabir, S. Dafrallah, A. Amine, "A Novel Solution to Prevent Accidents using V2I in Moroccan Smart Cities," in 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 621–625, Dubai, UAE, 2019, doi:10.1109/ICCIKE47802.2019.9004397.

[2] F. Losilla, A.-J. Garcia-Sanchez, F. Garcia-Sanchez, J. Garcia-Haro, Z. J. Haas, "A Comprehensive Approach to WSN-Based ITS Applications: A Survey," Sensors, **11**(11), 10220–10265, 2011, doi:10.3390/s111110220.

[3] L. a. W. Ministry of Equipment, Transport, D. of Roads, "Recueil Des Statistiques Des Accidents Corporels De La Circulation Routière," 2016.

[4] N. C. for the Prevention of Traffic Accidents, Guide référenciel pour la planification de sécurité routière dans les zones urbaines, Editions & Impressions Bouregreg, Rabat, 1st edition, 2017.

[5] J. A. Guerrero-Ibáñez, C. Flores-Cortés, S. Zeadally, "Vehicular Ad-hoc Networks (VANETs): Architecture, Protocols and Applications," in N. Chilamkurti, S. Zeadally, H. Chaouchi, editors, Next-Generation Wireless Technologies, 49–70, Springer, London, 2013, doi:10.1007/978-1-4471-5164-7_5.

[6] Z. Sabir, A. Amine, "Connected vehicles using NDN for intelligent transportation systems," in International Conference on Industrial Engineering and Operations Management, volume 2018, 2433–2441, Paris, France, 2018.

[7] Z. Sabir, A. Amine, "NDN vs TCP/IP: Which One Is the Best Suitable for Connected Vehicles?" in Recent Advances in Mathematics and Technology, Applied and Numerical Harmonic Analysis, 151–159, Springer International Publishing, Cham, 2020, doi:10.1007/978-3-030-35202-8_9.

[8] R. Brookes, P. Pagani, "What becomes a car," in Proposed Paper for: BIT 2014 Conference Workshop-Technology Enabled Business Models: Platforms, Analytics and Performance, 2014.

[9] H. Ullah, N. Gopalakrishnan Nair, A. Moore, C. Nugent, P. Muschamp, M. Cuevas, "5G Communication: An Overview of Vehicle-to-Everything, Drones, and Healthcare Use-Cases," IEEE Access, **7**, 37251–37268, 2019, doi:10.1109/ACCESS.2019.2905347.

[10] K. Kuramoto, K. Fujimura, T. Hasegawa, "The Multi-Class Zone ITS Communication Scheme for Real-time Communications in Intersections," in 2007 IEEE Intelligent Transportation Systems Conference, 431–435, 2007, doi:10.1109/ITSC.2007.4357804.

[11] N. Taherkhani, S. Pierre, "Centralized and Localized Data Congestion Control Strategy for Vehicular Ad Hoc Networks Using a Machine Learning Clustering Algorithm," IEEE Transactions on Intelligent Transportation Systems, **17**(11), 3275–3285, 2016, doi:10.1109/TITS.2016.2546555.

[12] J. Santa, A. F. Gomez-Skarmeta, "Sharing Context-Aware Road and Safety Information," IEEE Pervasive Computing, **8**(3), 58–65, 2009, doi:10.1109/MPRV.2009.56.

[13] K.-C. Su, H.-M. Wu, W.-L. Chang, Y.-H. Chou, "Vehicle-to-Vehicle Communication System through Wi-Fi Network Using Android Smartphone," in 2012 International Conference on Connected Vehicles and Expo (ICCVE), 191–196, 2012, doi:10.1109/ICCVE.2012.42.

[14] K. C. Dey, A. Rayamajhi, M. Chowdhury, P. Bhavsar, J. Martin, "Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication in a heterogeneous wireless network" Performance evaluation," Transportation Research Part C: Emerging Technologies, **68**, 168–184, 2016, doi:10.1016/j.trc.2016.03.008.

[15] X. Huang, D. Zhao, H. Peng, "Empirical Study of DSRC Performance Based on Safety Pilot Model Deployment Data," IEEE Transactions on Intelligent Transportation Systems, **18**(10), 2619–2628, 2017, doi:10.1109/TITS.2017.2649538.

[16] Y. Du, M. Chowdhury, M. Rahman, K. Dey, A. Apon, A. Luckow, L. B. Ngo, "A Distributed Message Delivery Infrastructure for Connected Vehicle Technology Applications," IEEE Transactions on Intelligent Transportation Systems, **19**(3), 2018, doi:10.1109/TITS.2017.2701799.

[17] N. G. Ghatwai, V. K. Harpale, M. Kale, "Vehicle to vehicle communication for crash avoidance system," in 2016 International Conference on Computing Communication Control and automation (ICCUBEA), 1–3, 2016, doi:10.1109/ICCUBEA.2016.7860118.

[18] S. A. Hadiwardoyo, S. Patra, C. T. Calafate, J. Cano, P. Manzoni, "An Android ITS Driving Safety Application Based on Vehicle-to-Vehicle (V2V) Communications," in 26th International Conference on Computer Communication and Networks (ICCCN), 1–6, 2017, doi:10.1109/ICCCN.2017.8038486.

[19] S. Tayeb, M. Pirouz, S. Latifi, A Raspberry-Pi Prototype of Smart Transportation, IEEE Computer Society, 2017, doi:10.1109/ICSEng.2017.25, pages: 176-182.

[20] V. Vibin, P. Sivraj, V. Vanitha, "Implementation of In-Vehicle and V2V Communication with Basic Safety Message Format," in 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 637–642, 2018, doi:10.1109/ICIRCA.2018.8597311.

[21] S. Dafrallah, A. Amine, S. Mousset, A. Bensrhair, "Will Capsule Networks overcome Convolutional Neural Networks on Pedestrian Walking Direction?" in The 22nd IEEE International Conference on Intelligent Transportation Systems (ITSC), Auckland, New Zealand, 2019, doi:10.1109/ITSC.2019.8917019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 205-212 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Effectiveness and Suitability of the Automotive EHPS Software Reliability and Testing

Yanshuo Wang[1,*], Jim (Jinming) Yang[2], Ngandu M. Mbiye[2]

[1]*Department of Quality and Reliability, Dare Auto Inc., Plymouth, 48170, U.S.A.*

[2]*Department of Software Engineering, Dare Auto Inc., Plymouth, 48170, U.S.A.*

ARTICLE INFO

Article history:
Received: 04 January, 2021
Accepted: 04 May, 2021
Online: 23 May, 2021

Keywords:
Effectiveness
Reliability Growth
Software Reliability
Design for Reliability
Testing

ABSTRACT

The effectiveness and suitability for the reliability and test of the embedded software of the automotive EHPS (Electrical Hydraulic Power Steering) pump are extensively explored in this paper. The Crow-AMSAA analysis has been applied to evaluate the embedded software reliability growth based on the failure data collected in the prototype phase and in the field. The slope β of the Crow-AMSAA plot is smaller than 1 which indicates that the reliability of the embedded software is increasing and failure rate is decreasing. The field performance and reliability of the embedded software, which is the key indicator to evaluate the effectiveness and suitability of the reliability management and testing methods used in design and development, are also summarized in this paper. Using the real field and zero mileage data to evaluate the effectiveness and suitability of DFR (Design for Reliability) is also beneficial for the company to make the continuous improvement for the future embedded system/software design and development.

1. Introduction

This paper is an extension of work originally presented in paper entitled "An Automotive EHPS Software Reliability and Testing" that was published in 2020 Annual Reliability and Maintainability Symposium (RAMS) [1]. The previous paper was addressing the methodologies for Design for Reliability (DFR) and testing of the embedded software. This paper is using the data collected in the prototype build and in the field to address the reliability growth, the effectiveness and suitability of the reliability management and testing which were used in design and development. (Figure 1 shows the scopes of these two papers).

The vehicle steering system consists of steering wheel, column, EHPS pump, steering gear, pipes, and linkage etc.. EHPS pump has a hydraulic pump, BLDC (brushless direct current) motor, and ECU (electronic control unit). The hydraulic pump is driven by an electrical motor, and motor is controlled by ECU. EHPS pump provides the hydraulic flow to the steering gear in the power steering system. The control software is embedded in the MCU (Microcontroller Unit). (Figure 2 shows the EHPS pump and its interface).

1.1. Advantage of EHPS pump

Comparing to the conventional engine belt driven hydraulic power steering pump, the EHPS pump is powered by vehicle alternator and controlled on demand by the algorithm, the conventional engine belt driven hydraulic power steering pump operates continuously while the engine is powered. The EHPS pump provides *70%* energy saving over a comparable conventional power steering pump with constant displacement volume.

1.2 Control Software of EHPS pump

Control software quality is the key for providing reliable and safe power steering system operations. The application of systematic processes and techniques ensures software reliability. This includes the DFR (Design for Reliability), component and sub-system level reliability test, vehicle level test, and feedback from the field during development and the life of the product. The quality of the control software is also a concern for the OEMs (Original Equipment Manufacturer). In [2], the author illustrated the embedded software in crisis with examples from the automotive industry. In [3], the authors edited the Automotive Embedded Systems Handbook which provides an introduction and the scientific challenges to the automotive embedded system. The

*Yanshuo Wang, 47548 Halyard Drive, Suite B, Plymouth, MI 48170, U.S.A., 1 (734) 927 2072 & yanshuo.wang@dare-auto.us

dependability and performance of an embedded system is also depending on communication networks and protocols. Hence, certain characteristics on communication network have to be verified. In our case, the CAN (Controller Area Network) communication is used.

In addition to applying the DFR and testing methodologies in the software design and development, the effectiveness and suitability of DFR and testing of the embedded software need be evaluated [Figure 1].

1.3 Literature review

In [4], the author stated that software reliability is different from hardware reliability, it does not follow bathtub curve, it follows revised bathtub curve (shark teeth curve). This matched our experience when new features were added in earlier prototypes. During development, we placed a strong emphasis on regression testing. In [5], the authors described the enormous potential for defect prevention that can be achieved before the software is even tested. We saw that the quality difference between the concept prototypes and the production software which used different processes. Our evaluation is ongoing regarding the use of identical processes for concept prototypes and production software. The ASPICE process [6] matches [7] which proposed the software reliability assurance approach during its life cycle, and in [8] the authors is to address the software complexity. In [9], the authors presented a new embedded software reliability growth model. We are still evaluating the most practical calculations for future projects. In [10], the author proposed a verification strategy which enhances the effectiveness of integration testing of the distributed software functions. In [11], the author described the organization to adapt the standard classification scheme to and describes a methodology for comprehensively evaluating defect classification schemes. In [12], the authors describes the impact of strategical decisions on the software quality. In [13], the authors described the model based verification for embedded software. In [14], the authors introduced a method for constructing a software reliability evaluation framework based on historical data. They establish a software reliability evaluation model based on code metrics. In [15] proposed two approaches to use metric to analyze the large amount of measurement data generated during the software development process. In [16], the author proposed a technique to guide for the selection of an appropriate software reliability model for an ongoing software development project. In [17], the authors performed a review of currently available quantitative software reliability methods. In [18], the authors proposed a hazard-based effect analysis method to assess the distribution of the hazard degree of a remaining requirements fault. In [19], the authors described that the full automation of evolutionary testing method can improve the effectiveness and efficiency of the test process.

We apply the Crow-AMSAA (CA) model and field return data analysis methods to evaluate the effectiveness and suitability of the software reliability. The advantage of our methods is to use the system level or field failures on the vehicle to assess the component level embedded software reliability.

The CA and field data analysis methodologies are introduced in section 2. The data analysis of CA and field data is described in section 3. The lessons learned and effectiveness of DFR methods are also summarized in section 3. In section 4, the lessons learned and effectiveness of the testing are summarized. Finally, the work of this paper is summarized in section 5.

2. Methodologies

2.1. Crow-AMSAA Reliability Growth

Dr. Crow proposed that the Duane model can be represented as non-homogeneous Poisson process (NHPP) model under Weibull intensity function ([20], [21], [22]).

When CA model applies, the cumulative failure *N(t)* can be calculated as following

$$N(t) = \lambda t^{\beta} \tag{1}$$

$$LnN(t) = Ln(\lambda) + \beta Ln(t) \tag{2}$$

The model intensity function $\rho(t) = \frac{dN(t)}{dt} = \lambda\beta t^{\beta-1}$ (3)

The cumulative event rate is to use the equation (1) divided by *t.* it is $C(t) = \lambda t^{\beta-1}$ (4)

Where *t* is the time in days, λ and β are constants, the scale parameter, λ, is the intercept on the y axis of *N(t)* when *t =1, (ln(1) =0)*; the slope β, is interpreted in a similar manner as a Weibull plot, If the slope $\beta > 1$, the failure rate is increasing, the failure rate is more rapid, if the slope $\beta < 1$, the failure rate is decreasing, the failure rate is slower, if the slope $\beta = 1$, the process is named the Homogenous Poisson Process (HPP), if the slope β is not equal *1*, the process is called Non Homogenous Poisson Process (NHPP). ([20], [21]). Weibull plot is for single failure mode, but CA model is for mixing failure modes.

IEC (International Electrotechnical Commission) MLE (Maximum Likelihood Estimation) solutions for interval or grouped data method is used for the fitting method in CA analysis ([23]). Using IEC 61164 methods, the Cramer-Von Mises statistic accepts the goodness of fit at a Fit-p% of 10% as indicated on the plots.

Based on the equation (1) and (2) above, the cumulative failures versus the cumulative days are used for CA analysis. We also apply the CA model to the cumulative quantities delivered in addition to the cumulative days for the failure event, the equation is as

$$N(t) = \lambda q(t)^{\beta} \tag{5}$$

here *q(t)* is the cumulative quantities delivered.

The cumulative reliability *R(t)* is calculated as following:

$$R(t) = 1 - F(t) \tag{6}$$

where *F(t)* is the failure rate,

$$F(t) = \frac{N(t)}{q(t)} \tag{7}$$

We also apply the CA plot to the cumulative reliability versus the cumulative days.

2.2. Field Data Approach

Not all the defects, faults and errors can be detected during software and system design and development. The field performance and reliability data has been collected and analyzed for this automotive EHPS pump. The product has been in the production for more than three years, we have sufficient data from the field (zero mileage [Table 1] and warranty [Figure 7,8,9,10]). The field complaints can be categorized to the following: supplier design manufacturing issue (SDMI), adjoining components' failure, and misdiagnosis by servicers [Figure 8]. The real root cause of the EHPS pump failures from warranty is listed in the Figure 10. Since the EHPS pump is an electrical mechanical/hydraulic component on the vehicle, we need determine the root cause by mechanical-hydraulic (pump), electrical-mechanical (BLDC motor), electronic and embedded software (ECU). The failure rate and reliability of embedded software are calculated by using the following formulas

$$WRR = \frac{NWPR}{NPS} * 100\% \quad (8)$$

$$SDMFR = WRR * \frac{NSDMI}{NWPR} * 100\% \quad (9)$$

$$SRFR = SDMFR * \frac{NSRI}{NSDMI} * 100\% \quad (10)$$

$$SR = 1\text{-}SRFR \quad (11)$$

where *WRR* is the warranty return rate. *NWPR* is the number of warranty parts replaced. *NPS* is the number of parts sold. *NSDMI* is the number of supplier design manufacturing issue. *SDMFR* is the supplier design manufacturing failure rate. *NSRI* is the number of software related issue. *SRFR* is the software related failure rate. *SR* is the software reliability.

The failures detected in the field can be traced back, and can be used to analyze the effectiveness and suitability for the reliability management and testing in design and development. The lessons learned is also beneficial for the future embedded software projects.

The aim of this paper is to assess the achievements, and to list all the anomalies found in the field, and to illustrate the effectiveness and suitability for all the methods listed in the previous 2020 RAMS paper [1].

3. Effectiveness and suitability for Design for Reliability

The CA software reliability growth is analyzed by using EHPS pump failure data. The following describes the embedded software reliability growth, the effectiveness and suitability by using ASPICE[6] design and development process.

3.1. Crow-AMSAA Data Analysis

The EHPS pump failure data has been collected since beginning of developing this project starting on 1/20/2014. There are four phases in this project: Phase I – Prototype, Phase II – Design Verification (DV) 1, Phase III –DV2 & Production Validation (PV), Phase IV – Production. The system failures and customer returns were counted at different phases. The root causes of the failures have been investigated and documented in the 8Ds (8 Disciplines). The failures which caused by the embedded software were recorded separately. The time period of different phases, the cumulated days, the cumulative quantities delivered, the cumulative failures (due to software) and the cumulative reliability are summarized in the Table 2. The SuperSmith package (developed by Fulton Findings) was used for CA plots.

The CA analysis has been conducted for the cumulative failures (due to software) versus the cumulative days, and plotted in Figure 3. From the plot, the β value is *0.415* which is smaller than *1*. Thus the failure rate (due to embedded software) is decreasing, the embedded software reliability is growing as the timing cumulated.

The CA analysis has been conducted for the cumulative failures (due to software) versus the cumulative quantity delivered, and plotted in Figure 4. From the plot, the β value is *0.085* which is smaller than 1. Thus the failure rate (due to embedded software) is decreasing, the embedded software reliability is growing as large quantities delivered to customer.

The total software reliability versus the cumulative days was plotted in CA format in Figure 5 and 6. Figure 5 is normal plot without the log-log, but Figure 6 is the log-log plot. From the plots, the overall reliability change by time is demonstrated. The slope (β_1= *0.422*) indicates the embedded software reliability change rate while the timing cumulated.

3.2. The effectiveness of understanding the requirements

Based on the zero mileage, warranty data and frequent DCRs (Design Change Request), the following statements are applying to the effectiveness and suitability of understanding the requirements.

- Up to now, we have shipped roughly more than *600,000* EHPS pumps to our customers. The *WRR* is 0.1% (equation 8) (the *0.07%* is replaced after misdiagnosis by servicers, only *0.03%* (equation 9) is EHPS pump related issue, only about *0.0025%* (equation 10) may be embedded software related). From the warranty data, we can see the customer original requirements were well understood and implemented in EHPS pump system and software design and development. The misdiagnosis by servicers (about *71%*) is a big concern, this was caused by the vehicle system level issues, or the other adjoining components with EHPS pump [Figure 9,10]. For example, for the sake of the safety concern, if the EHPS does not receive CAN communication from the controlling stability and braking module, the EHPS must shut down. But this requirement would make the end driver with no power steering assist when losing CAN communication but no information on the instrument panel. Consequently, the EHPS pump had replaced unnecessarily in warranty, and EHPS pump reliability score had been decreased because of this requirement.
- There are dozens of power circuitry failures [Figure 9] (about *7.5%* of warranty returns) and majority of these cases were caused by the high current on the vehicle. This is a vehicle system issue, and the requirements were not clearly defined by the OEMs. That means the situation on the vehicle which could have the high current can damage the ECU of EHPS pump, and the EHPS control software needs protect the power

circuitry in EHPS ECU in that situation. But unfortunately the situation and condition is not clearly defined yet.

- The zero mileage issues: (a) Un-necessary ignition startup/run DTC (Diagnostics Trouble Code) issue implies that the customer requirement especially in the OEMs manufacturing site was not discussed and understood. This is a lessons learned for the future embedded software design and development. In addition to the requirement from different stakeholders, the OEMs vehicle assembly requirements must be understood and implemented; (b) As we found in the warranty, there are power circuitry failures in zero mileage as well. These may be related to the high current spike situation on the vehicle as well.
- Almost all the OEMs would request software changes, and we have received dozens of DCRs for all our supplied pumps after starting the productions. DCRs had significant impacts on the final software. They changed the pump speed maps and fan motor speeds. These reflect the differences between the real data collected in the field and the original customer requirements. The EHPS software was designed to adapt to changed requirements.

3.3. The effectiveness of developing the embedded software requirements & reliability goals

According to the field data, the requirements of the embedded software of the EHPS pump were carried over very well from vehicle and system requirements, but some of the vehicle system requirements (section 3.2) were not well defined which caused the lacking of the embedded software and system requirements of EHPS pump. The reliability of the embedded software of the EHPS pump *SR* is *1-0.0025%=99.9975%* (equation 11), this indicates the performance of the embedded software in the field is excellent.

3.4. The effectiveness of designing the system architecture

According to the field data, the DTCs such as ignition start up/run had not well defined in the system and software design. This was caused by the lacking of understanding of the vehicle requirements in OEMs and their assembly plants. Also the power circuitry issues (both zero mileage and warranty) were not well understood and calibrated between hardware and software. These issues need be considered further in future embedded software design and development.

3.5. The effectiveness of fault tree and design failure mode effect analysis (DFMEA)

According to the zero mileage and warranty issue, we realize we need a vehicle/system DFMEA and fault tree analysis to catch the potential failure modes at the vehicle/system level. In this way, it would prevent a lot of vehicle/system level zero mileage and warranty issues (CAN bus short, loss CAN communication etc.). This vehicles/system fault or failure modes analysis could be done by OEMs engineers. Dare Auto system engineers need contact OEMs engineers to prevent this kind of vehicle/system level issue. Because of the vehicle/system CAN bus issues, the EHPS pumps were replaced without any sympathy since the driver feels there is non-assist from power steering. The EHPS pump is the victim of the "shut off the pump" requirement when the CAN is off (section 3.2).

Also the accuracy of the zero angle issues happened dozens of time during EHPS pump assembly. This caused the current and pump RPM were not accurate and out of the specification. The anomaly which we have experienced is to find the accurate zero-angle for the BLDC motor after motor assembly. This is caused by using a universal controller after motor assembly to find the zero angle, but after the whole EHPS assembly, the zero angle was changed since the new controller and motor are assembled together. In our software/system DFMEA and manufacturing PFMEA development, we need consider the accuracy of the parameters such as zero angle which could cause the performance degradant of EHPS pump. Hence we need require the motor/pump manufacturing to design and implement the correct processes in order to achieve the accurate parameters.

3.6. Software Reliability Growth & design change

We track the software reliability by using the software reliability growth matrix and four different levels of testing (Figure 11). The reliability performance of the embedded software in the zero mileage and warranty needs be updated in the reliability growth matrix. This will help us for future embedded software motor control project.

When we launched this product, the thorough reliability test had been performed in different design and development both in component and system level. Therefore, the software remained stable and without issues for a year. The software reliability is decreased when the OEMs requests a change. Generally, the software reliability after a change follows the shark teeth curve, not the bathtub curve [4]. After the first DCR, the software is considered suspect until its reliability is proven. For example, the OEMs requested a change in the DTC after the launch. Although, the DCR was successfully developed and validated according to our process, the requested change was found to be somewhat incompatible with the OEMs production line equipment. Under certain conditions at the factory, the DTC was set. The OEMs plant could not control the environment. This negatively impacted our reliability score. The lesson learned for the above issue is to certify the software only after the completion of the PER (Production Evaluation Run) at the OEMs factory.

Due to the customer requirements changes for the pump speed maps and fan motor speed requirements which mentioned in section 3.2, we had dozens of DCRs from customers for the software changes. One lessons learned is to cascade the customer CNs (change notice) to the cross functional team, the kickoff meeting is necessary to let every team members to understand the change request, and to implement it and validate it without any discrepancy.

4. Effectiveness and suitability of testing

4.1. Effectiveness of the component testing

As mentioned in section 3.2, some of the OEMs requirements were not clearly defined, this caused some of the system and software requirements were not well defined, hence the architecture design/coding and component testing were not well implemented.

4.2. Effectiveness of the SW integration testing

The comments of the effectiveness and suitability is the same as section 4.1.

4.3. The effectiveness of the sub-system level testing

The key output for the EHPS pump is to output the flow at a given pressure and pump speed. Each time we have the software release, we conduct the 5 points testing, the mapping test and pump bench testing (pressure vs. flow). The 5 points testing is described in the table below (Table 3). We also called the sub-system testing as "EHPS pump on the bench" test.

According to the zero mileage and warranty data, the simulated DTC code testing and the simulated CAN communication situation need be tested more at the system level. These need happen at EOL (end of line) test and system bench test. We have more work to do to define how we can simulated these testing without a vehicle in future.

4.4. The effectiveness of the vehicle level testing

A CAN bus is a robust vehicle bus standard designed to allow microcontrollers and devices to communicate with each other in applications without a host computer. It is a message-based protocol.

Without vehicle, it can only simulated the CAN message testing on the bench. But we do not have other electrical components on vehicle to communicate with. So the vehicle is needed in order to fully test the CAN network and communication timing. Some examples of tests include:

- CAN bus short
- Sleep/Wake up cycles

It is obvious that vehicle level CAN communication and electrical component harmony testing are needed according to the zero mileage and warranty data. *71%* EHPS pump replaced after warranty misdiagnosis by servicers is not acceptable due to the system or adjoining component related issues. The CAN and power mating connectors testing are needed before starting the production. In order to reduce the unnecessary cost of the warranty, the vehicle level test need be performed more by OEMs before starting the production.

5. Conclusion

The embedded software of the EHPS pump reliability growth has been modeled by using Crow-AMSAA method both in the cumulative timing and in the cumulative qualities delivered. The slope β value is < 1 which means the failure rate is decreasing and the embedded software reliability is growing.

Based on zero mileage and warranty performance & reliability data of the EHPS pump, the effectiveness and suitability of the reliability and test of the embedded software have been discussed in this paper. This is an extension of the 2020 RAMS paper [1]. Using the real field data to evaluate the effectiveness and suitability of ASPICE DFR is beneficial for the future embedded software/system design and development. The lessons learned have been documented and evaluated to improve the software reliability management and test processes. The ASPICE processes require the organization to collect the data to evaluate the effectiveness and suitability of each process. This paper summarizes the findings.

The main lesson is that the proper operations of the EHPS depend on the vehicle systems where it is installed. Our engineering staff need to involve the OEM's system engineering and service staff in order to prevent misdiagnosis of issues in the field that lead to unnecessary replacement of the properly operating units in the field.

As expected, the OEMs issue the change requests after the units have been in the field. The flexibility of the software and the regression testing reduce the occurrence of the shark tooth defects.

In summary, the paper is about the reliability of the EHPS pump software. Through multi-staged /multi-facet software testing procedures and methodologies, the zero-mileage failure rate of software has been significantly dropped. The rate has been dropped about *80%*. The zero-mileage issues are due to the failure of the process principally, (a) Incorrect software version loaded, (b) Incorrect calibration, (c) Incorrect process at the OEMs plant. The warranty issues are caused by mismatch of the specified behavior versus the real world conditions and the incorrect service instructions.

Table 1: List of zero mileage issues

Zero Mileage issues	**Total**
Unnecessary DTCs	10+
Power circuitry failures	2
Coupler broken	1
Inlet blockage	1
Residual oil	10+
Oil cap damaged	10+
Repeated bar code	4
Connector pin damaged	10+
Noise pump	2
Missing screws	2

Table 2: System failures due to software at different phases. The cumulative timing (total days), the cumulative quantities delivered, cumulative failures (due to the software) and cumulative reliability are documented in this table.

	Period		**Cum Days**	**Cum Quant Delivers**	**Cum Fail.**	**Cum Reliability**
	Starting	**Ending**				
I	1/20/2014	3/2/2015	406	168	32	0.809524
II	3/3/2015	4/29/2016	830	586	44	0.924915

III	4/30/2016	11/1/2017	1381	1964	49	0.975051
IV	11/2/2017	3/27/2021	2623	601964	69	0.999885

Table 3: 5 points test

Function and Performance 5 Points Test			
	T low (-40°C)	T normal (23°C)	T high (105°C)
V low (9V)	(5)		(5)
V normal (13.5V)		(5)	
V high (16V)	(5)		(5)

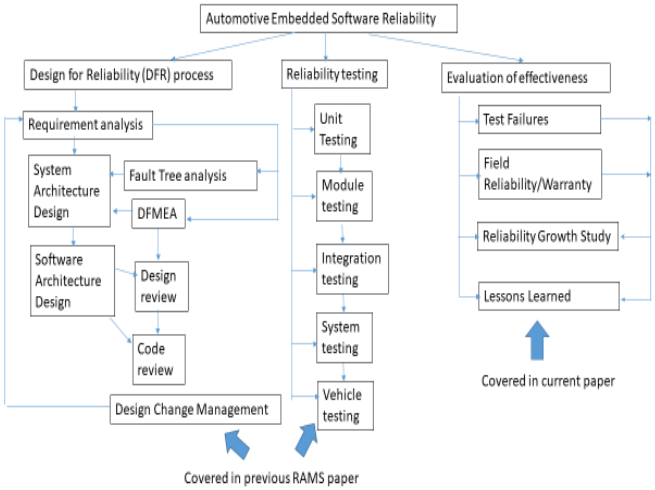


Figure 1: Flow chart of methods for automotive embedded software reliability. The DFR process and reliability testing were covered in previous RAMS paper. The evaluation of effectiveness is covered in current paper

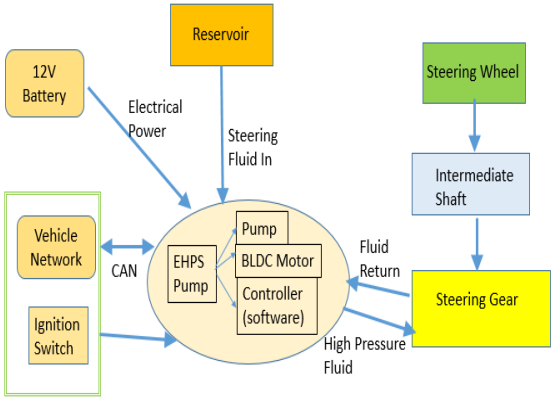


Figure 2: EHPS pump and its interface

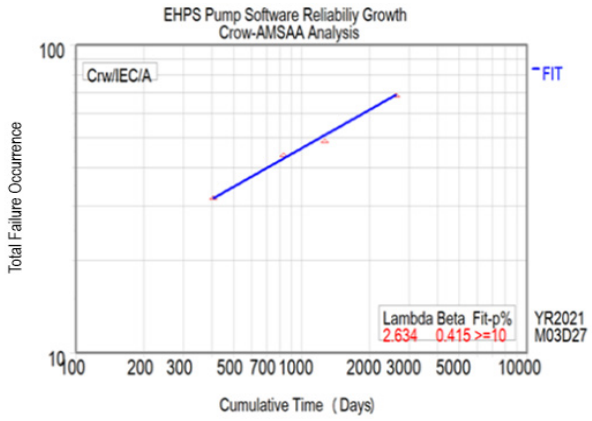


Figure 3: Crow-AMSAA Analysis (total failures vs. total days). The X axis is the total accumulative days since the project starting. The Y axis is the total accumulative occurrence of failures (due to software).

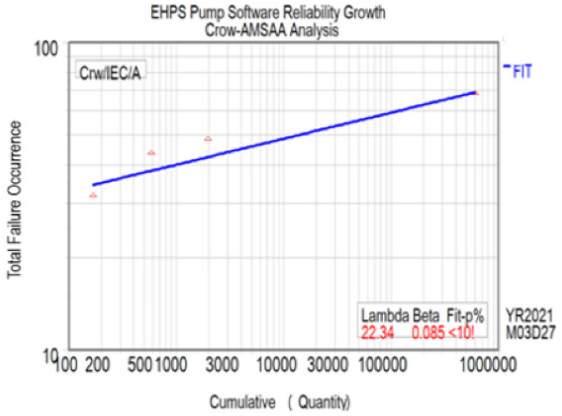


Figure 4: Crow-AMSAA Analysis (total failures vs. total quantities). The X axis is the total accumulative quantities delivered since the project starting. The Y axis is the total accumulative occurrence of failures (due to the software).

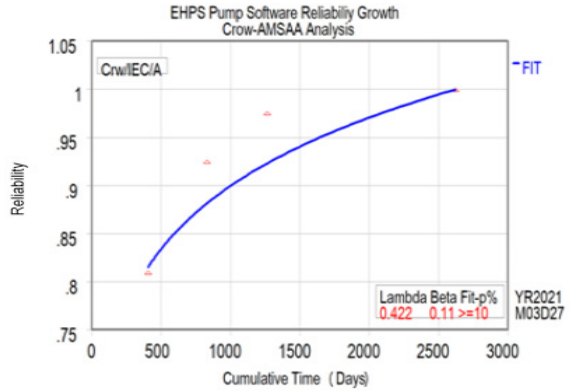


Figure 5: Crow-AMSAA Analysis (total reliability vs. total days). The X axis is the total accumulative days since the project starting. The Y axis is the total software reliability. (This is not the log to log plot).

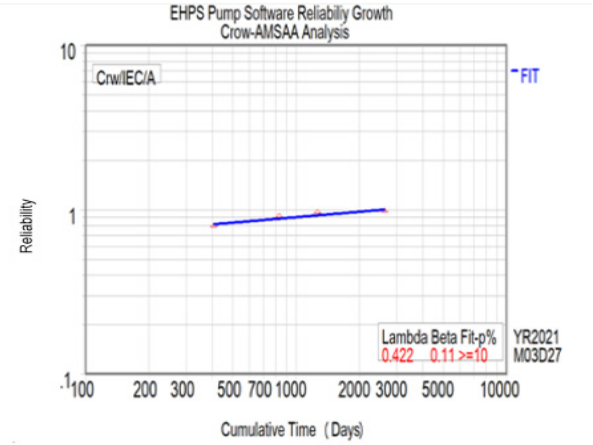


Figure 6: Crow-AMSAA Analysis (total reliability vs. total days). The X axis is the total accumulative days since the project starting. The Y axis is the total software reliability. (This is the log to log plot).

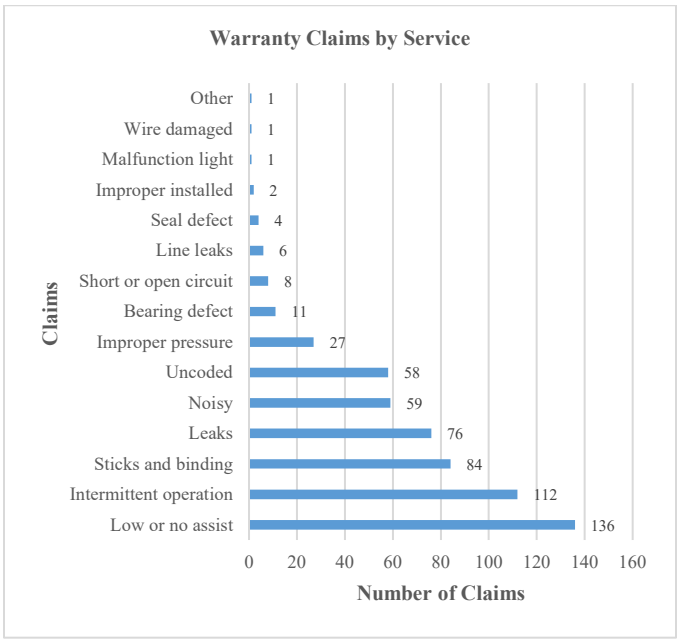


Figure 7: EHPS pump warranty complains by customer

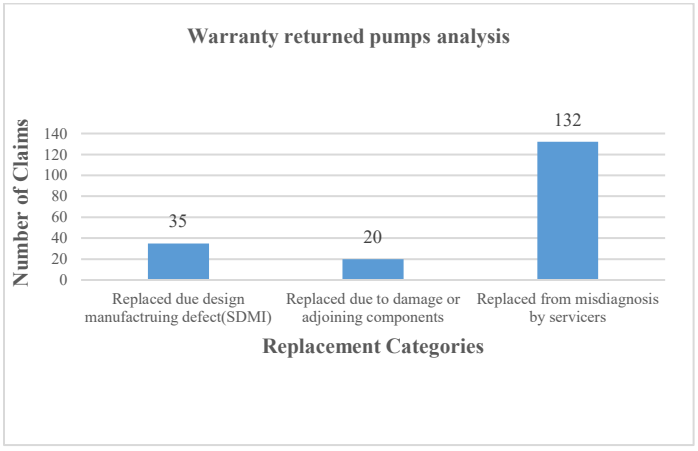


Figure 8: EHPS pump warranty replacement categories

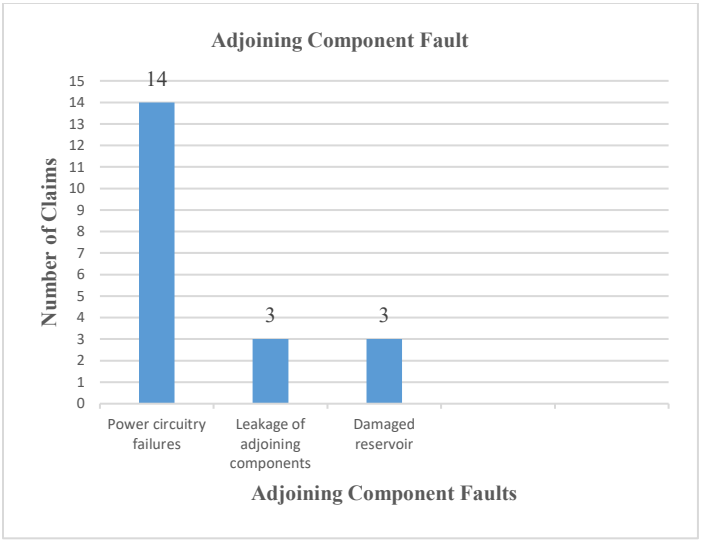


Figure 9: Warranty adjoining component faults

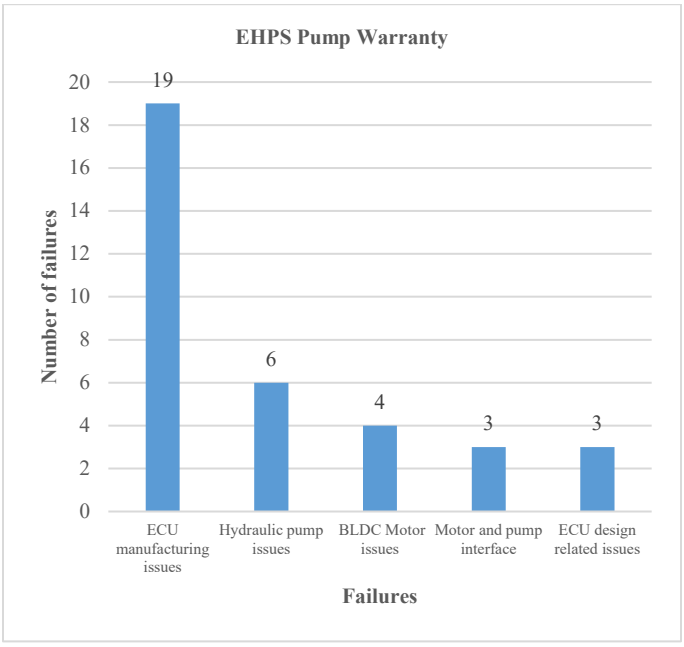


Figure 10: Total EHPS pump supplier design manufacturing issues

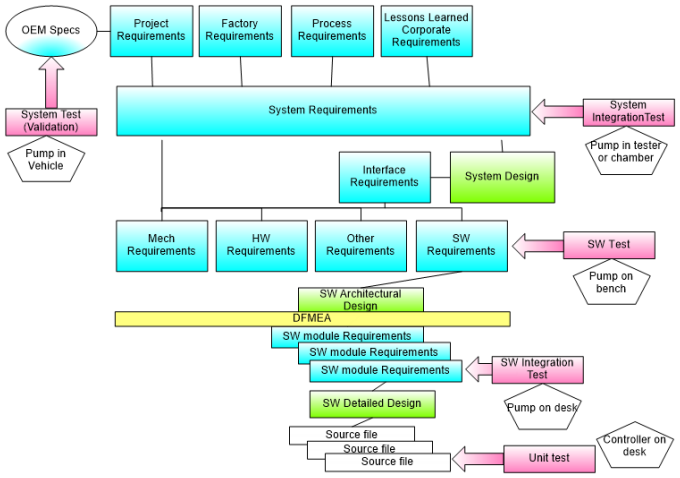


Figure 11: Requirements vs. test

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Thanks to the colleagues at Dare Auto Inc. in EHPS pump team for providing the data and implementing the embedded software, ECU, BLDC and pump. Also, thanks to the reviewers of the ASTES Journal for valuable advices for this paper.

References

[1] Y. Wang, J. J. Yang and N. M. Mbiye, "An Automotive EHPS Software Reliability and Testing," 2020 Annual Reliability and Maintainability Symposium (RAMS), Palm Springs, CA, USA, 2020, 1-6, doi: 10.1109/RAMS48030.2020.9153725.

[2] M. Wolf, "Embedded Software in Crisis," in Computer, vol. 49, no. 1, pp. 88-90, Jan. 2016, doi: 10.1109/MC.2016.18.

[3] N. Navet, F. Simonot-Lion, (Editors) "Automotive Embedded Systems Handbook, CRC Press , 2009 by Taylor & Francis Group, LLC.

[4] J. Pan, "Software Reliability," Carnegie Mellon University, 18-849 Dependable Embedded Systems, Spring 1999, https://users.ece.cmu.edu/~koopman/des_s99/sw_reliability/index.html

[5] M. Silverman,"Software Design for Reliability," George de La Fuente, Ops A La Carte LLC, 2008 Annual Reliability and Maintainability Symposium (RAMS).

[6] VDA QMC Working Group 13 / Automotive SIG, "Automotive SPICE Process Assessment / Reference Model", Version: 3.1, Date: 2017-11-01, Revision ID: 656, http://www.automotivespice.com/

[7] Y. Shi, P. Kalia, J. Evans, A. DiVenti, "An integrated life cycle-based software reliability assurance approach for NASA projects," 2010 Annual Reliability and Maintainability Symposium (RAMS).

[8] S. Saxena1, D. Agarwal, "A Systematic Literature Review on Software Reliability Estimation Model for Measuring the Effectiveness of Object Oriented Design," 2017 International Journal of Advanced Research in Computer and Communication Engineering, ISO 3297:2007, **6**(6), June 2017, doi:10.17148/IJARCCE.2017.6671

[9] R. K. Sharma, A. Kumar, "Analysis of Various Software Reliability Models and Proposing a New Model of Software Reliability for Embedded Systems," International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, **5**(3), May 2017 DOI: 10.21276/ijircst.2017.5.3.6

[10] A. Chunduri, "An Effective Verification Strategy for Testing Distributed Automotive Embedded Software Functions," Master of Science, Faculty of Computing Blekinge Institute of Technology, SE–371 79 Karlskrona, Sweden, 2016. Thesis no: MSSE-2016-08

[11] N. Mellegard, "Improving defect management in automotive software development," Ph.D thesis, Department of Computer Science and Engineering, Division of Software Engineering, Chalmers University of Technology | University of Gothenburg, Göteborg, Sweden 2013.

[12] C. Kugler, S. Kowalewski, J. Richenhagen, R. Maquet, A. Schloßer (2017) "Metrics-based strategies for quality assurance of automotive embedded software." In: Bargende M., Reuss HC., Wiedemann J. (eds) 17. Internationales Stuttgarter Symposium. Proceedings. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-16988-6_56, First Online08 April 2017, DOIhttps://doi.org/10.1007/978-3-658-16988-6_56

[13] H. Shokry and M. Hinchey, "Model-based verification of embedded software," Lero—the Irish Software Engineering Research Centre, Published by the IEEE Computer Society APRIL 2009, DOI:10.1109/MC.2009.125, Corpus ID: 17393996

[14] Z. Song, Y. Wang, P. Zong, L. Wang, G. Feng, and W. Kang, "An Empirical Study of Comparison of Code Metric Aggregation Methods and Software Reliability Evaluation," 2020 Annual Reliability and Maintainability Symposium (RAMS), Palm Springs, CA, USA, 2020, DOI: 10.1109/RAMS48030.2020.9153606

[15] S. Dahab, E. Silva, S. Maag, A. S. Cavalli and W. Mallouli, "Enhancing Software Development Process Quality based on Metrics Correlation and Suggestion," In Proceedings of the 13th International Conference on Software Technologies (ICSOFT 2018), pages 120-131, DOI:10.5220/0006856501540165

[16] N. Miglani, "On the Choice of an Appropriate Software Reliability Growth Model," International Journal of Computer Applications (0975 – 8887), Volume 87 – No.9, February 2014, DOI:10.5120/15237-3772

[17] T. Chu, M. Yue, G. Martinez-Guridi, and J. Lehner, "Review of Quantitative Software Reliability Methods," Brookhaven National Laboratory Letter Report, Digital System Software PRA, JCN N-6725,September 2010

[18] B. Li, C. Smidts, "Effects Assessment for Requirements Faults of Safety Critical Software in Nuclear Industry," 2020 Annual Reliability and Maintainability Symposium (RAMS), Palm Springs, CA, USA, 2020. DOI: 10.1109/RAMS48030.2020.9153594

[19] H. Sthamer, J. Wegener and A, Baresel, "Using Evolutionary Testing to improve Efficiency and Quality in Software Testing," 2002 In Proceedings of the 2nd Asia-Pacific Conference on Software Testing Analysis and Review (AsiaSTAR)

[20] L. H. Crow, "Reliability Growth Planning, Analysis and Management," Tutorial, 2011 Annual Reliability and Maintainability Symposium

[21] R. B. Abernethy, "The New Weibull Handbook," Fifth Addition, , 2006

[22] D. E. Tananko, "Practical Guide to Reliability Growth Analysis,." Tutorial at 2020 Annual Reliability and Maintainability Symposium, Palm Springs, CA, USA, 2020,

[23] IEC 61164 Ed. 2.0 en:2004, "Reliability growth - Statistical test and estimation methods," International Electrotechnical Commission, 2004.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 208-217 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Heuristic Techniques as Part of Heuristic Methods and Interaction of Personality Types in their Application

Viktor Ivanov*,[1], Lubomir Dimitrov[2], Svitlana Ivanova[3], Olena Olefir[4]

[1]*Odessa National Polytechnic University, Department Mechanical Engineering and Elements of Machine, Odesa, 65044, Ukraine*

[2]*Technical University of Sofia, Department Mechanical Engineering, Sofia, 1000, Bulgaria*

[3]*Department Mathematics and its Teaching Methods, South Ukrainian National Pedagogical University named after K.D. Ushynsky, Odesa, 65020, Ukraine*

[4]*Department Advanced Mathematics and Statistics of Affiliation, South Ukrainian National Pedagogical University named after K.D. Ushynsky, Odesa, 65020, Ukraine*

ARTICLE INFO

Article history:
Received: 08 November, 2020
Accepted: 24 December, 2020
Online: 15 January, 2021

Keywords:
Heuristic technique
Personality types
Generalized heuristic method

ABSTRACT

The widespread team project method is more effective when used in conjunction with heuristic methods. The large number of heuristic methods and the variety of their descriptions create a problem to prepare students for the use of these methods. A method based on two areas of knowledge - heuristics and psychology - is proposed. The personality types of students STEM specialties according to Myers-Briggs are considered. An analysis of interaction of personality types from the point of view of application of heuristic methods is performed. The survey for percentage composition personality types of student STEM specialties was carried out and predominantly types of student STEM specialties was determine. Heuristic methods are consideration as sum of heuristic techniques and procedure. It is shown that many methods involve the same heuristic techniques and differ only in procedures. A generalized method has been developed that allows replacing most of the methods based on collective discussion. This method included five heuristic techniques: collective discussion, pause between the presentation of ideas and their criticism, random associations, analogy, expert evaluation, using a matrix. This method is mainly aimed at teaching students of STEM specialties. A project team is formed to use the method. The composition of this team includes a discussion group, a criticism group and a expert evaluation group. These groups are formed in accordance with the personal types of participants. The method includes an algorithm for team members to interact when using heuristic techniques and procedures.

1. Introduction

Research in the field of heuristics is carried out within a many scientific disciplines. Different approaches and terminology are used. In philosophical works, heuristics is understood as a science, the subject of which is the process of solving a problem in conditions of uncertainty. Another aspect of heuristics is the branch of the science of thinking. Different interpretation is related to pedagogy, so heuristics are often understood as a way of learning. The concept of heuristic is also widely used in cybernetics, where it is interpreted as a heuristic algorithm. Researchers estimate the number of existing heuristic methods in different ways - from several to hundreds. There is no unambiguous idea about the structure of the heuristic method. In different scientific disciplines, a set of principles, a set of steps, or a set of heuristic techniques are considered to constitute a method. First of all, heuristic methods aimed at enhancing the creative abilities of students and involving them in collective creativity are relevant for teaching. The number of such methods is quite large, but there are many common features in the structure of these methods. When it comes to collective creativity, the question of

*Corresponding Author: Viktor Ivanov, ivv@opu.ua

https://dx.doi.org/10.25046/aj060123

performers immediately arises. How do personality differences between people affect their use of heuristic methods and how do these differences affect the organization of their interaction? This paper, that is an extension of work originally presented in International Conference on High Technology for Sustainable Development, is dedicated to solving this problem [1].

Consideration of the application of heuristic methods is associated with the personal differences of performers in two aspects: the effectiveness of the application of methods and teach students to use heuristic methods. Many heuristic methods are focused on collective creativity. Methods of this type are well suited for use in student learning. Heuristic methods are most widely used in teaching students studying the following specialties: science, technology, engineering and mathematics (STEM). When using heuristic methods formed by a group of students working together. The interaction of performers in such a group is well studied in project management. This kind of group of performers is called a project team. A large number of publications on psychology, management and education are devoted to organizing the work of the project team and the selection of performers. The project team can be organized for one seminar, or can work on a course project or diploma project during a semester or a year.

2. Review of heuristic methods

As methods of generating ideas used in teaching, the following are considered: Brainstorming, Method 6-3-5, C-Sketch, Gallery, Brainsketching. These methods differ in the form of communication: verbal - brainstorming; text - method 6-3-5; graphic - method C-Sketch; sketch, discussion and text - Gallery; sketch and text - Brainsketching [2]. In the 6-3-5 method, six participants write 3 ideas on a piece of paper in 5 minutes. Then they exchange their notes. This method resembles individual brainstorming in that each participant generates ideas individually, as well as electronic or online electronic brainstorming in which participants exchange ideas via e-mail. In Brainsketching, participants present their ideas in the form of sketches and then exchange them. As in the 6-3-5 method, the criticism of ideas is carried out without the participation of those who put forward the ideas. It is assumed that the graphical representation of ideas will help to find associations. In the Gallery method, the sketches created are discussed and selected together. In the C-Sketch method, the sketch is an imaginary construction. This method is targeted at mechanical engineers.

An experiment comparing Brainsketching and Brainstorming methods has been carried out. The experiment showed that during the Brainstorming process, the participants produced more ideas, and during Brainsketching, the ideas were more original. It is noted that Brainsketching is a type of Brainwriting. That is, the participants put forward ideas individually, as in the 6-3-5 method, only in the form of a sketch [3].

The methods can be divided into three groups [4]. Methods using graphic presentation of information: C-Sketch and Gallery. The methods of using the text include: Brainstorming, K-J Method, Checklists, Affinity Method, Storyboarding, Fishbone, and Method 6-3-5. In addition to the above two groups of methods, another group of Cross-representation methods is proposed [4].

In the Storyboarding method, the sketch is not static but is created in the presence of the discussion group members. The purpose of a sketch or series of sketches is to represent the origin and development of an idea. Affinity Methods operate with different types of information presentation: text, photo, sketch, etc. The group discussion task is sorting similar information into clusters. Affinity Method is, as many believe, a variation of the K-J method. In the Fishbone method, as a result of collective discussion, the factors are also grouped into clusters; only the grouping is visualized as a single diagram.

Cross-representation methods include Morphological analysis and Synectics [4]. Synectics uses collective discussion, the prohibition of criticism as in Brainstorming. Synectics encourages the use of analogies and introspection of the idea generation process. Morphological analysis requires the identification of a set of functions and a possible set of design solutions or physical principles. For this, the analogy of physical processes is widely used [5]. The comparison of possible combinations of properties and design solutions is performed using a matrix. Another name for the method is morphological matrix [5]. Note that the same method has different names Morphological analysis and Morphological matrix. The Affinity method is also called the Affinity Diagram; some researchers describe it as a separate method, others as another name for the K-J method. Brainwriting is an individualized brainstorming session and is very similar to the 6-3-5 method. The peculiarity of publications on heuristics is the use of different names of methods and different points of view on their classification, which is partly due to the authors' affiliation to different scientific disciplines, as well as different ways of presenting information by representatives of the humanities and engineering disciplines.

The comparison of the methods Brainwriting and S.C.A.M.P.E.R. was carried out [6]. In S.C.A.M.P.E.R. method the ideas are generated based on responses to prompts: Substitute, Combine, Modify, etc. The study did not obtain data that could indicate the superiority of one of the methods. It is noted that more creativity can be achieved if the participants are from different universities and different countries. This can be interpreted as a recommendation for using electronic Brainstorming.

The comparison of methods S.C.A.M.P. and TRIZ are performed; also data about Brainstorming was used [7]. It is noted that Brainstorming allows for more creative ideas, but S.C.A.M.P.E.R. and TRIZ provided ideas that could be implemented. Method TRIZ allowed getting more creative ideas than the S.C.A.M.P.E.R method. The S.C.A.M.P.E.R method is presented as intuitive in contrast to the logical TRIZ method.

It should be noted that TRIZ is aimed at solving inventive tasks, that is, it is focused at engineers. More precisely, the name is TRIZ methodology, not a method, since TRIZ contains a set of methods. Most of the TRIZ methods are not methods, but are recommendations for a mechanical engineer, for example: replace sliding friction with rolling friction or replace rectilinear motion with rotational. At the same time provides a method of analogy and decomposition.

A comparison of the methods used in the educational process is carried out; these methods include: Brainstorming, Method 6-3-5, Morphological matrix, Mind and Brain Mapping, Patent

Research, AIDA, Function Structure, House of Quality, Black Box, Decision, Matrix and Pugh Chart. The selection of methods raises questions - how and for what purposes these methods are used in the educational process [8]. The following design stages are highlighted: Task planning and clarification, Conceptual design, Embodiment design, Detailed design [9]. The advantages of the AIDA method are noted, which can be used both at the Task planning and clarification stage and at the Conceptual design stage. The possibility of using one AIDA method instead of using a number of methods is being considered: Brainstorming, Black Box, Morphological matrix, Decision Matrix and Pugh Chart [8]. Note that the feature of the AIDA method is the construction of an option graph, and then the established relationships are analyzed in the matrix. In this part, the method is similar to the Morphological matrix and Decision Matrix methods.

There are different ideas about the stages of the creative process and the stages of the engineering design process [10]. A large number of classifications have a lot in common. Most researchers indicate Analysis phase as the first stage or phase, and Generation phase as the second stage, which is also called Idea-finding, Generating ideas. Inspiration, Response generation. The third stage is the Evaluation phase, which is also called Looking Back, Solution-finding, Reinterpretation, Response validation. The fourth stage is Communication / implementation or Acceptance-finding, Developing an implementation plan [10].

Taking into account the division of the creative process into stages, it is possible to assess the applicability of various methods to these stages. Undoubtedly, of the above methods used in teaching for Generation phase, methods of collective creativity are suitable: Brainstorming, Method 6-3-5, S.C.A.M.P.E.R., Brainwriting, Brainsketching, C-Sketch и Gallery. The mind and brain map is suitable for use by a teacher. If it is used by students, it is transformed into methods with a graphical representation of the idea - Brainsketching, C-Sketch and Gallery.

The methods K-J Method, Checklists, Affinity Method, Storyboarding, Fishbone, AIDA are more suitable for the Evaluation phase, although they are also used in the second stage. The Morphological matrix and Decision Matrix methods are ideal for the third stage. In the fourth stage, Pugh Chart and Checklists can be used. Decision Matrix also has another name Decision support matrix (DSM) this method is aimed at solving engineering problems. The method DSM allows to establish the presence or absence of interactions between the nodes of the product. Based on the found interactions, elements are combined into a cluster [11]. Another method using a matrix is the Interaction Matrix method. This method also is aimed at identifying the interactions between structural elements and their systematization [12].

The TRIZ methodology contains a set of methods that are suitable for each stage of the design process. At the same time, the core of the method is also a matrix - a matrix of contradictions. The TRIZ method involves the formation of a matrix, in which the components and elements of the product are located horizontally, and functions or parameters — vertically. The contents of the matrix cells are a list of methods that can be used to resolve technical contradictions [13]. To this series of methods that use a matrix should be added - a Matrix Diagram (Quality Table). The matrix includes elements between which it is necessary to establish the interaction. The values in the cells of the matrix indicate the presence and strength of the interaction [14].

3. Myers-Briggs personality types and creative process

The effectiveness of the application of heuristic methods largely depends on the selection of participants and the organization of their work. The selection of participants is especially important when the project team is being formed for a long time. The interaction of participants in the project team is described by the team role methodology. A team role is a description of a behaviour pattern that defines how one team member interacts with other team members when working to accomplish assigned tasks. In project management, the most widespread is the classification of team roles in accordance with [15]. Along with the Belbin typology, there are other typologies of team roles, including MTR-I Team Dynamics [16] and Margerison-McKenna [17]. In psychology, the typology of K. Jung and his followers is generally accepted. Thus, A. Augustinavichute proposed her original modification of K. Jung's typology, which was called "socionics" [18]. The most widely known indicator of the Myers-Briggs types (MBTI) [19]. The MBTI divides individual differences into four opposite pairs, leading to sixteen possible personality types. The sixteen types are usually designated by a four-letter abbreviation, the initial letters of each of their four preference types. For example: INTJ: Introversion (I), Intuition (N), Thinking (T), Judgment (J) or ESFP: Extraversion (E), Sensation (S), Feeling (F), Perception (P) [19, 20]. Likewise, from a combination of four letters, abbreviations of the remaining fourteen personality types are formed.

It is important to note that the role of an individual in the project team may not correspond to his innate individual differences as determined by the Myers-Briggs test. Work experience, acquired qualifications affect the role of the individual in the project team and the results of the MTR-i and Belbin tests. The personality type, in accordance with the Myers-Briggs typology, reflects the individual differences of a person, to a lesser extent dependent on external circumstances and experience, and to a greater extent - on the innate characteristics of the nervous system. Innate human abilities are important in terms of the spontaneous expression of ideas on which most heuristic methods are based. Therefore, this typology is most suitable for describing the heuristic abilities of project team members and their interaction in the collective generation of ideas, critical discussion and expert evaluation.

In the first place it is of interest the possible role of a certain type of students in the process of collective discussion. Apparently for electronic brainstorming it will be important to note that the intuitive interface is preferred [21]. It is also remarked that for extroverts, the response of other participants in the discussion is important in order to clarify the problem [21]. Participants who have the constructs N - Intuition and E - Extraversion [22] are useful for brainstorming. It is also indicated that it is advisable for students with a combination of ES constructs to be entrusted with case studies based on real facts. And students with a combination of IS constructs are preferable to be entrusted with a critical analysis of real facts and conclusions obtained by other students. In [23], personality types with a combination of IS constructs are characterized as a thoughtful realist, and the practical orientation

of their thinking is also noted; they are better at solving problems related to real life and based on real facts.

Students with a combination of IN constructs are distinguished by a critical analysis of new ideas, possibilities, and they can also formulate ideas expressed by other participants in the discussion [23]. Personality types with an IN combination are thoughtful innovators, they generate new ideas and penetrate deeply into the essence of the problem, they are most successful at generalizing facts in the form of a concept or theory [23]. The personal type EN shows itself best in new situations; it is an innovator, action-oriented and sees an opportunity to achieve success in challenges [22]. The same characteristic is given in EN [23], it pointed out that this personality type is suitable for joint execution of tasks, including joint generation of ideas during brainstorming.

In [24] give detailed characteristic NT - "Self analysis exercises and structured controversy are more likely to appeal to NTs, who focus on principles and abstractions". In [25] believes that "…NTJ, the type that tends to have the highest grades in engineering and is commonly drawn to engineering research". In [26] supported this characteristic ENTJ "they also have strategic ability, including reasoning, creative problem solving and strategic thinking". In [27] point out "…the student with the ENTJ personality types, it is preferable to choose for the role of a moderator".

The NP personality type likes to participate in brainstorming [21]. In [28] believes that "Sensitive-Thinking (ST) types preferred hard data and logical analysis in reaching decisions" and the "Intuitive-Thinking (IT) types preferred logic and to test several premises before coming to a decision". In [24] describes that "…NT and SJ individuals who tend to be more linear and serial, more structured, more rational and analytical, and more goal-oriented in their approach to problem solving." In [24] also describes personality types SP and SJ - "Role-playing and simulations are especially appreciated by SPs and SJs, who generally take a more practical approach to problem solving". For SP advisable it is an iterative process of determining a suitable decision through action. "SJs, care should be taken to proceed in a step-by-step, orderly manner, with ample time for consideration of all details at each step." In [25] add to the characterization of SP and SJ that "…students leave things to the last minute (SP), why some prefer a rigid schedule (SJ), why some are intellectual perfectionists (NT)".

In [26] gives characteristic ENFP as "get involved with the people with whom they work, and are very capable of and willing to reconsider plans when others offer input". For two types ISTJ and ENFP has been described problem-solving characteristics [21]. "In problem solving, ISTJ will want a clear idea of the problem (I) and attack it by looking for the facts (S) and by relying on a logical, impersonal (T), step-by-step approach in reaching conclusions" In [21] and [25] are expressed even more definitely "ISTJ engineering and technology stereotype". An important observation is made that ST and NT is cognitive pairs, they provide the exact opposite of the problem [25]. In [26] notes about ISTJ "meeting deadlines and budgets, ensuring productivity and accountability are strengths of ISTI personality".

ESTJ have a leadership qualities, they characteristics are authority and orientation to results [26]. In [18] supported this think, ESTJ is characteristic as administrator which "possess capability to assimilate large amount detailed information". It is noted that ESTJ is logical and rational; also he knows how to make decisions [18]. It is remarks that the set of INT and EST constructs corresponds to the developer using the heuristic TRIZ method [29]. In particular, ESTJ is suitable "determination of corresponding design principles and reviewing and understanding examples showing appropriate application" and INTP is suitable "abstraction to general TRIZ problem" [29].

4. The structure of heuristic methods

The presented overview of heuristic methods does not cover all existing methods. Consideration of only a part of the methods showed that the same method can have different names, such as Affinity Method and K-J method. Either they may differ a little, like Brainwriting and Electronic Brainstorming. The excretion of the characteristic features of methods or their structural units can significantly reduce the number of these methods. From our point of view, the basis of such an analysis can be the excretion of heuristic techniques and procedures from the composition of method. For example, Brainsketching and Method 6-3-5 use the technique of random association. These associations are called through various procedures in the Brainsketching, this is a sketch of another participant, and in the Method 6-3-5 and Brainwriting, these are notes made by another participant. Technique of random associations is used in methods: Brainwriting, Method of focal objects, Method of garlands of accidents and associations, S.C.A.M.P.E.R., Lateral thinking, etc.

The number of heuristic techniques is much less than the number of methods. And within one stage, only a few techniques are used. Let us consider the structure of the methods used in the "Generation phase" stage from the point of view of the applied heuristic techniques. Methods: Brainstorming, Brainwriting, Brainsketching C-Sketch, Gallery, Method 6-3-5, S.C.A.M.P.E.R. be sure to use the collective discussion technique. An important feature of the Brainstorming method is the use of the technique "Pause between the presentation of ideas and their criticism". This technique is also used in methods: Garlands of accidents and associations, Gallery and Lateral thinking. Many methods are based on the use a technique of analogy. These are the following methods: Synectics, Analogy, Empathy and S.C.A.M.P.E.R., which with their questions force the participants to look for similar solutions. In the K-J Method, information is sorted on the basis of analogy, although the use of random associations is also possible. In most methods, the final selection of ideas is done collectively. At the same time, the participants act as experts and their work can be organized on the basis of well-known expert evaluation procedures.

A powerful analysis tool is a heuristic technique «Use of matrix". In different versions, this technique is used by methods: Morphological analysis, DSM matrix, Interaction matrix, Matrix diagram, AIDA. First of all, the methods are selected that are suitable for use in the "Generation phase" stage, are widely known and contain, if possible, several heuristic techniques. These methods are summarized in the table to analyze their structures (Table 1).

Table 1: Structure of heuristic methods

Heuristic techniques / Heuristic methods	Collective discussion	Pause between the presentation of ideas and their criticism	Random associations	Analogy	Expert evaluation	Using a matrix
Brainstorming	■	■			■	
Brainwriting		■			■	
Method 6-3-5	■		■		■	
S.C.A.M.P.E.R.	■			■	■	
Brainsketching	■		■		■	
Gallery	■	■	■		■	
Storyboarding	■		■			
C-Sketch	■		■		■	
Lateral thinking		■	■	■		
Method of focal objects	■		■			
Method of garlands of accidents and associations	■	■	■	■		
Synectics	■			■	■	
Affinity Method	■		■		■	
Lateral thinking		■	■			
Morphological Matrix						■
DSM matrix						■
Interaction matrix						■
Matrix diagram						■
AIDA	■				■	■

To develop a generalized heuristic method aimed at solving problems within the framework of the stage, we formulate a theorem.

Theorem: for each stage of the creative process one generalized heuristic method can be applied replacing all methods used for this stage if it includes all heuristic techniques used within this stage.

The generalized method for "Generation phase" stage includes the following heuristic techniques: collective discussion, pause between the presentation of ideas and their criticism, random associations, analogy, expert evaluation, use of matrix. The use of all these techniques in one method is mandatory. The order of using the techniques and the procedures applied in the method may be different depending on the branch of knowledge and the problems being solved. For example, the “Collective discussion” technique can be applied twice. Then one gets a variation of the brainstorming method - double brainstorming. The Brainsketching and Gallery methods use the “Random association” technique. These methods use different procedures to enhance creativity. In the Brainsketching method, participants can flip the image upside down, and in the Gallery method, they can observe the sketching process. The methods of the DSM matrix and the Matrix Diagram use the same heuristics "Using a matrix" and different procedures. In the DSM matrix, the procedure is to establish the presence or absence of an interaction between elements, and in the Matrix Diagram, the procedure is to assess the conformity of an element to certain requirements.

Let's point out the procedures typical for the methods used in teaching. These are procedures: selection of a moderator, selection of participants in a discussion group, selection of members of a group of critics, selection of members of the group of expert evaluation. Also, the procedures are - defining the rules of discussion and the method of forming an expert evaluation. Important procedures are the choice of the physical meaning of the columns and rows of the matrix and the rules for evaluating combinations of row and column elements.

Based on the analysis of heuristic techniques and procedures, it is possible to form an algorithm of the generalized method. This method must be suitable for use in education, also for design. That is, it should be a method of increasing creativity and at the same time a method suitable for use in the design process. The latter is important for students engineering specialties. To combine such an algorithm and data on the predisposition of students of different personality types to the use of heuristic techniques and procedures in the generalized method, it is important to establish what personality types are typical for students studying in STEM specialties.

5. Typical personality types of students STEM specialties

It is known that there are more introverts than extroverts among students studying STEM specialties. Thus, among students studying agricultural engineering, 83% have construct I, and only 17% have construct E [21]. In the ratio of the S and N constructs, there is no unambiguity: the construct S prevails in mechanical engineering (64 S%, 36% N), while for aerospace engineering the

construct N prevails. The most obvious is the predominance of the construct T over F and J over P [21].

It is claimed those students with the following personality types: ENTJ, INTJ, INTP, ESTJ, ISTJ, ESTP and ISTP, have the ability to work in engineering positions [30]. The personality type of ISTJ is especially highlighted, which is predominant among students of engineering specialties and makes up 26% of their total number [26]. Based on research carried out in [32], it is also stated that the personality type ISTJ occurs most often, the second most common is one of the following personality types ESTJ, INTP and INTJ. Testing of students according to typology (MBTI) was conducted in eight universities [21]. The percentage of students of a certain personality type is established, it is indicated in brackets. The study showed that personality types: ISTJ (16.46%), ESTJ (12.75%), ENTJ (9.43%), INTJ (9.43%), INTP (8.46%) and ENTP (7.43%), most often found among students studying STEM disciplines [21]. In total, personality types ISTJ, ESTJ, INTJ and INTP make up more than 50% of the students who passed the test, and personal types ESFP, ESFJ and ENFJ are less common among students of STEM specialties [31]. The most detailed study of the representation of certain personality types among students of STEM specialties is presented in [10]. First of all, we note that the most widespread personality type is ISTJ (16.9%), and the second place is occupied by INTJ (12.3%), which corresponds to the research data [32]; not much less than INTJ is represented by ENTP (11.8%); for other personality types the following data were received: ESTJ (9.2%), ISTP (7.7%) and INTP (7.2%) [10]. Thus, only four personality types out of sixteen ISTJs, INTJs, ENTPs and ESTJs account for 50.2% of the total number of students. The data of another survey [33] also revealed the prevalence of ISTJ (18.1%), and the second place in prevalence was taken by ESTJ (10.3%), as in [21]. Other personality types ranked in the following sequence: INTP (9.4%), INTJ (8.5%), ISTP (8.2%). It is noted that these five personality types make up more than half of the students [33]. It is noted that these five personality types account for more than half of the students [33]. Similar data were obtained in [32] - the following personality types are prevalent among students of STEM specialties: ISTJ, INTP, ESTJ, INTJ, ENTP, ENTJ and ISTP. Obviously, in further research, it is necessary to obtain data on the possibility of participation of precisely these personality types at different stages of heuristic search. According to the sources discussed above, the following personal types are typical for students STEM professions: ISTJ, ESTJ, ENTJ, INTJ, INTP и ENTP. It is also found that four personal types can account for more than half of the total. Three of them are ISTJ, ESTJ, INTJ, as the fourth, both INTP and ENTP are mentioned. For developing a heuristic method that uses personal types in a particular way, these facts are very important.

Therefore, a study of the distribution of personal types among students speciality "Mathematics and Physics" at the South Ukrainian National Pedagogical University named after K. D. Ushynsky (SUNPU) was carried out. The same test was performed for students speciality "Dynamics and durability of machines" Odessa National Polytechnic University (ONPU) in 2014-2016 years. In each semester, two academic groups numbering from 15 to 20 students were tested. The test was conducted twice, at the beginning and at the end of the semester. The coincidence of test results was observed in the range of 76–81%, which is within acceptable limits (75–90%) [34].

There were no students who refused to participate in the survey. But among the students of ONPU there were also those who did not take it seriously. Their questionnaires were not taken into account. Students of SUNPU treated the survey more scrupulously when the purpose of the survey was clarified - it was the selection of students into project team. They showed a great interest in testing, as it is directly related to their prospective teaching profession.

The distribution of personality types was calculated twice, first for all students, then separately, only for men majoring in "Dynamics and Durability of Machines" (ONPU), where the proportion of men is 87%, and only for women majoring in "Mathematics and Physics" "(SUNPU), where the share of women is 72%. The average value of data of the first and retesting were calculated. In that case, if the student equally corresponded to two personality types, for example, the constructs T and P scored the same number of points, then such people were taken into account twice, in each of the personality types. There were students who did not care about the test and gave mutually exclusive answers; such questionnaires were not taken into account in further calculations. The test results are summarised at Table 2.

Table 2: Personality types of ONPU and SUNPU students

Personality types	Students ONPU (%)	Students SUNPU (%)
ISTJ	22,1	18,1
INTP	11,9	15,5
ESTJ	11,7	14,7
INTJ	11,6	10,4
ENTP	7,8	8,4
ENTJ	5,3	4,4
ISTP	4,2	3,3
ENFP	2,0	2,5

ISTJ personality types are more common in ONPU, but INTP and ESTJ personality types are more common in SUNPU, as shown in Figure 1.

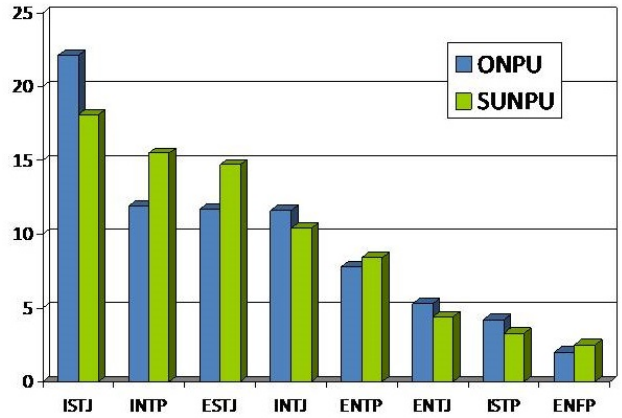


Figure 1: Percentage distribution of personality types

The relationship between personal types can be most clearly represented in the form of a pie chart in which the axes correspond to certain cognitive functions. The abscissa corresponds to the cognitive functions N and S. The construct N is located on the left, and the constructs S is located on the right along the abscissa. The

ordinate corresponds to the cognitive functions T and F. The construct T is placed below the abscissa, and the construct F above the abscissa is along the ordinate. The personality types within the inner circle are introverted; personality types outside the inner circle are extroverts. For each personality type, the average value of students of this type in ONPU and SUNPU was calculated: ISTJ (20.1), INTP (13.7), ESTJ (13.2). INTJ (11.0), ENTP (8.1), ENTJ (4.9), ISTP (3.8) and ENFP (2.3) (Figure 2).

The obtained average values of personality types were plotted on this pie chart. The size of the circle corresponds to the percentage of this personality type among all students.

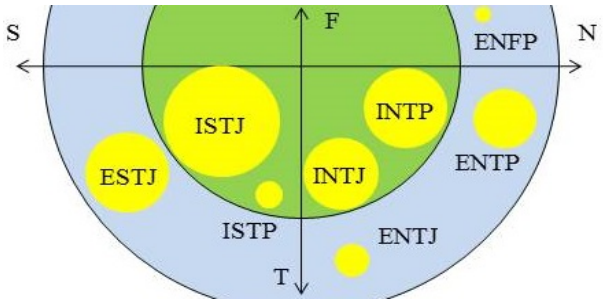


Figure 2: Distribution of personality types relative to the pie chart depicting cognitive functions

The small number of students surveyed does not allow making any conclusions about gender differences. The repeated survey for women in ONPU and men in SUNPU gave an error higher than acceptable. We cannot fully rely on the absolute value of a particular personality type. Because when surveying 50 students, one incorrectly completed questionnaire gives an error of 2%. The data of students at six technical universities in the United States and Canada shows that only one university has the represented ENFP type among students [32]. For this reason, we excluded the ENFP personality type from further consideration because the data obtained correspond to one representative from two academic groups.

We obtained the following order of personal types by their prevalence: ISTJ, INTP, ESTJ, INTJ, ENTP, ENTJ and ISTP. The first is ISTJ, as in the sources [21, 26, 31]. The first six personal types are the same as in [21], but their order is different. The resulting distribution of personal types is close to [33]. In our study, the INTP (13.7%) is in second place and the ESTJ (13.2%) is in third place. the difference between them is within the margin of error. In [33] type personal ESTJ (10,3%) at the second location, and INTP (9,4%) for the third, but the difference between them is small. In [31], also the personal type ESTJ (10.9%) is in second place, and INTP (9.9%) is in third place and the difference between them is about 1% as in the article [33]. The first four personal types in our study make up 58%, which is also more than half of the total number as in sources [31, 33]. According to our data, more than 75% of students STEM specialties belong to the following personality types: ISTJ, INTP, ESTJ, INTJ, ENTP, ENTJ and ISTP. The personality types in the table are arranged in descending order, in accordance with the number of representatives of this type (Table). Refer that the most common personality types include types containing the construct T. These personality types have a predisposition to STEM professions. The only personality type between of those which contain a construct T and which is ranked among the common - this ESTP. This type is also not mentioned among those common in papers [21, 33]. This is due to the fact that students of this type neglect the details and do not finish what they started. Equally represented are constructs Sensing (S) and Intuition (N).

6. The generalized method

The generalized method contains a heuristic technique *using a matrix*. Rows and columns of a matrix can be functions, units, a number of products with similar functions, design techniques, and so on. These recommendations are intended to solve engineering problems. The range of tasks solved by the technique *using a matrix* is much wider. It is proposed that the purpose of the technique *collective discussion* is to find ideas about the contents of the columns and rows of the matrix. After the first heuristic technique, there is a pause between the presentation of ideas and their criticism. In further, the criticism of the ideas put forward and their evaluation are providing. That is, the technique of *expert evaluation* is used. The semantic content of rows and columns is searched for by the technique of *analogy* and *random associations*. For example, for rows, the search is carried out by *analogy*, and for columns using *random associations*. If the search did not give satisfactory results, it is repeated, but the technique of *random associations* is already used for rows, and *analogy* for columns. The results are again *collectively discussed* and undergo *expert evaluation*. Based on the results of the heuristic analysis, the technique of *using a matrix* is applied - in this way the matrix is formed. Based on the results of the heuristic analysis, a technique *using a matrix* is applied – therefore the matrix is analyzed. Each cell of the matrix corresponds to a possible variant of a problem solution. These options are subject to *collective discussion* again. Then, *expert evaluation* and documenting of the results are performed.

The heuristic techniques and procedures described above require performers with a certain personal type. An effective *collective discussion* technique will not work if all participants are introverts. Extroverts with artistic abilities should not be entrusted with documenting matrix analysis results. The complete development of the method requires specifying the personality type in the performers for each of the techniques and procedures.

The ENTJ personality type has leadership qualities, strategic vision, logic, and therefore can be a moderator of the discussion. ESTJ - has organizational skills and the ability to keep track details of discussion. A student with the ENTJ personality type is selected as the moderator. If there was no such student in the group, then the student with the ESTJ personality type plays the role of the moderator. The moderator participates in the *collective discussion*.

Collective discussion is an extroverted process. Intuition is also important for generating ideas. In the process of collective discussion, there is a search for patterns, the construction of schemes and models, and a search for the relationship between known facts. It is important that the participants discard well-known dogmas, abandon stereotypes. Extroverted personality types (construct E) and intuitive personality types (construct N) take part in the *collective discussion*. Preference is given to personality types ENTJ, ENTP.

The logic is inherent by personality T. The personality NT is inventive, looking for regularities. The personality types INTJ and

INTP can generate ideas, however, they are introverts and the collective discussion is a problem for them. The personality type INTP is actively generating ideas; he has logic (T) and the ability to evaluate the logic of ideas expressed by other members of the team. The personality type INTJ is a researcher and has good intuition, but due to communication problems with others, he is not always possible to clearly express their ideas. The personality type INTP can be tactless and arrogant, which hinder a spontaneous expression of ideas. Based on the foregoing, the types of personality INTP perform the function of criticism.

The understanding an importance of the knowledge is inherent of a personality type TE. This quality, as well as extraversion, allows the personality type ESTJ to participate in the discussion, although ESTJ can consider the brainstorming ineffective. Based on the analysis of the above, students with personality types ENTJ, ENTP and INTJ are necessarily selected for the *discussion group*.

For the critical analysis of the proposed ideas, the following qualities are important: distrust, searching for flaws in everything, attention to detail, logicality. The critic is, without a doubt, an introvert type. The personality types ISTP, ISTJ, INTP, and INTJ are logical introverts. Critical analysis is the strong point of IS and IN. Sensitive-Thinking personality types ISTP and ISTJ prefer to focus on practical issues. These personality types prefer to carefully compare several options before making a final decision. In doing so, they rely on facts and logical analysis. Personality types INTP and INTJ are perfectionists that important for critics. Based on the foregoing, the students with the types of personality ISTP, INTP and ISTJ can function as critics.

Expert evaluation and critical analysis require similar personality types. Personality types ST and NT are a cognitive pair. For example, these are the ISTJ and INTJ personality types. They provide a different vision of the problem, this is important for the evaluation. The final decision on the meaning of rows and columns is made primarily by "experts" in the first place; these are representatives of the personality types ISTJ, and also ENTJ, ESTJ. The inclusion of personality types ENTJ and ESTJ in the number of experts is due to the fact that they led a collective discussion. The moderator can choose the final variant basing not only on his own opinion, but also on the opinion of other members of the team whose personal qualities he knows and appreciates. Therefore students with personality types ISTJ, INTJ, ESTJ and ENTJ can participate in the *expert evaluation*.

Note that the ENTJ and ENTP personality types have a good imagination and are useful when using the technique of random associations, and the INTJ personality type is endowed with perfect intuition. The ISTJ personality type is an ideal candidate for processing the results of the discussion and organizing them in the form of a morphological matrix of proposed solutions. The personal INTP type can do the job as well. It is also advisable to entrust the documentation of the final results obtained using the generalized method to the personal types of ISTJ and INTP.

If the proposed heuristic method does not work, a Feedback Group can be created. It is advisable to include representatives of each of the 16 personality types in this group, if students with such personal types are available. A group of students who performed a certain heuristic technique told the Feedback Group about their experience of participation, as well as problems that arose. After the discussion, all or part of the heuristic techniques are repeated.

Consider the algorithm of the generalized heuristic method, taking into account the heuristic techniques and the content of the considered procedures (Figure 3).

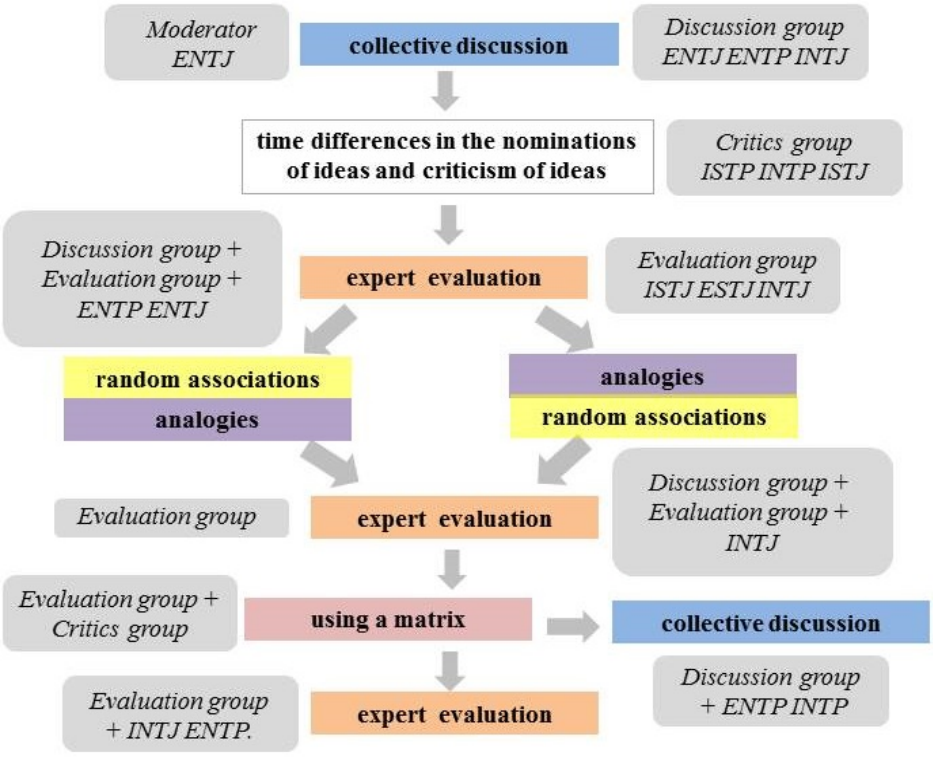


Figure 3: The algorithm of generalized method

The first procedure is procedure for selecting a moderator. A student with personality types ENTJ is selected as moderator for the collective discussion group. If there are two representatives of this type, then the student with higher academic achievements is chosen as the moderator. If there are no representatives of this personality type, then ESTJ is chosen as the moderator. The project team should be divided into several parts: discussion group, critics group, evaluation group. The one student can be by participant of few groups.

The second procedure is procedure for selecting participants for the discussion group; these are personality types ENTJ, ENTP and INTJ.

The third procedure for selecting the participants of critics group is as follows - students with the ISTP, INTP and ISTJ personality types are selected as critics group.

The fourth procedure for selecting the participants of expert evaluation group: personality types ISTJ, ESTJ, and INTJ.

This is followed by the first heuristic technique - collective discussion. The collective discussion group consisted of five to ten people. Personality type ESTJ participated in discussions and idea generation provided he is a moderator. This personality type joins the discussion group also in the case of its small size.

The second heuristic technique is a pause between the presentation of ideas and their criticism. After a pause, the group of critics set to work. A group of critics is consisted of three to five people. The composition of the group can be increased so that all members of the project team can participate in the work. In small project teams with a lack of "critics", the ISTJ is served as a critic; they usually predominate numerically over representatives of other personality types. In this case, the INTP is joined the group of discussion and generation of ideas. Several members of the collective discussion group join the group of critics. First of all, this is the personal type of INTJ, Thus, a critical analysis is carried out both by the project team members who participated in the discussion and those who did not participate in the discussion; they represent a different point of view on the problem.

The third heuristic technique is expert evaluation. The group of experts should not be large. Not all participants with personality types ISTJ, ESTJ, ENTJ should be members of this group. This group primarily includes students with high academic achievement. Moderator is always included in this group.

The fifth procedure is the selection of groups for heuristic techniques of *analogy* and *random associations*. These groups are mainly formed from members of the discussion group and expert evaluation group. But members of the critics' group also participated, so almost all members of the project team participated in the two groups. Personality types of ENTJ and ENTP should be included in the group of *random associations*, and the group of *analogy* - personal type INTJ.

The fourth heuristic technique - *analogy* and the fifth heuristic technique - *random associations* were used in parallel and repeated if necessary.

The sixth heuristic technique is *expert evaluation*. The group can be updated. Other members of the personality types ISTJ, ESTJ and INTJ may be included in this group.

The seventh heuristic technique is the *using a matrix*. The matrix is formed based on the results of applying the previous heuristic techniques. For the technical work on the formation of the matrix, personality types ISTJ and INTP are involved. An *evaluation group* and a *critics group* analyze each cell of the matrix. Decisions that do not make sense and are not feasible under the given conditions are discarded. Promising solutions are brought up for discussion.

The eighth heuristic technique is a *collective discussion*. The discussion group can be expanded. You can add personality types with good logic - ENTP and INTP. Note that INTP was involved in the matrix developing.

The ninth heuristic technique is an *expert evaluation*. This is final heuristic technique and final evaluation. In this case, the evaluation group should be expanded to include representatives of *discussion group*: INTJ and ENTP. As a result of considering the options detailed in the process of collective discussion, one best is chosen.

Final procedure is a documenting the results. For this the personality types ISTJ and INTP. To do this, personality types INTJ and INTJ are involved.

7. Conclusion

The analysis of heuristic methods based on their structure is carried out. The heuristic methods which belong to the generation phase were considered. As part of the methods highlighted heuristic techniques and procedures. It is shown that two dozen methods belonging to the generation phase use only five heuristic methods. Many methods use the same heuristic technique and differ only in procedure. The theorem was formulated: for each stage of the creative process one generalized heuristic method can be applied replacing all methods used for this stage if it includes all heuristic techniques used within this stage.

The study of the personality type's distribution of students STEM specialties was carried out at the South Ukrainian National Pedagogical University named after K. D. Ushynsky (SUNPU) and Odessa National Polytechnic University (ONPU). The main part of students belongs to seven personality types. Data for SUNPU students on the percentage of each type to the total number of students: ISTJ (18.1), INTP (15.5), ESTJ (14.7). INTJ (10.4), ENTP (8.4), ENTJ (4.4), ISTP (3.3) and ENFP (2.5). Data for ONPU students on the percentage of each type to the total number of students: ISTJ (22.1), INTP (11.9), ESTJ (11.7). INTJ (11.6), ENTP (7.8), ENTJ (5.3), ISTP (4.2) and ENFP (2.0). The generalized method was developed. This method focused to teaching students STEM specialties. Method based on five heuristic techniques collective discussion, pause between the presentation of ideas and their criticism, random associations, analogy, expert evaluation, using a matrix. The heuristic process is considered from the point of view of the personality types of the participants. Personal types are considered in terms of how they are suitable for performing the indicated heuristic techniques. The recommendation about performers: moderator, discussion group, critics group, evaluation group, for each heuristic technique was developed.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

This work has been accomplished with financial support by the Grant No BG05M2OP001-1.002-0011 "MIRACle (Mechatronics, Innovation, Robotics, Automation, Clean technologies)", financed by the Science and Education for Smart Growth Operational Program (2014-2020) and co-financed by the European Union through the European structural and Investment funds.

References

[1] V. Ivanov, L. Dimitrov, S. Ivanova, O. Olefir, "Creativity enhancement method for STEM education," in 2019 IEEE International Conference on High Technology for Sustainable Development (HiTech), 1-5, 2019, doi: 10.1109/HiTech48507.2019.9128255.

[2] J.S. Linsey, M.G.Green, J.T. Murphy, K. L. Wood, A.B. Markman, "Collaborating To Success: An Experimental Study of Group Idea Generation Techniques," in International Design Engineering Technical Conferences and Computers and Information in Engineering, **4742**, 277-290, 2005, https://doi.org/10.1115/DETC2005-85351.

[3] R. Van der Lugt, "Brainsketching and how it differs from brainstorming," Creativity and innovation management, **11**(1), 43-54, 2002, https://doi.org/10.1111/1467-8691.00235.

[4] F.L. McKoy, N. Vargas-Hernández, J.D. Summers, J.J. "Shah, Influence of design representation on effectiveness of idea generation," in Proceedings of the ASME design engineering technical conference, 4, 39-48, 2001.

[5] M. Fargnoli, E. Rovida, R.Troisi, "The morphological matrix: Tool for the development of innovative design solutions," in 2006 International Conference on Axiomatic design (ICAD), 1-7, 2006.

[6] Y.J. Park, C. Kim, J. Yoon, "Creativity and design method in idea generation: the comparison between intuitive approach vs structured approach," in 2019 IASDR Conference, 1-12, 2019.

[7] V. Chulvi, M.C. González-Cruz, E. Mulet, J. Aguilar-Zambrano, "Influence of the type of idea-generation method on the creativity of solutions," Research in Engineering Design, **24**(1), 33-41, 2013, https://doi.org/10.1007/s00163-012-0134-0.

[8] A. Weas, M. Campbell, "Rediscovering the analysis of interconnected decision areas," AI EDAM, **18**(3), 227-243, 2004, https://doi.org/10.1017/S0890060404040168.

[9] W. Beitz, G. Pahl, K. Grote, "Engineering design: a systematic approach," MRS Bulletin, **21**(8), 71-71, 1996, doi:10.1557/S0883769400035776.

[10] T.J. Howard, S.J. Culley, E. Dekoninck, "Describing the creative design process by the integration of engineering design and cognitive psychology literature," Design studies, **29**(2), 160-180, 2008, https://doi.org/10.1016/j.destud.2008.01.001.

[11] S. Buzuku, A. Kraslawski, "Use of Design structure matrix for analysis of critical barriers in implementing eco-design initiatives in the pulp and paper industry," Procedia Manufacturing, **11**, 742-750, 2017, https://doi.org/10.1016/j.promfg.2017.07.175.

[12] V. Ivanov, G. Urum, S. Ivanova, G. Naleva, "Analysis of matrix and graph models of transmissions for optimization their design," Eastern - European Journal of Enterprise Technologies, **4.1**(88), 11-17 2017, https://doi.org/10.15587/1729-4061.2018.131101.

[13] C. Pokhrel, C. Cruz, Y. Ramirez, A. Kraslawski, "Adaptation of TRIZ contradiction matrix for solving problems in process engineering," Chemical engineering research and design, **103**, 3-10, 2015, https://doi.org/10.1016/j.cherd.2015.10.012.

[14] Vezzetti, F. Marcolin, A.L. Guerra, "QFD 3D: a new C-shaped matrix diagram quality approach," International Journal of Quality & Reliability Management, 33 (2), 178-196, 2016, https://doi.org/10.1108/IJQRM-07-2013-0112.

[15] D. Lukianov, K. Bespanskaya-Paulenko, V. Gogunskii, O. Kolesnikov, A. Moskaliuk and K. Dmitrenko "Development of the markov model of a project as a system of role communications in a team," Eastern-European Journal of Enterprise Technologies, **3**(3), 21-28, 2017, https://doi.org/doi:10.15587/1729-4061.2017.103231.

[16] M. Farhangian, M. Purvis, M. Purvis, T.B.R. Savarimuthu, "Agent-based modeling of resource allocation in software projects based on personality and skill," Advances in Social Computing and Multiagent Systems. MFSC 2015. Communications in Computer and Information Science, **541**, 130-146, 2015, https://doi.org/10.1007/978-3-319-24804-2_9.

[17] N. Lehmann-Willenbrock, S.J. Beck, S, Kauffeld, "Emergent team roles in organizational meetings: Identifying communication patterns via cluster analysis," Communication Studies, **67**(1), 37-57, 2016, https://doi.org/10.1080/10510974.2015.1074087.

[18] J. Jankowski, The 16 Personality Types in a Nutshell, LOGOS MEDIA, 2016.

[19] M.H. McCaulley, "The Myers-Briggs type indicator: A Jungian model for problem solving." New Directions for Teaching and Learning, 1987(30), 37-53, 1987, https://doi.org/10.1002/tl.37219873005.

[20] S. Rushton, J. Morgan, M. Richard, "Teacher's Myers-Briggs personality profiles: Identifying effective teacher personality traits," Teaching and Teacher Education, **23**(4), 432-441, 2007, https://doi.org/10.1016/j.tate.2006.12.011.

[21] M.H. McCaulley, "Myers-Briggs Type Indicator: A bridge between counseling and consulting," Consulting Psychology Journal: Practice and Research, **52**(2), 117–132, 2000, https://doi.org/10.1037/1061-4087.52.2.117.

[22] D. Keirsey, M.M. Bates, Please understand me: Character & temperament types, Del Mar, CA: Prometheus Nemesis Book Company, 1984.

[23] C. Soles, L. Moller, "Myers Briggs type preferences in distance learning education," International Journal of Educational Technology, **2**(2), 2001.

[24] W. Huitt, "Problem solving and decision making: Consideration of individual differences using the Myers-Briggs Type Indicator," Journal of Psychological Type, **24**(1), 33-44. 1992.

[25] C.F. Yokomoto, R. Ware, "Applications of the Myers-Briggs Type Indicator in engineering and technology education - part II," Indiana University - Purdue University Indianapolis, **2230**, 4.88.1- 4.88.10, 1999.

[26] V.R. Montequín, J.V. Fernández, J.V. Balsera, A.G. Nieto, "Using MBTI for the success assessment of engineering teams in project-based learning," International journal of technology and design education, **23**(4), 1127-1146, 2013, https://doi.org/10.1007/s10798-012-9229-1.

[27] S. Ivanova, L. Dimitrov, V. Ivanov and G. Naleva, "An Experiment on the Joint use of the Heuristic and Project Methods at the University," in 2019 IEEE International Conference on High Technology for Sustainable Development (HiTech), 1-5, doi: 10.1109/HiTech48507.2019.9128248.

[28] J.B. Murray, "Review of research on the Myers-Briggs type indicator," Perceptual and Motor skills, **70**(3_suppl), 1187-1202, 1990, https://doi.org/10.2466/pms.1990.70.3c.1187.

[29] M. Ogot, G. Okudan, "A student centered approach to improving course quality using quality function deployment," The International journal of engineering education, **23**(5), 916-928, 2007.

[30] P.D. Tieger, B. Barron, K. Tieger, Do what you are: Discover the perfect career for you through the secrets of personality type. Little, Brown Spark, 2014.

[31] A.J. Williamson, "Suiting library instruction to the Myers-Briggs personality types and Holland Vocational personality types of engineering students," Issues in Science and Technology Librarianship, **37**(2), 2003. https://doi.org/10.5062/F4154F1G.

[32] L.F. Capretz, "Is there an engineering type?" World Transactions on Engineering and Technology Education, **1**(2), 169-172, 2002.

[33] A. Thomas, M.R. Benne, M.J. Marr, E.W. Thomas, R.M. Hume, "The evidence remains stable: The MBTI predicts attraction and attrition in an engineering program," Journal of Psychological Type, **55**, 35-42, 2000.

[34] R.M. Capraro, M.M. Capraro, "Myers-Briggs Type Indicator score reliability across studies: A metaanalytic reliability generalization study," Educational and Psychological Measurement, **62**(4), 590-602, 2002.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 3, 213-230 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

XMDD as Key Enabling Technology for Integration and Organizational Collaboration: Application to E-Learning Based on NRENs

Salim Saay [*,1], Tiziana Margaria[2]

[1]*Department of Computer and Software Engineering, Engineering Faculty, Athlone Institute of Technology, Athlone, N37 HD68, Ireland*

[2]*Department Computer Science and Information System, Faculty of Engineering, University of Limerick, Limerick, V94 T9PX, Ireland and Lero, the SFI Research Centre on Software*

ARTICLE INFO

Article history:
Received: 10 January, 2021
Accepted: 04 March, 2021
Online: 23 May, 2021

Keywords:
Collaboration
e-learning
XMDD
DIME
NREN
AfgREN
SSO

ABSTRACT

Collaborative E-learning is highly dependent on the well functioning of a complex socio-technical system that comprises information technology and various social processes. Large scale infrastructures like the National Research and Education Networks (NRENs) provide access to research and educational resources and provide collaboration between educational and research organizations, thus providing a set of essential services for e-learning. Currently, the lack of data integration between e-learning systems is still a problem in NREN domains, and a hurdle to collaborative e-learning. We address systematic cross-organizational collaboration and data integration between large-scale e-learning systems by designing an architecture for NREN e-learning systems to support open access education and learning. In particular, we design and provide a reference implementation for an e-learning broker that can provide the needed data integration and processes, and takes into consideration the strategies and policies for open access in education and training. We develop the architecture and reference implementation applying the eXtreme Model-Driven Development (XMDD) paradigm for software design and development, using the DIME low-code development environment for modelling data, processes, and user interface. We consider here two specific application settings: The national network of e-learning collaboration in AfgREN, centered on the Kabul University in Afghanistan, and the newly started collaboration between the Athlone Institute of Technology (AIT) and Limerick Institute of Technology (LIT) in Ireland, that are forming a new consortium under the newly introduced Technological University structure.

1 Introduction

Socio-Technical Systems are collaborative interacting systems where one or more social systems and a number of technical system work together to accomplish a goal meaningful for a community or user group. The technical part of the system is concerned with the processes, tasks, and technology needed to transform some inputs to corresponding outputs. The social part of the system is described by key attributes of people, such as people's skills, values, responsibilities, and their roles in the system [1]. In advanced collaborative e-learning, our application domain, the relationships among the administration of a university, the students, and lecturers are mediated by technical systems, as these people all cross-interact within individual and federated e-Learning systems. A large part of these interactions are mediated by the institutional e-learning platform(s), that act as a medium-to-large scale socio-technical system, and their underlying communication and federation infrastructure.

National Research and Educational Networks are a collaborative large scale scientific infrastructure that provides high quality and cost-effective services to academic and research organizations [2], providing cloud services such as Infrastructure as a Service (IAAS), Software as a Service (SAAS) and Platform as a Service (PAAS), that are new trends for NRENs [3]. Eduroam [4], EduGATE [5], eduGAIN [6] and Shibboleth [7] are widely adopted infrastructural building blocks used to provide a wide range of national services by NRENs and international services by so called regional NRENs. Specifically, Eduroam is a global wireless network that allows academics staff and researchers to access the services of their own institution with their roaming profile [4] from any participating institution. EduGATE is an identity manager used in GÉANT to provide participating users with access to the different applications with a Single-Sign-On technique, effectively providing a connection

[*]Corresponding Author: Dr. Salim Saay, Email: ssaay@ait.ie

between participating applications of the different NRENs that are members of GÉANT [5]. EduGAIN is also a national level identity management services provider, and it is used in the Irish Research and Education Network (HEAnet) [6]. Shibboleth supports collaboration and data integration between different learning management systems used in Internet2 and many NRENs [8]. Other platforms are directly or indirectly also service providers for collaboration, and provide widely adopted services for access and interoperability. For example, Office 365 provides a solution that integrates different applications, it uses Active Directory Federation Services (ADFS) and DirSync to provide a Single Sign-On password integration for a domain [9]. Similarly, the Security Assertion Markup Language (SAML) and Open Authentication (OAuth) [10] focus on providing and managing user Identities for SSO, so that the NREN users use the same Identity for all member applications. A secure SSO provides secured and uninterrupted services by keeping one authentication credential for each user so that, once authenticated, users do not need to provide their credentials again when accessing different web services that share the same SSO. Kamal et al. claim that Shibboleth is the most suitable SSO in the higher education domain. However, no global log-out, the infancy of the technology, implementation complexity, high reliability on assumptions that cause security risks are the main challenges of the Shibboleth technology [10]. The existence of duplicated data and duplicated profiles in the same domain is a serious problem for access, cost, and management. How to integrate the duplicated data requires further investigation. In this context, NREN infrastructures provide an opportunity to provide for uniform security, reliability, and reduction of the complexity for data integration across the participating e-learning system within the served NREN domain.

In this paper, that significantly extends work originally presented in ICALT 2020 [11], we focus on data integration in inter-institutional cross-organizational collaborations as well as in intra-institutional multi-campus collaboration, two challenging situations with which we have direct experience.

1.1 Data Integration of E-Learning Systems In Large Scale Cross-Organizational Collaborations

This is a challenge because they use different e-learning platforms that do not interoperate in a native way. For instance, Figure 1 shows that Moodle, eDX and the national Higher Education Management Information System (HEMIS) are implemented in a single domain at the Kabul University, Herat University and Balkh University in Afghanistan. The shortcomings of the current situation are due among other causes to the lack of support of Single Sign-On in the three running applications. Accordingly, the applications implemented in this national collaboration domain [12] do not currently exchange data and they reside on three different platforms. EdX is implemented at six universities across the Afghanistan Research and Education Network (AfgREN) for national collaboration and accessibility [13]. Moodle is the second popular e-Learning platform. It is widely implemented in universities of Afghanistan [14], and in fact lecturers and students used Moodle en masse during the recent universities lockdowns. The Higher Education Management Information System (HEMIS) is a traditional students management system. It was introduced by the National Higher Education Strategic Plan (NHESP) of the Ministry of Higher Education of Afghanistan to support in a standardized way the decision-making process and control management tasks in the adopting organizations, which impact the organization's functions, performance, and productivity in all its aspects [15].

In our collaborative e-learning context, for example, the enrolment of students to universities is managed through this system. While it is mandatory for all universities to use HEMIS, also HEMIS does not support Single Sign On. As a consequence, each student has three profiles in three different e-Learning platforms, as shown in Figure 1. This distribution and partial data replication are challenging for the administration and maintenance of all these systems and their collaboration.

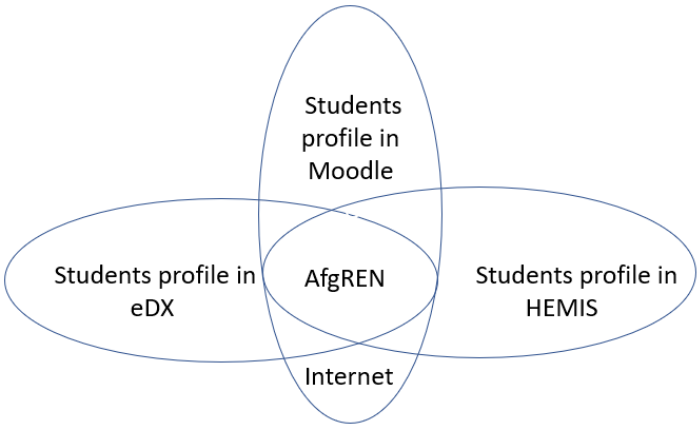


Figure 1: Student profiles in e-Learning systems of a single organization: The case of AfgREN

1.2 Data Integration of E-Learning Systems In Intra-Institutional Cross-Campuses

This second challenge is illustrated with an application in Ireland. Through the Technological Universities Act the Irish Department of Further and Higher Education, Research, Innovation (the Irish Ministry) established in 2018 a process whereby consortia of local Institutes of Technology[1] may submit an application for Technological University designation. De facto, this mechanism creates regional Technological Universities of larger size and diversity than the previous ITs. For example, in January 2019 the Technological University Dublin (TU Dublin) was formed by the amalgamation of three institutes of technology in the Dublin area (Dublin Institute of Technology, Institute of Technology Blanchardstown, and Institute of Technology Tallaght), yielding the second-largest university in Ireland based on the combined student population of 28,500. Other five TUs consortia have meanwhile applied, among which in 2021 the AIT-LIT consortium[2]. This consortium is now establishing a (technical) University for the Midlands and the Mid-West by combining the Athlone Institute of Technology (AIT) in Athlone (in the Midlands) and the Limerick Institute of Technology (LIT) in

[1] An Institute of Technology is similar to colleges/Universities of Applied Science in other countries
[2] https://aitlitconsortium.ie

Limerick (in the Mid-West). Accordingly, the new Technological University will be located in five different campuses across the Midlands and Mid-West of Ireland. In addition to the organizational and structural combination, many technical and infrastructural changes are also required: the data integration, resources sharing, and collaboration between technical systems of AIT and LIT need further investigation. To give an idea of the complexity of the challenge, of the six Working Groups established by the Professional Services Steering Group, five concern the socio-technical ecosystem of the e-learning platform[3].

Figure 2 shows the current state of the student profiles: they currently reside in different, independent e-learning systems of AIT and LIT, that need to be combined, and additionally also the financial system and many other information systems need data integration. The Higher Education Authority Network (HEANet) is the national research and education network (NREN) of Ireland: Similarly to AgfREN it provides a high speed secure and reliable backbone for the connection between these e-Learning systems, and users additionally also use public internet.

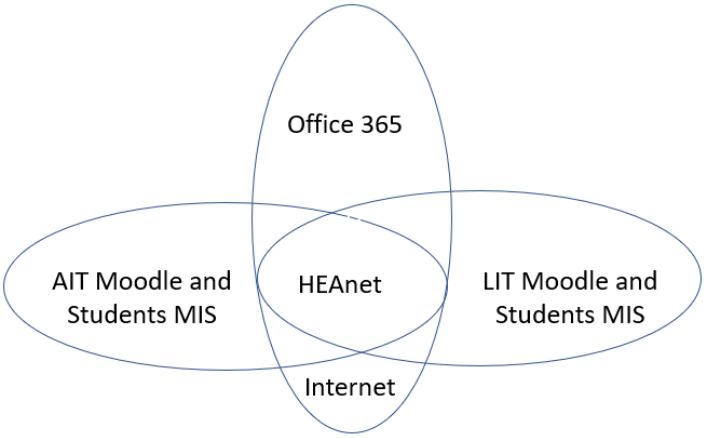


Figure 2: Currently interorganisational students profiles in the new AIT-LIT Technological University

In this paper, the contributions are

- to design an *e-learning architecture* based on NREN for the federated learning collaboration;
- in its pursuit, we apply the XMDD technology to design and implement a prototype of a *e-learning reference implementation* for data integration between different e-learning systems at three different levels: single organization, cross-organization and cross-nation.
- taking into consideration the *diverse organizational strategies, policies, law and procedures* of the collaborating parties.

Our solution and prototype go well beyond the SSO: we design a domain-specific e-learning broker for NRENs as a data integration application that focuses on an organization's reputation management, identity management, process management, service brokering, with a search engine to lead users to the services available in the NREN domain, like for example an NREN local course management system. The prototype is easily accessible to designers, managers, developers, customers, and users, and it can be used as a showcase of the concrete application [16]. In the prototype and later in the developed system, users of the e-learning system access the resources located in the NREN domain based on their role in their local organization. We define which type of data is accessible to students, and at which level and which types of data are accessible for the academic staff and admin staff of each university.

The architecture and the XMDD model-driven design are our methodological contributions. The practical realization of a service-based architecture on top of a variety of e-learning systems is the concrete proof of concept, that we applied to AfgREN as a case study.

Overall, we use advanced Information Technology and leading edge methods and tools for advanced Web application design and implementation, to produce computer software and applications that use the Internet and the World Wide Web to solve long term and large scale inter- and intra-institutional collaboration issues for e-learning.

The rest of the paper is organized as follows: Section 2 discusses the background of the domain technologies and the two case studies, Section 3 describes the methods and technologies adopted to produce the design and the solution, Section 4 illustrates the architecture design of the e-learning broker, Section 5 presents in detail the Model-driven design of the prototype reference system, in particular the Data model, Process models and GUI models, we explain our lesson learned in section 6, and the current status of the project explained in section 7, finally in Section 8 we conclude with a summary, some reflections and future work.

2 Background

Educational organizations need higher speed intranet and internet connections to access the heterogeneous traffic, and demand novel services [17]. Educational institutes at all levels suddenly closed in most countries worldwide due to the outbreak of COVID-19 early this year [18], and in many cases they have not yet returned to on-site education. According to the UNESCO report, due to the COVID-19 pandemic 850 million school children worldwide have been locked out of their schools, and have turned to online learning to continue their education, with added high pressure on e-learning systems and remote communication tools. While we hope the pandemic will soon pass, online education will always be needed for many aspects of high-quality education. Fortunately, over 120 countries [19] in the world established National Research and Education Networks (NRENs) to provide research and educational services, establishing a shared Information Technology infrastructure and providing it at a reduced cost with higher quality connection than regular and commercial Internet Service Providers (ISPs).

[3]https://aitlitconsortium.ie/our-journey/

2.1 AfgREN

Concerning the background in the specific case studies considered in this paper, Figure 3 shows the current physical layout of the Afghanistan Research and Educational Network (AfgREN). Its Network Operation Center (NOC) is physically located at Kabul University and currently most of the AfgREN connectivity is provided by Afghan Telecom fiber optic links. Most public universities (in green) have a direct connection to the AfgREN NOC, but some like Herat University are connected indirectly by fiber. They use the same group of Internet Protocol (IP) address of AfgREN and its services, but the fiber network is from other ISPs. A group of provincial universities is connected to the AfgREN NOC via wireless WiMAX technology. Internationally, AfgREN is connected with the TEIN network by a 155 Mbps fiber optic link via Pakistan. TEIN covers mostly South Asian countries, with 21 NRENs of South Asian countries [3].

The services provided by federated members of TEIN include internet access, e-learning, tele-medicine, and support for research collaborations [2]. For the global connection, TEIN connects Asia to the European regional NREN GÉANT, which includes HEANet, the Irish Higher Education Authority net. GÉANT provides global identification via eduGAIN [6].

Various E-learning platforms such as Moodle, edX, Cisco Networking Academy, and many traditional applications including HEMIS are implemented in the AfgREN domain, providing a wealth of the most popular services [20]. The Higher Education Management Information System and the University entrance exam management information systems are specific purpose web-based applications developed by the Ministry of Higher Education of Afghanistan and they are used by all the public Universities in the country.

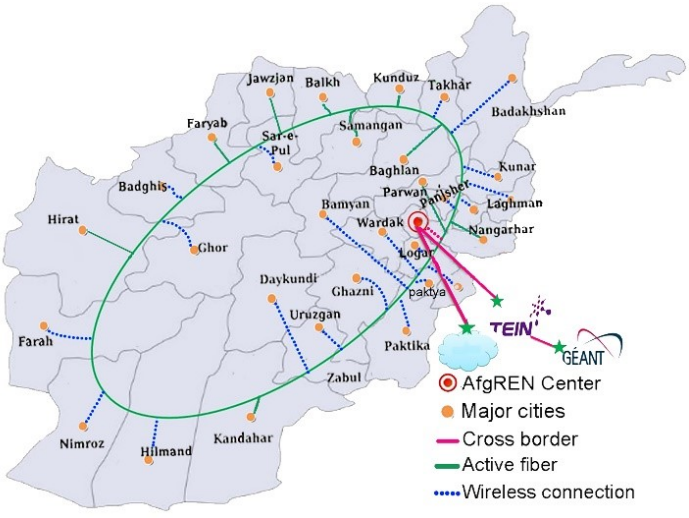


Figure 3: The AfgREN Layout and Connection Structure

Moodle [21] is an open-source e-learning platform that can be installed on a local server or accessed via the Internet. It provides a variety of communication facilities such as a discussion forum, message systems and wiki [21]. It is used for learning, assessment and management, and it provides various types of reports, courses, assignments and knowledge assessments.

edX [22] is also open-source, it was developed by the MIT and Harvard universities in the USA and it is mostly popular for MOOC courses and online certification. edX provides facilities for discussion between students and teachers, online video lectures, course materials, various online exams, and virtual libraries [23]. A comparison of edX with five other e-learning platforms in [21] (Coursera, Google Course Builder, Class2Go, Udemy, and Lernanta) shows that edX has more educational tools than the other four, each of these e-learning platforms has many good learning resources for the educational organization and avows that the integration of the platforms would provide many more learning opportunities for the students.

Cisco Networking Academy was the first e-learning system widely applied in Afghanistan universities. Kabul University is the main regional academy and trains instructors for other local academies [24]. The Cisco e-learning system is implemented in the Computer Science faculties in Afghanistan, providing course materials, laboratory tools, and an online examination system.

Moodle provides more features than edX and other e-learning systems used in the AfgREN domain. It offers more possibilities to customize and connect with other e-learning systems through its Learning Tool Interoperability [25] technology, and thus appears to be the most suitable platform for organizations that require a customized learning management system. However, there is a gap between these software applications and NRENs, and we endeavor to prototype an application to bridge the gap.

2.2 Technological University Consortium

The second case study in this paper concerns the new Technological University currently being established in Ireland from the combination of the Athlone Institute of Technology and the Limerick Institute of Technology.

AIT was established in 1970 under the structure of Regional Technical College (RTC) in Athlone, a city in the geographical centre of Ireland, to provide technician level courses. In 1993 AIT became an autonomous institution, in 1998 AIT was officially redesignated to Athlone Institute of Technology. Since then, AIT established itself as a center of academic excellence with an applied, industry-focused offering of engineering courses. AIT has more than 6,000 undergraduate, postgraduate and Ph.D. students from 84 nations around the world.

The Limerick Institute of Technology (LIT) is a multi-campus institute that was originally established in 1852 in the School of Ornamental Art and later restructured several times. Currently, LIT has three campuses located in Limerick, Tipperary, and Clare. LIT has more than 7,000 undergraduates, postgraduates, and Ph.D. students [26].

Based on the strategic plan of the Irish government [26] the new Technological University will start in September 2021 and operate in multiple campuses covering 4 counties (Limerick, Clare, Tipperary and Westmeath) as shown in Figure 4. The campuses will be connected by the HEAnet backbone. HEAnet [4] is the Irish national education and research network, providing e-infrastructure services

[4]https://www.heanet.ie/

for approximately 218,000 students and staff and connecting around 67 educational and research organizations at the national level. In total, around one million users in the research and academic organizations are using HEAnet. The main strategy of HEAnet [27] is to provide collaborative partners, trusted provider services, common, reparable and shareable solutions, innovative solutions, identity federation, brokerage, Key Advisor services, conduit to Europe services and make the country's institutions an excellent place to work.

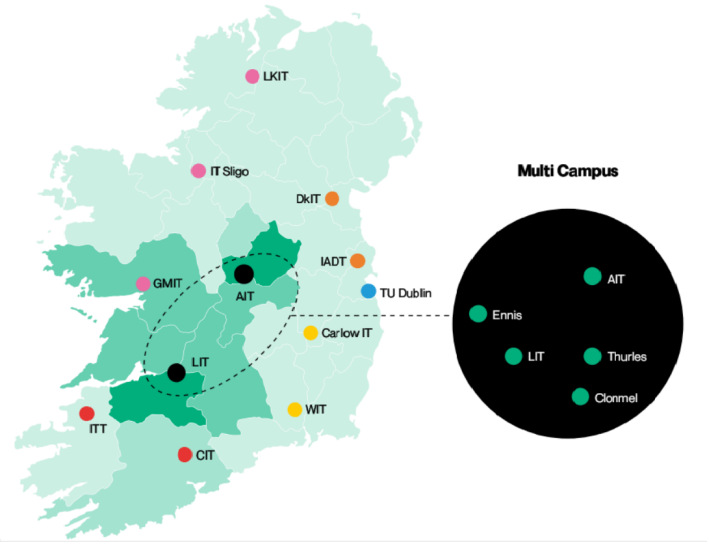


Figure 4: The Technological University Multi-campuses (AIT-LIT), from [26]

The campuses will be connected by the HEAnet backbone. HEAnet [5] is the Irish national education and research network, providing e-infrastructure services for approximately 218,000 students and staff and connecting around 67 educational and research organizations at the national level. In total, around one million users in the research and academic organizations are using HEAnet. The main strategy of HEAnet [27] is to provide collaborative partners, trusted provider services, common, reparable and shareable solutions, innovative solutions, identity federation, brokerage, Key Advisor services, conduit to Europe services and make the country's institutions an excellent place to work.

In this research, we target e-learning services brokerage in HEAnet as the means to provide collaboration between the multi-campuses of the new Technological University. Currently, HEAnet provides the served institutions with four types of brokerage including hardware and equipment, software and license, services and consultancy, and students and staff offers. However, e-learning brokerage is also very important for users. Instead of searching the websites of all institutes, users prefer to refer to the website of HEAnet and search there for their required courses[6].

3 Design and Implementation Method

For the analysis, design and implementation we adopted the eXtreme Model Driven Development paradigm (XMDD) [28]. This way we work primarily on models that are compiled to code and automatically deployed, achieving an agile approach along the entire process and an extended DevOps technique for the prototype, that starts from models instead of code. Specifically, we used a three cycles scrum approach [29] to design the architecture [30] and the prototype of this e-Learning collaboration system. We used the Requirements Bazaar tool [31] to collect user stories and prioritize the requirements. Altogether, the applied design and development methodology is participative, agile, model-driven and low-code.

3.1 Participative Design

Collection of the user stories and the prioritization of the requirements are important steps in software development. We used the Requirements Bazaar concept and web-based tool [31] to collect, analyze, and then prioritize user stories. In addition to collecting the user stories, researchers and system developers can prioritize requirements based on the number of votes and comments provided by the users. Requirements Bazaar provides interaction facilities with the stakeholders from the beginning until the end of the project, facilitating the co-design and co-development approach we wished to adopt. We posted the list of elicited functional requirements and system/architecture features on the Requirements Bazaar portal as a project that we shared with the stakeholders. This enabled the various groups of contributors to Vote on requirements, Select requirements, add comments, add new requirements, and maintain continuous communication with the researchers. A vote means that this requirement is important and it has a high priority, a Select means that this item is required. Through comments, contributors can provide more detail to existing requirements, request changes to system features, and add sub-requirements for the system features. In this work, we combine votes and select in the same column because several stakeholders used them quite indifferently. When contributors selected both options we counted it only once [32].

We also distributed a questionnaire concerning the current state of the e-learning systems. Figure 5 shows that currently lecturers are using many different tools for online teaching and to share course materials. We also asked students about the same and their needs: data scattering is the main challenges that students and lecturers identified as a key problem for integration of the data between different applications.

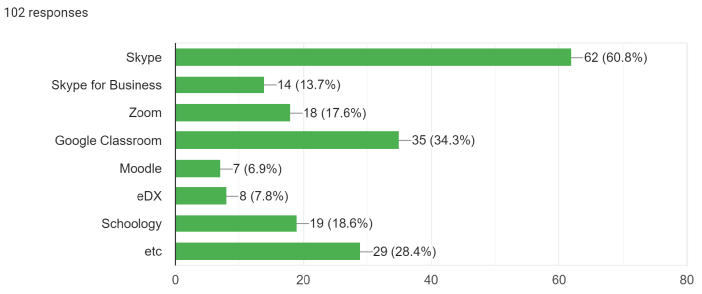


Figure 5: Applications in use for e-learning in the AfgREN domain

[5]https://www.heanet.ie/
[6]https://www.heanet.ie/

3.2 Agile Design and Development

Figure 6 summarizes the agile development of the application: the user stories and requirements are collected directly from the users by using Requirements Bazaar. After the planning we designed the prototype using DIME, an Integrated Modelling environment that supports a model driven development and a fast-turn around DevOps cycle adequate for rapid prototyping. Currently we have the first version of the functional prototype of the reference system.

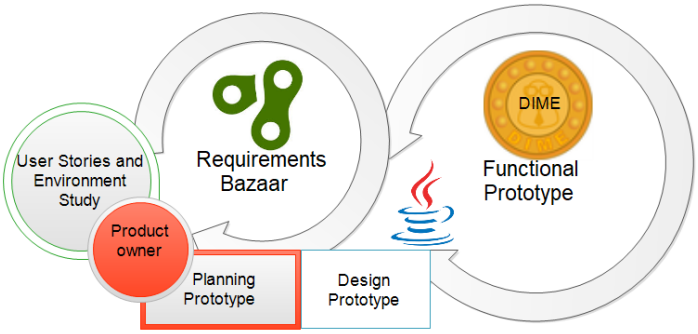


Figure 6: Methodology: the Agile development of the application and implementation tools

Figure 6 also shows the overall method from the tools perspective: we use Requirements Bazaar to collect the requirements, discuss with the users and collect the votes of the user and we prioritize the requirements based on the user's votes. We use DIME for modeling and prototyping. DIME automatically generates Java, JavaScript and DART code, so we export the generated code and proceed to the deployment to a Java EE server side and an Angular 2 Web execution environment. We use GitHub and Bitbucket for the version control and the technical review.

The fast turn-around time for a prototype release is enabled by the model driven approach combined with the adoption of low-code development.

3.3 Model Driven Development

XMDD [33] offers a fast, understandable and easily modifiable paradigm for the co-development of complex systems, and it supports the agile development of the project in a modern and efficient DevOps fashion. The quick and efficient design and deployment of a web application with the XMDD paradigm provides an opportunity for developers to get continuous feedback from the customers/users, and improve the application in agreement with the users until the end of the project life cycle. XMDD is an evolution of the traditional model-driven engineering method (MDE) [34] that in particular supports service-oriented integration, a platform- and feature-based construction and reuse, and continuous integration and evolution.

MDE is a domain-specific modeling methodology that formalizes via models the application structure, behavior and the requirements and properties of IT systems within a specific application domain. It also facilitates formally analyzing specific aspects of models, easing the correctness and performance checking on the models to detect errors early, before investing deep in the development life cycle.

XMDD focuses on **process models** as a primary artifact, understood as descriptions of what the system or application should do. It provides model-level artifacts that are executable even before the full application design and implementation, thus it is accessible also to people that are either less technically skilled, or even not interested at all in usual programming. The benefit for our application is that the adoption of XMDD reduces the time needed for the implementation of the application [34] by adding a coordination layer to the system architecture that describes and then implements the coordination logic between the different components of the architecture.

3.4 Low-code Development

Using DIME [35, 36] as a low-code development environment for the XMDD paradigm, we have a ready graphical modeling tool for the entire architecture, including the data model, control flow and data flow, GUI (interface model) and security/roles and rights model. These models are easier to understand than code, thus supporting the interaction and co-design between IT professionals, domain experts, business experts, and users better than at the code level or the traditional (more technical) modelling level as in a typical UML [37] approach. This approach is extensively validated: it has proven successful in the development of web applications, telecommunication systems, game development, robotics, and bioinformatics systems [34]. In particular, XMDD and DIME support Domain Specific Languages, intended as domain specific libraries of services that provide reusable functionalities [38] and features [39, 40].

4 E-Learning Broker Architecture Design

We designed a domain-specific service-oriented e-learning system for systematic data integration and service brokering between different e-learning systems in NRENs. We refer here to the specific case of AfgREN, but the architecture is general, as we aim at standardization for compatibility beyond specific regional consortia.

In the specific case study, AfgREN is established at Kabul University and it provides services to Afghanistan's public universities including Balkh University and Herat Universities. Figure 8 shows the architecture from the point of view of Kabul University: it has its own three layers (on the left) and it provides services to the external layer of the other universities (on the right) through the e-Learning Broker (center). The same layers and components exist for the other collaboration parties, which are symmetric.

In Afghanistan, the legal framework and procedures are the same for all the public Universities. This simplifies the collaboration because the three layers and several components in Kabul University's local architecture (shown in Figure 8 left) are also applicable to the other public Universities in Afghanistan. The second case study exemplifies how much of the architecture and the broker carries over to the Irish TU setting.

4.1 National, Cross-Organizational Collaboration

4.1.1 Requirements

Prior to the design of architecture, we performed an architectural requirements prioritization using Requirements Bazaar concept and tool with an expert group of stakeholders that comprised the e-learning committee of Kabul University and AfgREN network operation center engineers.

Figure 7 shows the functional requirements of system architecture together with the number of Votes/Select each attribute received in a Requirement Bazaar-supported session. We used Requirements Bazaar from the very early stages till the end of the development process to be in continuous contact with the remotely located expert groups, and negotiate and validate requirements and design decisions in each stage of the development.

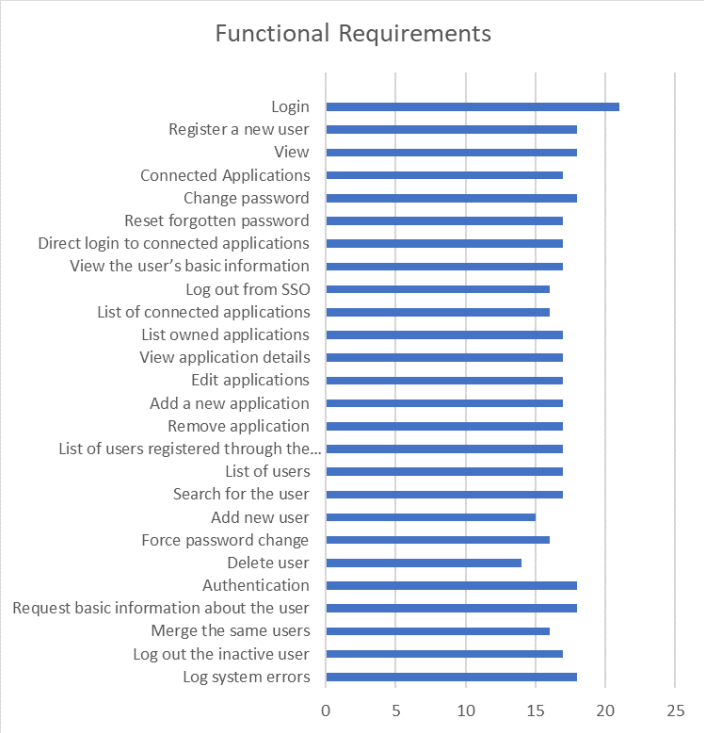


Figure 7: Functional Requirements of the System Architecture

4.1.2 Choosing the Architecture pattern

Different architecture patterns exist for different applications, such as Layered architecture, Event-Driven Architecture, Microkernel Architecture, Micro-services Architecture, and Space-Based Architecture [41]. We carefully analyzed the application domain and requirements to decide which architecture pattern is the most suitable for the target application.

Layered Architectures consists of several layers addressing individual concerns [42]. The required number of layers depends on the complexity of the application: some complex applications, such as the Internet of Vehicle (IoV), have five layers [43]. Most of the layered architectures have four layers: a presentation layer, a business logic layer, a persistence layer, and a database layer. Sometimes in sample architectures the architects combine the persistence layer with the business logic layer. The process of passing data between layers needs to be very systematic. Layers need to be closed or open, only after having been processed by a specific layer the next layer is made accessible for its data processing. In a layered architecture, the components are accordingly also divided into these layers. The classification of components into different layers helps for easier separation of concerns in the development, test, and maintenance. Layers also provide better modifiability, offering a form of virtualization: if a component of one layer is changed, it does not affect components of the other layers. However, layers and tight coupling of components have a negative effect on the performance, agility, deployment and scalability of a system [41].

Event-Driven Architectures are used for highly scalable systems. Its two main typologies are the mediator topology and the broker topology.

A mediator topology is commonly used when the architect centralizes the events. The typical mediator topology has four main components (event queues, event mediator, event channel, and event processor) and two types of events: initial events and process events.

The broker topology is best used when the architect chains the events and does not need to centralize them. The broker topology has two main components: a broker component and an event processor component [45]. The Event-Driven Architecture pattern is a complex pattern, and the architect needs to consider the remote process availability issue, the lack of responsiveness and the broker reconnection logic in the broker event or face mediator failure [41].

Plug-in Architectures are used as a way of supporting relatively independent third party products. A Plug-in architecture pattern that is also called **Microkernel pattern** consists of two main components including the core system and one or more plugin modules. This architecture pattern is mostly used for product-based applications that have downloadable files.

Microservice Architectures are the oldest services-oriented pattern [50] and widely adopted in distributed and module-based systems. Currently, it is used as the predominant service-oriented architecture [50]. The microservice architecture pattern is a specialized alternative to a more general service-oriented architecture. It is a distributed architecture, where all components are separated and can be controlled with a remote access protocol. The distributed nature of this pattern provides more scalability and ease of deployment attributes [41]. However, microservice architectures have challenges of security [47].

The **Cloud Architecture** pattern is also called space-based architecture, and addresses the problem of scalability. An application that is accessible through a web interface is a suitable application for this pattern. However, the cloud architecture pattern is expensive and it is not suitable for a traditional database with a large number of operational services. The advantage of a cloud based architecture is that we can implement different architecture components in this pattern. The cloud-based patterns simplify the deployment of applications [48, 41].

Table 1 summarizes the comparative characterization of the described architecture patterns based on their quality attributes relevant to our application. We qualitatively indicate the extent to which each quality attributes is provided as Low (*), Medium (**) and High (***).

Table 1: Analysis of Architecture Patterns based on Quality Attributes [44, 42, 43, 41, 45, 46, 47, 48, 49]

Architecture Pattern	Agility	Deployment	Testability	Performance	Scalability
Layered Architecture	*	*	***	*	*
Event-Driven Architecture	***	***	*	***	***
Plug-in Architecture	***	***	***	***	*
Microservice Architecture	***	***	***	*	***
E-Learning Broker Architecture	**	***	***	***	***

4.1.3 The Hybrid NREN e-Learning Architecture

Given the needs of our applications, we adopted a layered, event-driven architecture with a broker technology that provides a high level of agility, ease of deployment, high level of performance and scalability [12, 49, 30, 20, 51]. In this system, the broker needs to publish events to different collaborative systems that are in different domains and follow different roles. Additionally, the XMDD paradigm automatically enforces a service-oriented approach [34] at the implementation level [52], so that we also reap the benefits of this paradigm in the application design and implementation. XMDD is compatible with the SCA service-oriented architectures standard [53], and is a service-oriented evoilution of the more traditional component based architectures in software design [54] and of the feature oriented architectures in early telecommunications services design [55]. It would even support higher order processes, allowing for full reconfigurability of the processes at runtime as shown in [56].

The NREN e-learning architecture shown in Figure 8 is composed of three layers: Internal Layer, Conceptual Layer, and External Layer (on the left). Each layer has several components and sub-components that provide the features identified as a core for any NREN e-learning system architecture.

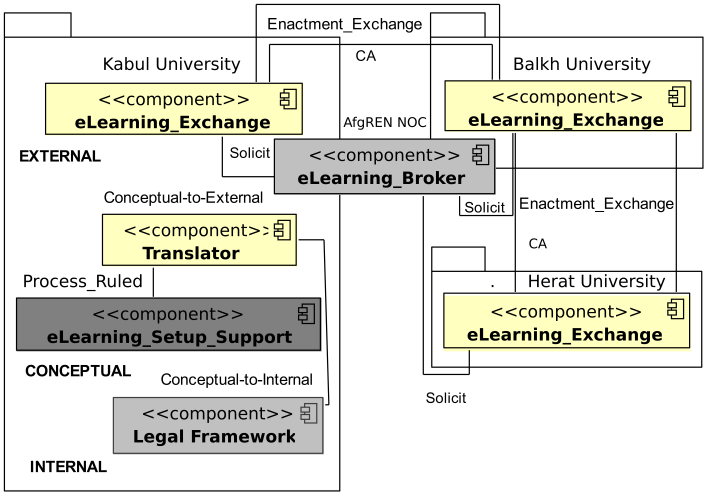


Figure 8: The Component diagram of eLearning System Architecture for a subset of AfgREN

Figure 8 is adapted from the reference architecture for managing dynamic inter-organizational business processes (eSRA) [57] and shows the abstract level of the AfgREN e-learning architecture [49].

- Looking at the structure in the main node, that in AfgREN is at Kabul University (Figure 8 left), the e-learning platforms are located in the **External Layer**.
- The **Conceptual Layer** comprises an `e-learning_Setup_Support component` in which e-learning services are composed collaboratively in a crowdsourcing way. These e-learning services are process aware, have rules attached. Additionally, the Translator component converts heterogeneous data formats between the external and internal layers.
- The **Internal Layer** contains a Legal framework component in which local process enactment engines, legal process and orchestrate information technology infrastructure that is wrapped as a Web-service. Furthermore, in the Internal Layer local rules engines and local database systems capture the roles of the respective departments and user groups, including academic, administration and students. This structure and subdivision correspond to the essential features identified in various workshops with stakeholders of these groups.

4.1.4 The Broker Aspect: System-Level Features

The design of the e-Learning broker is based on the following elicited system-level features and requirements:

1. The NREN system supports the conceptual formulation of e-learning services between collaboration parties based on the accepted agreement and the established procedures.
2. Translation of the existing legal framework to a technical system, and the mapping of these between collaboration parties.
3. Projection of the e-learning services from the conceptual layer of each collaboration party to the external layer.
4. Brokering capability for the projected services. The broker needs to have access to all e-learning systems of the collaboration parties, the Broker provides an interface that users access different e-learning tools from a single interface, the Broker interface provide SSO services ad also integrate the data of used e-learning applications.
5. Verification of collaboration parties and exchangeable components is another task of Broker,
6. Announcement of the verified result to users, the new services, new members, and any possible issues also announce by the Broker component.
7. Stepwise and systematic construction of interoperable e-Learning architecture based on NRENs.

8. Technology independent e-Learning architecture based on NRENs, easy to port to more modern technologies.

9. Interfaces for the separation of heterogeneous data, to protect private information and to enable a separation of internal and external processes. Collaboration between systems by using Single-Sign-On technology

10. Loose coupling for the interoperability of the e-learning platforms.

11. No central role for any member of the collaboration parties, each collaborative party should act as a party.

12. The Broker application will be installed in the network operation center of the NREN.

4.1.5 Data Management

For **data management**, the NREN e-learning architecture adopts an abstract data repository. This style keeps e-learning service consumers and providers of shared data from knowing of each other's existence and the details of their respective internal implementations. The abstract data repository implements a layering style by interposing an intermediary protocol between the producer and consumers of the shared e-learning services. Its interface further reduces the coupling between data producers and consumers. Global and local process enactment and rules engines in the external- and internal layers are dedicated components for example for the technology translation, data mediation, and policy enforcement.

4.1.6 Broker-Based Service Mediation

The broker pattern in the AfgREN is implemented in the `eLearning_Broker` component, a separate component that mediates between collaboration parties within the architecture, facilitating the rapid and performant matching of e-learning needs with requests from users. Its purpose is the redirection and bundling of communication among the parties, preventing parties from finding, contacting, and investigating every potential collaborating party separately.

The service oriented architecture of the `eLearning_Broker` component is shown in Figure 9, and it is adapted from an NREN e-learning reference architecture [49]. The broker uses a publish/subscribe style in which publisher institutes submit new e-learning services, and all subscribers that are members of the NREN receive notifications automatically. The broker works as a notifier between the collaboration parties. The subscription contract between collaboration parties can be a collaboration contract with free cost or with a price. In this star-like topology the publishers and subscribers are the leaves. This style is advantageous in a multi-party collaboration environment with large numbers of potential e-learning service consumers and providers. It provides a good system performance because of the reduced communication overhead and it enhances the flexibility and ease of integration of additional (national) e-learning systems. In this type of topology, users can easily access the many types of e-learning services provided by collaboration parties in the NREN domain.

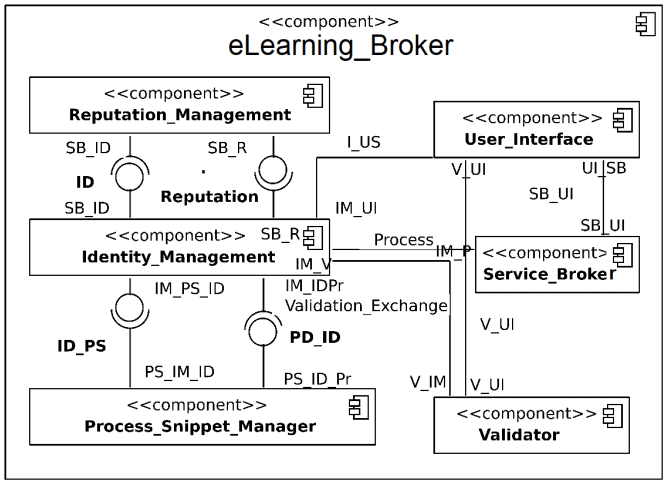


Figure 9: The e-Learning_Broker in Detail

4.1.7 The Core Services

The collaboration between organizations is based on the existing rules and procedures, and each party needs to specify their services and data. Conceptual patterns need to be set up in a technical system in such a way that collaborative parties access only the legal and correct exchangeable data.

Internally, the `eLearning_Broker` includes sub-components for `Reputation_Management` and `Identity_Management`, as well as a `Process_Snippet_Manager`, a `Validator`, the proper `Service_Broker` and a `User_Interface`.

The `Reputation_Manager` stores detailed information about collaboration parties including the list of exchangeable data/services and the level of service access. The `Identity_Manager` stores detailed information about the users, user groups, their access capabilities, and it exchanges information with all sub-components of the `eLearning_Broker`. The `Process_Snippet_Manager` stores the business processes and collaborates with the `Service_Broker` to provide the composed processes for the specific heterogeneous e-learning systems that each collaboration partner provides. There can be in fact various e-learning platforms, like Moodle or edX, that support different processes.

The `Validator` component consists of a `Process_Communicator`, a `Verifier`, a `Translator`, and an `Interface_Checker`. When this component receives a verification request, it processes it, verifies the legitimacy w.r.t the contractual sphere, the access to services via an interface, and sends back the results.

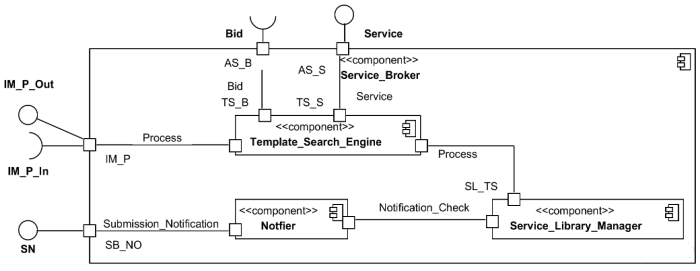


Figure 10: The e-Learning Service_Broker Details

The Service_Broker_Manager consists of three other sub-components: the Service_Library_Manager, a Template_Search_Engine, and a Notifier. The Service_Library stores the business processes of collaborative parties and provides information for the Search_Engine and Notifier. The business process specifications stored in the Service_Library_Manager works closely with the external layers of the collaboration parties. At the top-level users can search available services through the Template_Search_Engine interface of the eLearning_Broker. For instance in Figure 10 the Search_Engine exchanges data with the bid manager component for the brokering services. The User_Interface of the Broker provides access to all NREN e-Learning services through this search engine.

Finally, the Interface component provides the user access interface for the users, so that through SSO users can access many participating applications with a single login.

In Figure 8 the Balkh University and Herat University on the right side of the system architecture have the same components as Kabul University, they receive the services and data through their own E-Learning_Exchange components, and their internal architecture.

The eLearning_Broker component is an interface that facilitates the collaboration of e-learning systems through a rapid matching of e-learning needs with requests from users. Concretely, the broker works based on the collaboration agreement between the parties sharing e-learning services. Collaborating parties may have different agreements, e.g. Some exchangeable data may be offered for free while some services may need to be paid by the users or collaboration parties. The e-Learning broker also sends notifications to collaboration parties.

The central advantage of the broker in a multi-party collaboration environment with large numbers of potential e-learning service consumers and providers is high system performance, supporting efficient access to many e-learning services provided by collaboration parties. This is due to reduced communication overhead and enhanced flexibility and ease of integration of additional e-learning platforms, which were among the top quality attributes requested by the experts.

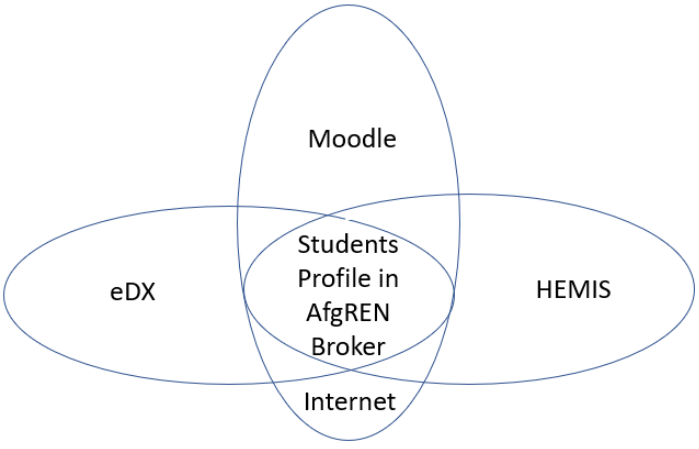


Figure 11: Successful Data Integration with the AfgREN Broker

The NREN e-learning architecture implements an abstract-data repository style. This style shields e-learning service consumers and providers of shared data from the knowledge of each other's existence and the details of their respective internal implementations. Figure 11 shows the result of the solution provided by the new broker architecture; it solves the problem of data replication in a different system located in the same NREN domain shown in Figure 1.

Our proof of concept system in the reference implementation covers all the system features listed in Figure 7. These features were discussed with domain experts and prioritized. We prototyped the application Based on the designed system architecture, and shared it with the domain experts for further feedback and enhancement.

While the prototype is designed only for the AfgREN case, it is nicely applicable also for the Technological University.

4.2 *Intra-Institutional Cross-Campus Collaboration*

Cross-campuses collaboration within a federated institution requires to focus more on the resources sharing and data integration between different different applications: the e-learning setup must filter the exchangeable data based on the respective procedures and responsibilities of the collaboration units, which usually differ. In such a context, the legal framework and procedures, security and privacy are the top priorities.

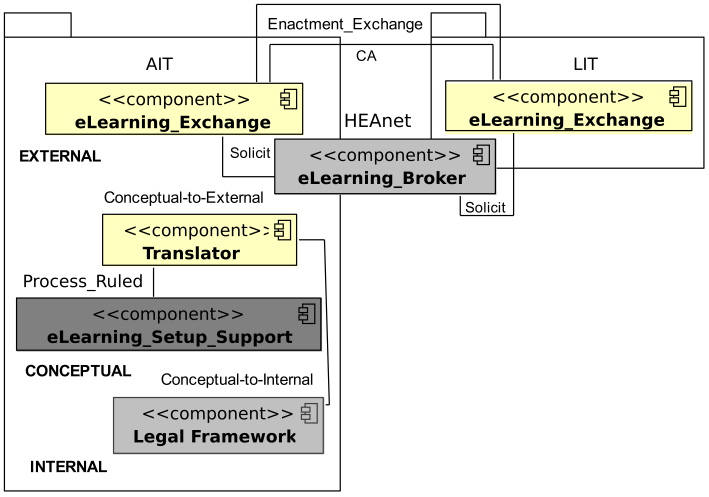


Figure 12: Intra-institutional cross-campuses collaboration Architecture

Considering that the various campuses in Figure 12 all connect through HEAnet, the collaboration setting of AIT and LIT under the new Technological University structure needs less filtering than other case studies, but still the sharing of the financial management systems and other resources sharing can be performed through the NREN Broker. The architecture proposed in Figure 12 shows that the three-layer NRENs cross-organization collaboration is applicable again. This time the focus of the translation and consistency check is between different units of the same (federated) organization and connects different e-learning applications that work in different campuses by considering the respective procedures, data security policies, and filters to apply to the exchangeable data.

The same three layers illustrated in AfgREN apply and are

replicated also for the cross-campuses collaboration. The formal move to the TU structure, with a start of operations as TU, will happen before the two ITs fully merge their policies and IT systems. The alignment of decisional and operational policies and regulations (that in Ireland are strongly local to each institution) and the de-duplication of IT systems to reach a single platform for every function will therefore happen gradually over time, if ever.

The TU will therefore face a transition time of likely several years where it will need to operate formally as one entity, but with regulations, policies and IT systems that start disjoint and will gradually converge, in some phased manner. The advantage of a broker based architecture as we propose is clear, due to the many components supporting the mapping of all these vital functions, it can evolve along the legal and operational changes, eliminating the need for costly point-to-point integration.

5 Model Driven Design of the E-learning Broker

Using DIME as a tool for prototyping and supporting XMDD and model-driven engineering method (MDE)[58] the web application development of the broker does not require extensive programming skills and it provides through the models an opportunity to involve a wide range of project stakeholders. This is a key factor for the success of a project.

We describe now the design of the AfgREN Collaboration Bridging Application, a web application that embodies the prototype for the AfgREN case study. We used the described XMDD approach and the DIME Integrated Modelling Environment as a MDD layer on top of the Eclipse IDE (integrated development environment).

For ease of understanding we describe the application design aspect by aspect, along the DIME model types: Data models for the data structures, Processes for the business logic, GUI for the web application presentation and interaction layer and Security for the Roles and rights management. We present the four aspects individually, but the prototype was developed in an agile and collaborative way, by co-designing and co-evolving the 4 types of models along the growth of the prototype in successive sprints.

5.1 Data Model

The data model in Figure 13 shows the structure of the e-Learning Broker that is an abstract model and organized elements of data that is relate to one another and to the properties of entities. The Base User, User, and Role were preexisting in DIME. We add the Broker component with its many sub-components: Thread, Category, SSO, Register, Search, Language, LegalFramwork, and detail of legal framework and many other sub components are required in the parts of the prototype. The Roles foreseen for the users of the application are a simple User group, an Admin group, a PowerUser group, and a GuestUser group. DIME creates a powerful data model for web applications with many advanced modeling features including AbstractType, ConcreteType, UserType, and EnumType, each of them have different attributes. The models are managed with bidirectional associations, association, and inheritance connection.

Three different Model Component Types are used in this data model: *Concrete types (C)* are used for the User, Broker, Thread, RegisterO, Category, SSO, Search LegalFramework, and Detail. The *User type (U)* is used for the Base User and the *Enum type (E)* is used for the Role and Language. We consider a User association connection between the User and BaseUser (in orange), a bidirectional connection between User and Thread (solid line), and an association connection between Broker and other components (arrows). In the associations as well as in the individual data types, we use square brackets to indicate lists or collections. For example, in Figure 13 the User association associates every User with the same BaseUser, but the BaseUser is associated to many concreteUsers, indicated as [concreteUser]. Similarly, a User can be associated to many Brokers and Threads.

The Model components have a rich set of typed attributes. For each attributes, the model shows whether they have no associations (yellow dot), unidirectional connection as the User.broker attribute with the Broker (yellow dot with one black dot), or bidirectional connection like the threads attribute of the User with another thread model component (yellow dot with two black dots).

All these model elements are formally defined, they are used by the DIME syntax and semantic checkers to ensure that the use of the components and data in the processes is syntactically and semantically correct. This integration of knowledge and checks across various model types is one of the advantages of XMDD as implemented in DIME in comparison with traditional model based approaches to software development, where these checks are not available, or not connected with a correct-by construction generation of code.

As we see, very little in this model is specific to a case study. While the languages will change (English and Persian in the AfgREN case, English and Irish in the TU case), as well as the specific Legal regulations (e.g., GDPR for privacy is mandatory in the EU but not in Afghanistan, which follows national laws instead) the vat majority of the fields is present and very similarly instantiated in both case studies. All the e-learning application fields for example are instantiated to Moodle in both cases, and the user categories and roles so far seem to be shareable as well.

In this sense, we think that a good part of the architecture, data model, implementation and ontology are domain specific to a federated collaboration architecture, thus widely reusable across various kinds of supervised and regulated collaboration, and that a good part of the domain specific aspects, like the e-learning applications and parts of the student management can also be shared at the platform level across all the deployment instances. Only a small part of the data model, processes, GUI etc will need to be customised to the specific collaboration and to the specific institution. For example, the fact that some institutions are connected through WiMax is taken care of at a lower layer (the underlying network and pure connectivity layer) and does not reflect itself on this architecture and models.

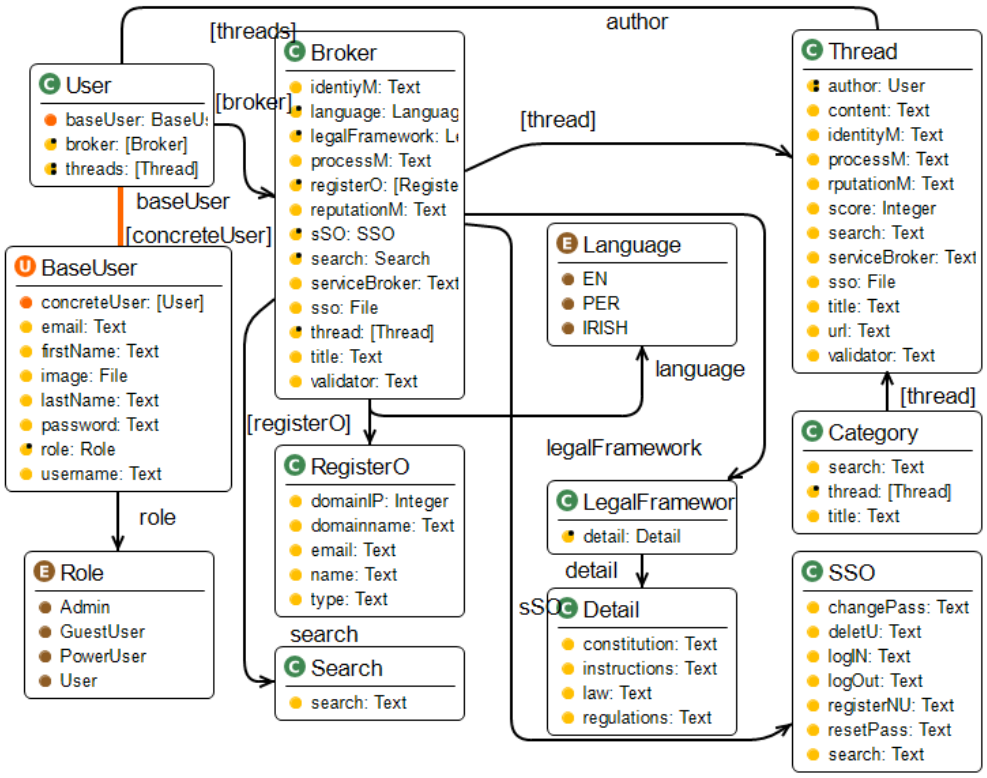


Figure 13: Data Model of the Broker Application (in DIME)

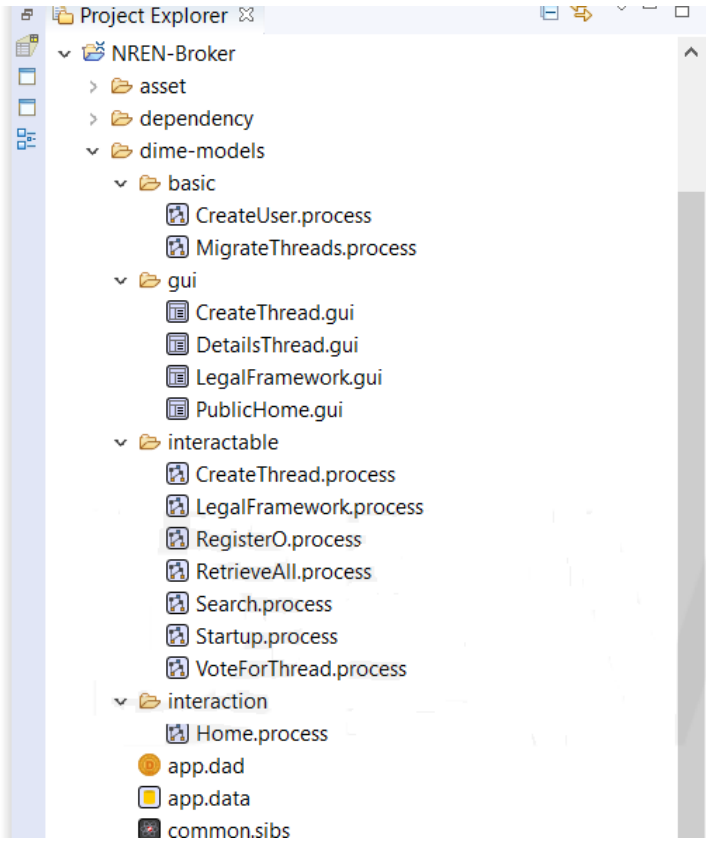


Figure 14: The e_Learning Broker Process Library (in DIME)

5.2 *Process Models*

The DIME process models support a complex business logic, with the usual patterns (sequential composition, alternative choice, parallel fork and join, and hierarchy), advanced dynamic access control and long-running processes. The process models are interdependently connected with the data model and the user interface models. A web application like the bridging application requires the collaboration of several model types. Figure 14 shows the palette of models we developed in this case study. We now describe some of them.

5.2.1 *Home Process*

The main process is the `Home process`. As shown in Figure 15, the Home Process model integrates the Public interface, Private interface and all required options for the Detail process of the e-Learning Broker application.

The process starts with the `RetrieveAllThreads` subprocess: it collects all the current learning materials including course title, learning support document, and library support materials, and then displays them on the `PublicHome` page. This is the public home page of the application, which is here represented by its GUI model. From the public home page, users can search for courses and materials. To do so, the system searches (`Search` process) the content metadata from all e-learning platforms registered in the NREN domain and collates the information in the `RetrieveAll` process. Filters can be successively applied.

Open-access material and courses that will be categorized by the Universities can be used publicly without system login and with no

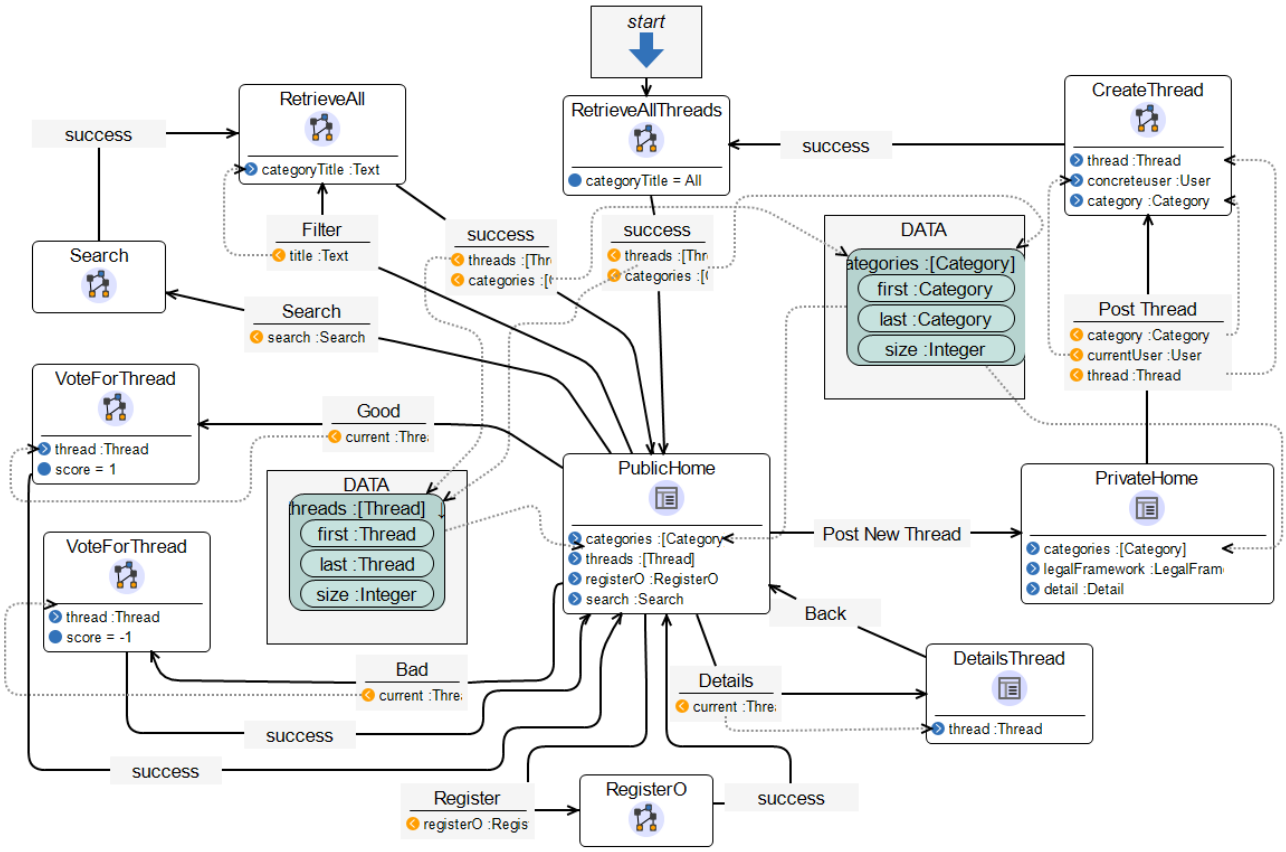


Figure 15: The e_Learning Broker Home Process: Control Flow and Data Flow in DIME

limitation. If a course requires login credential, the user is required to have a user account in the e-learning platform used for course management in their organization, and their e-learning platforms should be a member of the NREN SSO that is located in the Broker component. After logout, the user returns to the public page of the Broker application, and at the same time user logout from all collaboration e-Learning systems.

5.2.2 *Startup Process*

Figure 16 (left) shows the `Startup` process model that adds users and threads to the collaboration system in a role based fashion. The roles Admin and PowerUser can add threads. PowerUsers are involved in creating a course, learning content, and informative materials. The subjects are categorized by the name of the departments, and PowerUsers of the e-Learning application can add a list of subjects under the name of each department. The logic of this part of the process model leads the system users to a local bridging application server. The Moodle platform is also located on the same server. A title, a short description, and a URL are required for all threads (see the GUI model in Figure 17) (right). Then a category is selected and the thread is submitted. The system records the author and date of the thread creation, duplicate threads are notified to their author.

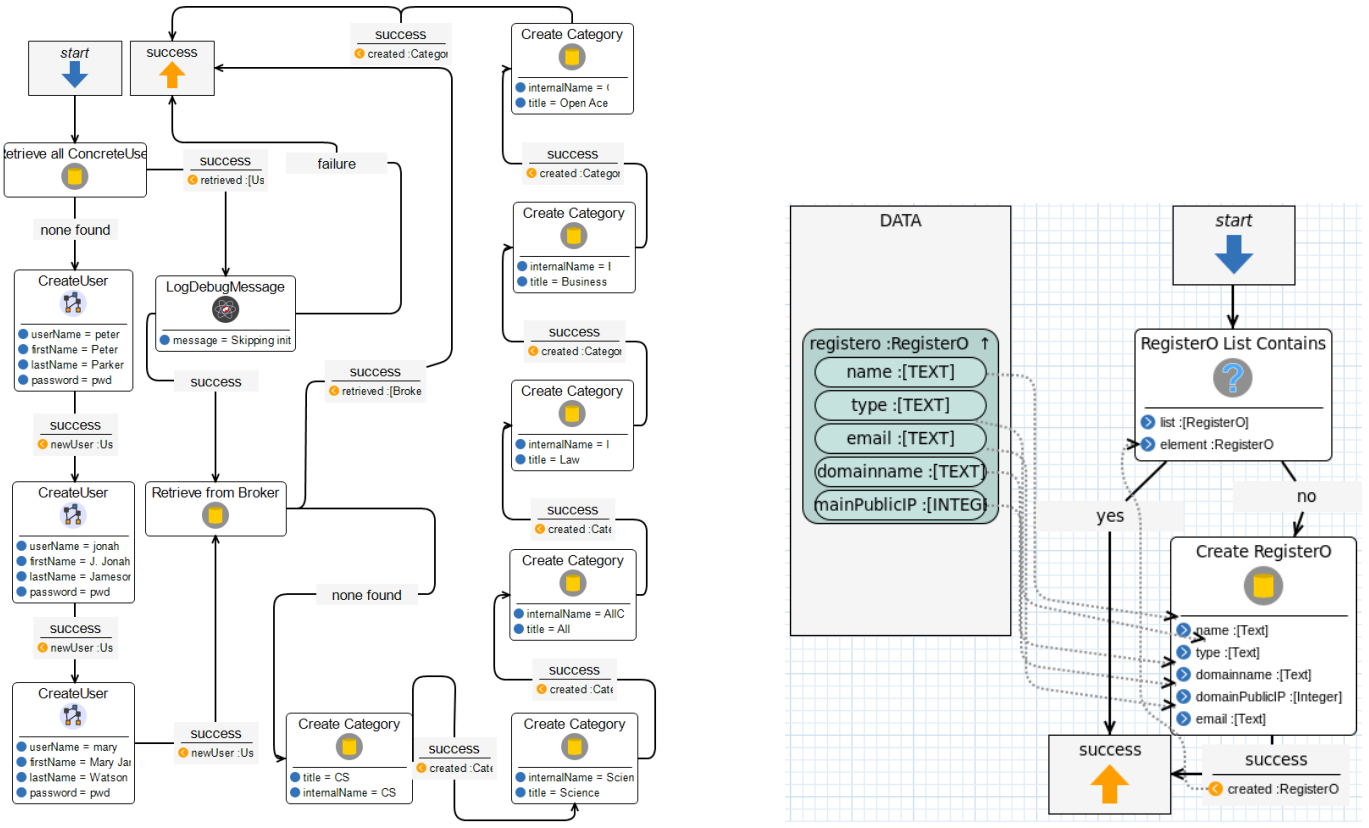
5.2.3 *Register Organization Process*

Registering new organizations to the list of collaboration parties is a core feature of the Broker, taken care of in the `RegisterO` process. Any user can add the reputation of the organization to the system, that needs approval by the Admin. The reputation of an organization is accessible through the registration button in the public interface. Figure 16(right) shows the process model for registering an organization on the basis of the name, type and official email of the organization, its domain name, and IP. This is a low-administration way to provide new facilities for the collaboration parties.

As evident from the Process Library in Figure 14, we defined and used more processes to perform other sub-tasks. From the process models briefly introduced we see that all the process models, the data model, and the interface models are integrated with each other and work together to generate a fully functional web application.

5.3 *User Interface Models*

The dynamic and responsive graphical user interface of the Collaboration Application is built by means of DIME's GUI process models, that are connected seamlessly to the business logic via hierarchical modeling. The Broker Web application has both a **public user interface** accessible to anyone, that shows a list of options and a list of subjects based on their category, and also a **private interface** accessible only to registered users. Figure 17(left) shows the top



(a) Startup Process: Control Flow

(b) RegisterO Process: Control and Data Flow

Figure 16: The e_Learning Broker Startup and Register Organization Processes in DIME

level DIME GUI model of the the public interface, and Figure 18 shows how the public interface of the deployed and running web application presents to the users.

The public user interface is intentionally kept simple. All the elements appearing on this interface (input fields, buttons, menus etc) are part of the GUI element library that comes with DIME, so there is no need to develop code for the GUI design, which is done in a purely zero-code drag and drop and configuration manner. This feature allows in particular a very efficient turn-around cycle for modifications of the look and feel of the web application.

The public interface also provides a good opportunity for the software developers to share with the reviewer team, product owner, and other stakeholders the prototype of the application, to get feedback and keep them committed to the project: these users can test and assess the application and give timely feedback, while most of them would not be able or willing to review the code.

Users of the participating institutions to the e-learning collaboration use the services of the NREN through the broker and the local e-learning systems in a secure way. They log in once to the server with their username and password, then to access other participating applications they are required at the first login a secure key. After the first login, they are enabled to use their applications without further formalities.

Once logged in, users reach their private user interface. From there they reach all the applications they are entitled to see in the NREN, and those where they can add new threads. Figure 17 shows the private interface model (right). The private interface is role-specific, currently covering course developers, system administrators, auditors, and maintainers. The administrator role of the system can define different roles for the users.

For **security purposes**, we use the Elytron[59] technology contained in WildFly 14 to provide authentication, authorization, and data encryption.

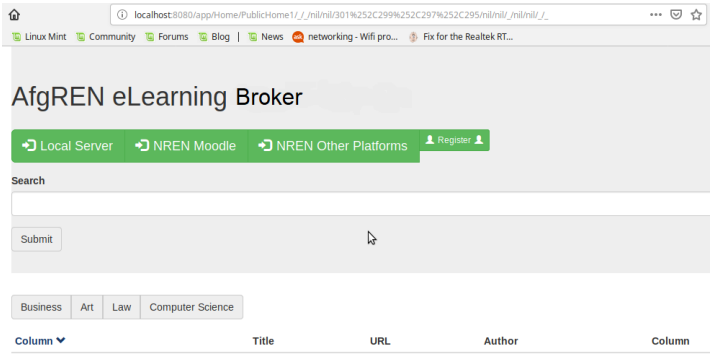
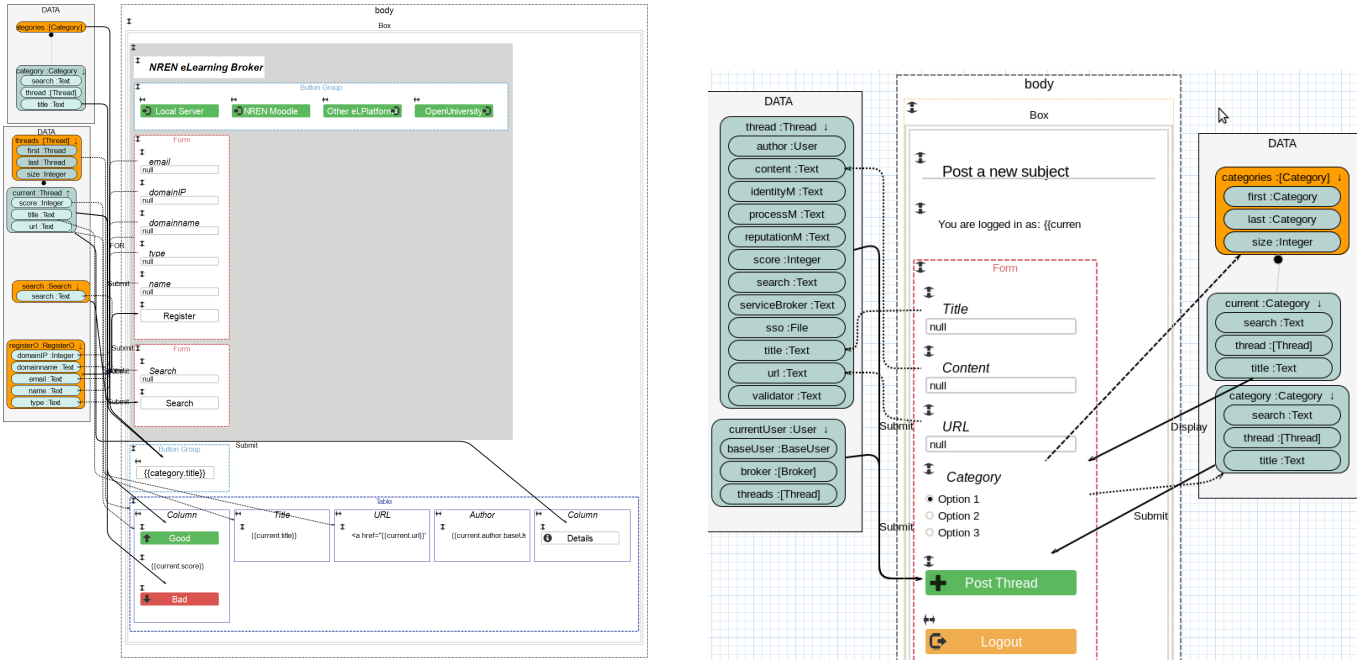


Figure 18: The e_Learning Broker Web Page

For the development of the current prototype, shown in Figure 18, we considered three options:

- In the first, a Moodle instance is installed on the server that runs the bridging application, and the Moodle application uses its Shibboleth instance to reach e-learning systems that support Shibboleth. In this case, the GUI model provides



(a) Public Interface: GUI Model

(b) Private Interface: GUI Model

Figure 17: The e_Learning Broker Public and Private Interface in DIME (GUI Models)

users with a button to login to Moodle. Once logged into this Moodle, users can access any other participating application that supports Shibboleth.

- A second option uses instead the Wildfly 14 contained in DIME. To configure the SSO, we replace the original DIME configuration with a modified ¡standalone-ha.xml¿. A key for using the SSO is created and shared with each member's application. We use this option for the applications that support wildfly Web Single Sign-On.

- Additionally, to access other e-learning platforms such as edX, and in-house developed applications, the GUI model of bridging application provides the third option for users. Currently, the button of the third option is linked to edX Afghanistan a customized e-learning application for Afghan public Universities. All three options are available in the public interface shown in figure 14 and users can select any option they required.

6 Lesson Learned

A central benefit of DIME and XMDD is that once a model is designed, due to the low-code approach to process development that reuses many preexisting components and a zero-code approach to the data design and GUI design, having the models is equivalent to having built a reference implementation. By code generation from the collection of models and components, as supported in DIME, we obtain a ready (prototype) application and this application is automatically deployed and fully functional for user testing or usage. This approach thus supports a more widely accessible collaborative DevOps style of design and development that improves over the currently standard agile approaches, that are code-based.

This way, it is possible to share and reuse the generated (native Java, Javascript etc) code as-is, but also to change the models (in the DIME graphical editor, or alternatively in the standard Xpand Eclipse textual format) and re-generate and deploy a new version. This feature in particularly attractive in an environment that is still evolving, like for the TU intra-organisational collaboration. That collaboration is actually establishing an inter-organisational collaboration between two up to now independent institutions. It is likely to start as a very lightweight federation of distinct tools and systems, joined by a properly designed and managed access (SSO, security, and roles and rights), but it will over time substitute the 2 or more systems in place up to now for each aspect of the TU's e-learning and students management with potentially only one system - either one of the current locally adopted solutions or a new one. In this context, it is important to be able to quickly experiment with alternatives, for example in a staging configuration, with quick turn around of modifications and easy and immediate access to the successive versions of the various features and services.

We already assessed that the current reference prototype is portable with minor modifications also for the Technical University use case. The detailed design of the Technological University collaboration is still ongoing and the development of the full application is accordingly future work. Along the agile development style, a first beta-release will be followed by two more increments in the coming months.

Additionally, the prototype works as generic reference platform for collaborative e-learning and student management, including for example lighter-weight collaborations with associated institutions, for example for ERASMUS+ international programs or inter-institutional national courses of study as the Irish Higher Education

Authority increasingly promotes, for example in Digital Health Transformation. The central feature of the architecture, reference application and concrete Web Application is that we can evolve and extend them with ease, integrating more functionalities and features. Specifically, the DIME models are both intuitive and formal, and designing systems with these models does not require computer programming knowledge. This provides a real opportunity to have prototypes co-developed by real users, including student administration employees, for a true co-design by those who deeply know the requirements of the organization.

7 Current Status

From the list of features and the data model of Figure 7, the SSO, registration, search engine, user accounts, legal framework, thread, and categories are already actively working. The multilingual language option and the user role management will be active in the next version of the prototype. At this stage, validation has taken place technically in the developer test and continuous feedback has been gathered from representatives of the user groups addressed in the workshops.

In the specific AfgREN context, there is a higher uniformity than in general NRENs because the members are public Universities that follow the same policy, law, and procedures defined by the Ministry of Higher Education of Afghanistan, while international NRENs are more heterogeneous. The network infrastructures, e-learning systems, and other applications are the same as in other NRENs. We provided AfgREN with an SSO by configuring the AfgREN Broker application (designed and code-generated with DIME) to serve also as an SSO server. We installed Moodle and enabled Shibboleth on it in the same server. We have currently prototypes of the searching engine, identity manager and reputation manager. We are working to develop the remaining subcomponents of the broker (validator, translator) and to add an intelligent agent in the Broker that provides a fully transparent SSO service for the different e-learning platforms implemented in the NREN domain.

8 Conclusions

We provided a modular and extensible architecture for NRENs as a modification of a previous reference architecture for business collaboration, now ported and extended to the e-learning data integration, as well as a novel architecture of an e-Learning Broker that is technology independent and thus helps furthering open education and learning collaboration. We adopted a model-driven design paradigm for the development of a reference prototype of a Web-based Collaboration Application that supports open education collaboration and data integration between different e-learning systems co-existing in an NREN domain.

In particular, we adopted AfgREN as a case study, considering the specific policy and strategy legally in place in the AfgREN domain. Based on the proposed NREN e-Learning architecture, we designed the data model, several process models and the interface models of the NREN e-learning broker. These models and processes are going to be provided open source, both in a DIME-specific format and in export formats reusable independently of DIME.

The current prototype of the reference implementation is functional, and we have indeed made use of the rapid turn-around time of design and DevOps iterations enabled by the chosen XMDD and low-code development technologies.

Acknowledgements This work was supported, in part, by Science Foundation Ireland grant 13/RC/2094 to Lero, the SFI Research Centre on Software (www.lero.ie). It was also supported in part by ALECS, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754489. A CC BY or equivalent licence is applied to the Author Accepted Manuscript, in accordance with the grant's open access conditions

References

[1] K. Upadhyaya, D. Mallik, "E-learning as a socio-technical system: An insight into factors influencing its effectiveness," Business Perspectives and Research, **2**(1), 1–12, 2013, doi:https://doi.org/10.1177/2278533720130101.

[2] A. C. Risdianto, P. M. Thet, A. Iqbal, N. A. M. Shaari, H. K. Atluri, G. N. Nurkahfi, A. Wantamanee, R. Hakimi, U. Javed, M. Ahmad, et al., "Deploying and Evaluating OF@ TEIN Access Center and Its Feasibility for Access Federation," Proceedings of the Asia-Pacific Advanced Network, **42**, 34–40, 2016.

[3] N. Suresh, J. Mbale, A. Terzoli, T. K. Mufeti, "Enhancing cloud connectivity among NRENs in the SADC region through a novel institution cloud infrastructure framework," in 2015 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), 179–184, IEEE, 2015, doi:10.1109/ETNCC.2015.7184830.

[4] L. Ssentongo, R. Kimera, B. Kakeeto, M. Mubiru, B. K. Moyer, M. Economou, C. J. Whalen, M. Tartakovsky, "Deploying Educational Roaming (eduroam) in a Rural Research Institution in Rakai, Uganda; Challenges and Lessons Learned," 2016.

[5] U. Pietrzyk, A. Zakhnini, M. Axer, S. Sauerzapf, D. Benoit, M. Gaens, "EduGATE–basic examples for educative purpose using the GATE simulation platform," Zeitschrift für Medizinische Physik, **23**(1), 65–70, 2013, doi:http://193.48.81.146/sites/default/files/dx.doi.org/10.1016/j.zemedi.2012.07.005.

[6] S. Michael, Z. J. Anna, "An Identity Provider as a Service platform for the eduGAIN research and education community," in 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), 739–740, IEEE, 2019, doi:978-3-903176-15-72019IFIP.

[7] M. Li, C.-H. Chi, C. Ding, R. Wong, Z. She, "A Multi-protocol Authentication Shibboleth Framework and Implementation for Identity Federation," in International Conference on Security and Privacy in Communication Systems, 81–101, Springer, 2018, doi:https://doi.org/10.1007/978-3-030-01704-0_5.

[8] N. Hamamoto, H. Ueda, M. Furukawa, M. Nakamura, T. Nishimura, S. Yokoyama, K. Yamaji, "Toward the Cross-Institutional Data Integration From Shibboleth Federated LMS," Procedia Computer Science, **159**, 1720–1729, 2019, doi:https://doi.org/10.1016/j.procs.2019.09.343.

[9] M. Katzer, D. Crawford, "Office 365 DirSync, ADFS, Single Sign On and Exchange Federation," in Office 365, 545–630, Springer, 2013, doi:https://doi.org/10.1007/978-1-4302-6527-6_11.

[10] P. Kamal, S. Mustafiz, F. M. A. Rahman, R. Taher, et al., "Evaluating the efficiency and effectiveness of a federated sso environment using shibboleth," Journal of Information Security, **6**(03), 166, 2015, doi:10.4236/jis.2015.63018.

[11] S. Saay, T. Margaria, "Xmdd as key enabling technology for integration of large scale elearning based on NREN," in 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT), 45–46, IEEE, 2020, doi:10.1109/ICALT49669.2020.00020.

[12] S. Saay, A. Norta, "An architecture for e-learning infrastructures on a national level: a case study of the Afghanistan Research and Education Network," International Journal of Innovation and Learning, **23**(1), 54–75, 2018, doi: https://doi.org/10.1504/IJIL.2018.088790.

[13] U. K. Hazra, R. Das, S. Mukherjee, "e-Learning Platform in SAARC Countries," Library Philosophy and Practice, 1–9, 2019, doi:https://digitalcommons.unl.edu/libphilprac/2882.

[14] A. Khaliq, M. S. Shams, M. M. Niazi, "COVID-19 Pandemic: Challenges and Opportunities for Higher Education in the Afghanistan context," doi:https://kardan.edu.af/data/public/files/KJEMS-3.3-2020-03%2014102020083157.pdf.

[15] M. Z. Sana, "Common Distributed Data Storage for Higher Education Management Information System in Afghanistan," International Journal of Scientific Engineering and Science, 32–38, 2020, doi:http://ijses.com/wp-content/uploads/2020/05/120-IJSES-V4N4.pdf.

[16] M. Beer, M. C. Meier, B. Mosig, F. Probst, "A prototype for information-dense it project risk reporting: an action design research approach," in 2014 47th Hawaii International Conference on System Sciences, 3657–3666, IEEE, 2014, doi:10.1109/HICSS.2014.456.

[17] I. Turus, J. Kleist, A. M. Fagertun, "Evaluation of flex-grid architecture for NREN optical networks," in TERENA Networking Conference 2014, 2014, doi:https://orbit.dtu.dk/en/publications/evaluation-of-flex-grid-architecture-for-nren-optical-networks.

[18] S. Mahmood, "Instructional Strategies for Online Teaching in COVID-19 Pandemic," Human Behavior and Emerging Technologies, 2020, doi:https://doi.org/10.1002/hbe2.218.

[19] S. L. J. Villalón, C. C. L. Hermosa, "The Role of National Research and Education Networks Providing Connectivity and Advanced Network Services to Virtual Communities in Collaborative R&E Projects. CUDI: The Mexican Case," in International Conference on Supercomputing in Mexico, 24–47, Springer, 2015, doi:https://doi.org/10.1007/978-3-319-32243-8_3.

[20] S. Saay, M. Laanpere, A. Norta, "Requirements for e-testing services in the AfgREN cloud-based e-learning system," in International Computer Assisted Assessment Conference, 133–147, Springer, 2016, doi:https://doi.org/10.1007/978-3-319-57744-9_12.

[21] C. Costa, H. Alvelos, L. Teixeira, "The use of Moodle e-learning platform: a study in a Portuguese University," Procedia Technology, **5**, 334–343, 2012, doi:https://doi.org/10.1016/j.protcy.2012.09.037.

[22] D. Volchek, A. Romanov, D. Mouromtsev, "Towards the semantic MOOC: Extracting, enriching and interlinking e-learning data in open edX platform," in International Conference on Knowledge Engineering and the Semantic Web, 295–305, Springer, 2017, doi:https://doi.org/10.1007/978-3-319-69548-8_20.

[23] M. Blagojević, D. Milošević, "Massive open online courses: EdX vs Moodle MOOC," in Proc. 5th International Conference on Information Society and Technology, Kopaonik, Serbia, 346–351, 2015, doi:https://aosus.org/uploads/default/original/1X/97ba7ad9a50b2a3171ebaaa64d8cf2943af37807.pdf.

[24] H. Sokout, T. Usagawa, "Analyzing the Current Situation of E-learning at Kabul Polytechnic University," in Proceedings of the 2nd International Conference on Education and Multimedia Technology, 49–53, 2018, doi: https://doi.org/10.1145/3206129.3239428.

[25] A. Naim, M. R. Hussain, Q. N. Naveed, N. Ahmad, S. Qamar, N. Khan, T. A. Hweij, "Ensuring interoperability of e-learning and quality development in education," in 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 736–741, IEEE, 2019, doi:doi:10.1109/JEEIT.2019.8717431.

[26] AIT – LIT Consortium, "The context for the new TU," in Application for designation as a technological university, 16–18, `https://www.aitlitconsortium.ie/`, 2020.

[27] A. Harding, "Design and Implementation of Service Level Agreements at HEAnet," www.heanet.ie, 2–14, 2007.

[28] T. Margaria, B. Steffen, "Agile IT: thinking in user-centric models," in International Symposium On Leveraging Applications of Formal Methods, Verification and Validation, 490–502, Springer, 2008, doi:https://doi.org/10.1007/978-3-540-88479-8_35.

[29] S. Sharma, N. Hasteer, "A comprehensive study on state of Scrum development," in 2016 International Conference on Computing, Communication and Automation (ICCCA), 867–872, IEEE, 2016, doi:10.1109/CCAA.2016.7813837.

[30] S. Saay, A. Norta, "Designing a Scalable Socio-Technical Method for Evaluating Large e-Governance Systems," in Advanced Computational and Communication Paradigms, 571–580, Springer, 2018, doi:https://doi.org/10.1007/978-981-10-8240-5_64.

[31] D. Renzel, M. Behrendt, R. Klamma, M. Jarke, "Requirements bazaar: Social requirements engineering for community-driven innovation," in 2013 21st IEEE International Requirements Engineering Conference (RE), 326–327, IEEE, 2013, doi:doi:10.1109/RE.2013.6636738.

[32] S. Saay, T. Margaria, "Model-Driven-Design of NREn Bridging Application: Case Study AfgREN," in 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), 1522–1527, IEEE, 2020, doi:doi:10.1109/COMPSAC48688.2020.00-39.

[33] T. Margaria, B. Steffen, "eXtreme Model-Driven Development (XMDD) Technologies as a Hands-On Approach to Software Development Without Coding," Encyclopedia of Education and Information Technologies, 732–750, 2020, doi:https://doi.org/10.1007/978-3-030-10576-1.

[34] T. Margaria, B. Steffen, "Service-orientation: conquering complexity with XMDD," in Conquering Complexity, 217–236, Springer, 2012, doi:https://doi.org/10.1007/978-1-4471-2297-5_10.

[35] S. Boßelmann, M. Frohme, D. Kopetzki, M. Lybecait, S. Naujokat, J. Neubauer, D. Wirkner, P. Zweihoff, B. Steffen, "DIME: a programming-less modeling environment for web applications," in International Symposium on Leveraging Applications of Formal Methods, 809–832, Springer, 2016, doi: https://doi.org/10.1007.

[36] J. Neubauer, M. Frohme, B. Steffen, T. Margaria, "Prototype-driven development of web applications with DyWA," in International Symposium On Leveraging Applications of Formal Methods, Verification and Validation, 56–72, Springer, 2014, doi:https://doi.org/10.1007/978-3-662-45234-9_5.

[37] N. Moreno, P. Fraternali, A. Vallecillo, "WebML modelling in UML," IET software, **1**(3), 67–80, 2007, doi:10.1049/iet-sen:20060067.

[38] T. Margaria, A. Schieweck, "The Digital Thread in Industry 4.0," in International Conference on Integrated Formal Methods, 3–24, Springer, 2019, doi:https://doi.org/10.1007/978-3-030-34968-4_1.

[39] S. Jörges, A.-L. Lamprecht, T. Margaria, I. Schaefer, B. Steffen, "A constraint-based variability modeling framework," International Journal on Software Tools for Technology Transfer, **14**(5), 511–530, 2012, doi:https://doi.org/10.1007/s10009-012-0254-x.

[40] M. Karusseit, T. Margaria, "Feature-based modelling of a complex, online-reconfigurable decision support service," Electronic Notes in Theoretical Computer Science, **157**(2), 101–118, 2006, doi:https://doi.org/10.1016/j.entcs.2005.12.049.

[41] M. Kircher, P. Jain, Pattern-oriented software architecture, patterns for resource management, volume 3, John Wiley & Sons, 2013.

[42] A. Patwardhan, R. Patwardhan, S. Vartak, "Self-Contained Cross-Cutting Pipeline Software Architecture," arXiv preprint arXiv:1606.07991, 2016, doi: arXiv:1606.07991v1.

[43] O. Kaiwartya, A. H. Abdullah, Y. Cao, A. Altameem, M. Prasad, C.-T. Lin, X. Liu, "Internet of vehicles: Motivation, layered architecture, network model, challenges, and future aspects," IEEE Access, doi:doi:10.1109/ACCESS.2016.2603219.

[44] M. Richards, Software architecture patterns, volume 4, O'Reilly Media, Incorporated 1005 Gravenstein Highway North, Sebastopol, CA ..., 2015.

[45] O.-A. Schipor, R.-D. Vatavu, J. Vanderdonckt, "Euphoria: A Scalable, event-driven architecture for designing interactions across heterogeneous devices in smart environments," Information and Software Technology, **109**, 43–59, 2019, doi:https://doi.org/10.1016/j.infsof.2019.01.006.

[46] A. Balalaie, A. Heydarnoori, P. Jamshidi, "Microservices architecture enables devops," London: Sharif University of Technology, 2014.

[47] N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, L. Safina, "Microservices: yesterday, today, and tomorrow," Present and ulterior software engineering, 195–216, 2017, doi:https://doi.org/10.1007/978-3-319-67425-4_12.

[48] F. J. García-Peñalvo, J. Cruz-Benito, C. Maderuelo, J. S. Pérez-Blanco, A. Martín-Suárez, "Usalpharma: A cloud-based architecture to support Quality Assurance training processes in health area using Virtual Worlds," The Scientific World Journal, **2014**, 2014, doi:https://doi.org/10.1155/2014/659364.

[49] S. Saay, "A Reference Architecture for A National E-Learning Infrastructure," in Ph.D Thesis, 20–80, Tallinn University, 2018.

[50] L. Chen, "Microservices: architecting for continuous delivery and DevOps," in 2018 IEEE International conference on software architecture (ICSA), 39–397, IEEE, 2018, doi:doi:10.1109/ICSA.2018.00013.

[51] S. Saay, A. Norta, M. Laanpere, "Towards an architecture for e-learning infrastructures on a national level: a case study of AfgREN," in International Conference on Web-Based Learning, 98–107, Springer, 2015, doi:https://doi.org/10.1007/978-3-319-32865-2_11.

[52] M. Karusseit, T. Margaria, "Feature-based modelling of a complex, online-reconfigurable decision support service," Electronic Notes in Theoretical Computer Science, **157**(2), 101–118, 2006, doi:https://doi.org/10.1016/j.entcs.2005.12.049.

[53] G. Jung, T. Margaria, R. Nagel, W. Schubert, B. Steffen, H. Voigt, "SCA and jABC: Bringing a service-oriented paradigm to web-service construction," in International Symposium On Leveraging Applications of Formal Methods, Verification and Validation, 139–154, Springer, 2008, doi:https://doi.org/10.1007/978-3-540-88479-8_11.

[54] T. Margaria, "Components, Features, and Agents in the ABC," in Objects, Agents, and Features, 154–174, Springer, 2004, doi:https://doi.org/10.1007/978-3-540-25930-5_10.

[55] T. Margaria, B. Steffen, M. Reitenspieß, "Service-oriented design: the roots," in International Conference on Service-Oriented Computing, 450–464, Springer, 2005, doi:https://doi.org/10.1007/11596141_34.

[56] J. Neubauer, B. Steffen, T. Margaria, "Higher-order process modeling: product-lining, variability modeling and beyond," arXiv preprint arXiv:1309.5143, 2013, doi:,doi:10.4204/EPTCS.129.1.

[57] A. Norta, P. Grefen, N. C. Narendra, "A reference architecture for managing dynamic inter-organizational business processes," Data & Knowledge Engineering, **91**, 52–89, 2014, doi:https://doi.org/10.1016/j.datak.2014.04.001.

[58] S. Boßelmann, A. Wickert, A.-L. Lamprecht, T. Margaria, "Modeling directly executable processes for healthcare professionals with xmdd," in Service Business Model Innovation in Healthcare and Hospital Management, 213–232, Springer, 2017, doi:https://doi.org/10.1007/978-3-319-46412-1_12.

[59] F. Marchioni, WildFly Administration Guide: The ultimate and most up-to-date guide to manage WildFly appication server, ITBuzzPress, 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 218-223 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Analysis of Real-time Blockchain Considering Service Level Agreement (SLA)

Minkyung Kim[1], Kangseok Kim[2,3], Jai-Hoon Kim[*,3]

[1]*Dasan College, Ajou University, Suwon, 16499, South Korea*

[2]*Department of Artificial Intelligence and Data Science, Ajou University, Suwon, 16499, South Korea*

[3]*Department of Cyber Security, Ajou University, Suwon, 16499, South Korea*

ARTICLE INFO

Article history:
Received: 15 October, 2020
Accepted: 21 December, 2020
Online: 15 January, 2021

Keywords:
Consensus Algorithm
Byzantine Fault Tolerance
Blockchain
Internet of Things
Decentralized Framework

ABSTRACT

The Blockchain technologies enable decentralized networking consisting of large number of nodes. To determine the shared states and failures of all nodes in a fully distributed peer-to-peer system, the appropriate consensus algorithm needs to be selected for each Internet of Things system. In this paper, a novel hierarchical voting-based byzantine fault tolerance (HBFT) consensus algorithm is proposed. The proposed HBFT algorithm utilizes a typical PBFT algorithm hierarchically to guarantee low latency. The message complexity of HBFT shows that our proposed algorithm has better scalability. We also mathematically calculate the optimal number of groups based on the total number of nodes to determine the ratio of allowable faulty nodes per group. In addition, we analyze the reliability of byzantine fault tolerance to compare the reliability of group case with the reliability of non-group case. Finally, we introduce the methods of real-time Blockchain considering the service level agreement (SLA). The real-time processing performance of transactions is analyzed for the service level agreement (SLA).

1. Introduction

A Blockchain is a peer-to-peer distributed ledger , in which the process of transaction verification and recording is continuously executed [1, 2]. Participant nodes constantly verify a set of time-stamped transactions at a given time using a built-in consensus algorithm. The Blockchain technology improves reliability in addition to availability by storing and sharing data in a distributed manner. The use of Blockchain in Internet of Things (IoT) enables secure, decentralized networking for IoT data privacy and security [3]. Therefore, an efficient consensus algorithm is crucial to verify transactions and adjust interactions among IoT devices [4]. IoT data are validated through the consensus mechanism and then recorded securely in a distributed ledger.

This paper introduces a novel hierarchical voting-based byzantine fault tolerance (HBFT) consensus algorithm. The proposed HBFT algorithm operates the group level's consensus rather than the node level's consensus. It is to overcome limited scalability and high latency in a typical practical byzantine fault tolerance (PBFT) algorithm. Thus, the overall exchanging messages are significantly reduced even if the large number of nodes are participated in consensus operation. Despite the exponential growth in the number of nodes in the network, the scalability of HBFT is ensured compared with PBFT. Also, we calculate the optimal number of groups depending on the total number of nodes for grouping because the reliability can be affected by the number of groups. If faulty nodes are well distributed into groups, the probability of reaching consensus is higher even with an increased number of faulty nodes. Therefore, two cases are calculated to understand the influence on the number of faults per group.

However, since the participant nodes mutually verify and store the data, considerable number of computing resources are required, and concurrently the processing time decreases [5-8]. The same method of participant nodes (all nodes or certain nodes allowed to participate in Blockchain) has been used to process transactions on the Blockchain [9]; nevertheless, it is possible to apply the service level agreement (SLA) on Blockchain by determining the number of participant nodes according to the user's requirements. Therefore, we propose methods to control the number of participant nodes adaptively considering user requirements and computing environments. The transaction processing speed and throughput can be improved by using adaptive control method. Figure 1 is the diagram of real-time

*Corresponding Author: Jai-Hoon Kim, Department of Cyber Security, Ajou University, Suwon, 16499, Korea, jaikim@ajou.ac.kr

Blockchain considering service level agreement(SLA). In this diagram, the reliability of a node with respect to time is periodically collected by the Blockchain monitor. It depends on the failure rate of a node determined intentional or unintentional faults. The Blockchain control also reads user requirements and Blockchain status. According to the user's request, a transaction is processed for generating a smart contract of Blockchain. The user can request reliability, maximum time and the cost for block generation. The reliability of a participant node can affect the reliability of the transaction processed.

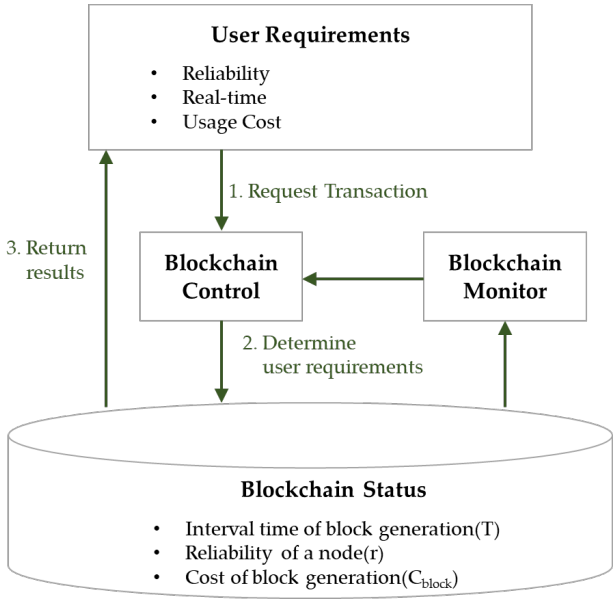


Figure 1: Schematic of Real-time Blockchain considering Service Level Agreement(SLA)

The remainder of this paper is organized as follows. Section 2 overviews relevant background about various consensus algorithms. The details of HBFT algorithm are presented in Section 3. Here, how the HBFT algorithm differs from previous consensus mechanisms has been described. In Section 4, we analyze the reliability of byzantine fault tolerance, and present the algorithm of real-time block generation in Section 5. Finally, Section 6 concludes the contributions with future research perspectives and additional improvements.

2. Related Work

There are various consensus algorithms adopted by the popular Blockchain-based platforms such as *Bitcoin*, *Ethereum* and *Hyperledger Fabric*. In [10-11], the different advantages and disadvantages of each existing consensus algorithm are introduced. Bitcoin utilizes Proof of Work (PoW), which the high computational power is required to solve a cryptographic puzzle for mining operation. Proof of Stake (PoS) is designed to overcome the disadvantages of PoW algorithms. However, PoS suffers from a problem called Nothing at Stake because each participant can vote to both blocks. Also, Delegated Proof of Stake (DPoS) achieves consensus by delegates (called block producers) elected by nodes instead of participation by all nodes. Practical Byzantine Fault Tolerance (PBFT), implemented in *Hyperledger Fabric*, has a primary node to manage the other nodes. Even if some untrusted nodes participate in the consensus operation, it is considered an agreement if more than 2⁄3 of the results are the same. Also, a variant of byzantine fault tolerance consensus algorithm called Delegated Byzantine Fault Tolerance (DBFT) [12, 13] is introduced for scalable participation. The randomly elected delegates (called bookkeeping nodes) of each group can only participate in the operation of the consensus process. Furthermore, PBFT is an permissioned protocol in which the identities of all nodes are exposed to the network [14]. Therefore, PBFT can be applied in permissioned platforms that enable autonomous, commercial, and financial application services in IoT environments. It is also suitable for applications that do not require tokens and incentives. That is why we developed a novel and efficient consensus algorithm based on PBFT for decentralized IoT.

3. Hierarchical voting-based Byzantine Fault Tolerance (HBFT) Consensus Algorithm to reduce System Resources

In this section, a novel hierarchical voting-based byzantine fault tolerance (HBFT) consensus algorithm is proposed. Also, we analyze the message complexity of the proposed novel consensus algorithm. Additionally, the best case and worst case are calculated to identify the influence on the number of faulty nodes per group in our proposed algorithm.

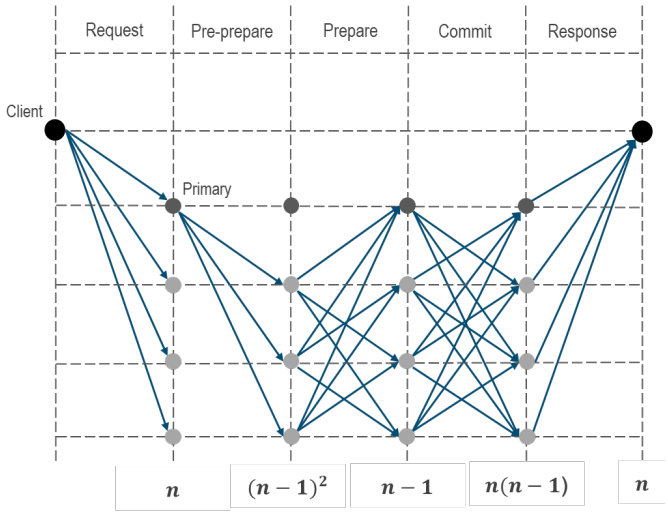


Figure 2: Flow Chart of PBFT

Practical Byzantine Fault Tolerance (PBFT) consensus algorithm has five processes for consensus processing [15-17]. Figure 2 shows the flow chart of PBFT. If the number of faulty nodes (f) is not more than $f = \lfloor (n-1)/3 \rfloor$, a consensus can be reached. Even though faulty nodes participate in the network, PBFT consensus algorithm can prevent up to $n/3$ of faulty nodes since more than $2n/3$ nodes must be reliable nodes. The overall message complexity is $2n^2$ based on the messages generated in each process. In case of the PBFT algorithm supported by *HyperLedger Fabric*, the reliability of the consensus is improved because every node participates in the consensus process through the interactions with each other. However, the message complexity

increases exponentially as the number of participating nodes increase. Therefore, if the number of nodes increases, the problem of scalability can be generated by high network traffic and latency.

3.1. Proposed HBTF Consensus Algorithm for Scalability

Based on our previous researches [18, 19], the operation of HBFT consensus algorithm, which utilizes a typical PBFT algorithm hierarchically, is re-designed effectively. Our proposed HBFT consensus algorithm has two phases to validate transactions. The flow chart of HBFT is shown in Figure 3.

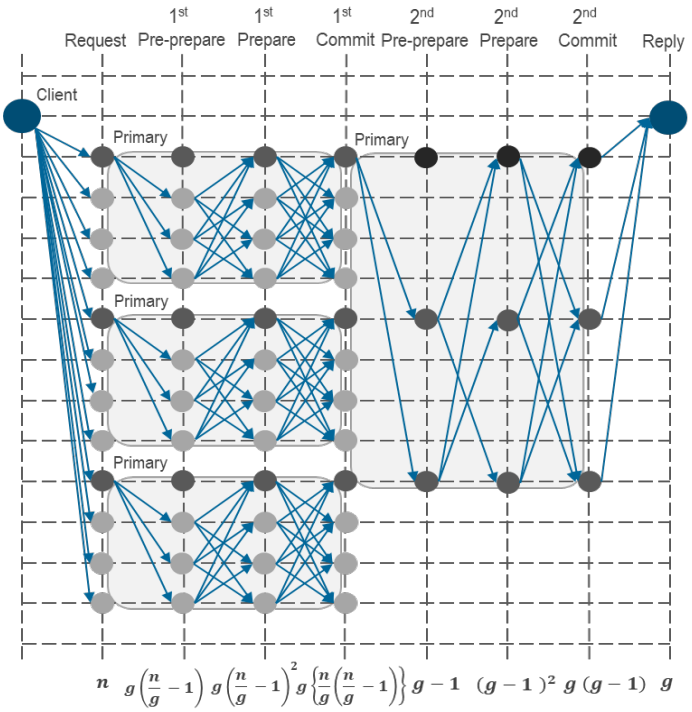


Figure 3: Flow Chart of HBFT

Firstly, the first phase start to request a verification about new transactions to all replica nodes. In the pre-prepare process, all nodes are randomly grouped into some groups and then, a primary node in each group is randomly selected to send a pre-prepare message to the other nodes. Each node compares a pre-prepare message with a request message. If the results match, a primary node responds a prepare message to the other nodes. In the commit process, each node reaches an agreement in the group depending on whether it received more than $2n/3g$ of prepare messages. Next, the second phase is executed in the same processes as the first phase to verify the voting results of each group. Finally, if more than $2g/3$ of prepare messages are received, the consensus is reached. Each primary node broadcasts the results to a client.

3.2. Message Complexity of HBFT Consensus Algorithm

The overall message complexity generated in each process of HBFT is $f(g) = 2g^2 - g + 2n^2/g - n$. The overall message complexity of HBFT is lower than that of PBFT by increasing the total number of participant nodes in Figure 4. Therefore, the processing overhead is significantly decreased. The lower latency and better scalability can be guaranteed because the overall overhead is reduced.

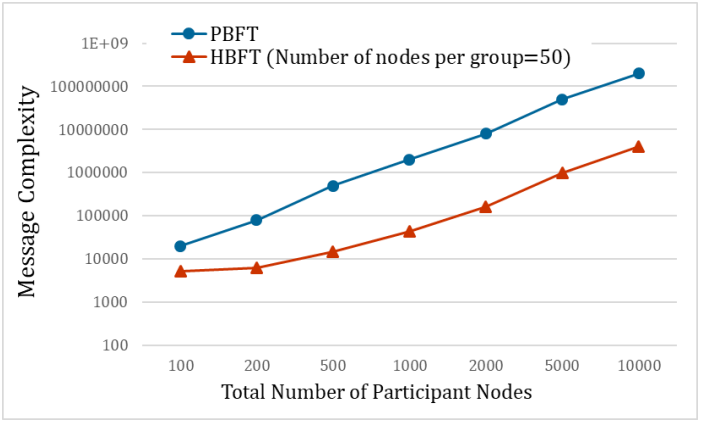


Figure 4: Overall Message Complexity Comparison between PBFT and HBFT

Additionally, we try to figure out the impact on the message complexity by increasing the number of groups in Figure 5. If the total number of nodes is 1000, the message complexity depends on the number of groups. The message complexity increases as the number of groups is too small or too large. Since the number of groups can affect the reliability of the agreement, the optimal number of groups according to total number of participant nodes needs to be calculated for facilitating efficient grouping. The optimal number of groups based on total number of nodes is calculated as follows. Eq. (1) is an equation for the overall message complexity of HBFT. By differentiating Eq. (1), we can achieve the number of groups with minimal message complexity. The optimal number of groups is given by Eq. (2). If the total number of nodes is 1000, the calculation result of Eq. (2) is about 79.45 as shown in Figure 5.

$$f(g) = 2g^2 - g + \frac{2n^2}{g} - n \tag{1}$$

$$\rightarrow \quad f'(g) = 4g - 1 - \frac{2n^2}{g^2} = 0$$

$$g = \frac{1}{12}\left(1 + \sqrt[3]{1 + 432n^2 - \sqrt{432n^2(432n^2 + 2)}} + \sqrt[3]{1 + 432n^2 + \sqrt{432n^2(432n^2 + 2)}}\right) \tag{2}$$

3.3. Fault Tolerance of HBFT Consensus Algorithm

In our proposed HBFT algorithm, consensus can be achieved to add a new block when the number of groups is more than two-thirds the total number of groups. Two cases are calculated to determine the influence on the number of faulty nodes per group. In the first case, the maximum number of allowable faults to reach consensus is calculated in (a). Additionally, a case where the consensus cannot be reached with the minimum number of faulty nodes is calculated in (b). When it is detected as a faulty node, a primary node is changed in the typical PBFT algorithm. Therefore, all primary nodes are assumed to be non-faulty nodes, similar to the PBFT algorithm. Additionally, we assume that the number of total groups is more than two, and the number of nodes per group

is at least four in our proposed HBFT algorithm for normal operation.

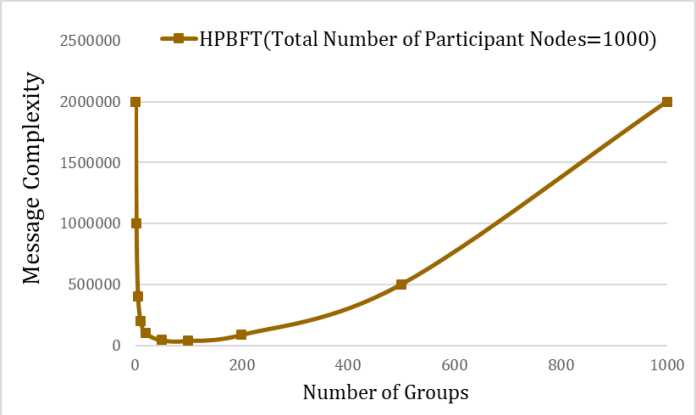


Figure 5: Message Complexity of HBFT depending on the Number of Groups

3.3.1. *Best case: The maximum number of allowable faulty nodes in order to reach a consensus*

One-third of total number of groups consists of faulty nodes and cannot reach a group consensus. The other two-thirds have only one-third faulty nodes in each group, leading to a group agreement. Consequently, the best case is that the number of faulty nodes is high; however, a consensus can be reached since faulty nodes are well distributed into groups. Eq. (3) shows the equation of the best case.

$$\left(\left\lfloor\frac{g}{3}\right\rfloor \times \left(\frac{n}{g}-1\right)\right) + \left(\left\lceil\frac{2g}{3}\right\rceil \times \left\lfloor\frac{n}{3g}\right\rfloor\right) \quad (3)$$

3.3.2. *Worst case: A consensus cannot be reached with the minimum number of faulty nodes*

If the number of faulty nodes in each group exceeds one-third, each group agreement fails. The worst case is in which the number of groups that do not attain group agreement exceeds one-third. As a result, even though the number of faulty nodes occupies a relatively small number, the consensus cannot be reached according to the ratio of faulty nodes per group. An equation of this worst case is described as follows.

$$\left(\left\lfloor\frac{g}{3}\right\rfloor + 1\right) \times \left(\left\lfloor\frac{n}{3g}\right\rfloor + 1\right) \quad (4)$$

Both equations are calculated to assess whether a consensus can be achieved according to the ratio of faulty nodes per group. The influence of the ratio of faulty nodes per group is shown in Figure 6. From these calculations, the number of allowable faulty nodes is up to approximately $5n/9 - g/3$ in the best case. In the worst case as already discussed, a consensus cannot be reached according to the ratio of faulty nodes per group. In both calculations, whether a consensus is reached or not is affected by the ratio of faulty nodes per group. Although the number of faulty nodes participating in consensus is high, a consensus can be reached depending on how all faulty nodes are well-distributed into groups as described in the best case. Finding the optimal ratio of faulty nodes per group to reach consensus in HBFT algorithm will be an important research area. In addition, it is required that the verification regarding the faulty nodes is executed at the beginning of consensus process. As a result, the number of allowable faulty nodes can be low compared to PBFT algorithm depending on the ratio of faulty nodes per group in the worst case. On the contrary, the number of allowable faulty nodes can be more than that in the PFBT algorithm depending on the ratio of faulty nodes per group in the best case.

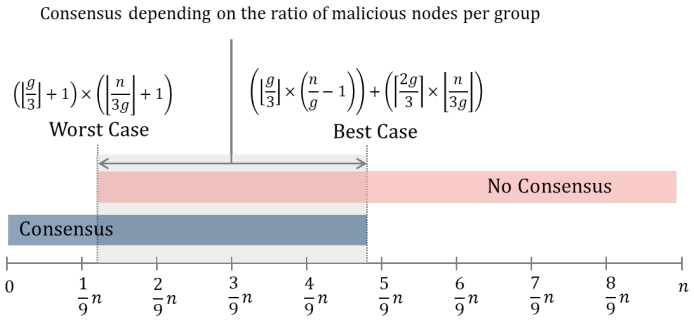


Figure 6: Influence on the Ratio of Faulty Nodes Per Group

4. Analysis of Reliability

We analyze the reliability of Blockchain operation in this section. It leads to the analyzing of the probability of normal operation in unit time. Assuming that the reliability of each participant node is equal to *r*, the calculations below are defined when *n* nodes participate in the processing a transaction. The reliability of the fail-stop faults can be calculated by the probability that at least one node is operating normally. A fail-stop means the occurrence of an unintentionally generated failure. It is assumed that a fail-stop at a node happens when a node is faulty.

$$1-(1-r)^n \quad (5)$$

In the case of the reliability for consensus, even if there are the fail-stop faults, the consensus can be reached if more than half of the number of nodes agree. Subsequently, the reliability of the byzantine faults can be calculated by the probability that less than half of faulty nodes are included on Blockchain as follows.

$$R(n,r) = 1 - \sum_{k=ceil(\frac{n}{2})}^{n} {}_nC_k(1-r)^k r^{n-k} =$$

$$\sum_{k=0}^{ceil\left(\frac{n}{2}-1\right)} {}_nC_k(1-r)^k r^{n-k} \quad (6)$$

In the case of the reliability for consensus in byzantine faults, if less than one-third of faulty nodes are in the operation, the consensus is reached. The calculation is shown in Eq. (7)

$$R(n,r) = 1 - \sum_{k=ceil(\frac{2n}{3})}^{n} {}_nC_k(1-r)^k r^{n-k} =$$

$$\sum_{k=0}^{ceil\left(\frac{2n}{3}-1\right)} {}_nC_k(1-r)^k r^{n-k} \quad (7)$$

In addition, the reliability of the byzantine faults in a hierarchical structure as in the proposed HBFT consensus algorithm can be defined as *R*(*g*, *R*(*n*/*g*, *r*)). It is assumed that the total number of nodes is n, and the number of nodes per a group is *n*/*g*. Additionally, when all nodes are divided into several groups for consensus, the number of groups is g. Therefore, the reliability of group is calculated as *R*(*n*/*g*, *r*) based on the reliability of a node. In case of the reliability for consensus, even if there are the

byzantine faults, this same definition is applied. If the more than two-third the number of normal nodes are in the consensus operation, a consensus is reached to add a new block.

Assuming that the number of nodes is 400 and less than half the faulty nodes are included, the reliability based on that of each node is shown in Figure 7. Furthermore, if the number of groups is 20 with the total number of nodes, the reliability of hierarchical byzantine fault tolerance by grouping is shown in Figure 7. When the consensus operation executes hierarchically in the group level, the reliability of group case is similar or lower than the reliability of non-group case. Although the difference in the case of reliability of hierarchical byzantine fault tolerance is lower than in the case by the node's reliability, the message complexity is much better described in the Section 3.2; so it can be said that it is good to be grouped in the proposed algorithm. When the number of faulty nodes is less than one-third of the total nodes, the consensus operates hierarchically according to the reliability of each node.

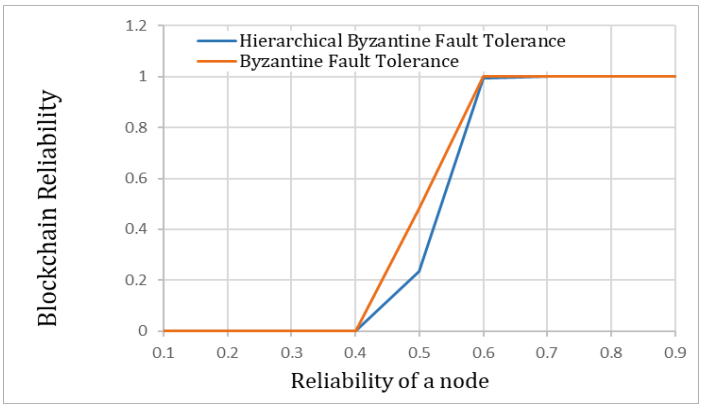


Figure 7: Reliability in the Byzantine Fault

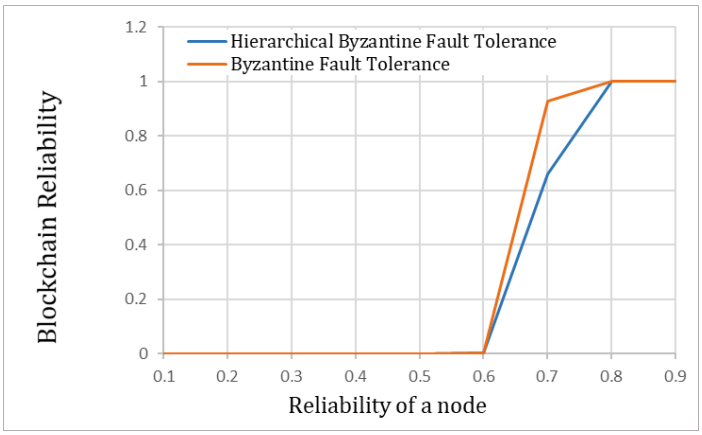


Figure 8: Reliability for Reaching Consensus in the Byzantine Fault

Figure 8 shows that the reliability of the consensus in the case of group level and in the case of non-group level consensus. Although the range with a little low reliability exists in the reliability of hierarchical byzantine fault tolerance for reaching consensus, there is no notable difference in the reliability of the overall ranges, similar to Figure 7. Even if the reliability of each node is different, or an incorrect answer is generated by an intentional failure or byzantine faults, the reliability can be calculated. In addition, even if the participant nodes have different conditions such as reliability, the possibility of cyber infringement, computing power, and data collection, it is possible to apply the proposed method of using Blockchain resources minimally according to various Blockchain configuration conditions and the user's requirements.

5. Real-time Performance Analysis using Block Generation Algorithm

Although the Blockchain has various advantages such as reliability, security, traceability and transparency, the excessive computing resources and communication bandwidths required to reach consensus and maintain consistency between duplicated ledgers are disadvantageous. Additionally, it is necessary to consider the reduction of the deadline meet ratio according to the laxity time for consensus in computing environments required real-time transaction processing. In this paper, the real-time processing performance of transactions is analyzed according to the redundancy of nodes. There are the applicable methods for maximizing real-time performance and satisfying user's requirements. The first method is real-time block generation algorithms. The second method is to use off-chain for reliable real-time performance although the reliability is decreased. A method to obtain the priority for block generation exists, as it pays more gas costs. Subsequently, the delay time is controlled by adjusting the number of nodes participating in consensus. In this section, we propose real-time block generation algorithm to adjust block generation according to the laxity time of the transaction. The two algorithms of block generation are compared in Figure 9 and Figure 10.

5.1. Algorithm of block generation at regular intervals (Conventional Blockchain)

Assuming that the time interval of block generation is T, the transaction arrival time is t, and the laxity time for block generation is *D*. When each block is generated at regular intervals, the deadline meet ratio, the average response time and the cost for block generation can be calculated as follows. A block is generated when a transaction arrives in the period *D*. Therefore, the deadline meet ratio is calculated as *D/T*. In Figure 9, if the period of *D* is greater than the period of *T*, a block is generated at any time within the period of block generation. The average response time is defined as *D/2*. The period *D* refers to the laxity time to wait for block generation. The cost for block generation in period of T is calculated as C_{block}/T. The graph of (a) algorithm for the cost is shown in Figure 10.

5.2. Algorithm of block generation by the shortest deadline of the arrived transactions (Real-time Blockchain)

If the transaction processing request arrives, and other transactions arrive at *1/f* intervals, all blocks about the arrived transactions within the laxity time of the first requested block generation are generated. Therefore, all transactions in the laxity time are processed based on the laxity time of the initially arrived transaction. It is assumed that the arrival frequency of transaction is *f*, and a new transaction arrives at equal intervals. Then, the arrival time of a new transaction is *1/f* (= *P*), and the block generation time of the new transaction can be calculated as *1/f* (= *P*) + *D*. All transactions are always processed when the transactions arrive in the laxity time. Therefore, the deadline meet

ratio of (b) algorithm is 1 as shown in Figure 9. The average response time for block generation is calculated as $(P + D)/2$. The graph in Figure 10 shows the cost for block generation of (b) algorithm calculated as $C_{block}/P + D$.

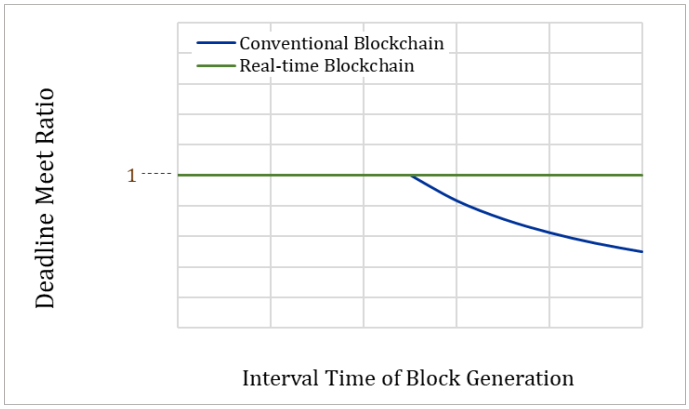


Figure 9: Deadline Meet Ratio

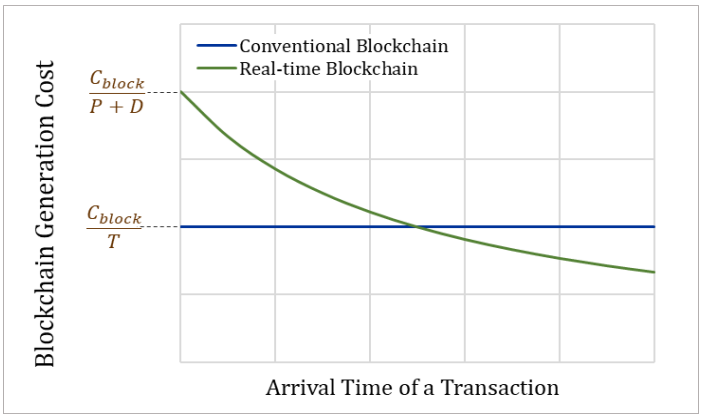


Figure 10: Blockchain Generation Cost

6. Conclusion

We propose a novel HBFT consensus algorithm required for transaction generation to reduce the processing throughput and increase scalability in a typical PBFT algorithm. Since our proposed algorithm utilize PBFT algorithm hierarchically, the number of exchanging messages in the consensus processing is significantly reduced. It means that the HBFT consensus algorithm ensures the scalability even though large numbers of nodes are involved in the consensus process. Furthermore, we calculate the optimal number of groups to identify the influence on the overall message complexity depending on the number of groups. In addition, both the best and worst cases are calculated to determine the ratio of allowable faulty nodes per group. Additionally, the reliability of byzantine fault tolerance is analyzed to compare the reliability of group case with the reliability of non-group case. Finally, it is necessary to consider the reduction of the deadline meet ratio according to the laxity time for consensus in computing environments required real-time transaction processing. To compare with the conventional Blockchain algorithm to consider the service level agreement (SLA), the real-time processing performance of transactions is analyzed according to the redundancy of nodes. In the future work, we plan to demonstrate HBFT algorithm through simulations. Furthermore, we would also implement Blockchain-based application services to effectively manage the decentralized IoT data.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2018R1D1A1B07040573).

References

[1] O. Novo, "Blockchain Meets IoT: An architecture for Scalable Access Management in IoT," IEEE Internet of Things Journal, **5**, 1184-1195, 2018, doi: 10.1109/JIOT.2018.2812239.

[2] R. Casado-Vara, P. Chamoso, F. De la Prieta, J. Prietoa, J. M. Corchado, "Non-linear adaptive closed-loop control system for improved efficiency in IoT-blockchain management," Information Fusion, **49**, 227-239, 2019, doi: /10.1016/j.inffus.2018.12.007.

[3] P. Brody, V. Pureswaran, "Device democracy - Saving the future of the internet of Things," IBM institute for business value (2015) Available online: https://www.ibm.com/downloads/cas/Y5ONA8EV (accessed on 15 July 2020)

[4] Y. Li, W. Susilo, G. Yang, Y. Yu, D. Liu, M. Guizani, "A Blockchain-based Self-tallying Voting Scheme in Decentralized IoT," eprint arXiv:1902.03710, 2019.

[5] K. Košťál, P. Helebrandt, M. Belluš, M. Ries, I. Kotuliak, "Management and Monitoring of IoT Devices Using Blockchain," Sensors, **19**, 1-12, 2019, doi: /10.3390/s19040856.

[6] G. Sagirlar, B. Carminati, E. Ferrari, E. Ferrari, J. D. Sheehan, E. Ragnoli, "Hybrid-IoT: Hybrid Blockchain Architecture for Internet of Things - PoW Sub-blockchains," 2018, doi: 10.1109/Cybermatics_2018.2018.00189.

[7] M. Ruta, F. Scioscia, S. Ieva, G. Capurso, E. D. Sciascio, "Semantic Blockchain to Improve Scalability in the Internet of Things," Open Journal of Internet of Things (OJIOT), **3**, 46-61, 2017.

[8] R. Jayaraman, K. Salah, N. King, "Improving Opportunities in Healthcare Supply Chain Processes via the Internet of Things and Blockchain Technology," International Journal of Healthcare Information Systems and Informatics (IJHISI), **14**. 1-20, 2019, doi: 10.4018/IJHISI.2019040104.

[9] S. Gupta, M. Sadoghi, "Blockchain Transaction Processing," in Encyclopedia of Big Data Technologies, 1-11, 2019, doi:10.1007/978-3-319-63962-8_333-1.

[10] A. Panarello, N. Tapas, G. Merlino, F. Longo, A. Puliafito, "Blockchain and IoT Integration: A Systematic Survey," Sensors, **18**(8), 1-37, 2018, doi: 10.3390/s18082575

[11] A. Baliga, "Understanding blockchain consensus models," Persistent Systems Ltd. White paper, 2017, URL: https://www.persistent.com/wp-content/uploads/2018/02/wp-understanding-blockchain-consensus-models.pdf

[12] L. M. Bach, B. Mihaljevic, M. Zagar, "Comparative analysis of blockchain consensus algorithms," in Proceedings of 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 1545-1550, 2018.

[13] E. Zhang, "A Byzantine Fault Tolerance Algorithm for Blockchain," NEO White paper, 2018, URL: https://docs.neo.org/en-us/basic/consensus/whitepaper.html

[14] Y. Xiao, N. Zhang, W. Lou, Y. T. Hou, "A Survey of Distributed Consensus Protocols for Blockchain Networks," in IEEE Communications Surveys & Tutorials, **22**(2), 1432-1465, 2020, doi: 10.1109/COMST.2020.2969706.

[15] M. C.Igor, N. C.Vitor, P. A.Rodolfo, Y. Q.Wang, D.R.Brett, "Challenges of PBFT-Inspired Consensus for Blockchain and Enhancements over Neo dBFT," in future internet, **12**(8), 129, 2020

[16] M. Castro, B. Liskov, "Practical Byzantine Fault Tolerance," in Proceedings of the Third Symposium on Operating Systems Design and Implementation, 173-186, 1999.

[17] M. Castro, B. Liskov, "Practical Byzantine Fault Tolerance and Proactive Recovery," ACM Transactions on Computer Systems, **20**(4), 398-461, 2002, doi: 10.1145/571637.571640.

[18] M. Kim, J-H. Kim, "Hierarchical Voting-based Byzantine Fault Tolerance Consensus Algorithm," in Proceedings of International Conference on APIC-IST 2020, 183-185, 2020

[19] M. Kim, J-H. Kim, "Decentralized Data Management Schemes for IoT Blockchain," in Ph.D. Dissertation, Department of Computer Engineering Graduate School of Ajou University, 2019

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 228-238 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Vehicle Rollover Detection in Tripped and Untripped Rollovers using Recurrent Neural Networks

Kailerk Treetipsounthorn[1], Thanisorn Sriudomporn[2], Gridsada Phanomchoeng[*,2,3], Christian Dengler[4], Setha Panngum[1], Sunhapos Chantranuwathana[2,5], Ali Zemouche[6,7]

[1]*Department of Computer Engineering, Chulalongkorn University, Bangkok, 10330, Thailand*

[2]*Department of Mechanical Engineering, Chulalongkorn University, Bangkok, 10330, Thailand*

[3]*Applied Medical Virology Research Unit, Chulalongkorn University Bangkok 10300, Thailand*

[4]*Chair of Automatic Control, Technical University of Munich, Munich, Germany*

[5]*Smart Mobility Research Unit, Chulalongkorn University, Bangkok 10300, Thailand*

[6]*University of Lorraine, CRAN UMR CNRS 7039, Cosnes et Romain, 54400, France*

[7]*EPI Inria DISCO, Laboratoire des Signaux et Systèmes, CNRS-CentraleSupélec, Gif-sur-Yvette, 91192, France*

ARTICLE INFO

Article history:
Received: 01 September, 2020
Accepted: 01 November, 2020
Online: 10 November, 2020

Keywords:
Rollover Index
Tripped Rollover
Untripped Rollover
Recurrent Neural Network
Feedforward Neural Network
Tanh
LSTM
GRU

ABSTRACT

Comparing to other types of vehicle accidents, fatality rate of tipped rollover accidents shows significant number. Thus, tripped rollover prevention systems are important in order to keep driver safe. In other hands, different rollover indices are defined to handle the risk. The variable unknown parameters of each index, for instance, current load of the vehicle or center of gravity, are considered as a difficulty. In this work, the recurrent neural networks, which are designed to work on sequential data in order to provide data estimation without additional estimation algorithm, are investigated in purpose to estimate the tripped and untripped rollover index. The vehicle simulation software with industrial standard CarSim is applied to validate the result. The Tanh recurrent neural network is stated in the result to be the most accurate tripped rollover index estimator for the uncertain parameters, for example, sprung mass and the height of the center of gravity. The suitable input features for tripped and untripped rollover index and neural network structure are verified. To prevent and provide warning of rollover, an advance future prediction can also be designed for the future tripped and untripped rollover prediction.

1. Introduction

This work is an extension based originally on the groundwork presented in the 9th IEEE International Conference on Robotics, Automation and Mechatronics (CIS-RAM 2019) [1]. This paper will focus on development of recurrent neural networks for tripped rollover while Ref. [1] focus on only untripped rollover. The number of rollover accidents has grown continuously according to the change in vehicles' size, which are both weight and dimension, over the last 10 years. The majority type of fatality on the road is pointed to rollover accidents. In 2010, 9.1 million cars, which includes passenger cars, SUVs, pickups and vans, crashed and 2.1% of those crashes were caused by rollovers collected by National Highway Traffic Safety Administration's records (http://www.safercar.gov). From those 2.1% of crashes, there were over 7,600 deaths which was up to 35% of total deaths from vehicle accidents [2]. In addition, 20% of vehicle-crash deaths in Europe are from the rollover [3]. In Germany, 2-5% and 90% of vehicle accidents and deaths from the accidents are caused by vehicle rollover accidents in 2005 respectively [4].

The vehicle rollover accidents are categorized as tripped and untripped rollovers. A rollover accident that have contacted with an external obstacle, such as a pothole or guardrail, is called a tripped rollover. Untripped rollovers are defined the other way

*Corresponding Author: Gridsada Phanomchoeng, Department of Mechanical Engineering, Chulalongkorn University, Bangkok, 10330 Thailand, gridsada.phanomchoeng@gmail.com

round. It happens because of the effect of high lateral acceleration on a dangerous turn, i.e. sharp turn, and it is unrelated to external input [5-6]. Tripped rollovers tend to happen more frequently than untripped rollovers with 95% and 5% of total rollovers according to the statistic provided by NHTSA [2]. From the significant number of deaths from rollovers compared to the total number of high-way traffic deaths, vehicle safety must be improved to prevent the vehicle from rollover accidents.

Hence, several automotive manufacturers, e.g., Ford and Volvo have advanced with active rollover prevention systems which aim to prevent untripped rollover accidents [7]. Analysts and researchers likewise take interest largely in differential braking systems e.g. Anti-lock Braking Systems (ABS), Electronic Braking Systems (EBS) and Electronic Stability Programs (ESP) [3, 7-13]. The baking systems perform vehicle's yaw rate and speed reduction by adjusting each wheel's sequence of brake patterns. Also steer-by-wire technology [13-17] e.g. active front steering (AFS), active rear steering (ARS), or four-wheel steering (4WS) are to be used in case of the balancing of performances among yaw stability, violence, and rollover prevention and an active suspension system [18-20] can be used to avoid rollover when the wheels lifted off the road. Then, an active roll bar stabilization system [20-22] uses an active stabilizer for active torque distribution between front and rear axles of the vehicle. This can reduce the roll motion. In addition, active torque management [23] has the capability to prevent rollover accidents. Many control techniques are developed to enhance these systems.

Precise detection of vehicle rollover is mandatory in order for these active rollover prevention systems to fully function. To the present, there are multiple impending vehicle rollover detections developed with computational rollover index. The rollover indices are developed by various approaches. For example, Static Stability Factor (SSF) was developed to measure the rollover resistance of a given vehicle design configuration [24]; ref. [25] introduced rollover indices that consider kinetic energy, then Rollover Prevention Energy Reserve (RPER) is started off to look into rollover risks by taking energy function as the main account [7]; ref. [26] introduced a dimensionless rollover index that takes roll dynamics into consideration by utilizing phase plane analysis; zero-moment point also is developed to predict the possibility of rollover [9]; lateral Load Transfer Ratio (LTR) is popularly discussed in numerous studies [5, 7, 14-18]; to name a few. It is seen that a lot of rollover indicators are developed based on LTR such as Multiple Rollover Indices (MRIs) for trailer vehicles and Predictive Lateral Transfer Ration (PLTR) [9]. Lateral Load Transfer Ration is actually one of the approaches that take several factors into account, for example, roll angle, roll center and height of center of gravity of a vehicle. Many techniques have been established to estimate these factors. For example, extended Kalman filter technique or nonlinear observer/estimator has been built up to improve the accuracy of the lateral Load Transfer Ratio. Other than that, ref. [27] introduced an algorithm that utilizes the time-to-rollover (TTR) metrics to assess the time until rollover as a warning option. With this approach, an advance real-time prediction of rollover risk can be provided. Moreover, this technique can be used with several other rollover indices simultaneously. Despite a number of rollover detection techniques developed until now, only untripped rollovers are detected.

Ref. [5, 6, 28] has introduced a new rollover index which incorporates external road inputs as a factor to detect not only the untripped but also tripped rollover conditions. The external obstacle road inputs are estimated based on a vehicle dynamic model and nonlinear observer. The simulation and experimental results are presented using a 1/8th scaled vehicle.

Since the indicator to detect rollover risk is very important for improving the rollover prevention systems and rollover warning system [29-31], it is necessary to develop the rollover indicator for detecting both tripped and untripped rollovers. It is shown that the traditional rollover index is not effective enough to identify tripped rollovers and there are only few focused on tripped rollovers that are published. In addition, most accurate rollover indicators relay on dynamic models which require known parameters such as vehicle mass or height of center of gravity of a vehicle. Therefore, here proposed is a neural network indicator developed to identify tripped and untripped rollovers. Neural networks known for adaptable and nonlinear information processing capability perform fittingly in the areas of predication, expert system, and mode identification [32-37]. By taking the approach to be introduced afterward into account, the estimation algorithms to estimate unknown parameters such as roll angle, height of center of gravity of a vehicle, or vehicle mass are not required. These parameters may be changed all the time due to the number of passengers and load. Also, the technique can be applied in many situations. This research is the extended technique in [1, 38] which is focused on untripped rollover to detect tripped rollover.

This research has the following outline: Section 2 will introduce the vehicle rollover and the rollover index for detecting tripped and untripped rollovers. Next, the recurrent neural network technique for detecting tripped and untripped rollover is presented in Section 3, and the methodology to apply the recurrent neural network to the rollover problem is presented in Section 4. The process of evaluation of the suitable neural network, the suitable input features, the suitable neural network structure and the future tripped rollover prediction are included in this section. Then, Section 5 will present the tripped rollover results. Finally, Section 6 will discuss the conclusion to the findings.

2. Vehicle Rollover Index

Traditional rollover index posts as most popular technique to detect vehicle rollover among many techniques [5-7, 14-18, 28-33] used nowadays.

2.1. Traditional Rollover Index

In order to detect lift-off condition of the wheels [6, 7], the rollover index is used as a real-time indicator. Lateral Load Transfer Ratio (LTR) is known as another name of the index. The standard rollover index is designed for predicting incidents that are caused with the lift-off conditions of the wheels for untripped accidents. A number of researches use the basic model to derive their rollover indices. A scaled lateral acceleration, a function of roll angle [12], or a function of lateral acceleration and roll angle [9, 18] is considered as a base argument for these rollover indices.

The rollover index by the conventional definition is portrayed by

$$R = \frac{F_{zr} - F_{zl}}{F_{zr} + F_{zl}}, \quad -1 \le R \le 1 \tag{1}$$

where F_{zr} is right normal tire force and F_{zl} is left normal tire force. $R = 1$ ($F_{zl} = 0$) or $R = -1$ ($F_{zr} = 0$) is the value when a rollover incident is defined. F_{zr} equals to F_{zl} and $R = 0$ is when the vehicle travels straight forward. If $F_{zr} = 0$, then $R = 1$ and this will cause the lift-off occurrence of the left wheel. If $F_{zl} = 0$, then $R = -1$ and this will cause the lift-off occurrence of the right wheel.

Based on the one-degree-of-freedom model, the summation and difference of tire forces $F_{zr} + F_{zl}$ and $F_{zr} - F_{zl}$ can be calculated. An implementable rollover index processed with the one-degree-of-freedom model is demonstrated by [1, 6]

$$R = \frac{F_{zr} - F_{zl}}{F_{zr} + F_{zl}} = \frac{2m_s a_y h_R}{mgl_w} + \frac{2m_s h_R tan\phi}{ml_w} \tag{2}$$

where m_s is Sprung mass, $m = m_s + m_u$, m_u is unsprung mass, a_y is lateral acceleration, h_R is Center of gravity (C.G.) height from roll center, g is acceleration due to gravity, l_w is track width, and ϕ is roll angle.

Lacking influence of the vertical road input and other external inputs limit the rollover index in (2) with only the untripped rollover detection. Lateral acceleration and roll angle influence this rollover index. However, the roll angle, which is complicated to measure, is removed in some published paper's rollover index calculation.

2.2. Rollover Index for Tripped and Untripped Rollover

The influence of road inputs and other external forces are considered in rollover index both for tripped and untripped rollovers in order to indicate the tripped rollovers caused by external input tripping. Figure 1. shows a vehicle rollover model where the influence of right road inputs, z_{rr}, left road input, z_{rl}, and an unknown lateral force input, F_{lat}, at an arbitrary height, h_{lat}, from the roll center are presented. It also demonstrates the normal tires forces, F_{zr} and F_{zl}.

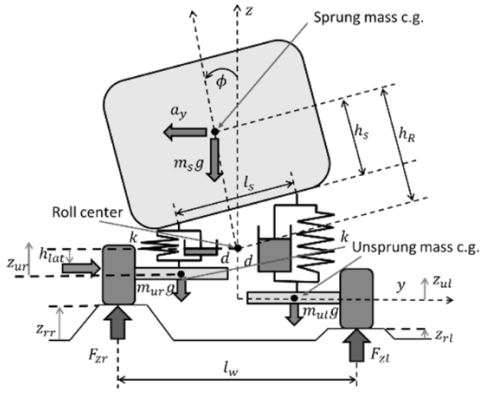


Figure 1: Tripped and Untripped Rollover Model

The four-degree-of-freedom vehicle model provides the derivation of the rollover index in the case of tripped rollover and the derivation is presented in [6]. Equation (3) shows the rollover index for tripped and untripped rollover which is shown in [6].

$$R = \frac{m_u(\ddot{z}_{ur} - \ddot{z}_{ul}) - \frac{2}{l_s^2}(I_{xx} + m_s h_R^2)(a_{zl} - a_{zr})}{m_u(\ddot{z}_{ur} + \ddot{z}_{ul}) + m_s\ddot{z}_s + mg} + \frac{\frac{2}{l_s}m_s a_y h_R cos\phi + \frac{2}{l_s}m_s g h_R sin\phi}{m_u(\ddot{z}_{ur} + \ddot{z}_{ul}) + m_s\ddot{z}_s + mg} \tag{3}$$

where $\ddot{z}_{ur}$ is right unsprung mass acceleration, $\ddot{z}_{ul}$ is left unsprung mass acceleration, $\ddot{z}_s$ is sprung mass acceleration, l_s is length between the left and right suspensions, I_{xx} is roll moment of inertia, a_{zl} is right vertical accelerometer measurement, a_{zr} is left vertical accelerometer measurement, and $m = m_s + 2m_u$.

The second term of (3), the function of lateral acceleration and roll angle, is similar to (2) when comparing the traditional rollover index in (2) and rollover index for tripped and untripped rollovers in (3). To be more clarify, the second term of (3) is the traditional rollover index for detecting untripped rollover, and the first term of (3) is the additional term used for detecting tripped rollover.

There are some difficulties in implementing the (3) to determine tripped rollover due to the parameters and variables that are unable to identify, e.g. the location of roll center, center of gravity (C.G.) height or roll angle. The estimations of these unknowns are required to complete the estimation algorithms. Furthermore, the difference of dynamics between an actual vehicle and the four-degree-of-freedom model might lead to errors. The recurrent neural network is implemented to detect tripped and untripped rollovers in order to optimize the estimation. After training the data, the traditional estimation algorithm to estimate parameters and detect rollover are not required in the model anymore.

3. Recurrent Neural Networks (RNNs)

Recurrent neural network is a popular special type of neural networks. In this paper, sequential data will be sensor data over a period while in other cases it can be, for example frames in video, or words in a text. When a new input is introduced to the sequence, a hidden state vector, h, which is an added memory of the recurrent neural network, is updated. To avoid features losses while training neural networks, the crucial data of the previous sequence is maintained in this hidden state. The recurrent neural network has the ability to work with various lengths of sequence data in addition to the benefit in compression of past information.

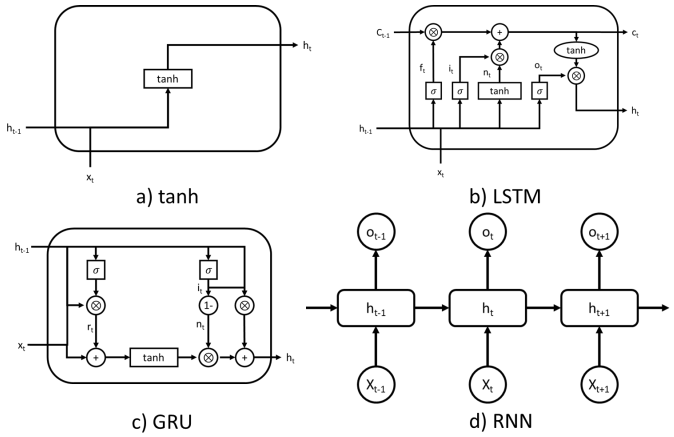


Figure 2: The Structural of Recurrent neural network a) tanh architecture b) LSTM architecture c) GRU architecture d) RNN Connection

There are four types of neural networks used to evaluate and create the neural network models for detecting tripped and untripped rollover. The four types of neural networks are 1) feedforward neural network (FNN), 2) Tanh neural network, 3) long short-term memory (LSTM) neural network [39], and 4) gates recurrent unit (GRU) neural network [40]. The detail of implementing these neural networks can be found in [39-46] and the short introduction of them is shown in [1]. The structural of recurrent neural network is shown on Figure 2.

4. Methodology

The data which is not normally determined in a vehicle such as the wheel's forces exerted on the ground by the vehicle wheels is mandatory to compute the precise rollover index in (1). Moreover, Ref. [6, 28] determine the tripped and untripped rollover index based on half vehicle model (4 Degree of Freedom). The variables, which are used, are right unsprung mass vertical acceleration, $\ddot{z}_{ur}$, left unsprung mass vertical acceleration, $\ddot{z}_{ul}$, right vertical accelerometer measurement, a_{zr}, left vertical accelerometer measurement, a_{zl}, lateral acceleration, a_y, sprung mass acceleration, $\ddot{z}_s$, roll angle, ϕ, roll rate, $\dot{\phi}$, and right suspension compression, $(z_s - z_{ur} + \frac{l_s}{2} sin\phi)$. They also assume that the parameters such as sprung mass, m_s, unsprung mass, m_u, C.G. height from roll center, h_R are constant and not changed.

To detect tripped rollover with various conditions, the good measurement inputs are needed to verify for recurrent neural networks. Also, a feedforward neural network is compared with recurrent neural networks in this case to validate the advantages of employing recurrent neural networks rather than a static neural network. The feedforward neural network uses the same inputs as the recurrent neural networks for processing the data and a comparable number of parameters. Moreover, the structure of neural networks is verified for determining the suitable one.

There is a reconstruction of the relationship between the inputs and the rollover index from a dataset processed in a simulation constraint. The next part is presented with the details on the dataset used for training, also the structure of the neural networks and the training setup will be described

4.1. Dataset generation for tripped and untripped rollover

The software CarSim, a standard industry vehicle simulation software, is used to process all the data concerned. In a first step, 22,000 simulations with non-selectively sampled height of the center of gravity and sprung mass of the car are simulated with a variety of tripped rollover scenarios. The height of the center of gravity is uniformly studied between 0.6 and 1.2 meters, based on [47] as a background which also take an estimate of additional loads into account. The sprung mass is uniformly generated between 2100 to 3000 kg, a mass comparable to loaded SUVs. Next, a trajectory of the scenarios is generated based on Fig 2. The trajectory consists of the 50 meters of straight path and the circular arc path of the random radius in the range from 50 to 200 meters based on highway data of Bangkok, Thailand. Samples 80 to 180 km/h uniform distribution will determine the target velocity of the vehicle. Then, the CarSim is going to control the real velocity over the trajectory in a feedback loop to remain adjacent to the speed that is set goal. Also, the steering angle of the vehicle is close-loop controlled to follow the trajectory.

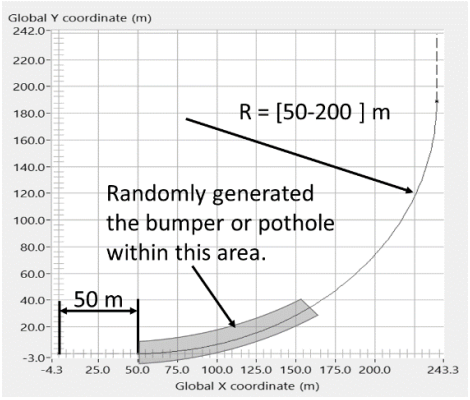


Figure 3: CarSim Trajectory Setup

To create tripped rollover scenarios, the bumper and pothole are randomly generated along the circular arc path of the trajectory as shown in Figure 3. The shape of the bumper and pothole is randomly created from the sine function with the amplitude between -0.15 to 0.15 m and the frequency between 0.625 to 1.25 Hz. Only one bumper or pothole is created for each trajectory. Also, it is created only half lane of the trajectory to make only half side of the vehicle strike the bumper or pothole. For the left turn trajectory, if the bumper is generated, the bumper is created on the left side of the vehicle and if the pothole is generated, the pothole is created on the right side of the vehicle. Then, for the right turn trajectory, if the bumper is generated, the bumper is created on the right side of the vehicle and if the pothole is generated, the pothole is created on the left side of the vehicle. These increase the chance for tripped rollover.

The normal forces on the left and right wheel, F_{zl}, F_{zr}, are provided by a simulation, so each estimation is able to process to determine its rollover index. The rear wheel vertical forces are used to identify the rollover index. However, the front wheels, employing the said methodology together with a small variation from elastic deformations in the wheels and wheel suspensions, can also specify the rollover index as well. The rollover index computed with (1) is the neural networks' training target, calculated according to (1). From there, when two wheels are lifted off road, the sequences stop. That is because the rollover index is defined only for the case of at least one wheel lifting the ground. And it will contribute to a sequence range of length between 0.325 to 19.925 seconds. In every simulation case, there will be a collection of data of the vehicle variables at the sample rate of 40 Hz. However, each simulation dataset also possesses insignificant and irrelevant information of the subject vehicle or those that are unable to be scaled in an actual subject. So, the unnecessary data are needed to be discarded. Then, for the 22,000 files of h5 files (1.64 GB), they are splitted into 3 groups. There are a training set of 16,000 files, a validation set of 4000 files, and

a test set of 2,000 files. The training set is used for training. Other than that, it is mandatory for the validation set to detect possible overfitting of the model while the test set is used to evaluate the error of the neural networks.

4.2. Training of the recurrent neural networks for tripped rollover

There are three major factors of neural networks for detecting tripped and untripped rollover needed to be verified. The first factor is which neural network should be suitable, the second factor is which inputs should be good for neural networks, and the last one is which structure of the neural networks should be implemented. To verify these three factors, three training sets are simulated. Then, future tripped and untripped rollover prediction also is verified with simulation.

4.2.1 Training for determining a suitable neural network type

For this task, four types of neural networks will be compared and given information. Table 1 is presented with the detail of the neural networks which are constructed to take into account approximately the same quantity of parameters. About 5,000 parameters are to be taken to comparison between each neural network. In Table 1, "$N°$ neurons" shows the number of neurons in each layer in the horizontal axis, and also takes into account the input layer (n inputs) and output layer (1 output). The recurrent neural networks, the second entry in "$N°$ neurons" represents the number of hidden states. "$N°$ params" specifies the number of parameters in the neural network. In this column, the added number will be those of all elements in its weight matrices and bias vectors.

Firstly, a feedforward neural network (FNN) will be introduced. FNN is made up of three hidden layers using the $tanh$ activation function and a linear output layer as shown within [1]. Secondly introduced is a neural network (Tanh) which is a recurrent neural network, that employs a $tanh$ recurrent layer. And followed the recurrent layer aforementioned are a $tanh$ hidden layer and an output layer. Thirdly, LSTM neural network which utilizes a long short-term memory layer, again followed by a $tanh$ layer and linear output layer. And for the fourth one, GRU neural network that adopts a static Tanh layer and a linear output layer, works with the gated recurrent unit. Then, the neural networks training diagram is shown on Figure 4.

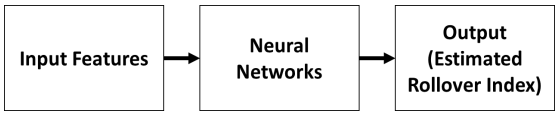


Figure 4: The Neural Networks Training Diagram

Table 1: Summary of The Neural Networks' Structure

NN Type	$N°$ neurons	$N°$ param
FNN	n 48 48 42 1	48n+4501
Tanh	n 46 46 1	46n+4417
LSTM	n 30 16 1	120n+4353
GRU	n 32 32 1	96n+4353

Table 2: Input Variables for Neural Networks

Symbolic	Descriptions
$v_{z_{wc}}$: $[v_{z_{wc_{L1}}}; v_{z_{wc_{L2}}}; v_{z_{wc_{R1}}}: v_{z_{wc_{R2}}}]$	Vertical speed of each wheel: Left front, $v_{z_{wc_{L1}}}$, Left rear, $v_{z_{wc_{L2}}}$, Right front, $v_{z_{wc_{R1}}}$, Right rear, $v_{z_{wc_{R2}}}$, (km/h)
A_{xyz}: $[A_x; A_y; A_z]$	Longitude acceleration, A_x, Lateral acceleration, A_y, Vertical acceleration, A_z, (g)
AA_x	Roll Acceleration of sprung mass body (rad/sec^2)
AV_{xyz}: $[AV_x; AV_y; AV_z]$	Roll rate, AV_x, Pitch rate, AV_y, Yaw rate, AV_z, (deg/sec)
$A_{z_{sm}}$	Vertical acceleration of sprung mass C.G. (g)
v_x	Longitude speed (km/h)
$Steer_{DM}$	Steer angle from driver (deg)
$CmpS$: $[CmpS_{L1}; CmpS_{L2}; CmpS_{R1}; CmpS_{R2}]$	Compression of spring each wheel: Left front, $CmpS_{L1}$, Left rear,$CmpS_{L2}$, Right front, $CmpS_{R1}$, Right rear,$CmpS_{R2}$, (mm)

Table 3: Input Combination Training for Determining a Suitable Neural Network Type

No.	Input Features	n	FNN	Tanh	LSTM	GRU
1	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}]$	8	4885	4785	5313	5121
2	$[v_{z_{wc}}; A_{z_{sm}}; AA_x]$	6	4789	4693	5073	4929
3	$[v_{z_{wc}}; A_{z_{sm}}; AV_{xyz}]$	8	4885	4785	5313	5121
4	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AA_x]$	9	4933	4831	5433	5217
5	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}]$	11	5029	4923	5673	5409
6	$[v_{z_{wc}}; A_{z_{sm}}; AA_x; AV_{xyz}]$	9	4933	4831	5433	5217
7	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}; AA_x]$	12	5077	4969	5793	5505

For the inputs to the neural networks, it is still unclear to select them to predict the tripped rollover. Thus, to initially investigate this task, the inputs are selected based on (3) [6] and the full vehicle suspension model. In addition, the selected inputs can be measured by sensors [6, 28]. The list of the inputs is shown on Table 2. The input combinations are also evaluated in this case. The input combinations for training in this section is shown on Table 3.

Next, the neural networks are trained over 3000 epochs using backpropagation through time. When the process is completed, since the result beyond that came out slower on the validation and training loss. The mean squared error is used as the loss function:

$$L = \frac{1}{N_{traj}} \sum_{i=1}^{N_{traj}} \frac{1}{T_i} \sum_{t=1}^{T_i} \left(R_{ro,i,t} - y_{i,t}\right)^2 \quad (4)$$

with $R_{ro,i,t}$ the rollover index as calculated using the cases data in trajectory i at time step t and $y_{i,t}$ the scalar output of one of the neural networks for the same trajectory and time step.

The leading software that is used for computing the neural networks is the Knet module [49] by the Julia programming language. The program runs on the notebook, Acer Nitro, with CPU i7-7700HQ (2.8 GHz), 8 GB Ram, GeForce GTX1050, and

Window 10 (64 bits). In the meantime, Training is done on a consumer level GPU is used to perform the network training for about 1 hour per model by using average quantity of parameters. This altogether resulted in more than 240 hours of training time. The Ram usage during the training is approximately 6 to 7 GB. And there also implemented the Adam optimizer along with the suggested settings in [50]. Then, the gradient was computed on mini-batches of 2,048 sequences.

4.2.2 Training for determining suitable inputs for the neural network

Since the performance of the neural network is required and the sensor cost is needed to be minimized, the inputs to the neural network need to be optimized. The extra inputs, suspension compressions as using in [28] for each wheel, are included to the list for evaluation. Then, the input combinations are evaluated to determine the suitable inputs for predicting rollover. The input reduction is also included in this test for reducing number of using sensors. The lists of the inputs are shown in Section 5.2 Table 5.

Next, only the best neural network type from the section 4.2.1 are trained. The training process and the loss function are implemented similar to the section 4.2.1.

4.2.3 Training for determining suitable neural network structure

In this task, the neural network improvement is investigated. Eight different structures of the best neural network from the section 4.2.1 are compared. The neural network structure is modified by increasing number of the layers and the neurons. Also, the optimized inputs from the section 4.2.2 are used in this task. Then the neural network structures are also summarized in Table 6. Next, the training process and the loss function are implemented similar to the section 4.2.1.

4.2.4 Training for future tripped and untripped rollover prediction

The future tripped and untripped rollover perdition uses the vehicle parameter measurement to prediction rollover index of a vehicle in the future time based on the training model. It can provide an extra time to prevent or warn the danger of the rollover.

In this case, the output of each data set is shifted by time while the input is reduced to identify with to the length of the output. The shifted time varies from 0.1 second up to 1 second. Then, the best model from Section 4.2.3 are used and, the training process and the loss function are implemented similar to the section 4.2.1

5. Results

The training as descript in Section 4 are applied to the data set. The final losses on the train, validation and test datasets are presented in this section.

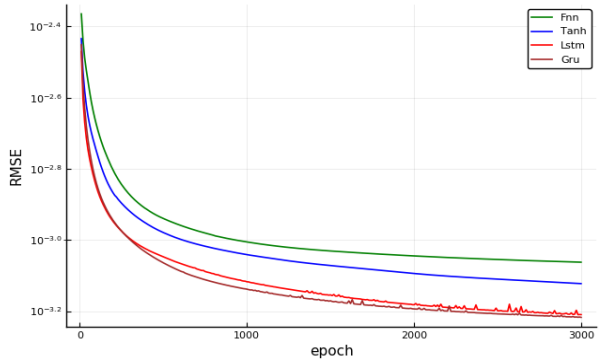
5.1. The suitable neural network type

The test losses in different neural networks on the different input features are shown on Table 4. Also, the example of loss on the train dataset and validation dataset over the epochs are presented in Figures 5a and 5b, subsequently. The training and validation losses of all different neural networks and all input features shows that the overfitting does not occur for all cases because the validation loss decreases as the training loss decreases and the losses from validation dataset are small as the losses from training dataset. Also, the underfitting did not occur for all cases because the neural network models can detect the rollover index as shown in Figures 6 and 7.

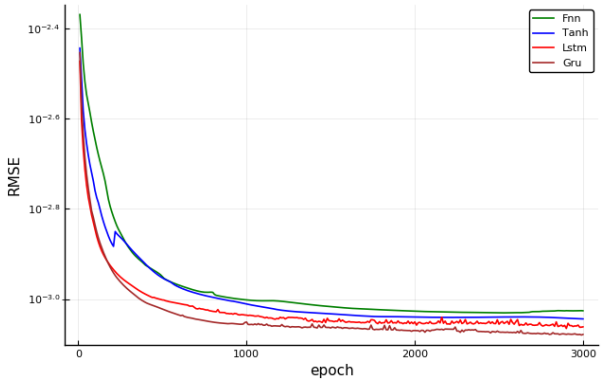
Table 4: Remaining Loss of Different Neural Networks on Test Data Set

No.	Input Features	n	Average RMS error of 2000 Data Set $\times 10^{-4}$			
			FNN	Tanh	LSTM	GRU
1	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}]$	8	15.9	10.5	11.3	27.2
2	$[v_{z_{wc}}; A_{z_{sm}}; AA_x]$	6	228	33.3	96.3	156
3	$[v_{z_{wc}}; A_{z_{sm}}; AV_{xyz}]$	8	17.7	13.2	14.4	28.9
4	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AA_x]$	9	12.5	10.3	11.1	17.4
5	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}]$	11	9.82	7.91	9.98	22.8
6	$[v_{z_{wc}}; A_{z_{sm}}; AA_x; AV_{xyz}]$	9	17.1	22.7	13.3	43.0
7	$[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}; AA_x]$	12	9.09	7.54	9.47	21.5

Note: the highlighted cell of the table represents the smallest value of the table.



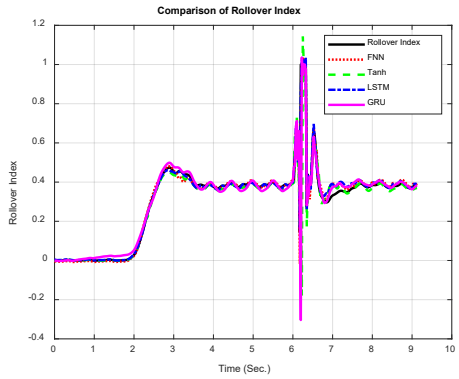
a) Training Loss of No. 7 Table 4



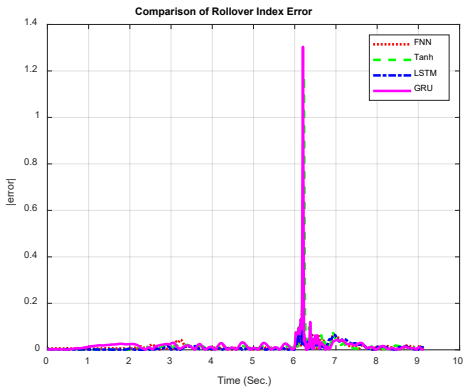
b) Validation of No. 7 Table 4

Figure 5: The Example of Training and Validation Loss

By comparing four types of the neural networks on Table 4, GRU recurrent neural network and FNN seems have the large test losses for all different input features. Since the tripped rollover happen in the very short time So, there is few tripped data occurred. This makes GRU which has a too long-term memory cannot learn the model well. Also, FNN contributes to a large outstanding error because the formula only processes the valuation with the current input x_t.

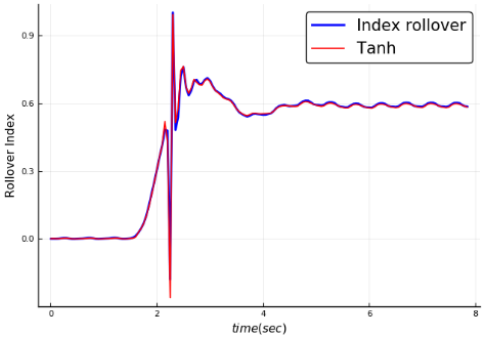


a) Comparison of Rollover Index and Rollover Index from Neural Networks

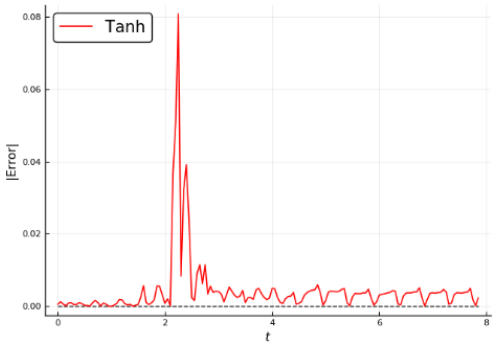


b) Comparison of Rollover Index Error

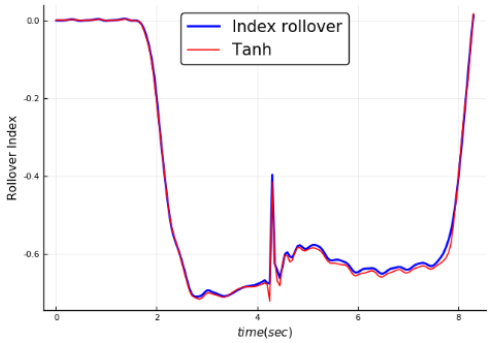
Figure 6: The Example of Rollover Index in the Case of ($m_s = 2722\ kg$, $h_R = 0.847\ m$, $v_x = 82\ km/h$)



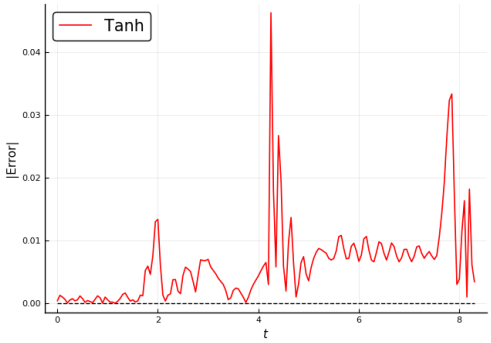
a) Comparison of Rollover Index in the Bumper Case



b) Comparison of Rollover Index Error in the Bumper Case



c) Comparison of Rollover Index in the Pothole Case



d) Comparison of Rollover Index Error in the Pothole Case

Figure 7: The Example of Rollover Index in the Bumper Case ($m_s = 2331\ kg$, $h_R = 0.777\ m$, $v_x = 118\ km/h$) and Pothole Case ($m_s = 2912\ kg$, $h_R = 0.730\ m$, $v_x = 141\ km/h$)

However, Tanh recurrent neural network has the lowest final loss on test loss excepted for No. 6 Table 4 with the input features, $[v_{z_{wc}}; A_{z_{sm}}; AA_x; AV_{xyz}]$ because it has not too long term memory. While the test loss from LSTM recurrent neural network is larger than it from Tanh recurrent neural network except for No. 6 Table 4.

For an empirical comparison, the rollover index (black solid line in Figure 6a) according to (1) and the neural network estimation are shown in Figure 6a for the selected section of the test trajectory. The vehicle takes a cornering at about 2 seconds and during cornering it strikes a bumper at about 6 seconds. This caused a rear wheel to leave the road while its counterpart stayed intact on the ground. The estimated rollover index for the LSTM, Tanh, and FNN are similar and barely distinguishable, whereas the estimation of GRU shows a clear offset and a large error at around 6 seconds. The absolute value of the estimation error is presented subsequently in Figure 6b. As shown, the highest values in the estimation error for the rollover index is seen following the subject vehicle returns from rollover condition back to its normal condition that all of the wheels are on the ground at around 6.1 seconds in the trajectory. However, situational peaks are more prominent, while those of GRU are even more grand.

Figure 7 shows the rollover index and rollover index error of the Tanh recurrent neural network of different vehicle parameters on the bumper and pothole trajectories. The Tanh recurrent neural network can predict the rollover index even though the vehicle parameters are changed because the neural network can handle the vehicle parameter changing. It can learn and adapt the parameters inside the neural network and then predict the rollover index.

Figure 5a shows the case that the vehicle with the parameters, $m_s = 2722\ kg$, $h_R = 0.847\ m$, $v_x = 82\ km/h$, takes a cornering at about 1.8 seconds and during cornering it strikes a bumper at about 2.4 seconds. The maximum error of the estimation happens at about 2.4 seconds. Next, Figure 5b shows the case that the vehicle with the parameters, $m_s = 2912\ kg$, $h_R = 0.730\ m$, $v_x = 141\ km/h$, takes a cornering at about 1.8 seconds and during cornering it strikes a pothole at about 4.4 seconds. Then, the maximum error of the estimation happens at about 4.4 seconds.

In summary, the suitable neural network for tripped rollover is Tanh recurrent neural network since it has ability to retain information of the past and handle the vehicle parameter changing. Also, the lowest test loss happens when using Tanh recurrent neural network with No. 7 Table 4. with the input features, $[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}; AA_x]$.

5.2. The suitable input features for the neural network

Since the error of the Tanh neural network is needed to be improved, the suspension compressions for each wheel, longitude speed, and steering angle are added and evaluated. The test losses of these cases are shown in Table 5. The results show that the suspension compressions for each wheel, longitude speed, and steering angle are significant input features for the neural network. They can reduce the test losses up to 2.93×10^{-4}. Also, once the the suspension compressions for each wheel, longitude speed, and steering angle have been included in the input features, the input, AA_x, seems not effect to the test losses.

In case that the number of suspension compressions is reduced to be only front or rear side, the test losses is increased as shown in No. 3 and 4 Table 5. No. 3 Table 5. has higher the test losses because the actual rollover index is determined from the rear wheel vertical forces.

In summary, the suitable input features for tripped rollover for Tanh neural network is $[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}; v_x; steer_{DM}; CmpS]$ which are agree with [28]. The example of rollover index in this case is shown in Figure 8.

Table 5: Remaining Loss of the Tanh Neural Network on Different Input Feature on Test Data Set

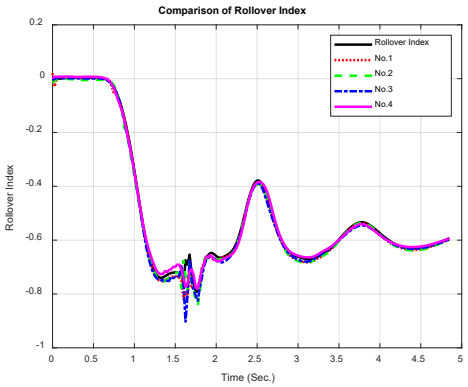
No.	Input Features	n	$N°$ param	Average RMS error of 2000 Data Set $\times 10^{-4}$
1	$[v_{z_{wc}}; [A]; CmpS]$	17	5,199	4.61
2	$[v_{z_{wc}}; [A]; AA_x; CmpS]$	18	5,245	4.82
3	$[v_{z_{wc_{L1}}}; v_{z_{wc_{R1}}}; [A]; CmpS_{L1}; CmpS_{R1}]$	13	5,061	7.41
4	$[v_{z_{wc_{L2}}}; v_{z_{wc_{R2}}}; [A]; CmpS_{L2}; CmpS_{R2}]$	13	5,061	6.58

*$[A] := [A_{z_{sm}}; A_{xyz}; AV_{xyz}; v_x; steer_{DM}]$

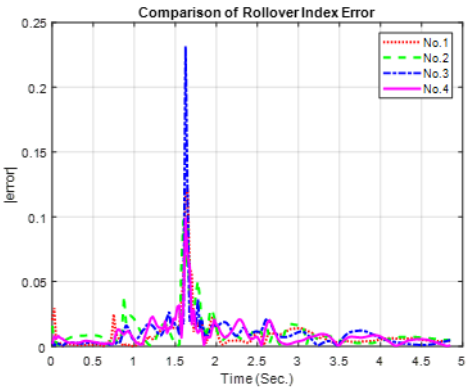
Note: the highlighted cell of the table represents the smallest value of the table.

5.3. The suitable neural network structure

The error of the Tanh neural network is still needed to be improved. The different types of neural network as shown in Table 6. are evaluated with 3 set of input features, case No. 1, 3, and 4 from Section 5.2.



a) Comparison of Rollover Index



b) Comparison of Rollover Index Error

Figure 8: Example of Rollover Index ($m_s = 2583\ kg$, $h_R = 0.829\ m$, $v_x = 101\ km/h$)

The results show that the neural networks in the case no. 7 Table 6 with the input #1 have the lowest test losses of 3.66×10^{-4}. In addition, the neural networks in the case no. 5 Table 6 with the input #3 has the test losses of 4.33×10^{-4} and the neural networks in the case No. 8 Table 6 with the input #2 has the test losses of 7.30×10^{-4}.

In summary, for the lowest test losses, the suitable Tanh neural network structure is the case No. 7 Table 6 which is (n 46 46 50 30 1). Moreover, for the low test losses and low number of input features, the suitable Tanh neural network structure is the case No. 5 Table 6. The test losses in this case is different from the test loss from the best one only 0.67×10^{-4}.

Table 6: Remaining Loss of Different Tanh Neural Network Structure on Test Data Set

No.	Neural Networks' Structure	Average RMS error of 2000 Data Set $\times 10^{-4}$			
		$N°$ param	Inputs #1	Inputs #2	Inputs #3
			n=17	n=13	n=13
1	n 12 12 1	12n+337	7.87	12.30	9.18
2	n 24 20 1	24n+1145	5.76	9.91	7.11
3	n 46 46 1	46n+4417	4.61	7.41	6.58
4	n 46 46 30 1	46n+5811	5.54	7.90	7.35
5	n 46 46 30 30 1	46n+6741	4.34	10.30	4.33
6	n 46 46 30 50 1	46n+7381	4.43	7.32	6.14
7	n 46 46 50 30 1	46n+8281	3.66	7.83	5.71
8	n 46 46 50 50 1	46n+9321	4.14	7.30	5.43

* Inputs #1: $[v_{z_{wc}}; [A]; CmpS]$

Inputs #2: $[v_{z_{wc_{L1}}}; v_{z_{wc_{R1}}}; [A]; CmpS_{L1}; CmpS_{R1}]$

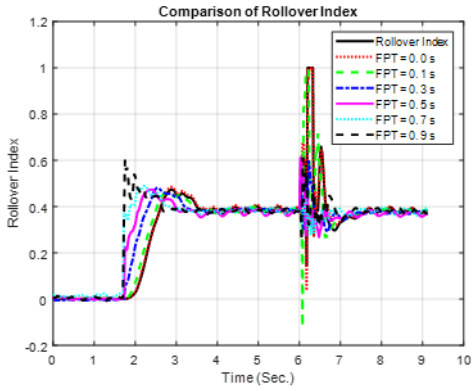
Inputs #3: $[v_{z_{wc_{L2}}}; v_{z_{wc_{R2}}}; [A]; CmpS_{L2}; CmpS_{R2}]$

Note: the highlighted cell of the table represents the smallest value of the table.

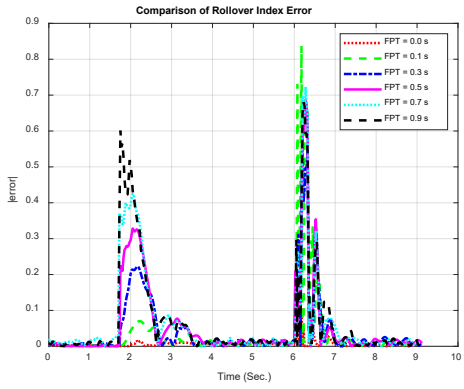
5.4. Future tripped and untrippped rollover prediction

By using the Tanh neural network in No. 7 Table 6, the error of the models shown in Table 7. Also, the example of the rollover index perdition is shown on Figure 9. The results show that once the shift time increases, the error increases. The difference between shift time 0.0 and 0.1 second is approximately 25×10^{-4}. The error seems large. However, if consider the future tripped rollover prediction plot on Figure 9a, the plot shows that they seem correctly predict the trend of rollover index up to 0.5 seconds in future. The prediction lines look like the actual rollover index but it is shelfed by the time.

In summary, this predict is very useful and can provide an extra time prediction for rollover prevention and warning system.



(a) Comparison of Rollover Index



(b) Comparison of Rollover Index

Figure 9. The Example of the Rollover Index in the Case of ($m_s = 2722\ kg$, $h_R = 0.847\ m$, $v_x = 82\ km/h$)

Altogether, the results exhibit that the Tanh recurrent neural network is able to predict the tripped rollover index with a larger preciseness than others once it is considered with the undetermined parameters, for instance, the height of the center of gravity and sprung mass. So, once processing the training data, it is necessary to take the said undetermined values into consideration. The suitable input features for tripped rollover index should be $[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}; v_x; steer_{DM}; CmpS]$ with Tanh neural network structure of (n 17 46 46 50 30 1). Moreover, the future tripped rollover prediction can provide an extra time prediction for rollover prevention and warning system.

Table 7: Remaining Loss of the Tanh Neural Network on Test Data Set

Future Perdition Time (Sec)	Average RMS error of 2000 Data Set $\times 10^{-4}$
0.0	3.66
0.1	29.2
0.2	35.6
0.3	43.4
0.4	55.5
0.5	83.4
0.6	133.4
0.7	214.1
0.8	328.2
0.9	464.5
1.0	616.6

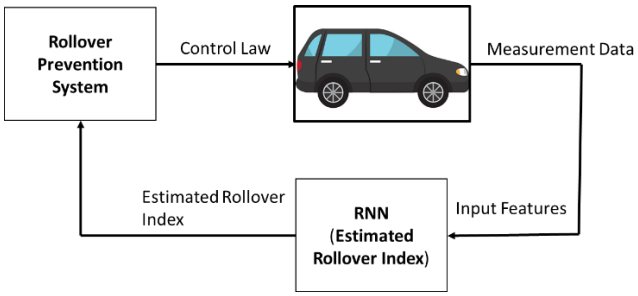


Figure 10: Application Diagram of The Neural Network

The application of the developed neural network is shown on Figure 10. The measurement data from a vehicle is sent to the recurrent neural network which can compute the estimated rollover index within 0.03 ms. Then, the estimated rollover index is sent to rollover prevention/warning system to calculation the control law to make the vehicle action and reduce the roll risk.

6. Conclusions

An essential factor of the carrying out the rollover preventive systems and warning ones is the assessment of the rollover risk. The rollover risk pointer or the rollover index is estimated by employing the feedforward neural network together with three other recurrent neural networks. It is seen that when in occurrence of any uncertain parameters, for instance, undetermined height of the center of gravity of the subject, the Tanh recurrent neural network shows higher performance juxtaposed with others. Calculations of the loss on a test set and an empirical evaluation on test trajectories support in verifying the estimation quality. The outcomes of the test are encouraging for further research, and the performance assessment of the recurrent neural networks in an actual vehicle would be an interesting subject of further research.

The sensory inputs that are chosen play a crucial part in the estimation of the rollover index. The results show that the suitable input features for tripped rollover index should be $[v_{z_{wc}}; A_{z_{sm}}; A_{xyz}; AV_{xyz}; v_x; steer_{DM}; CmpS]$ which are agree

with the previous research. Also, the suitable Tanh neural network structure of is (n 46 46 50 30 1). Moreover, the future tripped rollover prediction can correctly predict the trend of rollover index up to 0.5 seconds in future. This will be useful for rollover prevention and warning system.

Acknowledgment

This research was funded by Ratchadaphiseksomphot Endowment Fund Chulalongkorn University, grant number: CU_GI_62_05_21_01.

References

[1] C. Dengler, K. Treetipsounthorn, S. Chantranuwathana, G. Phanomchoeng, B. Lohmann, and S. Panngum, "Vehicle rollover detection using recurrent neural networks," in The 9th IEEE International Conference on Robotics, Automation and Mechatronics (CIS-RAM 2019), Bangkok, Thailand, Nov. 2019.

[2] US Department of Transportation, Type of Rollovers, National Highway Traffic Safety Administration (NHTSA), Washington, DC, August 2017. [Online] Available: http://www.safercar.gov

[3] W. Xinye, GE. Xiaohong, L. Shuyou, and H. Hongwu, "Research on Vehicle Rollover and Control," in 2010 2nd International Conference on Advanced Computer Control, 2010.

[4] R. L. Martin, C. Ilse, "Robust indirect-defined envelope control for rollover and lateral skid prevention," Control Engineering Practice, **61**, 149-162, 2017.

[5] J.J. Rath, M. Defoort, and K.C. Veluvolu, "Rollover index estimation in the presence of sensor faults, unknown inputs, and uncertainties," IEEE transactions on intelligent transportation systems, **17**(10), 2949-2959, 2016.

[6] G. Phanomchoeng, and R. Rajamani, "New rollover index for the detection of tripped and untripped rollovers," IEEE Transactions on Industrial Electronics, **60**(10), 4726-4736, 2013. DOI: 10.1109/TIE.2012.2211312

[7] R. Rajamani, and P. Damrongrit Neng, "New paradigms for the integration of yaw stability and rollover prevention functions in vehicle stability control," IEEE Transactions on Intelligent Transportation Systems, **14**(1), 249-261, 2013. DOI: 10.1109/TITS.2012.2215856

[8] K. Warwas, and A. Krzysztof, "An Application of Neural Networks to Control Stability of an Articulated Vehicle in Real Time," in Information Systems Architecture and Technology: Proceedings of 36th International Conference on Information Systems Architecture and Technology–ISAT 2015–Part I, 135-146. Springer International Publishing, 2016.

[9] B. Zhu, P. Qi, Z. Jian, and G. Litong, "Integrated chassis control for vehicle rollover prevention with neural network time-to-rollover warning metrics," Advances in Mechanical Engineering, **8**(2), 2016. https://doi.org/10.1177%2F1687814016632679

[10] K. Warwas, and K. Augustynek, "An Application of Neural Networks to Control Stability of an Articulated Vehicle in Real Time," in Information Systems Architecture and Technology: Proceedings of 36th International Conference on Information Systems Architecture and Technology – ISAT 2015 – Part I, 135-146, 2016.

[11] S. Solmaz, M. Akar, R. Shorten, "Adaptive Rollover Prevention for Automotive Vehicles with Differential Braking," in Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea, July 6-11, 2008.

[12] S. Lee, Y. Fitri, K. Misawa Kasahara, and M. Yasuchika, "Rollover prevention with predictive control of differential braking and rear wheel steering," in Robotics, Automation and Mechatronics (RAM), 2013 6th IEEE Conference on, 144-149. IEEE, 2013.

[13] Z. Yu, J. Li, F. Dai, and S. Li, "Anti-Rollover Control Based on Fuzzy Differential Braking for Heavy Duty Commercial Vehicle," in Conference: 2016 4th International Conference on Machinery, Materials and Computing Technology, 2016.

[14] S. Yim, "Design of a preview controller for vehicle rollover prevention," IEEE Transactions on Vehicular Technology, **60**(9), 4217-4226, 2011. DOI: 10.1109/TVT.2011.2169687

[15] S. Solmaz, C. Martin, and S. Robert, "A methodology for the design of robust rollover prevention controllers for automotive vehicles with active steering," International Journal of Control, **80**(11), 1763-1779, 2007. https://doi.org/10.1080/00207170701473987

[16] R.A. Azim, and F.M. Malik, "Rollover mitigation controller development for three-wheeled vehicle using active front steering," Mathematical Problems in Engineering, 2015.

[17] N. Zhang, G.M. Dong, and H.P. Du, "Investigation into untripped rollover of light vehicles in the modified fishhook and the sine maneuvers. Part I: Vehicle modelling, roll and yaw instability," Vehicle System Dynamics, **46**(4), 271-293, 2008. https://doi.org/10.1080/00423110701344752

[18] S. Yim, P. Yongjin, and K. Yi. "Design of active suspension and electronic stability program for rollover prevention," International journal of automotive technology, **11**(2), 147-153, 2010.

[19] D. Konik, "Dynamic drive-the new active roll stabilization system from BMW group-system description and functional improvements," in Proc. 5th Int. Symp. Avec 2000, Ann Arbor, MI 2000.

[20] S. Yim, K. Jeon, and K. Yi, "An Investigation into Vehicle Rollover Prevention by Coordinated Control of Active Anti-roll Bar and Electronic Stability Program," International Journal of Control Automation and Systems, **10**(2), 2012. https://doi.org/10.1080/002071707014739887

[21] D.J. Sampson, and C. David, "Active roll control of single unit heavy road vehicles," Vehicle System Dynamics, **40**(4), 229-270, 2003.

[22] J. Yoon, W. Cho, and J. Kang, "Design and evaluation of a unified chassis control system for rollover prevention and vehicle stability improvement on a virtual test track," Control Engineering Practice, **18**(6), 585-597, 2010.

[23] D. Piyabongkarn, J.Y. Lew, R Rajamani, and J.A. Grogg, "Active driveline torque-management systems," IEEE Control Systems, **30**(4), 86-102, 2010.

[24] M.C. Walz, Trends in the static stability factor of passenger cars, light trucks, and vans. No. HS-809 868. 2005.

[25] S.B. Choi, "Practical vehicle rollover avoidance control using energy method," Vehicle System Dynamics, **46**(4), 323-337, 2008.

[26] J. Yoon, K. Dongshin, and Y. Kyongsu, "Design of a rollover index-based vehicle stability control scheme," Vehicle system dynamics, **45**(5), 459-475, 2007.

[27] B.C. Chen, "Warning and control for vehicle rollover prevention," No. 30-00934 UMI. Ann Arbor, MI: University of Michigan, 2001.

[28] G. Phanomchoeng, and R. Rajamani, "Real-Time Estimation of Rollover Index for Tripped Rollovers with a Novel Unknown Inputs Nonlinear Observer," IEEE/ASME Transactions on Mechatronics, **19**(2), 743-754, 2014.

[29] R. Rajamani, and N. Piyabongkarn, "New Paradigms for the Integration of Yaw Stability and Rollover Prevention Functions in Vehicle Stability Control," IEEE Transactions on Intelligent Transportation Systems, **14**(1), 249-261, 2013.

[30] X. Cai, and J. Wu, "Hazardous Chemicals Vehicles Rollover Pre-warning System Based on CPS," in Proceedings of the 2012 IEEE Sixth International Conference on Software Security and Reliability Companion, 2012.

[31] J.R. Kerrigan, N.J. Dennis, D.P. Parent, and S. Purtsezov, "Test system, vehicle and occupant response repeatability evaluation in rollover crash tests: The deceleration rollover sled test," International Journal of Crashworthiness, **16**(6), 583-605, 2011.

[32] Bo.C. Chen, and H. Peng, "A Real-time Rollover Threat Index for Sports Utility Vehicles," in Proceedings of the Amencan Control Conference, 1999.

[33] Bo.C. Chen, and H. Peng, "Rollover Warning Of Articulated Vehicles Based On A Time-To-Rollover Metric," in Proceedings of the 1999 ASME International Congress and Exposition, Knoxville, TN, November 1999.

[34] B.L. Boada, M.J. L. Boada, L. Vargas-Melendez, and V. Diaz, "A robust observer based on H1 filtering with parameter uncertainties combined with Neural Networks for estimation of vehicle roll angle," Mechanical Systems and Signal Processing, **99**, 611–623, 2018.

[35] X. Glorot, and A. Bordes, and Y. Bengio, "Deep Sparse Rectifier Neural Networks," in Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS) 2010.

[36] L. Wang, N. Zhan, and H. Du, "Real-time Identification of Vehicle Motion-Modes using Neural Networks," Mechanical Systems and Signal Processing **50-51**, 632–645, 2015.

[37] E.N. Sanchez, L.J. Ricalde, R. Langari, and D. Shahmirzadi, "Rollover Prediction and Control in Heavy Vehicles via Recurrent Neural Networks," 43rd IEEE Conference on Decision and Control, December 14-17, 2004.

[38] K. Treetipsounthorn, and G. Phanomchoeng, "Real-Time Rollover Warning in Tripped and Untripped Rollovers with A Neural Network," in The 4th International Conference on Control Science and Systems Engineering (ICCSSE 2018), Wuhan, China, Aug. 2018.

[39] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, **9**(8), 1735–1780, 1997.

[41] R. Dey and F. M. Salemt, "Gate-variants of gated recurrent unit (gru) neural networks," in 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). IEEE, 1597–1600, 2017.

[42] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks,"

in Proceedings of the fourteenth international conference on artificial intelligence and statistics, 315–323, 2011.
[43] P. J. Werbos et al., "Backpropagation through time: what it does and how to do it," in Proceedings of the IEEE, **78**(10), 1550–1560, 1990.
[44] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in International Conference on Machine Learning, 1310–1318, 2013.
[45] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555, 2014.
[46] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in International Conference on Machine Learning, 2342–2350, 2015.
[47] C. Winkler, K. Campbell, and C. Mink, "Center of gravity height: a round-robin measurement program," A Technical Report to The Motor Vehicle Manufacturers Association MVMA Project No. 0167A, The University of Michigan, Transportation Reserch Institute, 1991.
[48] C. E. Rasmussen, Gaussian Processes for Machine Learning. MIT University Press Group Ltd, 2005.
[49] D. Yuret, "Knet: beginning deep learning with 100 lines of julia," in Machine Learning Systems Workshop at NIPS 2016, 2016.
[50] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," CoRR, abs/1412.6980, 2014. [Online] Available: http://arxiv.org/abs/1412.6980

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 5, 237-244 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

FPGA Acceleration of Tree-based Learning Algorithms

Haytham Azmi*

Microelectronics Department, Electronics Research Institute, Cairo, Egypt

ARTICLE INFO

Article history:
Received: 21 June, 2020
Accepted: 02 September, 2020
Online: 14 September, 2020

Keywords:
Acceleration
Machine learning
Decision tree
FPGA

ABSTRACT

Machine learning classifiers provide many promising solutions for data classification in different disciplines. However, data classification at run time is still a very challenging task for real-time applications. Acceleration of machine-learning hardware solutions is needed to meet the requirements of real-time applications. This paper proposes a new implementation of a machine learning classifier on Field Programmable Gate Arrays (FPGA). The proposed implementation utilizes the MicroBlaze soft-core processor on FPGA and uses the Advanced eXtensible Interface (AXI) bus to integrate the MicroBlaze with hardware peripherals. Experimental results shows that hardware-software co-design is a promising solution as it saves silicon area and provides a flexible configuration of decision tree algorithms at run time.

1 Introduction

The continuous growth of Internet of Things (IoT) applications that need real-time data analysis, such as video cameras in surveillance applications, live streaming videos on social media networks and sensor data from smart cities, is driving the demand for an efficient platform that can process this huge amount of data and discover hidden patterns.

When designers explore hardware options for big-data analysis, Graphics Processing Units (GPUs) and Field Programmable Gate Arrays (FPGAs) appear as two promising solutions for the implementation of machine learning applications. GPUs have been dominating the parallel computing market for a long time and their hardware solutions have been thoroughly tested in different industrial domains. However, FPGA solutions have been recently considered in many high performance Artificial Intelligence (AI) applications and showed an improved power consumption, compared to GPUs.

To choose the optimum implementation, designers need to answer several design questions and explore possible solutions at early design phases. Examples of these questions are: what is the maximum clock frequency that can be reached in the target hardware platform? How many General Purpose Input/Output (GPIO) pins that can be utilized by the target hardware platform? Does the target hardware platform support parallel processing and concurrency or not? what is the average power consumption of similar implementations on the target platform? etc. Designers must study these questions carefully at early deign phases. The choice of the hardware platform has a significant impact on the performance of the target applications.

FPGA-based implementations of machine learning classifiers provide an efficient solution for real-time applications that use deep learning algorithms for classification [1]. FPGA-based implementation helps system engineers design modular systems and execute several tasks concurrently, which helps in processing real-time data more efficiently. The implementation of machine learning classifiers on FPGA can be done in many different ways. This paper discusses these alternative implementation options and presents a case study of hardware implementation of a machine learning algorithm on FGPA.

This paper has two main contributions.

- First, we present a comparative analysis of different design approaches of machine learning classifiers in the literature. The paper discusses technical implementation challenges and proposes practical solutions that help designers better implement efficient classifiers.
- Second, we propose a hybrid implementation of a machine learning, tree-based classifier. The proposed implementation utilizes hardware-software co-design approach that integrates a soft-core microprocessor on FPGA. The soft-core is used along with other digital blocks to build an area-efficient machine learning classifier. As a proof of concept, we use a MicroBlaze Xilinx IP core on Virtex-UltraScale FPGA to run the main tree-based algorithm and report its implementation details.

The novelty of the proposed implementation is that it bridges the

*Corresponding Author: Haytham Azmi, Microelectronics Department, Electronics Research Institute, Cairo, Egypt, haitham@eri.sci.eg

Table 1: Some of the proposed work in the literature that discusses the hardware implementation of machine learning classifiers.

N	Dataset	Classification	Target Platform	Performance	Reference
1	IBM DataQuest Generator	Decision Tree (Gini calculation)	PowerPC	5.58 times	**[2]**
2	UCI repository	Adaboost algorithm	VHDL Generator	5 to 10 ns per decision	**[3]**
3	Spambase UCI repository	Decision Trees Majority voting	Xilinx Virtex 5 XC5VLX110T	N/A	**[4]**
4	EEG data in house	Soft Decision Tree	Xilinx Spartan 3	98.1%	**[5]**
5	From **[6]**	Decision Tree (image-based)	software tool	75 million samples per second	**[7]**
6	From **[8]**	C4.5 classifier	Xilinx Virtex-6 XC6LX670	97.92% accuracy	**[9]**

gap between hardware and software when dealing with machine learning classifier implementation. Our solution addresses several challenges related to creating a customized machine learning classifier on FPGA from a raw dataset file. We discuss these challenges in detail in this paper and explain how researchers can use our proposed implementation to build low-cost, high-performance classifier on FPGA.

The rest of this paper is organized as follows. Section 2 presents the related work in the literature. Section 3 discusses the hardware implementation of the decision tree algorithm. Section 4 presents the results of the experimental work. Section 5 draws our conclusions and suggests new directions for future work.

2 Related Work

Utilizing hardware solutions for machine learning classification has been a challenging task for many researchers and system engineers. Researchers targeted different platforms, such as microprocessor-based systems, FPGA-based designs and parallel computing techniques to design efficient hardware classifiers for real-time applications. Table 1 highlights some of the proposed work in the literature that discusses the hardware implementation of machine learning classifiers. In this section, we highlight examples of the related work in more details.

In [2], the authors used *Gini* score calculation to develop a decision tree model. Authors went through two phases to build a decision tree model. First, they chose the splitting attribute and a split index for the root. Then, they split the children records based on the the first phase's decision. The proposed implementation in [2] recursively repeated this process until a stopping criterion is met [2]. In practical applications, Conventional Neural Networks (CNN) are trained off-line using training datasets through the back-propagation process. Following that, the trained CNN models are used to recognize images using the feed-forward process. Since there is currently a trending approach in recognizing images in real-time applications, the execution time of the feed-forward process is the most significant factor when it comes to performance evaluation [10]. In [11], the authors presented a software tool that can automatically generate a tree classifier Verilog code from python scripts. Authors used SciKit-Learn machine learning library to build the trained model and calculate the threshold values for the tree nodes. Following that, the tool is used to generate a Verilog code for decision trees and random forests [11].

In [12], the authors presented an implementation of Q-learning with Artificial Neural Networks (ANN) on FPGA for real-time applications that have latency and power constraints. The proposed implementation reduced processing time by utilizing the parallel structure of FPGA logic units. In [12], the authors demonstrated the implementation of a single neuron Q-learning accelerator as well as Multilayer Perception (MLP) Q-learning accelerator on a Virtex 7 FPGA. The authors discussed the architecture of the implementation in their paper [12]. However, when they evaluated the performance, it was compared to a CPU-based implementation. The speedup in performance was completely expected because of the concurrency achieved by FPGA. The paper did not provide any information about the speedup in performance when compared to GPU implementation [12].

In [13], the authors proposed a method to optimize CNN by utilizing four techniques: 1) fixed-point quantization to minimize calculations, 2) approximation of activation function to reduce the complex mathematical operations, 3) pipelining and parallelization of loops and tasks to speedup execution time, and 4) memory reorganization to enhance fetch time. Authors targeted the LeNet-5 architecture, which is a neural network architecture for handwritten and machine-printed character recognition [14]. They implemented their proposed accelerator on a Xilinx FPGA and used Zynq-7000 platform. The paper reported an operating frequency at 166MHz, which is not a useful measure for high performance applications. However, the paper reported that the proposed implementation can reduce the consumed energy by 93.7% compared to a GPU implementation [13]. The authors in [15] designed a decision tree classifier to recognize letters and digits. Xilinx Zynq SoC is used by the authors as a target hardware platform and Vivado high level synthesis is used as a development tool with C/C++ synthesis. Authors verified the correctness of the generated HDL code using C/RTL co-simulation [15].

In [16], the authors proposed a pipelined design that is partitioned into two stages. The first stage of the proposed design determines the conditions of the decision tree for which the classification can be done. The second phase of the proposed design used a pipelined data path for parallel execution.

Other implementations have been also proposed in the literature. In [17], the authors presented an automatic modulation classifier to identify the signal modulation format for electronic applications. In [18], the authors presented a hybrid classification engine using CPU and FPGA shared memory.

Although there are several proposed solutions in the literature for the hardware implementation of machine learning classifiers, the FPGA design and implementation of such classifiers were not discussed in proper details. We try to address this gap in our paper to present a hardware-software co-design implementation of a machine learning classifier on FPGA.

3 Hardware Implementation of Decision Trees

Designers explored various ways to implement decision tree algorithms on hardware platforms. In this section, we first highlights the basic concepts of implementing a single node on FPGA and how the synthesis tool interprets the HDL model and convert it into logic gates. Then, we discuss the common methodologies used in hardware implementation of such algorithms.

Classifying instances based on a training dataset has been used extensively to mine hidden patterns within detests [19]. One of the most popular classification algorithms is the tree-based classifier, where the main goal is to construct a tree from a labeled training dataset. The decision tree consists of a root node and children nodes. Each node represents a test on a feature, whereas each branch represents the outcome of a test, and each leaf node represents a class label. Figure 1 shows an example of a decision tree.

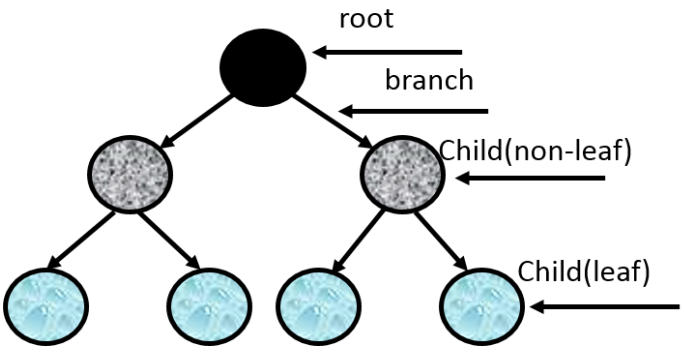


Figure 1: An example of a decision tree.

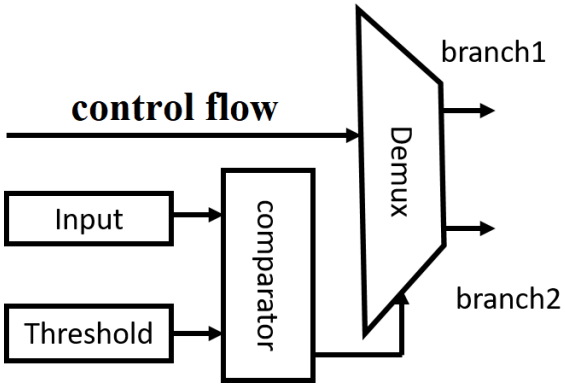
Decision trees can be constructed using different algorithms. The most critical design parameter is how to choose the feature at each node that best splits the dataset. The classification algorithms in the literature use different evaluation metrics to perform two design decisions.

1. Which feature should be used to split the dataset records at each node?
2. What is the the threshold value that should be used to split the dataset records at each node?

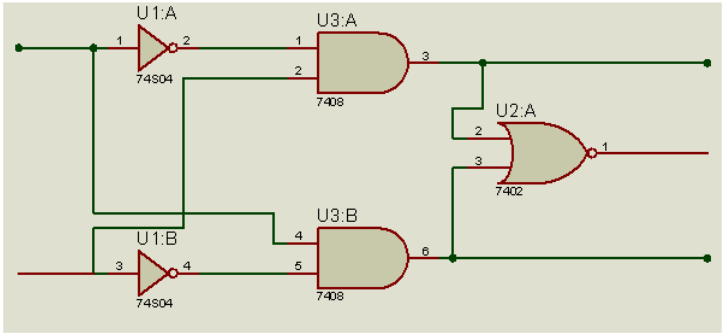
Examples of these evaluation metrics include using Gini index [20] and information gain, which is based on the concept of entropy [21]. Examples of decision tree algorithms include ID3, C4.5, CART, CHAID, and MARS [22]. ID3 is used in this paper as an example of a decision tree implementation. Other algorithms, such as C4.5 and CART, can be implemented using the same approach

A node in a decision tree algorithm represents a branching condition. If the input meets the condition, the flow of the tree goes through a certain path. Otherwise, a different path is selected. The tree can have a binary node or a multi-threshold node. Binary nodes branch out into two paths only, whereas multi-threshold nodes can branch out into more than two paths.

Figure 2 shows an example of the internal structure of a simple binary node. Figure 2a shows a block diagram of a branching circuit for a single node. The node first checks the input value and compares it to a given threshold value. The comparator then sends the result to a demultiplixer circuit that directs the flow of the control signals into one of two branches. A simple implementation of a one-bit comparator circuit is shown in Figure 2b. The comparator is usually designed to deal with n-bit values but Figure 2 is used here to simplify the concept.



(a) A block diagram of a branching circuit for a binary node



(b) Equivalent schematic of a branching circuit of a binary node

Figure 2: An example of a branching circuit of a binary node.

Since we are targeting Xilinx UltraScale architecture, we conducted an extensive study of the Configurable Logic Block (CLB) architecture of Xilinx UltraScale family. The CLB is the main building block that is used to implement all combinational and sequential logic of our proposed tree-base classifier. Each CLB unit consists of a group of logic elements along with an interconnect resource to connect all logic units. We followed the UltraScale Architecture Configurable Logic Block user guide in [23] to make sure that the synthesis tool will interpret the HDL code correctly and will generate an optimum logic design for the target application.

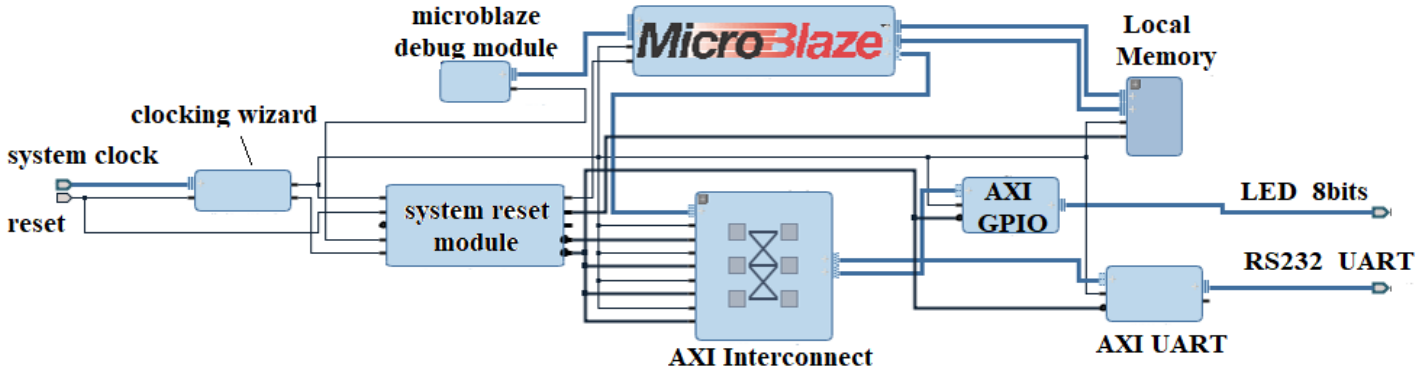


Figure 3: An example of a MicroBlaze interface circuit.

3.1 Implementation Challenges

Researchers utilized different techniques to implement machine learning algorithms on FPGA. Two common approaches are mostly cited in the literature. The first approach is writing the machine learning algorithm in a high level language, such as C, and then using a synthesis tool to convert from a high level language to a hardware description language. The second approach is targeting a structure-based design. Then, building all components in hardware and connecting these modules using port mapping. Each one of these two techniques has its advantages and disadvantages. Moreover, there are different implementation challenges associated with each approach. In this section, we discuss these challenges in details and explore different solutions to address them.

High Level synthesis (HLS) tools provide a mechanism to auto-generate HDL codes from high level languages, such as C, C++ and SystemC. HLS tools can save development time when implementing complex algorithms that can be easily represented in C language, for example. However, the majority of these tools do not provide complete support to generate recursion function in hardware description languages. Recursion comes with a difficult synthesis challenge by nature because the synthesis tool does not know when the exit condition will happen and hence can not predict the complexity of the hardware circuit that should be generated to implement the recursion function. For example, to generate a tree-based classifier, it is important that the synthesis tool knows the maximum depth of the tree. Otherwise, the tool can't support the generation process of the tree hardware logic at run time if the depth is unknown.

This limitation is blocking the usage of HLS tools in such scenarios. Researchers created various techniques to address this problem by developing recursive functions as an Embedded Domain Specific Language (EDSL) in C++ [24]. Authors in [24] addressed the recursion limitation by utilizing the C++ front-end of an HLS compiler [24]. However, this approach is not easy to integrate with HLS tools, such as Vivado, as the library needs to be re-customized every time the tool is used. This limitation makes it difficult to rely on this approach as a permanent solution, especially for real-time applications that require recursions to operate at run-time. Another proposed solution is having a maximum estimated depth of the tree and generating the digital logic for any tree-based classifier based on that maximum depth. However, this solution might fail if the target application requirements exceed the maximum depth. Moreover, if the requirements were within the maximum depth, the generated logic will have extra gates that are not used in the design, which causes an overhead area cost.

Another approach is utilizing structural implementation by dividing up the machine learning classifiers into separate blocks. These blocks are designed to deal with data-input, feature engineering, dataset training, and classification or prediction. For each block, designers write the hardware description code that implements the functionality of the intended block. The design is then connected together through port mapping as one complete design and used to process input data and solve machine learning problems.

This approach is proven to provide an optimized design for the target problem. However, the design process forces engineers to make very rigid design decisions to handle the target training set. For example, when generating a tree, designers can build a simple tree design and then re-instantiate this block through a pipeline structure. The challenge with this design method is that it causes a significant delay during the training phase because it executes the learning tasks sequentially. On the same context, designers can expand the tree to the required depth using hardware generators. This approach has a technology-based limitation as the maximum tree depth must be defined based on the target available silicon area.

3.2 Proposed Hardware/Software Co-design

Our proposed design aims at addressing the limitations of previous designs that are discussed earlier. We utilize the hardware-software co-design approach which gives designers the ability to execute several tasks in parallel while benefiting from the conventional implementation of complex recursive functions. This approach has been adopted previously by several researchers. The authors in [25] presented various implementations of decision trees using a sequence of universal nodes.

We explored soft and hard cores offered by Xilinx technology and we chose to use the the MicroBlaze IP core as our main processor [26]. MicroBlaze is built as a soft processor, RISC-based architecture, with a rich set of instructions that are optimized for embedded applications [27]. The MicroBlaze solution offers complete flexibility to select the needed peripheral, bus interfaces for memory and GPIO, and the ability to implement complex recursive functions that are needed in the machine learning training phase [27].

To complete the machine learning system design, we utilized other IPs to be able to train the network using different types of datasets. The soft-core is wrapped up by a Verilog code so that it can interact with the other components in the system. The soft-core is configurable in terms of number of ports, interrupt vectors, programmable timers, size of memory, and the maximum operating frequency.

Figure 3 shows a schematic diagram of the MicroBlaze interface circuit. The IP comes with default ports for interrupt, debug, clock , reset, DLMB, ILMB and AXI interfaces. Designers have several options when it comes to the configuration of the MicroBlaze interface. Designers can choose to integrate AXI or ACE bus and can also enable or disable the debug interface. To speed up the performance, designers can use instruction and data caches. The MicroBlaze is connected to its peripherals through an AXI interconnect.

Figure 4 shows that AXI interconnect can be configured in two different ways. First, the number of master/slave connections can be programmed to generate the required interfaces. Second, the interconnect optimization strategy can be set to minimize the area or maximize the performance, which gives the designer the ability to choose various strategies based on the application requirements.

Having a standard interface like AXI bus allows designers to easily integrate other hardware peripherals/IPs to the design easily. Since AXI is part of the Advanced Microcontroller Bus Architecture (AMBA) family, it will also easy to integrate any peripheral/IP that can be connected to AXI through bridges, such as APB, ACE, etc.

We conducted an extensive study of the architecture of the target FPGA to analyze the propagation delay through each LUT regardless of the function implemented, which is done for all classification functions.

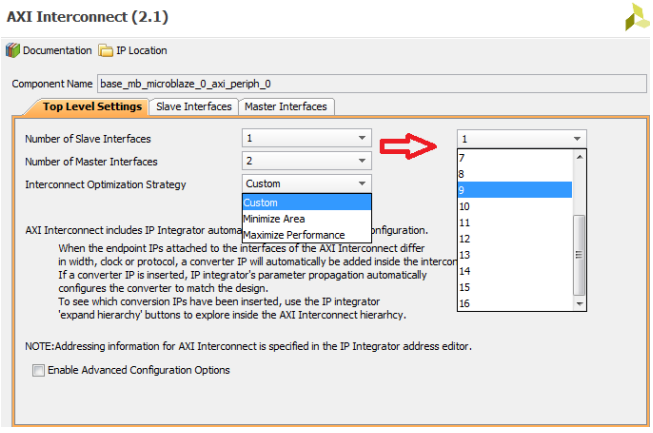


Figure 4: A programmable AXI interconnect control window from Xilinx.

To implement a hardware/software co-design on Virtex UltraScale, we use Vivado software tool. After completing the software programming in the Software Development Kit tool, we had to generate the .elf file and associate it with the MicroBlaze IP in Vivado to complete the integration process. The association process had several technical challenges. For example, the default block RAM size was one big limitation. We addressed this issue by allocating extra memory for the MicroBlaze software code.

The integration between hardware logic design and software code that runs on a microprocessor achieved several performance enhancement goals. First, the hardware logic of the entropy calculation and tree splitting is done using combinational logic gates that execute in parallel, which speeds up the processing of the tree generation during training and the classification during testing. The software code that runs on the MicroBlaze soft-core addressed the recursion task during the generation of the tree structure. Therefore, this integration was a major factor in improving the overall performance of the proposed design.

4 Experimental Work

The proposed system is designed to deal with different types of datasets that might need further processing. Figure 5 shows various data types that need to be considered before reading dataset for processing. Data can be qualitative or quantitative. Qualitative data includes

- nominal data, such as variables with no ranking sequence or inherent order like race and gender, etc.
- ordinal data, such as variable with an order like blood group, etc.
- binary data that has only two options, such as yes and no.

On the other hand, quantitative data includes

- discrete data, such as finite numbers of values that can not be classified meaningfully, etc.
- continuous data is information that can be measured on a scale and can be classified meaningfully, etc.

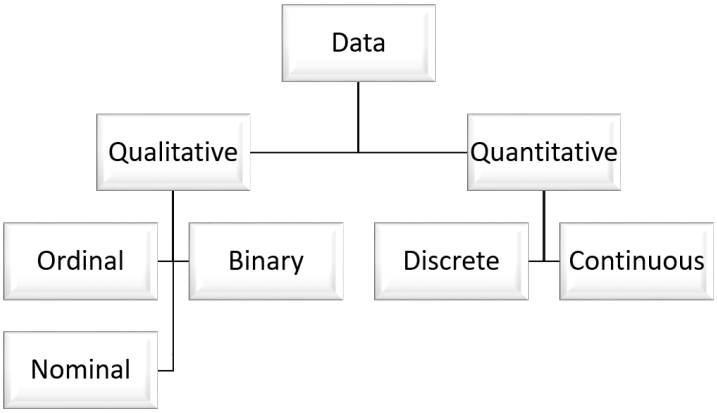


Figure 5: An explanation of the different data types that should be handled by the proposed design.

A pre-processing step must be executed first in order to convert the input data, into a standard format that is easy to process by the machine learning engine. One possible solution is to convert all various data types into binary format. This conversion is a very significant step to ensure that all input data types are converted to a machine-friendly format.

Figure 6 shows an example of the conversion process from nominal and numerical types of data to binary data type. The conversion

process is done as follows. First, each feature is examined to extract its list of all possible values. For example, the first column in Figure 6 (i.e. currency) can take three values: 1A, 5A and 6A. For each one of these values, we create a separate feature in the dataset in binary format. The new feature indicates whether each specific instance in the dataset has the currency set to that value or not. Since the currency feature has three values. Then, three new binary features are created to replace it. As shown in the example in Figure 6, the currency feature is replaced in the conversion by three binary features: currency(1), currency(5), and currency(6).

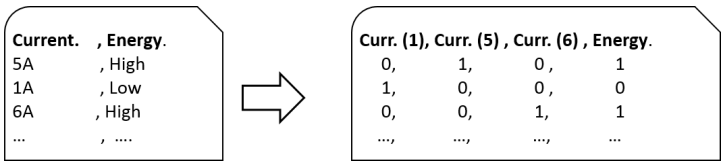


Figure 6: An example of converting nominal to binary format.

The number of features in Figure 6 for the output set will most likely increase based on how many nominal values are found in each feature. This conversion process helps researchers get an early estimation of the required memory blocks that should be allocated to host and process the training set.

4.1 Choosing the most significant feature

The most important decision in building a decision tree is choosing the root element that best distinguishes different patterns base on a certain threshold. The tree structure helps in building a model that classifies instances by navigating through various nodes in the tree until reaching one of the leaf nodes [28]. We use the information gain in this paper as our target evaluation metric so that the feature with the highest information gain is selected as the root element. The children nodes are then generated recursively. We continue to split the tree until we reach the decision nodes. The information gain calculations are based on the impurity level calculations in a set of samples. The formula used to calculate the entropy is given by [28]:

$$Entropy = -\sum_{i} P_i log_2 P_i \quad (1)$$

Where P_i is a probability of the class i relative to the total number of samples in the dataset.

The MicroBlaze processor is used to complete the calculations of the information gain as part of the software implementation of the target classifier. Feature selection plays a major role in improving the success rate of the decision tree classification. We conducted a comparative analysis to check which features are highly correlated and which features are causing overlaps that reduces the classification accuracy.

4.2 FPGA Implementation

FPGA implementation goes through various steps. First, we compile the code, then we simulate the design to make sure that the design functions correctly. Following that, we perform design synthesis to convert the design into equivalent logic gates. After synthesis, we perform post-synthesis simulation to ensure that the setup, hold, and propagation delays of the used gates don't affect the functionality of the design. Then, we perform the translation, mapping, place and route steps. Finally, we perform post-layout simulation and generate the bitstream file that is used to program the target FPGA.

Table 2 summarizes the results of the FPGA implementation of the proposed design. Due to the utilization of the MicroBlaze soft-core, we used limited resources from the target FPGA. Figure 7 presents a snapshot of the software tool that shows the placement and layout of the generated cells on the target FPGA. Figure 7 highlights the layout utilization percentage when implementing the proposed design in the target FPGA. The layout confirms the synthesis results that such a design doesn't require a lot of silicon area when implemented using the proposed SW/HW co-design approach. This is due to the utilization of the MicroBlaze software core, which saves a lot of silicon area during the implementation process. Figure 8 shows the statistics of on-chip components used from slice logic to implement the interface circuit with the MicroBlaze soft-core. It is clear from the figure that look-up-tables and registers constitute the majority components used from the slice logic in the FPGA.

Table 2: Implementation results

N	Item	Measure
1	Total On-Chip Power (W)	0.232
2	Dynamic (W)	0.154
3	Device Static (W)	0.151
4	Effective TJA (C/W)	1.7
5	Max Ambient (C)	84.5
6	Junction Temperature (C)	25.5
7	Slice Logic	3542
8	Block RAM	9
9	MMCM	2
10	I/O	14

We studied several implementations in the literature to compare the implementation of our proposed solution to previous implementations in the literature. We found out that it is difficult to make a one-to-one comparison due to three reasons. First, the target FPGA in the published work in the literature is different from ours. Therefore, the hardware utilization and used resources will not use the same base-reference to report the results, which makes it difficult to compare one implementation to another.

Second, there is no standard benchmark for dataset/classifier pair that is used across different implementations in the literature. Therefore, we could not implement a reference dataset/classifier pair to make a fair comparison. Third, the results are reported in different formats in the published papers. Few authors used utilization ratios, whereas others created their own cost unit formulas. Hence, the performance evaluation criteria are not the same across different publications. Although these factors affect the correctness and acc-

Table 3: Comparison with other implementation in the literature.

N	Implementation	Used Hardware	Implementation Main Finding	Reference
1	Gini-based decision tree	Xilinx Virtex-II Pro-based	31% minimum utilization	[29]
2	Hyper-rectangle hardware implementation	Xilinx Virtex (unknown version)	minimum 18 custom cost unit	[30]
3	Various architectures	Xilinx Virtex 5	SMpL needs 56% less resources	[31]
4	Multiple classifier systems,	Xilinx Virtex 5	95.8% accuracy - custom performance	[32]
5	Random Forest	GPU GeForce GTX 280	GPU implementation evaluates the forest in about 1% of the time compared to CPU	[33]
6	Random Forest	NVIDIA Tesla C2050	speed increases at least 300x compared to CPU	[34]
7	Proposed HW-SW co-design	Xilinx UltraScale	Less than 1% utilization of resources	This paper

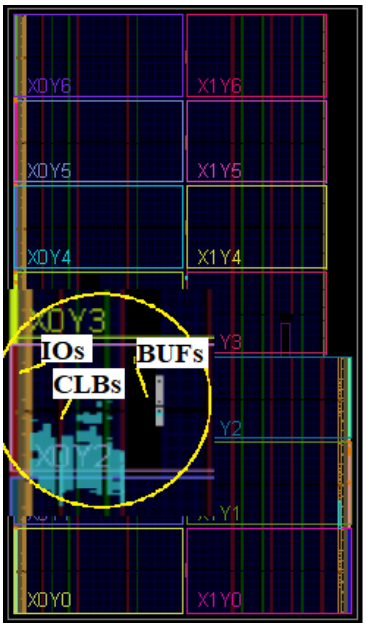


Figure 7: A snapshot of the FPGA placement and routing of the target design.

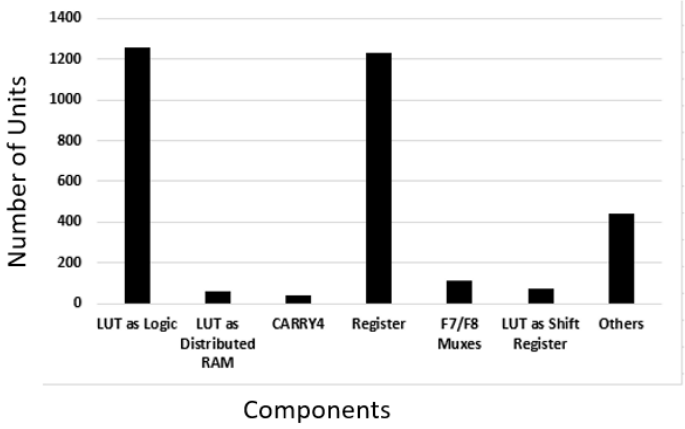


Figure 8: On-chip components used from slice logic.

curacy of the comparison, we reported the implementation results of various published work in Table 3. We presented these findings in Table 3 to show the different ways authors reported their implementation results of tree-based classifiers on FPGA. We also added GPU implementation results as examples of alternative hardware implementations. Experimental results show that the proposed design utilized less than 1% of the FPGA resources when implemented on Xilinx Virtex UltrScale FPGA.

5 Conclusion and Future Work

This paper presented a promising software-hardware co-design architecture for the implementation of machine learning algorithms. As a proof of concept, the paper discussed the implementation of a decision-tree algorithm based on ID3 implementation. The main functions of the ID3 algorithm ran on a MicroBlaze soft-core IP from Xilinx to utilize the processing units that are already available in the FPGA design flow. The integration of peripherals with the MicroBlaze is done through an AXI AMBA bus using a light weight version. Choosing AXI bus is an important design decision to make sure that the soft-core IP has a standard interface that simplifies the integration with other peripherals.

References

[1] A. Shawahna, S. M. Sait, A. El-Maleh, "FPGA-based Accelerators of Deep Learning Networks for Learning and Classification: A Review," CoRR, **abs/1901.00121**, 2019.

[2] R. Narayanan, D. Honbo, G. Memik, A. Choudhary, J. Zambreno, "An FPGA Implementation of Decision Tree Classification," in 2007 Design, Automation Test in Europe Conference Exhibition, 1–6, 2007, doi:10.1109/DATE.2007.364589.

[3] R. Narayanan, D. Honbo, G. Memik, A. Choudhary, J. Zambreno, "An FPGA Implementation of Decision Tree Classification," in 2007 Design, Automation Test in Europe Conference Exhibition, 1 – 6, 2007, doi:10.1109/DATE.2007.364589.

[4] M. Barbareschi, S. D. Prete, F. Gargiulo, A. Mazzeo, C. Sansone, "Decision Tree-Based Multiple Classifier Systems: An FPGA Perspective," International Workshop on Multiple Classifier Systems, 194–205, 2015.

[5] R.Harikumar, M.Balasubramani, "FPGA Synthesis Of Soft Decision Tree (SDT) For Classification Of Epilepsy Risk Level From Fuzzy Based Classifier Using EEG Signals," International Journal of Soft Computing and Engineering, **1**, 206–211, 2011.

[6] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), volume 1, 886–893 vol. 1, 2005, doi: 10.1109/CVPR.2005.177.

[7] M. FULARZ, M. KRAFT, "Hardware implementation of a decision tree classifier for object recognition applications," Measurement Automation Monitoring, **61**, 379–381, 2015.

[8] T. Traces, "TCP Statistic and Analysis Tool," http://tstat.tlc.polito.it/, accessed: 2019-08-06.

[9] K. K. M. Da Tong, Lu Sun, V. Prasanna, "High Throughput and Programmable Online Traffic Classifier on FPGA," Proceedings of the ACM/SIGDA international symposium on Field programmable gate arrays, 255–264, 2013.

[10] C. Zhang, P. Li, G. Sun, Y. Guan, B. Xiao, J. Cong, "Optimizing FPGA-based Accelerator Design for Deep Convolutional Neural Networks," in Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, FPGA '15, 161–170, ACM, New York, NY, USA, 2015, doi: 10.1145/2684746.2689060.

[11] T. Bhaaskar, K. Vishal, "Automatic generation of hardware Tree Classifiers," https://open.bu.edu/handle/2144/23688, 2017, accessed: 2020-09-07.

[12] P. R. Gankidi, J. Thangavelautham, "FPGA architecture for deep learning and its application to planetary robotics," in 2017 IEEE Aerospace Conference, 1–9, 2017, doi:10.1109/AERO.2017.7943929.

[13] Gan Feng, Zuyi Hu, Song Chen, Feng Wu, "Energy-efficient and high-throughput FPGA-based accelerator for Convolutional Neural Networks," in 2016 13th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), 624–626, 2016, doi:10.1109/ICSICT.2016.7998996.

[14] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, **86**(11), 2278–2324, 1998.

[15] R. Kulaga, M. Gorgon, "FPGA Implementation of Decision Trees and Tree Ensembles for Character Recognition in Vivado HLS," Image Processing and Communication, **19**, 71–82, 2015.

[16] F. Saqib, A. Dutta, J. Plusquellic, P. Ortiz, M. S. Pattichis, "Pipelined Decision Tree Classification Accelerator Implementation in FPGA," IEEE Transactions on Computers, **64**, 280–285, 2015.

[17] J. Grajal, O. Yeste-Ojeda, M. Sanchez, M. Garrido, M. Lopez-Vallejo, "Real-time FPGA Implementation of an Automatic Modulation Classifier for Electronic Warfare Applications," The 19th European Signal Processing Conference (EUSIPCO), 1514–1518, 2011.

[18] M. Owaida, H. Zhang, C. Zhang, G. Alonso, "Scalable Inference of Decision Tree Ensembles: Flexible Design for CPU-FPGA Platforms," 27th International Conference on Field Programmable Logic and Applications (FPL), 1–8, 2017.

[19] S. Surekha, "A Comparative Study of Rough Set Theoretic Decision Tree Induction Algorithms," in 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), 1–6, 2018, doi:10.1109/ICCTCT.2018.8550978.

[20] H. Liu, M. Zhou, X. S. Lu, C. Yao, "Weighted Gini index feature selection method for imbalanced data," in 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 1–6, 2018, doi: 10.1109/ICNSC.2018.8361371.

[21] B. Chen, L. Xing, B. Xu, H. Zhao, J. C. Príncipe, "Insights Into the Robustness of Minimum Error Entropy Estimation," IEEE Transactions on Neural Networks and Learning Systems, **29**(3), 731–737, 2018, doi:10.1109/TNNLS.2016.2636160.

[22] R. C. Barros, M. P. Basgalupp, A. C. P. L. F. de Carvalho, A. A. Freitas, "A Survey of Evolutionary Algorithms for Decision-Tree Induction," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), **42**(3), 291–312, 2012, doi:10.1109/TSMCC.2011.2157494.

[23] "UltraScale Architecture Configurable Logic Block User Guide," *https://www.xilinx.com/support/documentation/user_guides/ug574-ultrascale-clb.pdf*, accessed: 2019-08-06.

[24] D. B. Thomas, "Synthesisable recursion for C++ HLS tools," in 2016 IEEE 27th International Conference on Application-specific Systems, Architectures and Processors (ASAP), 91–98, 2016, doi:10.1109/ASAP.2016.7760777.

[25] J. R. Struharik, "Implementing decision trees in hardware," in 2011 IEEE 9th International Symposium on Intelligent Systems and Informatics, 41–46, 2011, doi:10.1109/SISY.2011.6034358.

[26] H. Azmi, R. Sayed, "FPGA-based Implementation of a Tree-based Classifier using HW-SW Co-design," 2019 6th International Conference on Advanced Control Circuits and Systems (ACCS) & 2019 5th International Conference on New Paradigms in Electronics & information Technology (PEIT), 224–228, 2019.

[27] "Xilinx MicroBlaze Processor," *https://www.xilinx.com/products/intellectual-property/microblazecore.html*, accessed: 2020-08-26.

[28] Y. Zhong, "The analysis of cases based on decision tree," in 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), 142–147, 2016, doi:10.1109/ICSESS.2016.7883035.

[29] R. Narayanan, D. Honbo, G. Memik, A. Choudhary, J. Zambreno, "An FPGA Implementation of Decision Tree Classification," in 2007 Design, Automation Test in Europe Conference Exhibition, 1–6, 2007.

[30] J. Mitéran, J. Matas, J. Dubois, E. Bourennane, "Automatic FPGA based implementation of a classification tree," in SCS 2004 Signaux, Circuits et Systèmes Application to Informations Systems, IT 03, Workshop on Information Systems, 2004.

[31] J. R. Struharik, "Implementing decision trees in hardware," in 2011 IEEE 9th International Symposium on Intelligent Systems and Informatics, 41–46, 2011.

[32] M. Barbareschi, S. Del Prete, F. Gargiulo, A. Mazzeo, C. Sansone, "Decision Tree-Based Multiple Classifier Systems: An FPGA Perspective," in F. Schwenker, F. Roli, J. Kittler, editors, Multiple Classifier Systems, 194–205, Springer International Publishing, Cham, 2015.

[33] T. Sharp, "Implementing Decision Trees and Forests on a GPU," in D. Forsyth, P. Torr, A. Zisserman, editors, Computer Vision – ECCV 2008, 595–608, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[34] D. Marron, A. Bifet, G. D. F. Morales, "Random Forests of Very Fast Decision Trees on GPU for Mining Evolving Big Data Streams," in Proceedings of the Twenty-First European Conference on Artificial Intelligence, ECAI'14, 615–620, IOS Press, NLD, 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 239-246 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Hybridization of Improved Binary Bat Algorithm for Optimizing Targeted Offers Problem in Direct Marketing Campaigns

Moulay Youssef Smaili[*,1], Hanaa Hachimi[1,2]

[1]*Systems Engineering Laboratory, ENSA, KENITRA, 14020, Morocco*

[2]*Systems Engineering Laboratory LGS, BOSS Team, Sultan Moulay Slimane University, Av Med 5, Po Box 591, Beni Mellal, Morocco*

ARTICLE INFO

Article history:
Received: 19 August, 2020
Accepted: 21 October, 2020
Online: 10 November, 2020

Keywords:
Binary Bat Algorithm
Genetic Algorithm
Hybridization
Low level Hybridization
Customer Targeting

ABSTRACT

One of the biggest business problems of marketers, is to optimize the return on investments of direct Marketing campaign. This main purpose which can only be ensured by targeting the appropriate customer. The main challenge faced by companies when advertising, is to configure properly a campaign, by choosing the right target, so A high user acceptance rate is ensured to advertisements. However, when dealing with an important size of data, the important specification to consider is the combinatorial aspect of the problem and the limitation of the approach based on mathematical programming methods. In this work, and considering the optimization of targeting offers as an of NP-hard problems, we concluded that the use of a meta heuristic algorithm is more suitable to use a classical (exact) method. We choice to use an improved bat Algorithm hybridized with Genetic Algorithm. The results of computational experiments confirmed that the proposed algorithm gives competitive results.

1. Introduction

There are two axes which characterize research in metaheuristic: exploration and exploitation. Exploration guarantees a diversified search in the space regions; however, the exploitation seeks to ensure the intensification of the search to find a better solution. The goal is to have a balance between these two contradictory objectives [1]. In the literature, many works make a difference between local search methods, instead of single solution-based and population-based methods. In this work, and unlike previous works, we deal with the targeted offers problem in direct marketing campaigns using population-based methods, and we focus on two algorithms; the Genetic Algorithm which can easily escape from the local minimum, and the Bat Algorithm (BA) which has the characteristic to quickly converge towards a local solution. Based on the imitation the concept of natural choice and heredity, the Genetic Algorithm is considered one of the most famous evolutionary search techniques. It was effective in optimizing several types of problem where a deterministic polynomial solution is infeasible. The Bat Algorithm, inspired by the foraging behavior of micro-bats, has been developed by Yang in 2010 [2], and it can be considered as an effort to gather positive side of previous metaheuristic algorithms. Thanks to its simplicity, flexibility, and easy implementation, the Bat Algorithm works well with complicated problems. The main feature of this algorithm is the quick convergence at a very initial stage by keeping a balance when switching from exploration to exploitation. In order to maximize their benefits, firms use "marketing campaigns" as a fundamental tool, along two main axes; by attracting new customers, so-called acquisition or by prospecting additional revenues with their existing customers by proposing new products to them; and this is called retention. Our approach focuses on the second axis (retention). To maximize the marketing profit to determine the combination: which product to offer and for which customer, under a set of constraints by using a new improved Bat Algorithm hybridized with Genetic Algorithm. The final goal is to choose the right product, to offer to the right customer, at the right time to maximize the marketing Return on Investment (ROI). This goal is not easy to implement because of the fact that firms have multiple products to offer under a complex set of business constraints [3]. In this work, and due to the discrete aspect of this kind of problems, the use of a binary version of the Bat Algorithm is essential. The standard BA was originally designed for the optimization of non-discrete problems [4]. The objective is to find a solution with a discrete value. The sought solution is represented by a matrix which contains binary values; 1 if product j is offered to customer i, and 0 otherwise. This paper is organized as follows,

[*]Corresponding Author: Moulay Youssef Smaili, 00212664349580, Email: smailli.moulayyoussef@uit.ac.ma

section 2 describes some related works in the same area; works about the targeted marketing problem are mentioned first, followed by works which process the Bat Algorithm's binary aspect in a discrete search space. In section 3, we describe the problem of Market Targeting and the modeling adopted to search a solution for the problem. In section 4, we describe two metaheuristic algorithms; namely the Binary Bat Algorithm (BBA), and the Genetic Algorithm. In this section we present our proposed algorithm: Improved Genetic Binary Bat Algorithm (IGBBA) developed to encounter the shortcomings of the two first algorithms while searching for adapted solutions. Experimental results are presented in section 5 for the three algorithms with a comparison of their performance.

In [3], author presented an optimization problem and replaced the classic database Marketing problem intended to generate the optimal profit under a set of business constraints. Approximately 2.5 million customers, eleven unique offers, and five investments are considered. The solution provided an approximately optimal solution to the ideal capacity assignment problem. The result is a decision describing which customer to target and which product to offer to him. In [5], author presented eight algorithms to solve the problem of maximizing the Return on Investment of targeted marketing. Seven of these algorithms are considered as exact methods and the eighth one is a heuristic method: the Tabu Search (TS). In the results, the TS showed the best performances. In [6], author took up the problem as described by author in [5], focusing on solving the problem by developing an efficient technique. The approach followed by author in [6] consists of hybridizing two metaheuristics: Greedy Randomized Adaptive Search Procedure (GRASP), and General Variable Neighborhood Search (GVNS). Three steps characterize the method followed by in [6]; the first one verified the efficiency of the GRASP algorithm proposed to initialize a good solution. In the second step, author kept the same parameters, but this time, they called the VND procedure for a local search. In the third step the complete proposed algorithm GGVNS was tested and it provided better results. In [7], author proposed a multi-objective DMLS metaheuristic for searching for sets of non-dominated solutions, in order to maximize both the promotion campaign's total profit and the risk-adjusted return. In order to remedy to the concept of volatility, author, in [7] proposed a greedy randomized initial solution builder, able to perform local searches considering different neighborhood exploration techniques, and a genetic method is designed. All the works cited above used metaheuristic based on local search methods to ensure an optimization of the return on investment, while respecting the business constraints. On the other side, the use of a Binary Bat Algorithm will allow us to deal with the target marketing optimization problem by looking for solutions made up of discrete (binary) values instead of looking for continuous solutions. In designing the Bat Algorithm in its binary version, how to update velocity and position must be adapted in order to guarantee the binary aspect of the research. In [8], the goal is to use the concept of velocity and position of bats in real space to move each bat in a discrete space with 0 and 1 values only. The idea was to change the position of micro-bats with the probability of their velocity. To achieve this objective, the technique used is based on a transfer function which allows a transformation of the value of the velocity to a probability, to update the position. this amounts to saying that this transfer function calculates the probability of changing the position from 0 to 1 and vice versa.

2. Problem Description

Direct Marketing is a tool that allows firms to promote their products directly to customers, and measure results quickly. One of the most important benefits of Direct Marketing is "Upgrading firm's loyalty strategies" in order to maximize the companies' Return on Investment. In [5], and In [3] before, authors presented the formulation of the product targeting problem as a mixed-integer programming (MIP) problem including more business constraints, and it can be rewritten as follows:

Given a set of m customers $C = [c_1, c_2 ... c_m]$, and a set of n products' offers $P = [p_1, p_2 \ldots p_n]$, the objective is to maximize the Return on Investment under these business constraints:

- The corporate hurdle rate: each company defines its hurdle rate (HR) to make sure that the Return on Investment is equal, at least, to a value of HR.
- During the campaign, the budget of each product is limited.
- A limitation is imposed on the total number of products offered to each customer.
- And there is also a Minimum Quantity Commitment (MQC), which is the number of units of a product to be offered in order for that product to be part of the campaign. It means that no customer will receive an offer of a product which is not part of the campaign. If the product belongs to the proposed ones then at least $Pj > 0$ customer will receive an offer.

A solution is represented by a binary array $R_{|C|\times|P|}$, where C indicates the set of available costumers, and P represents the possible products to be used in the campaign. If a given cell $s_{i,j}$ |i $\in$ C, j $\in$ P is equal to "1" (true), the product j will be offered to the customer i; otherwise, the value would be "0" (false).
There are two basic parameters of the customer lifetime value's computation; p_{ij} is the probability that customer i responds positively to an offer of product j, and DFV_{ij} is the return to the firm when customer i responds positively to the offer of product j [11, 12]. the problem can be modeled as follows:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(r_{ij}-c_{ij})x_{ij} - \sum_{j=1}^{m} fc_j y_j \qquad (1)$$

where:

- r_{ij} is the expected revenue of the company from the offer of product "j" to customer "i", and $r_{ij} = p_{ij} DFV_{ij}$
- c_{ij} is the cost generated while offering product "j" to customer "i"
- M_i is the upper limit value of number of different products to offer to a customer "i"
- P_j is the MQC limit associated with product "j"
- B_j is the budget, associated to the product "j" during campaign.
- fc_j is the cost of using product j in the campaign
- O_j is the MQC bound of the product j,
- HR is the hurdle rate specific to each company

The goal is to maximize the fitness function (1), by finding the optimal combination of the two matrices X and Y:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \text{ and } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

where:

$$x_{ij} = \begin{cases} 1 \text{ If product j is offered to customer i,} \\ 0 \text{ Otherwise} \end{cases}$$

and,

$$y_j = \begin{cases} 1 \text{ If product j is used in the campaign,} \\ 0 \text{ Otherwise} \end{cases}$$

The constraints, to which the problem is subject to, can be modeled as follows:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} r_{ij}x_{ij} \geq (1+HR)\left[\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{j=1}^{n} fc_j y_j\right] \quad (2)$$

$$\sum_{i=1}^{m} c_{ij}x_{ij} \leq B_j \quad \text{for each j = 1…, m} \quad (3)$$

$$\sum_{j=1}^{n} x_{ij} \leq M_i \quad \text{for each i = 1…, m} \quad (4)$$

$$\sum_{i=1}^{m} x_{ij} \leq m y_j \quad \text{for each j = 1…, n} \quad (5)$$

$$\sum_{i=1}^{m} x_{ij} \geq P_j y_j \quad \text{for each j = 1…, n} \quad (6)$$

$$x_{ij}, y_{ij} \in \{0,1\} \quad \text{for each i =1…m and j=1 …, n} \quad (7)$$

Figure 1 below, illustrates an example of a problem with an advertising campaign optimization. The matrix represents a solution $S_{|C|x|P|}$ concerning 5 products to offer to 10 customers, so as to optimize the Return on Investment by taking into account business constraints (maximum products to offer for each customer, fixed cost and budget for each product). In summary, this figure represents the offers to be adopted by each customer. The first column on the left, represents the 10 customers who are the subject of the campaign with the maximum product values to offer to each Customer. The first line at the top represents the different product offers with the fixed cost of each product and the budget allocated during the campaign. the values 1 means that product j is offered to customer i during the campaign and 0 if not.

3. Methodology

There are several types of algorithm hybridization; The hybridization used in this paper contains both sequential and parallel hybridization. The sequential one concern using GA in the initialization of a feasible solution. The parallel hybridization is used for perfecting the solution found in each BA iteration by an application of the GA limited in terms of number of iterations.

3.1. Building initial solution using Genetic Algorithm (GA):

One of the very important phases when dealing with the customer targeting problem, is the creation of an initial good quality solution which must be feasible by respecting the business constraints.

$$(P1 - P2 - P3 - P4 - P5)$$

$$\begin{pmatrix} M_1 = 3 \\ M_2 = 3 \\ M_3 = 2 \\ M_4 = 2 \\ M_5 = 2 \\ M_6 = 2 \\ M_7 = 3 \\ M_8 = 2 \\ M_9 = 2 \\ M_{10} = 3 \end{pmatrix} \begin{pmatrix} 1\ 0\ 0\ 1\ 1 \\ 0\ 0\ 1\ 0\ 0 \\ 1\ 1\ 0\ 0\ 1 \\ 1\ 0\ 1\ 0\ 1 \\ 0\ 1\ 0\ 0\ 0 \\ 1\ 0\ 0\ 1\ 0 \\ 0\ 1\ 1\ 0\ 1 \\ 1\ 0\ 1\ 1\ 1 \\ 1\ 0\ 0\ 0\ 1 \\ 1\ 0\ 1\ 0\ 1 \end{pmatrix}$$

Figure 1: A solution example for the optimization problem of customer targeting

The classical initialization of the Bat Algorithm, and which is based on randomization, creates a non-feasible or poor-quality initial solution. To solve this problem, we opted to build the initial solution using GA with a reduced population, and a number of reduced chromosomes. Genetic Algorithms are considered as the most known heuristic technique of search and optimization, and are inspired by natural evolution. They have proven to be effective in several complex problems [3]. GA modeled the decision variable of a search problem by a string composed of alphabets with a certain cardinality. Each chain is considered a candidate solution, and it is called a chromosome, while the alphabets that make up the chromosomes represent the genes. In the problem of marketing campaigns' optimization, a chromosome is considered as a string representation of solutions to the problem. In this kind of problems, the classical GA is imposed and represents the solutions by a bit string. Chromosomes composed of bit strings is a representation formed of genes with binary values allele. These values take 1 if a product is offered to the customer and 0 otherwise. The gene may represent a decision to offer a product j to a specific customer i. The final goal is to optimize the fitness function. It is a calculation which determines to what extent the chromosome can represent a solution to a given problem. GA is also characterized by the notion of population; It treats an artificial population; these are strings in a binary alphabet usually [9]. GA is composed of 6 steps to find an optimal solution for the problem; the objective in targeted offers problem is to maximize the return on investment by finding an optimal solution of equation (1):

1. Initialization

   A randomized procedure is usually used when generating the initial population of the candidate solution. However, if we randomly choose the composition of the chromosomes with a weighting of 50% for each value 0 and 1, we will obtain a non-feasible solution which do not respect the business constraints. For this, we have opted in this paper for a weighting distribution of 80% when a product j is offered to customer i, and 20% when it is not. With this method we have the certainty that our initial solution respects the business constraints but it will not necessarily represent an initial good quality solution.

2. Evaluation

   When the population is initialized first time (population initialized), the fitness values of each candidate solutions are evaluated to determine how fit an individual is.

3. Selection

   The selection is a step that keeps the best solutions by memorizing copies of those solutions with the highest fitness values. This concept imposes the mechanism of maintaining the most suitable among the candidate solutions

4. Crossover

   Crossover is the most impactful phase in GA. It is a phase which consists in combining two or more parent solutions in order to generate a better solution.

5. Mutation

   The GA ensures via this stage a diversity of the population in order to escape a local convergence, by modifying the solution in an arbitrary way while acting locally.

6. Termination

   The algorithm terminates if there is no improvement in the solution quality, when time is reached, or when an acceptable solution is obtained.

The following is a pseudo code of Genetic Algorithm (GA):

Algorithm 1: Genetic Algorithm

```
Randomly create initial population
Calculate the fitness function for all individuals.
while Iteration < MaxGeneration do
        Select the best-fit individuals for reproductions;
        Proceeding to offspring by execution crossover and mutation phases;
        Evaluating the fitness function of new chromosomes;
end while
Return the best solution in the population
```

Figure 2: Pseudo Code of Genetic Algorithm

3.2. Binary version of Bat Algorithm

The Bat Algorithm (BA) is one of the most recent metaheuristic algorithms which uses a technique of imitation of micro-bats behavior to find an optimal solution. Previous work has proven the effectiveness of this algorithm compared to old metaheuristics like Genetic Algorithm (GA) or the Particle Swarm Optimization [8,10]. However, the continuous solutions generated by the BA are not valid for certain types of problem. The objective of our paper is to find a discrete solution (1 when the product "j" is offered to the customer "i" and 0 otherwise). For that reason, in this paper, the use of a binary version of Bat Algorithm is imposed [4]. In order to adapt this algorithm to the context of the Targeted Offers problem in Direct Marketing Campaigns, a binary version of Bat Algorithm will be used and which is proposed by authors [8,10]. This version is called Binary Bat Algorithm (BBA). By modifying the micro-bats positions values from 0 to 1 and vice versa, the BBA moves micro-bats in a binary search space [8, 10]. In [11,2], When designing the BA, author respected the two main techniques used by microbats to find their prey. Micro bats decrease loudness and increase the pulse rate of sound emitted during chase. These characteristics have been modeled as follows [11,2]: Each micro-bat i is characterized by its position pos_i^t, and its velocity vel_i^t at a specific iteration (time). Exploring new solutions is performed by updating positions of micro-bats as following:

$$vel_i^{t+1} = vel_i^t + (x_i^t - Glob^{best})Freq_i \quad (8)$$

$$pos_i^{t+1} = pos_i^t + vel_i^{t+1} \quad (9)$$

$Glob^{bes}$ represents the best solution achieved until this iteration, and $Freq_i$, which represents the frequency of the i-th microbat, is obtained following this equation:

$$Freq_i = Freq_{\min} + (Freq_{\max} - Freq_{\min})\beta \quad (10)$$

where β is randomly generated from a uniform distribution in [0,1]. Thanks to these equations the exploitability aspect is ensured in the Bat Algorithm. However, the exploitation side is guaranteed by a random walk procedure as follows:

$$pos_{new} = pos_{old} + \varepsilon L^t \quad (11)$$

L^t is the loudness of the bats' emitted sound, in order to perform an exploration instead of exploitation. L^t is the average loudness of all the micro-bats at this iteration step: $L^t = \langle L_i^t \rangle$, and ε is a random number in [-1,1] [2,8,9].

To keep balance between exploration aspect and exploitation aspect during the search process, loudness L^t, and pulse emission rate r_i should be updated only if the candidate solution is improved as the iterations proceed. They are updated as following [2,10]:

$$L_i^t = \alpha L_i^t \quad (12)$$

$$r_i^t = r_i^0 \left[1 - \exp^{(-\gamma t)}\right] \quad (13)$$

where α and γ are predefined constants. The search principal in BA, is once a micro-bat has found the prey's location, the loudness will be decreased while the pulse rate will be increased. However, The Bat Algorithm is designed, at the origin, to search in a continuous space of solutions. The algorithm, in its original version is not applicable when the set of solutions is discrete, because during an optimization in a binary search space, by treating only values 0 and 1, the Eq. (9) is no longer valid to update the position of the microbats. To encounter this obstacle, authors in [8, 14] proposed a new solution: the idea is to use a transfer function with the aim of updating the position of micro-bats with the probability of their velocities. In practice the transfer function has the role of mapping velocity values to probability values. In other words, the transfer functions are used to move the microbats in a binary space instead of a discrete space. The most important factor for designing this function is to force micro-bats with high velocities to switch their positions. To do this, a transfer function, called v-shaped function, and position updating rule are proposed in [8,10], and described in Equations (14) and (15) as follows:

$$VEL(vel_i^k(t)) = \left| \frac{2}{\pi} \arctan\left(\frac{\pi}{2} vel_i^k(t) \right) \right| \quad (14)$$

$$pos_i^k(t+1) = \begin{cases} (pos_i^k(t))^{-1} & \text{if rand} < VEL(vel_i^k(t+1)) \\ pos_i^k(t) & \text{if rand} \geq VEL(vel_i^k(t+1)) \end{cases} \quad (15)$$

$pos_i^k(t)$ and $vel_i^k(t)$ are the position and velocity of microbat "i" at the iteration "t", where k is the problem dimension. $(pos_i^k(t))^{-1}$ is the complement of $pos_i^k(t)$. Figure 3 presents the general steps of the Bat Algorithm in its binary version.

3.3. Proposed Hybridized Bat Algorithm

While dealing with the problem of targeted marketing maximization, we note that the Genetic Algorithm takes a considerable time to find a solution but with a minimum chance of blocking in a local minimum, while the Bat Algorithm has the advantage of quickly converging on a solution to the problem. However, BA may easily be trapped in a local minimum. Taking into account the random aspect of the initialization procedures, the two algorithms suffer from the poor quality of the initial solution. To overcome these obstacles, we propose to hybridize the Binary Bat Algorithm, to benefit from convergence in advanced iterations, with the Genetic Algorithm that generates better population from good parents. These results are close to a global optimum.

Algorithm 2: Standard Binary Bat Algorithm

Initializing the population of micro-bats: POS_i(i=1, 2, ..., n)=rand(0 or 1) and VEL_i=0
Define initial pulse frequency value $Freq_i$
while (num_iteration < Max generation iterations) **do**
  Update both frequency and velocities values
  Determining transfer function value by Eq. (14)
  Update micro-bats' positions by executing Eq. (15)
  if (rand > r_i) **then**
    Randomly choice a solution ($Glob^{best}$) among the best solutions.;
    Modifying some of the dimensions of position vector with some of $Glob^{best}$ dimensions;
  end if
  Randomly flying and Generating a new solution
  if (rand < Li and f(posi) > f(Globbest) and Constraint(Globbest)=True) then
  Accept the new solutions;
  end if
  Evaluating the fitness function for all microbats in the population
  Increase r_i and reduce L_i
  Getting $Glob^{best}$ after ranking the micro-bats
end while
Return the best solution.

Figure 3: Pseudo Code of Standard Bat Algorithm

The initial solution in the proposed algorithm is generated by performing the Genetic Algorithm for the first time, with a reduced number of chromosomal parents; this will quickly generate an initial solution of better quality than standard algorithms. The exploration and exploitation are controlled by pulse emission rate, in the case of standard Binary Bat Algorithm [12,13]. By proceeding iterations, this factor takes on larger values, which causes a progressive loss of the exploitation capacity of the algorithm. by introducing linear decreasing inertia weight factor, the exploitation capability of the algorithm is improved while the exploration capability will be improved by the hybridization of both algorithms. Figure 4 presents the steps of the proposed algorithm:

Algorithm 3: Hybridized Binary Bat Algorithm

```
Randomly create GA initial population
Calculate the fitness function for all individuals
while Iteration < MaxGeneration do
      Select the best-fit individuals for reproductions;
      Proceeding to offspring by execution crossover and mutation phases;
      Evaluating the fitness function of new chromosomes;
end while
Return Initial solution to execute Binary Bat Algorithm
Initializing the population of micro-bats: POS_i(i=1, 2, ..., n)=rand(0 or 1) and VEL_i=0
Define initial pulse frequency value Freq_i
while (num_iteration < Max generation iterations) do
   Update both frequency and velocities values
   Determining transfer function value by Eq. (14)
   Update micro-bats' positions by executing Eq. (15)
   if (rand > r_i) then
      Randomly choice a solution (Glob^best) among the best solutions.;
      Modifying some of the dimensions of position vector with some of Glob^best dimensions;
   end if
  Randomly flying and Generating a new solution
  if (rand < Li and  f(posi) > f(Globbest) && Constraint(Globbest)=True) then
  Accept the new solutions;
  end if
  Evaluating the fitness function for all microbats in the population
  Increase r_i and reduce L_i
  Getting Glob^best after ranking the micro-bats
end while
Return the best solution.
```

Figure 4: Pseudo Code of Hybridized Genetic Binary BA (IGBBA)

4. Computational Experiments and Discussion

The proposed algorithm was developed in Matlab 2015on a Lenovo T470s computer Core i7, with 2.6 GHz clock speed and 20 GB RAM. The linear solution was obtained using CPLEX (IBM solution), and the dataset, proposed and generated by author in [5], is used for comparing the three algorithms:

- The cost value c_{ij} may take 1,2 or 3 (Randomly)
- The return of the firm r_{ij} is an integer generated between 0 and 16 randomly.
- The corporate hurdle rate HR can take 5%, 10%, and 15% values.
- In this paper, there are three different values of the number of customers: 100 customers for small category S, 1000 customers for medium category M, and 10,000 customers for large category L.
- For each category of customers, we have two different numbers of product n: 5 and 10 products.
- For each combination (category of customer and number of products), a minimum-quantity commitment bound P_j is randomly generated between $\left[\frac{\sum_i M_i}{n}\right]$ and $\left[2\frac{\sum_i M_i}{n}\right]$.
- The fixed cost fc_j is generated randomly between $\frac{O_j}{2m(1+HR)}\sum_i\left[r_{ij}-(1+HR)c_{ij}\right]$ and $\frac{O_j}{m(1+HR)}\sum_i\left[r_{ij}-(1+HR)c_{ij}\right]$
- For budget B_j, three values are adopted: the two values of $\left[O_j\frac{\sum_i c_{ij}}{m}\right]$ $\left[2\frac{\sum_i M_i}{n}\frac{\sum_i c_{ij}}{m}\right]$ and a random integer between $\left[O_j\frac{\sum_i c_{ij}}{m}\right]$ and $\left[2\frac{\sum_i c_{ij}}{n}\right]$.
- The upper bound M_i is selected between 1 and n/5.

To test the performance of the proposed algorithm, we report the same dataset and conditions from computational experiments in [5] and [6]. The GAP is calculated when a feasible solution is found by the adopted algorithm. By analyzing Table 1, we clearly deduce that the proposed IGBBA algorithm gave better solutions compared to GA and BBA.

The GAPs of IGBBA are significantly lower, and the BBA appears as the second-best algorithm by comparing the GAPs. Figure 5, which is a graphic representation of the fitness function using the three discussed algorithms, shows the fast convergence of the BBA and the IGBBA compared to the Genetic Algorithm. As reported in Table 1, and for the three categories of the S3 format, IGBBA is clearly better than the methods used in the previous works, with a clear difference in performance based on the GAP calculation. However, for the medium and large categories, the IGBBA is considered very competitive and sometimes better than the other algorithms.

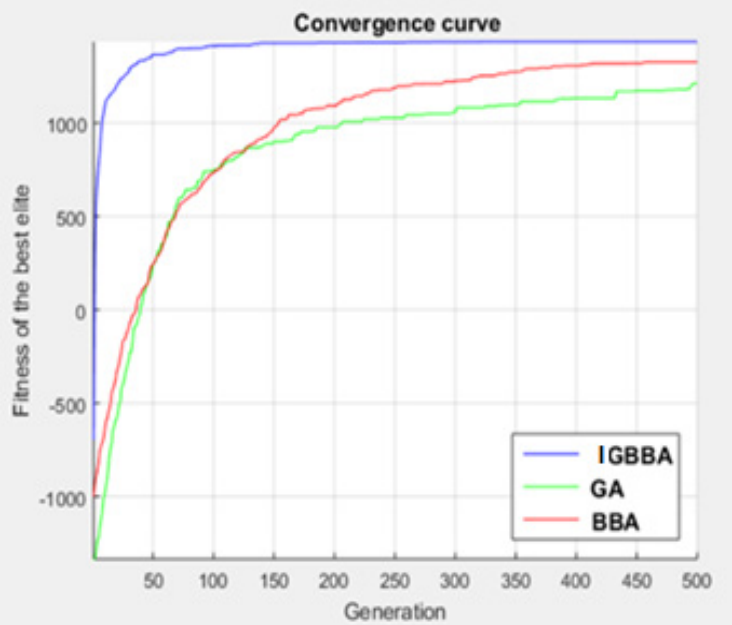


Figure 5: Convergence of fitness function

Table 1: Experiment results

	Small			Medium1			Medium2			Large		
	5	**10**	**15**	**5**	**10**	**15**	**5**	**10**	**15**	**5**	**10**	**15**
			Results experiments of [5]									
Initial Solution	19.2	17.85	18.34	17.9	16.2	18.36	18.92	19.23	17.84	26.262	24.866	25.655
Tabu Search	6.86	6.52	7.76	7.22	8.54	7.6	9.75	**9.58**	9.11	10.866	11.04	10.23
γ			**Results experiments of [6]**									
0	10.68	6.98	7.77	9.72	9.12	12.15	11.29	9.9	11.15	12.999	12.29	12.15
0.2	6.9	7.14	7.92	7.22	8.67	7.69	10.26	9.59	9.34	11.224	10.94	10.74
0.4	6.79	6.45	7.58	8.47	9.1	7.89	9.68	9.76	9.91	10.866	11.22	10.11
0.6	6.77	**6.5**	7.49	7.24	**8.42**	7.63	9.85	9.78	9.44	11.511	11.93	10.22
			IGBBA results									
Initial Solution (GA)	23.8	19.05	19.75	20.1	18.8	17.21	19.45	22.88	18.66	27.332	25.667	26.31
IGBBA	**6.69**	6.51	**7.01**	**7.11**	8.55	**7.12**	**9.45**	9.77	**9.07**	**10.866**	**11.01**	**10.11**

5. Conclusion

The BA which is one of the most efficient algorithms was appointed to deal with that type of problem; it was used in binary version given the discrete nature of the solution space. In order to be able to search in a discrete space, a transformation of the positions of the micro-bats, to discrete values, is essential. The hybridization can provide a more efficient behavior and a higher performance when dealing with large-scale problems. This combination carried out in this paper has shown its effectiveness by remedying the disadvantage of the risk of jamming in a minimum local. As extension of our work, new constraints can be proposed, like reducing products costs where customer respond positively to a big number of products or where the number of customers respond positively to a specific product j. On the other hand, an optimization of IGBBA algorithm can be proposed to reduce the execution time. We addressed a very important issue in the marketing strategy of companies in order to optimize their return on investment during advertising campaigns. We have taken up the modeling done by author in [5] with the business constraints that must be respected to maximize the profit of the company. To test its proposal, author, in [5], created a dataset randomly, respecting a certain logic of the upper and lower limits of each parameter; a job that will produce several dataset categories according to different sizes depending on the number of customers, number of products and the value of Hurdle rate. A dataset on which this paper was based to test our proposal and compare it with the results of [5] and the result of other works. Since the problem is considered an NP-hard one, the use of metaheuristics is required; we have treated the problem based on methods based-population differently to the works cited above and which are all based on neighborhood methods.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] R. Faia, T. Pinto, Z. Vale, J.M. Corchado, "Hybrid particle swarm optimization of electricity market participation portfolio," 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings, 2018-January (703689), 1–8, 2018, doi:10.1109/SSCI.2017.8285218.

[2] X.S. Yang, "A new metaheuristic Bat-inspired Algorithm," Studies in Computational Intelligence, **284**, 65–74, 2010, doi:10.1007/978-3-642-12538-6_6.

[3] M.D. Cohen, "Exploiting response models - Optimizing cross-sell and up-sell opportunities in banking," Information Systems, **29**(4), 327–341, 2004, doi:10.1016/j.is.2003.08.001.

[4] M.Y. Smaili, H. Hachimi, "Metaheuristic for optimizing targeted offers in Direct Marketing," 6th International Conference on Optimization and Applications, ICOA 2020 - Proceedings, 4–8, 2020, doi:10.1109/ICOA49421.2020.9094454.

[5] F. Talla Nobibon, R. Leus, F.C.R. Spieksma, "Optimization models for targeted offers in direct marketing: Exact and heuristic algorithms," European Journal of Operational Research, **210**(3), 670–683, 2011, doi:10.1016/j.ejor.2010.10.019.

[6] T.A. Oliveira, V.N. Coelho, M.J.F. Souza, D.L.T. Boava, F. Boava, I.M. Coelho, B.N. Coelho, "A hybrid variable neighborhood search algorithm for targeted offers in direct marketing," Electronic Notes in Discrete Mathematics, **47**, 205–212, 2015, doi:10.1016/j.endm.2014.11.027.

[7] V.N. Coelho, T.A. Oliveira, I.M. Coelho, B.N. Coelho, P.J. Fleming, F.G. Guimarães, H. Ramalhinho, M.J.F. Souza, E.G. Talbi, T. Lust, "Generic Pareto local search metaheuristic for optimization of targeted offers in a bi-objective direct marketing campaign," Computers and Operations Research, **78**, 578–587, 2017, doi:10.1016/j.cor.2016.09.008.

[8] S. Mirjalili, S.M. Mirjalili, X.S. Yang, "Binary bat algorithm," Neural Computing and Applications, **25**(3–4), 663–681, 2014, doi:10.1007/s00521-013-1525-5.

[9] J. McCall, "Genetic algorithms for modelling and optimisation," Journal of Computational and Applied Mathematics, **184**(1), 205–222, 2005, doi:10.1016/j.cam.2004.07.034.

[10] R.M. Rizk-Allah, A.E. Hassanien, "New binary bat algorithm for solving 0–1 knapsack problem," Complex & Intelligent Systems, **4**(1), 31–53, 2018, doi:10.1007/s40747-017-0050-z.

[11] S.S. Guo, J.S. Wang, X.X. Ma, "Improved Bat Algorithm Based on Multipopulation Strategy of Island Model for Solving Global Function Optimization Problem," Computational Intelligence and Neuroscience, **2019**, 2019, doi:10.1155/2019/6068743.

[12] X. Huang, X. Zeng, R. Han, "Dynamic Inertia Weight Binary Bat Algorithm with Neighborhood Search," Computational Intelligence and Neuroscience, **2017**, 2017, doi:10.1155/2017/3235720.

[13] S. Yilmaz, E.U. Kucuksille, "Improved Bat Algorithm (IBA) on Continuous Optimization Problems," Lecture Notes on Software Engineering, **1**(3), 279–283, 2013, doi:10.7763/lnse.2013.v1.61.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 241-247 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Estimation of Land Degradation Loss by Water Erosion: Case of the Site of Biological and Ecological Interest of Ain Asmama (Western High Atlas, Morocco)

Adnane Labbaci[*,1], Said Moukrim[2], Said Lahssini[3], Said Laaribya[4], Hicham Mharzi Alaoui[5], Jamal Hallam[6,*]

[1]*Geoscience and Geoenvironment Laboratory, Department of Geology, Faculty of Sciences Ibn Zohr University, Agadir, 80000, Morocco*

[2]*BIOBIO Research Center, Faculty of Sciences, Mohammed V University, Rabat, 1000, Morocco*

[3]*Forest Inventory Department, National School of Forest Engineer, Salé, 10000, Morocco*

[4]*Department of Geography, Faculty of humanities and social sciences Chouaib Doukkali University, El Jadida, 24000, Morocco*

[5]*Fire Forest Service, High Commission on Water Forest and Combating Against Desertification, Rabat, 10000, Morocco*

[6]*National Institute of Agricultural Research of Morocco (INRA), Avenue des FAR B.P. 124 Inezgane, Agadir, 1000, Morocco*

ARTICLE INFO

Article history:
Received: 18 December, 2020
Accepted: 19 May, 2021
Online: 23 May, 2021

Keywords:
SBEI of Ain Asmama
Water erosion
RUSLE
Map of erosion
Loss

ABSTRACT

Erosion affects large parts of Moroccan land, particularly in the mountains leading to soil quality deterioration and less vegetation cover. Located in the South-west of Morocco, the Site of Biological and Ecological Interest (SBEI) of Ain Asmama, where erosion threatens a major part of the region was investigated. The site has a terraced transitional bioclimate, between arid to sub-humid, of local conservation importance. The varied, dense, continuous, and well-preserved vegetation of the area is crucial to protect the soils against erosion. Qualitative observations show that soils are increasingly degraded, water erosion is developing, and sediments accumulate in dams and ponds. In this study we have used the Revised Universal Soil Loss Equation (RUSLE) to assess the erosion risk in this area. It helped to develop a synthetic map of erosion susceptibility. Our results show that the integration of the different parameters in the water erosion process estimated the loss of soil amounting 339,03 tons/ha/year over the whole site. This is equivalent to a value of soil lowering of 2,82 cm which is considered extremely high.

1. Introduction

Water erosion is an extremely complicated natural phenomenon. It has serious economic, human, and ecological consequences. Economic losses are estimated in millions of Dollars and are mainly associated to downslope damage to property, roads and other infrastructure, and destruction of plant cover trough grazing affecting the most productive agricultural and forest lands [1–3]. People, including women and youth, in the rural area whose livelihoods depend on natural resources for subsistence, are vulnerable to land degradation [4]. In this area, land degradation by water erosion leads to a persistent impoverishment of the populations, which causes their exodus towards cities.

Land degradation by water erosion impacts extend to ecosystems. It has significantly altered ecosystems functions and services, decreased the biological diversity of soils and forest, and reduced agricultural lands productivity [5].

Estimates of soil losses due to water erosion in Morocco, range from 500 to > 5,000 tons / km^2 / year depending on the region. Accumulation of sediments in natural and artificial waterbodies (siltation of dams and ponds) is estimated to average 75 million m^3, equivalent to an annual reduction of 0.5% in the storage capacity of dams. In addition to the degradation of the quality of the mobilized drinking water, the storage capacity decline leads to a significant loss of water that would irrigate 10,000 ha / year [6]. In the Souss-Massa region, the cost of erosion has been estimated overall at 20.7 million DH; the core contributions to this cost are

*Corresponding Author: Adnane Labbaci, Email: a.labbaci@uiz.ac.ma,
*Jamal Hallam, Contributed equally with first author

https://dx.doi.org/10.25046/aj060326

associated to dam reservoirs (76.5%), upstream of the basin (22.4%), and to the Souss-Massa plain (1.12%) [7].

This study forms part of a larger project which evaluates land degradation at the Sites of Biological and Ecological Interest (SBEI) at national level. The aim of the study detailed here was to estimate loss of soil by water erosion in the SBEI of Ain Asmama, an important site for nature conservation in the Souss-Massa region. The outputs would provide useful data to assist in the design of adequate facilities based on multi-criteria analysis using GIS and remote sensing.

2. Study area

The SBEI of Ain Asmama is located at the western High Atlas of Morocco (between 9.03° and 9.33° East Longitude and 30.74° and 31° North Latitude) on the upper parts of the Ait Moussa river right bank. The total site area is 23 564 ha (Figure. 1).

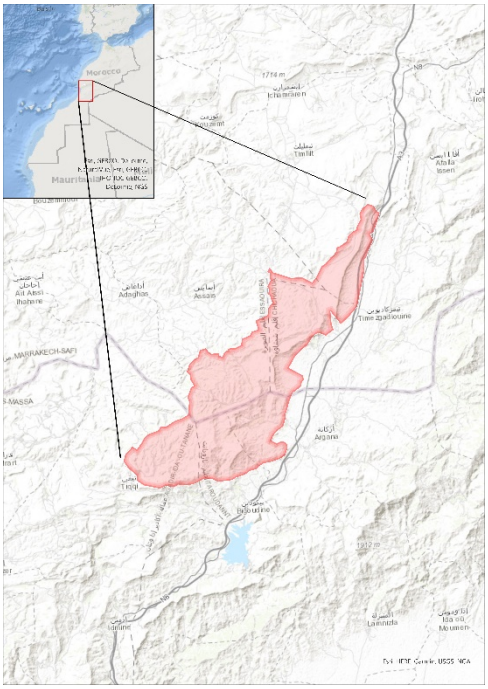


Figure 1: Location map of the SBEI of Ain Asmama

Storied stands plateaus dominate the SBEI. They decrease gradually at altitude, from 1744 m to 780 m, toward the west and north. Argana corridor is a valley depression of the Issen river bordering the SBEI of Ain Asmama. The geological massif of the area contains deep valleys, canyons, sheer cliffs, and steep relief cut by short ravines and giving a rugged hilly feature with an eventful and picturesque appearance.

From the geological view, we have found the domination of the Jurassic and Permo-Triassic classes, with the presence of Quaternary terracing of local alluvium cones and colluvium deposits of the recent piedmonts (Figure. 2).

The climate of the area is dominated by the arid and the sub-humid [8]. The annual rainfall average varies from 240 mm to 470 mm and the highest recorded rainfall falls between November and February. The site is characterized by a diverse vegetation including endemic trees of the region such as Argan, Thuija, Holm oak and Juniper.

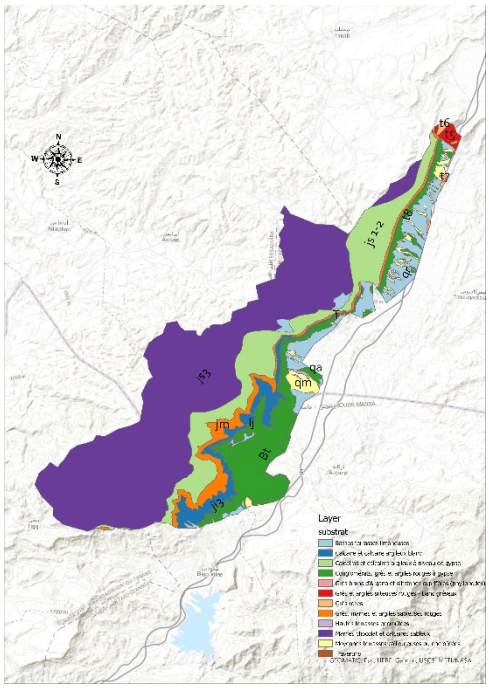


Figure 2: Geological classes' map of the Ain Asmama's SBEI

3. Materials and methods

Soil loss rates were estimated using the RUSLE factors were integrated into the GIS following the adjusted model to Moroccan conditions [9,10]:

$$A = R * K * LS * C * P \quad (1)$$

where:

A = average soil loss (ton/ha/year)

R = rainfall-runoff erosivity factor (MJ.mm/ha.h.year)

K = soil erodibility factor (t.ha.h/ha.MJ.mm)

L = slope length factor (m)

S = slope steepness factor (degree)

C = cover-management factor

P = support practice factor.

3.1. Rainfall erosivity factor (R)

The erosive potential of raindrops impact is determined by the R factor. Several studies quantifying erosion in Morocco, to estimate the R factor we have used the Arnoldus model [11]:

$$R = 1.735*10*(11.5 * \log \Sigma (pi2 /p) - 0.8188) \quad (2)$$

where

R: Rainfall erosivity factor (MJ.mm/ha.h.year)

Pi: Average monthly (mm), and

P: mean annual precipitation (mm).

To allow comparison with previous studies, the R factor was estimated using the same model. Monthly and annual rainfall data

was computed for each weather station within the SBEI to calculate the R index using.

3.2. Soil erodibility factor (K)

The K factor is the intrinsic susceptibility of a soil to resist to climatic aggression. Given that the soil units map of SIBE does not exist, we developed a homogeneous distribution map of lithological units. the lithological map assumes that soil units are intricately linked to the parent material e of the underlying geological formations. The map of lithological formations was in turn assigned to different homogeneous soil units. A field survey was then carried out to collect soils samples within each homogeneous soil unit. The organic matter and soil textures were measured in the laboratory and the K factor was computed using the following formula:

$$100*K = 2.1*10\text{-}4*M1.14\ (12\text{-}a) + 3.25\ (b\text{-}2) + 2.5\ (c\text{-}3) \quad (3)$$

where

K: soil erodibility factor (t.ha.MJ-1.mm-1)

M = (% sand + silt) * (100 -% clay)

a =% organic matter,

b = code of permeability,

c = code of the structure

3.3. Slope length (L) and steepness (S) factors

The slope length (L) and the slope steepness (S) factors give us an overview of the effect of topography on soil erosion. LS factor estimation was much easier and more accurate when using digital elevation model (DEM) in GIS. To evaluate the LS factor, we have used the following equation [12]:

$$LS= (L/22.15)\ m\ *\ (65.41*Sin^2(S) + 4.56\ *\ Sin(S) + 0.065) \quad (4)$$

where

L = Slope length
S = Slope degree
m = factor dependent on the value of S:
m = 0.5 if S > 5.0%
m = 0.4 if 3.5<= S =< 4.5%
m = 0.3 if 1.0 <= S =< 3.0%
m = 0.2 if S = 1.0%

The slope length was determined using the Horton (1945) model that relates the slope steepness and length:

$$B = S/2L \quad (5)$$

where

B: The average of the watershed slope length (m);

S: The catchment area (m^2) and

L: The length of the river (m).

3.4. Land cover factor (C)

The estimation of the land cover factor takes in consideration the soil degradation's conditions, the nature of the land cover (presence of a different type of vegetation) allowed a slow runoff and a better infiltration [13]. The C factor is of a significant importance in RUSEL equation and was estimated out of a developed land use map using Sentinel-2 satellite imagery, missions 2019 [14].

3.5. Support practice factor (P)

The (P) factor takes in consideration the different management related to the conservation measures on water and soil. In this study, we identified the cultivation practices at the SBEI for which the conservative value was assumed to be 0,50 [15].

4. Results and discussions

Erosion is a natural process that occurs worldwide and cannot be stopped [16]. The estimation of soil loss by erosion aims, however, to regulate human effects to maintain tolerable soil loss. For the very same purpose we mapped the water erosion of the Ain Asmama SBEI to help locate the impacts of the different influencing factors on the soil to manage them so that the erosion rate is within an acceptable range. In what follows, we present the results of each factor of the RUSLE equation and then the estimated soil loss of the watershed. Since there is no established threshold between acceptable and unacceptable erosion levels [16], we will convert our final estimation results to values of soil lowering [17] to verify the severity of erosion in the area. We will then discuss the loss of soil thickness relative to the literature.

4.1. Rainfall erosivity factor (R)

Rainfall data series from 1982 to 2019 (37 years) were collected for eight weather station within the SBEI. The data were used to estimate the Rainfall erosivity factor R following the Arnoldus method [11] (Table 1).

Table 1: Weather stations locations within the Ain Asmama SBEI's and the estimated Rainfall erosivity factor R

Station	X Easting (m)	Y Northing (m)	P (mm)	R
Argana	143 624	425 520	349.66	74.04
Ain Asmama	127 963	426 663	473.72	83.00
Imouzzer ida outanane	108 989	414 756	540.78	96.37
Amsoul	148 975	432 400	214.45	46.27
Aguenza	140 550	422 600	252.78	58.12
Bge Abdelmoumen	139 379	411 720	413.25	88.88
Tizguine	144 302	441 023	236.63	50.31
Iloudjane	175 648	469 482	332.15	54.89

The estimated R values (Table 1) within the SBEI range between 46,27 and 96,37 units / year. A linear interpolation model was built in ArcGIS to accurately differentiate between the climate zones around the area of a selected weather station. All the estimated Rainfall erosivity were then adjusted to the weather station altitude to account for elevations effect. Figure 3 shows the resulting R factor map.

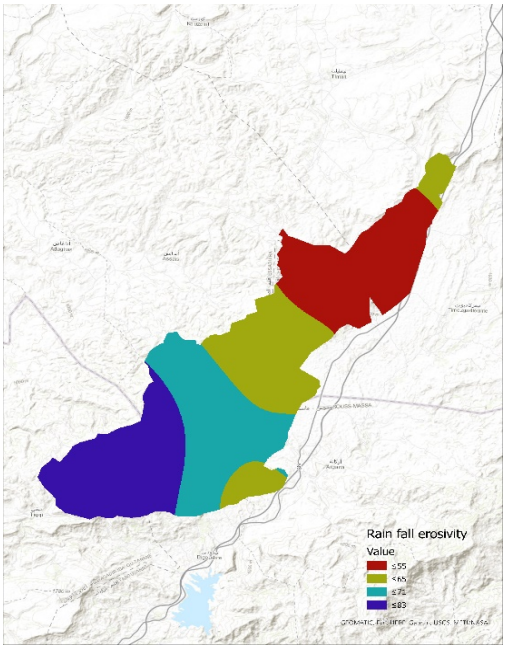


Figure 3: Rainfall erosivity factor (R) of the SBEI of Ain Asmama

4.2. Slope length (L) Slope and Slop steepness (S)

A topographic contour lines of a scanned and georeferenced standard topographic map of 1:50 000 were digitized to generate generated a digital elevation model (DEM). This resulted in an actual spot elevation of each selected point of the grid. The resulting data shows an elevation amplitude between 780 m to 1740 m.

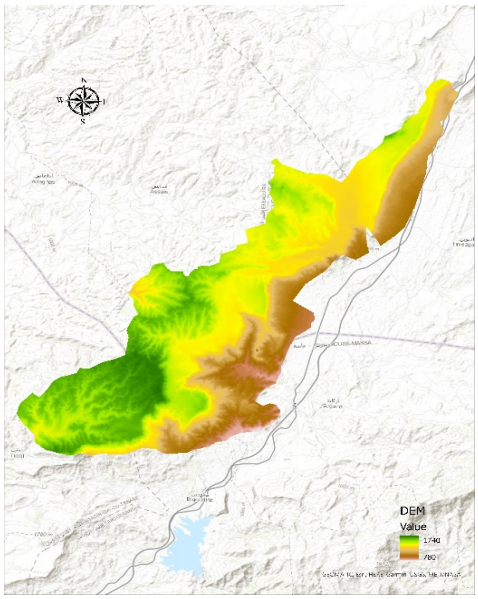


Figure 4: Digital Elevation Model (DEM) map of the SBEI of Ain Asmama

The processing of the developed DEM showed a varying topography within the SBEI of Ain Asmama. The extracted slopes were between 0 and 74 degrees. The reclassification of the DEM resulted in 5 slope classes (Table 2). The slope classes, 0 to 5 degrees, suitable for agronomic land use were only 29.25% of the total area. Thirty eight percent of the area has a slope greater than 15 % suggesting an extremely high vulnerability to water erosion. Our results show the percentages of steep slopes are large and the can increase the impact of the erosion in our study area which will in turn have an impact on the average soil loss downstream of the basin.

Table 2: Slope classes obtained from the reclassification of the Digital Elevation Model (DEM) of the SBEI of Ain Asmama

Slope in degree	Area (ha)	Percentage (%)
0 - 2	3 913.66	16.61
2 - 5	2 978.28	12.64
5 - 10	4 474.67	18.99
10 - 15	3 260.92	13.84
15 - 74	8 936.70	37.92
Total	**23 564,23**	**100.00**

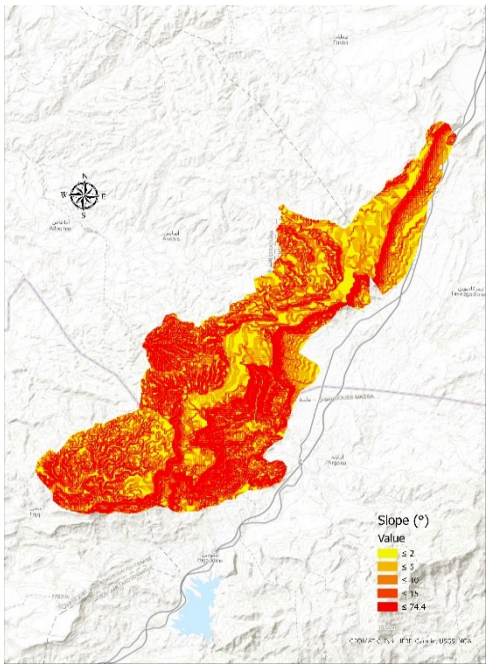


Figure 5: Slope map of the SBEI of Ain Asmama (classes values are in degree)

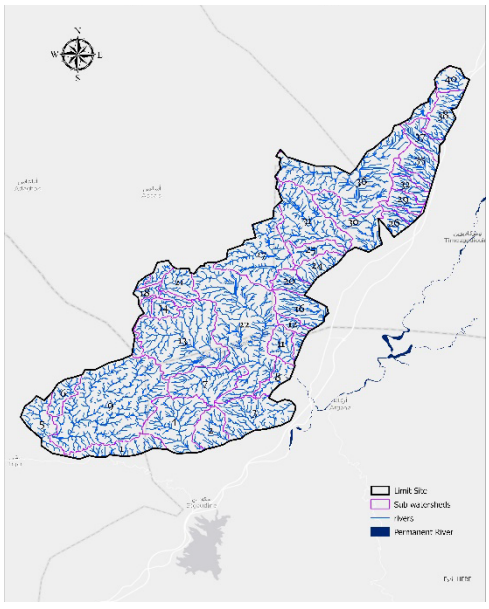


Figure 6: Sub watersheds map of the SBEI of Ain Asmama

The Hydrology tools were applied in GIS to create a stream network of 40 sub-basins within the watershed. The slope length (L) was measured for each delineated sub basin using Horton method. The L values were between 45 and 183 m and the average length was101.95 m.

The estimated LS values were between 0 to 201 units 66% of which are higher than 30 units. This % coverage is in accordance with the classification of Manrique [18].

Table 3: LS factor classes and their %coverage in the Ain Asmama SBEI

LS classes	% coverage
0 - 5	20.01
5 - 15	07.05
15 - 20	02.70
20 - 30	04.52
> 30	65.71

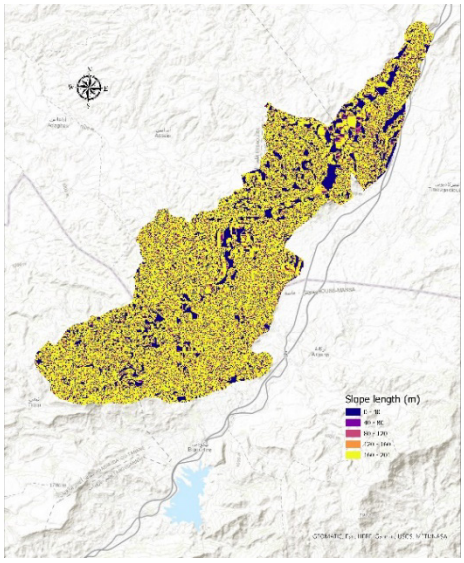


Figure 7: Map of the combined slope steepness and Slope length factor (SL) of the SBEI of Ain Asmama

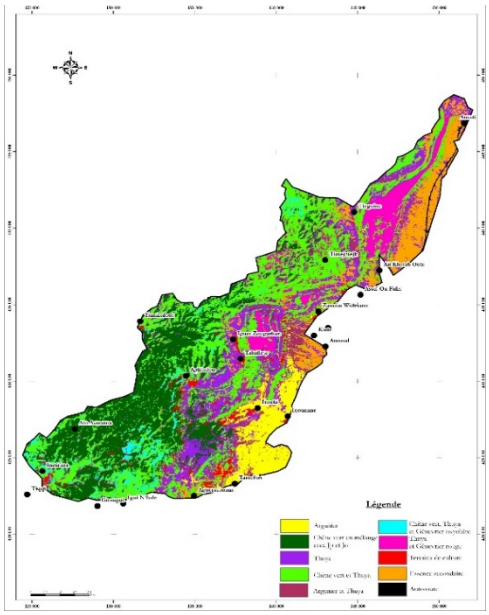


Figure 8: Land cover map of the SBEI of Ain Asmama

4.3. Land cover factor (C)

Supervised classification of land cover using Sentinel-2 images brought out the following groupings: Crop lands, Matorral, Argan trees, Thuija, Juniper, Holm oakand Wasteland.

Uncovered soil (bare soil with low canopy and residue) is likely to be exposed to water erosion. Combining the NDVI map together with the 2019 land cover gives an estimation of density of vegetation. Using this outcome in addition to the land cover map (Figure 8) allowed identification of five land vegetation classes (Table 4) for which C factor values were assigned.

Table 4: Land cover classes and Values of land cover classes in the site of Ain Asmama

Vegetation classes	C factor value
Dense forest	0.01
Clear forest	0.10
Matorrals	0.16
Cultivated land	0.50
Wasteland	1.00

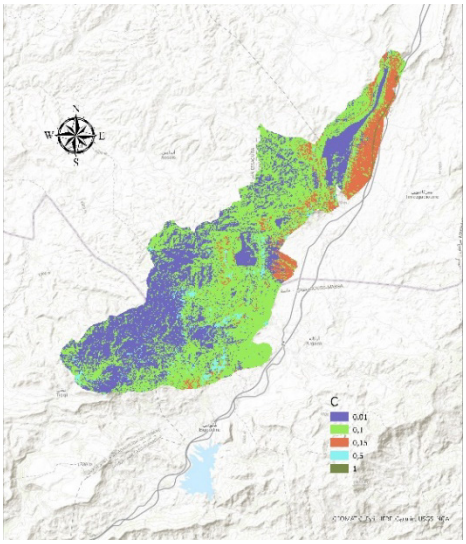


Figure 9: Land cover factor map of the SBEI of Ain Asmama

4.4. Soil erodibility factor (K)

Soil erodibility factor depends on several parameters such as cohesion strength between soil particles, soil and plant roots interaction, plants exudates, biological organisms (e.g., earthworms) and soil structure and its relation to soil moisture content [19]. Following the classification of Manrique [18] our results show that 95 % of the soils are highly to very highly erodible (Figure 10), in accordance.

Table 5: Soil erodibility factor (K) classes at Ain Asmama's site and the equivalent risk of erodibility following the classification of Manrique

K factor classes	Risk of erodibility	Area (ha)	Percentage coverage
K < 0.15	Very little erodible soils	0	0.00
0.15 < K < 0.25	Slightly erodible soils	474.29	02.01

0.25 < K < 0.35	Moderately erodible soils	762.98	03.24
0.35 <K < 0.45	Highly erodible soils	11 225.79	47.64
0.45 < K < 0.60	Very highly erodible soils	11 101.17	47.11

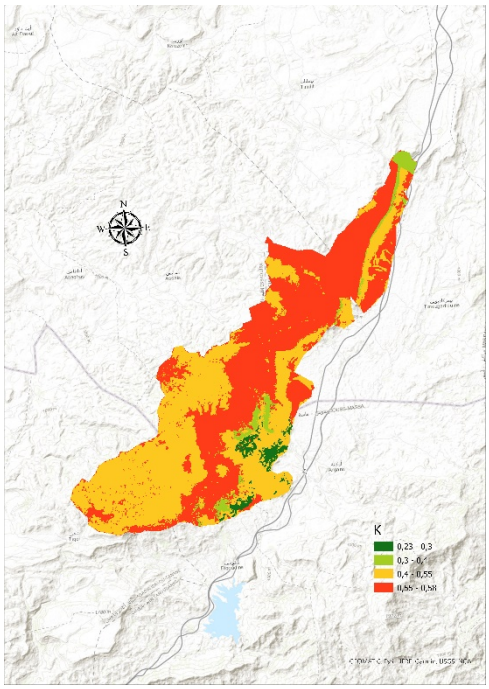


Figure 10: Soil erodibility factor (K) map of the SBEI of Ain Asmama

4.5. Support practice factor (P)

The factor P shows the effect of water and soil management and conservation measures on erosion. Qualitative observation made whilst identifying the land management practices of the SBEI indicate that 0.5 is the suitable P factor value to use for the soil loss assessment.

4.6. Soil loss rate estimation (A)

The overlapping of the thematic maps of RUSEL model factors, using GIS, has led to the establishment of the erosion map. This was in turn reclassified to produce the losses in soils by water erosion (tons / ha / year) (Figure. 11) of the SBEI of Ain Asmama: The estimated losses in soil regularly displaced downstream were between 0.1 and 339 tons/ha/year.

This range of rates of erosion alone is hard to interpret in terms of erosion severity and is meaningless because of the lack of an established threshold between acceptable and unacceptable. levels of erosion [16]. In the 1950s, the USDA considered values of soil-loss of ≈5 ‑ 12 tons/ha/year as tolerance values to evaluate rates of soil erosion [20]. Researchers have expressed concern that the suggested values by USDA pose uncertainty on how to interpret the differences between modern and geological erosion rates [17]. The estimate of soil erosion in previous studies and its interpretation is very dependent on 1) the scale of the study (From laboratory measures, Field-plots compilation, sub-basins to catchments estimation), 2) the approaches adopted to generate the estimates (Rainfall simulation, Splash cups, Stereo photography, RUSLE, USLE, etc), 3) and the aptitude of a soil to maintain its functions to provide ecosystem services [21].

To contrast the severity of erosion in the SBEI to previous studies using typical ranges of erosion, we converted our outputs to values of soil lowering [17]. For the conversion we used the standard soil bulk density of 1200 kg m^{-3} as suggested by Montgomery [17] in his widely cited paper. The results show that the soil in the SBEI is lowering by up to 2.82 cm per year, which is considered as an extremely high loss of soil. This soil lowering value is on average hundreds of orders of magnitude greater than the estimated soil lowering values in most of the previous literature summarized in [16], [17]. It is also several times greater than estimates of mean global soil production (from 0.06 to 0.08 mm per year) [22, 23].

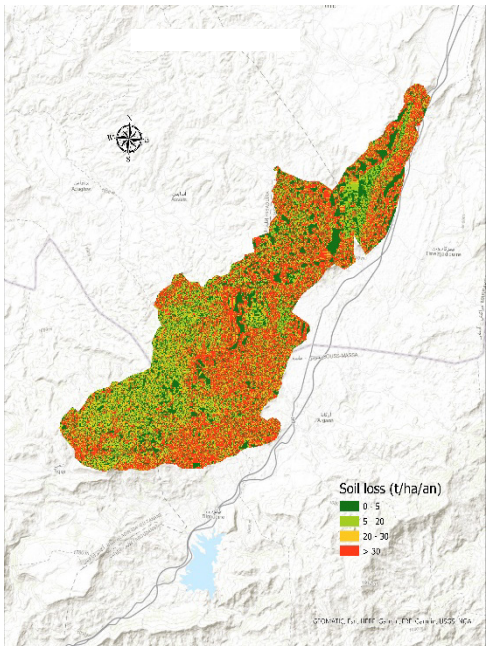


Figure 11: Soil loss map of the SBEI of Ain Asmama

5. Conclusion

This note presents the results of the hydric erosion assessment of the SBEI of Ain Asmama. We first established and implemented a cartographic database for each of the factors of the RUSLE model. which were in turn associated to thematic maps using GIS.

The outputs show that the rainfall an aggressiveness was up to 96.37 units / year. Most of the soils of the region are alluvium and colluvium depositions and weakly developed soils due to erosion on red clay. These soils are poorly permeable and prone to streaming development which in turn keeps the ravines actively widening. The estimated erodibility values were up to 0.58. The values of the slope length and steepness LS vary from 0 to 201. Land use map was established after a supervised classification of the Sentinel imagery associated to the field survey C factors were then assigned to each land cover class identified.

The thematic maps of each factor of the RUSLE equation; were overlapped to give the water erosion map [24]. The results show that the soil in the SBEI is lowering by up to 2.82 cm per

year, several times greater than estimates of average global soil production

References

[1] J. Boardman, "A short history of muddy floods," Land Degradation & Development, **21**(4), 303–309, 2010, doi:10.1002/ldr.1007.

[2] R. Evans, J. Boardman, "The new assessment of soil loss by water erosion in Europe. Panagos P. et al., 2015 Environmental Science & Policy 54, 438–447—A response," Environmental Science & Policy, **58**, 11–15, 2016, doi:10.1016/j.envsci.2015.12.013.

[3] P. Panagos, P. Borrelli, J. Poesen, C. Ballabio, E. Lugato, K. Meusburger, L. Montanarella, C. Alewell, "The new assessment of soil loss by water erosion in Europe," Environmental Science & Policy, **54**, 438–447, 2015, doi:10.1016/j.envsci.2015.08.012.

[4] L. Olsson, H. Barbosa, A. Cowie, K. Delusca, D. Flores-Renteria, K. Hermans, E. Jobbagy, W. Kurz, D. Li, D.J. Sonwa, L. Stringer, Land Degradation. In Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, Intergovernmental Panel on Climate Change (IPCC), 2019.

[5] R. Parveen, U. Kumar, "Integrated approach of universal soil loss equation (USLE) and geographical information system (GIS) for soil loss risk assessment in Upper South Koel Basin, Jharkhand," **4**(6), 9, 2012, doi:10.4236/jgis.2012.46061.

[6] S. El Khatri, T. El hairech, "Drought conditions and management strategies in Morocco," Direction de La Météorologie Nationale, Casablanca, 7.

[7] F. Elame, M.R. Doukkali, A. Fadlaoui, "La gestion intégrée de l'eau à l'échelle du bassin de Souss-Massa: développement d'un modèle intégré de Bassin," Al Awamia127, 91–111, 2013.

[8] S. Moukrim, S. Lahssini, M. Rhazi, A.H. Mharzi, A. Benabou, I. Wahby, M. El Madihi, M. Arahou, L. Rhazi, "Climate change impacts on potential distribution of multipurpose agro-forestry species: Argania spinosa (L.) Skeels as case study," Agroforestry Systems, **93**(4), 1209–1219, 2019, doi:10.1007/s10457-018-0232-8.

[9] R. Kalman, The Climatic Factor of Erosion in the Sebou Watershed Morocco. Sebou Project, 32 p, 1967.

[10] S. Schmidt, C. Alewell, K. Meusburger, "Monthly RUSLE soil erosion risk of Swiss grasslands," Journal of Maps, **15**(2), 247–256, 2019, doi:10.1080/17445647.2019.1585980.

[11] H.M.J. Arnoldus, "Methodology used to determine the maximum potential average annual soil loss due to sheet and rill erosion in Morocco," 1977.

[12] W.H. Wischmeier, D.D. Smith, Predicting rainfall erosion losses: a guide to conservation planning, Department of Agriculture, Science and Education

[13] S. El Khatri, T. El hairech, "Drought conditions and management strategies in Morocco," Direction de La Météorologie Nationale, Casablanca, 7.

[14] A.Y. Yesuph, A.B. Dagnew, "Soil erosion mapping and severity analysis based on RUSLE model and local perception in the Beshillo Catchment of the Blue Nile Basin, Ethiopia," Environmental Systems Research, **8**(1), 17, 2019, doi:10.1186/s40068-019-0145-1.

[15] A. Gayen, S. Saha, H.R. Pourghasemi, "Soil erosion assessment using RUSLE model and its validation by FR probability model," Geocarto International, **35**(15), 1750–1768, 2019, doi:10.1080/10106049.2019.1581272.

[16] D.J. Pennock, Soil erosion: The greatest challenge for sustainable soil management, Food and Agriculture Organization of the United Nations, Rome, Italy, 2019.

[17] D.R. Montgomery, "Soil erosion and agricultural sustainability," Proceedings of the National Academy of Sciences, **104**(33), 13268–13272, 2007.

[18] L.A. Manrique, Land erodibility assessment methodology (LEAM): using soil survey data based on soil taxonomy, Editorial and Publication Shop, Honolulu, Hawaii, 1988.

[19] J. Hallam, M.E. Hodson, "Impact of different earthworm ecotypes on water stable aggregates and soil water holding capacity," Biology and Fertility of Soils, **56**(5), 607–617, 2020, doi:10.1007/s00374-020-01432-5.

[20] D.L. Schertz, "The basis for soil loss tolerances," Journal of Soil and Water Conservation, **38**(1), 10–14, 1983.

[21] FAO, Revised World Soil Charter, Rome, Italy, 2015.

[22] F.R. Troeh, J.A. Hobbs, R.L. Donahue, Soil and water conservation for productivity and environmental protection., Prentice-Hall, Inc., New jersey, USA, 1980.

[23] T. Wakatsuki, A. Rasyidin, "Rates of weathering and soil formation," Geoderma, **52**(3–4), 251–263, 1992, doi:10.1016/0016-7061(92)90040-E.

[24] V. Prasannakumar, H. Vijith, S. Abinod, N. Geetha, "Estimation of soil erosion risk within a small mountainous sub-watershed in Kerala, India, using Revised Universal Soil Loss Equation (RUSLE) and geo-information technology," Geoscience Frontiers, **3**(2), 209–215, 2012, doi:10.1016/j.gsf.2011.11.003.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 248-256 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Parametric Study for the Design of a Neutron Radiography Camera-Based Detector System

Evens Tebogo Moraba[*,1,2], Tranos Zuva[3], Chunling Du[2], Deon Marais[1]

[1]*Department of Radiation sciences, The South African Nuclear Energy Corporation SOC. Limited, Pretoria, 0001, RSA*

[2]*Department of Computer Systems Engineering, Tshwane University of Technology, Pretoria, 0001, RSA*

[3]*Department of Information Technology, Vaal University of Technology, Vanderbijlpark, 1911, RSA*

ARTICLE INFO

Article history:
Received: 10 December, 2020
Accepted: 24 March, 2021
Online: 27 May, 2021

Keywords:
Camera-Based-Detector System
Field-of-View
Neuron Radiography
Spatial Resolution

ABSTRACT

The field of view (FOV) and spatial resolution (SR) are the major image quality parameters which are being optimized in neutron radiography (NR) technique. This requires effective components selection during the design of NR detector systems. The selection is a discouraging task owing to often having conflicting experimental requirements and related constraints. In this work, models were developed to study the relationship between detector system components. These allow the simulation of detector system components whilst taking cognizance of specific experimental requirements and constraints in order to aid the design. Various commercial detector system components combinations were simulated to evaluate their performance. Results were benchmarked with the result from secondary data. 100% agreement between these data demonstrated the accuracy of the models, allowing purposeful selection. The ≈90% negative correlation between SR and FOV was identified as a tradeoff between these variables. Currently selected best combination offers a monotonic SR range of 25.5 – 170.92 µm pixel size, over FOV range of 52.3 × 52.3 mm² – 350 × 350 mm². The results also show that components can be selected for design of desired detector system within constraints to manage the field of view effectively; thereby optimizing the SR.

1. Introduction

The demand for high spatial resolution (SR) in neutron radiography (NR) facilities has grown with the need to investigate components comprising smaller features for scientific and industrial applications. This necessitates NR facilities to develop their detection methods as for maximal SR resulting in increased image quality. SR as a key parameter determines the imaging performance of detector systems[1, 2]. This describes the degree which two features of the object can be distinguished or separated in the acquired image [3], as well as the ratio of the image sensor size to that of the field-of-view (FOV) [4]. The availability of a simple and efficient method to manage the FOV in optimizing the SR is of high importance for NR, especially for camera-based detector systems (CBDS), as most NR installations incorporate a variable FOV using this type of the detector [1, 2, 5, 6]. The optimum combination of a large FOV and high SR is a great technical challenge for neutron detectors [1, 2, 7] due to the trade-offs between the two. Sophisticated high resolution detectors are ideal for small specimens; however, they are not appropriate for experiments involving larger specimens owing to their small, fixed FOV [2, 8]. This is especially true where image stitching cannot be exploited for example in time dependent investigations such as liquid flow experiments. This requires detector development to meet the demands for various types of experiments that are performed using the NR technique [5, 9].

CBDS comprises a camera, lens and scintillator as basic components [2, 8]. They utilize prime lenses as part of the optics [4, 10]. These lenses are of great value in this type of detector system as their optical performance is usually much better than that of zoom lenses. Prime lenses are optimized for a single fixed FOV and most often have a wide maximum aperture setting which enable much faster shutter speeds and improved low-light functionality when compared to their zoom counterparts [4]. By employing a light tight detector box together with properly aligned components enables them to preserve the light illuminated from the back of the scintillator to yield high quality results [2, 5]. Prime lenses are significantly sharper than zoom lenses as they do not have additional internal elements that translate in order to enlarge

[*]Corresponding Author: Evens Tebogo Moraba, evens.moraba@necsa.co.za

https://dx.doi.org/10.25046/aj060327

the image [11]. Diffraction increases with the number of internal lens elements as in the case of zoom lenses, resulting in degraded optical quality (N. Mansurov, Prime vs Zoom Lenses - A Beginners Guide, Photography Life, 2011). Finally, prime lenses cost relatively less [2] than zoom lenses for such reasons as having fewer internal components.

The use of prime lenses requires that the object distance (OD) be varied to increase or decrease the FOV size [4]. This can be achieved by either moving camera or the object plane (scintillator surface in this case) on a linear path corresponding to optical path. The latter is however not possible due to a fixed scintillator configuration in conventional CBDS. Translating the camera is therefore the only viable option to change the OD without exchanging components [4]. This technique enables a variable FOV[7] within a fixed range which is limited by factors such as space available in experimental area, scintillator size and neutron beam dimensions. The diagram in Figure 1 shows motorized CBDS intrinsic components to demonstrate the FOV management concept.

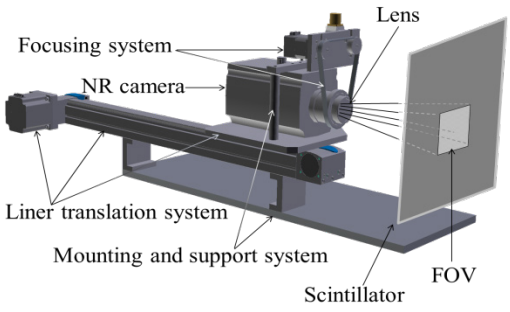


Figure 1: FOV management concept with a motorized camera translation

The other available method for consideration in changing the FOV is interchanging similar lenses of different focal lengths [5,12]. A different focal length will introduce a different angle-of-view (AOV) and therefore a different FOV. The number of achievable FOVs with this method is however limited by the number of lens units available and monotonic FOV variation across a desired range cannot be achieved. Methods such as these are expensive as they involve many components. The exchange process is time consuming and may also damage the fragile scientific components as a result of mishandling.

CBDS as used in NR requires careful selection of appropriate components during design based on established specific experimental requirements and constraints. Conventional component selection, configuration and calibration often lead to sub-optimal use with respect to the components maximum potential, limiting the ability to maximize SR.

The selection of components for design of CBDS can be a discouraging task, especially when constraints such as space available in the NR facility chamber and budget must be adhered to [13]. These together with experimental requirements involving investigations of specimens with various sizes and time dependent examinations, makes a choice of appropriate components combination for design of CBDS inevitable [2,4,12]. The detector system in NR facilities can be viewed as a portable instrument that can be employed at different NR facilities for experimental purposes [4]; therefore its technical requirements can be formulated independently of characteristics of the beam and other external factors [4]. In this research, pragmatic mathematical models were developed to study parameters of the detector components and their relationship. Arbitrarily selected components combinations were used to evaluate the models. Combinations which ensure versatility with respect to investigations of specimens with various sizes, and still obtain the best possible SR, are presented. The trade-offs that must be considered when designing a CBDS under given constrains and experimental requirements are also described.

2. Material and Methods

The mathematical models were developed deriving from the detector system's optical system represented in Figure 2. Different inherent properties of similar detector system components results in different responses during experiments. The inherent properties are referred to as primary or independent parameters in this work and are defined by parameter values as specified by the manufacturer of the specific component. Selection of components allows these parameters to be varied to produce different results. Variables that are being manipulated using primary parameters are referred to in this work as secondary parameters. These performance-related variables are used in the design of the CBDS in order to satisfy specific requirements and are a function of the primary parameters. Their values are derived from mathematical models which were developed based on the CBDS's optical system and geometric arrangements of components. This allows parametric studies to be performed in order to establish relationships between parameters.

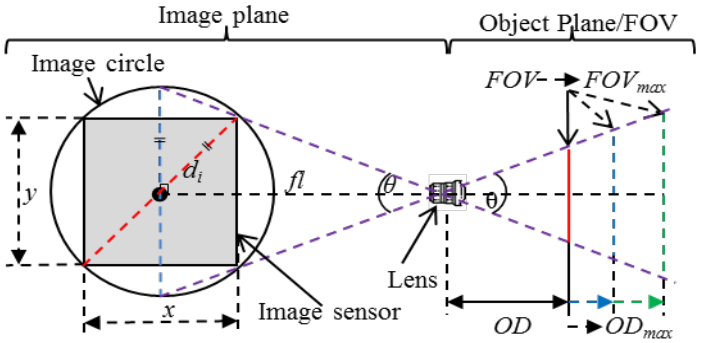


Figure 2: Schematic of the detector system's components geometrical arrangement and optical system

2.1 Material

Similar but different detector system components including scintillators, cameras and lenses were chosen arbitrarily to undertake a comparative parametric study through simulation in order to enable selection of components for the design of a suitable detector system based on requirements and constraints.

The following provides details for each category of components selected for simulation:

- *Cameras*

Three cameras with vastly different properties were selected from the Andor products range for comparison. Manufacturer's specifications show that connectivity is through a USB interface and all three feature a standard bayonet F-mount which enables lenses to be exchanged easily (Andor Technology, Belfast, Northern Ireland). The selected cameras are described below.

The Andor iKon-L 936 CCD-camera has a 2048 × 2048 array, aspect ratio of 1.0 and 13.5 µm pixel pitch which delivers a 27.6 × 27.6 mm^2 active image area. It features a 5-stage cooler which enables its large area sensor to be cooled down to -100 °C. This makes the iKon-L camera suitable for long exposure times and is ideal for low-light applications such as radiography

The Andor Zyla 5.5-HF front-illuminated CMOS-camera with a 2560 × 2160 array and 6.5 µm pixel pitch delivers a 21.8 mm diagonal FOV with a 16.6 × 14 mm^2 active image area giving an aspect ratio of 1.185. Its operating temperatures are from 0 ºC to 30 ºC ambient.

The Andor iKon-XL 230 model with back-illuminated CCD sensor offers a very large 61.4 × 61.4 mm^2 imaging area with a 4096 × 4108 array format, aspect ratio of 0.997 and 15 µm pixel size. The chip is thermoelectrically cooled down to -100 °C, and is ideal for applications such as Astronomy or X-ray/Neutron radiography.

- *Lenses*

Table 1 shows a summary of the four lenses selected from Nikon products and their properties from datasheets (Nikkor, Nikon Inc., Melville, NY, USA).

Table 1: Summary of Selected Lenses and their Properties

Lens	Focal Length	AOV	MFD
PC-E Micro NIKKOR f/2.8D ED	45 mm	51.4º	250 mm
Nikon AF-S NIKKOR f/1.4G	50 mm	46.8º	450 mm
Micro-NIKKOR f/2.8	55 mm	42.9º	250 mm
AF-S Micro NIKKOR F2.8G ED	60 mm	39.7º	185 mm

- *Scintillators*

Two, high light (blue emission) output, high resolution scintillators, namely a ^{6}LiF: ZnS Large Scintillator (LS) and a ZnS (Ag) -^{6}LiF Medium Scintillator (MS) were adopted for the object plane. The type, active size and the aspect ratio of these scintillators are shown in Table 2.

Table 2: Summary of Selected Scintillators for Simulation and their Properties

Scintillator	Type	Size	Aspect ratio
Larger	^{6}LiF: ZnS	400 × 400 mm^2	1
Medium	ZnS(Ag)-^{6}LiF	350 × 350 mm^2	1

3.1 Secondary Parameters Modeling

Figure 2 show that, the relationship between the image circle and the image sensor can be represented by the theorem of Pythagoras in equation (1).

$$d_i^2 = x^2 + y^2 \tag{1}$$

The sensor aspect ratio (δ) is the ratio of sensor width (x) to height (y) as given in the following equation as: height (y) as given in the following equation as:

$$\delta = x/y \tag{2}$$

Also in Figure 2, the focal length (fl) and AOV (θ) are fixed owing to the use of prime lenses. A change in FOV can therefore be achieved by varying the OD or by exchanging similar lens of different focal lengths; thus change in AOV. The latter cannot easily achieve monotonic FOVs similar to when a single lens is in use with a motorized translation systems. This leaves the change in OD as the only appropriate method to change the FOV.

Calculation of the FOV is established by considering the similar triangles theorems, namely the vertical opposite angles and the angle-angle similarity theorem. These indicate that the object plane, as observed through a lens, has the same properties as the image plane of the optical system. As the OD can be changed, various FOVs can be achieved by varying the OD as shown in Figure 2, producing an instantaneous FOV (IFOV). This is calculated by the width (FOV_w) and the height (FOV_h) at a given OD and taking the aspect ratio of the image sensor into account. The length of the rectangle representing the IFOV's diagonal line is the actual FOV at the sensor position, and is calculated according to Pythagoras' theorem as in equation (3).

$$FOV = \sqrt{FOV_w^2 + FOV_h^2} \tag{3}$$

The physical size of the FOV is defined using the OD and the lens AOV using a trigonometric equation (4), as follows:

$$FOV = 2 \times tan(\theta/2) \times OD \tag{4}$$

The maximum FOV (FOV_{max}) is achieved at the position where the maximum OD (OD_{max}) is realized, and vice versa. OD_{max} is given by (5).

$$OD_{max} = FOV_{max}/2 \times tan(\theta/2) \tag{5}$$

The camera image sensor properties such as aspect ratio; play a major role in defining how the scintillator surface would be imaged, and to what extents the image can be captured and resolved. By disregarding the neutron beam size, the maximum FOV (FOV_{max}) the detector system can image is equal to the size of the maximum rectangle subjected to the sensor aspect ratio that can fit into the area of the scintillator defined by the width (W) and the height (H). The aspect ratios of both the scintillator and the image sensor must match, for the sensor to image the entire scintillator at FOV_{max} as shown in Figure 3(a).

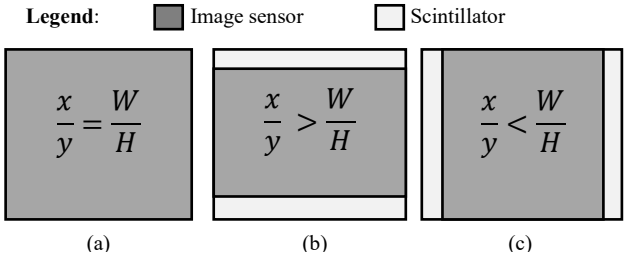


Figure 3: Effect of image sensor-scintillator aspect ratio matching at maximum FOV (a) Equal aspect ratios, (b) Large image sensor aspect ratio, causing vertical mismatch (c) Large scintillator aspect ratio, causing horizontal mismatch

If this is not the case, the mismatch in aspect ratios will result in the area of the scintillator not being imaged entirely at FOV_{max} [5], this phenomenon is illustrated in Figure 3 (b) and (c).

The mismatch as shown in the figures will cause either the horizontal side or vertical side to saturate first, curtailing any further increase in FOV, giving FOV_{max}. The following equations (6) – (8), based on Pythagoras' theorem were conditionally used to

calculate the FOV_{max}, given the scintillator and the image sensor aspect ratios:

$$FOV_{max} = \sqrt{W^2 + H^2} \quad ; \text{if } x/y = W/H \tag{6}$$

$$FOV_{max} = \sqrt{W^2 + (W/\delta)^2} \quad ; \text{if } x/y > W/H \tag{7}$$

$$FOV_{max} = \sqrt{(\delta H)^2 + H^2} \quad ; \text{if } x/y < W/H \tag{8}$$

The minimum focus distance (MFD) is a primary parameter which describes the minimum distance between the lens and the object plane, from which the lens is able to focus. This limits the extent to which the lens can be closest to the object plane, thereby defining the minimum OD (OD_{min}). The FOV_{min} the detector system can capture is calculated from this value by substituting the OD_{max} with OD_{min} and FOV_{max} with FOV_{min} in (5), as in (9).

$$FOV_{min} = 2 \times tan(\theta/2) \times OD_{min} \tag{9}$$

The FOV is changed primarily to achieve high level of SR. A change in FOV however changes the surface area of the object plane which the camera can image, thereby causing the area that the image sensor can resolve using a given lens to vary. This is equivalent to varying a SR. SR can be calculated either from the ratio of the sensor size to the FOV size, giving SR in line pairs per millimeter (lp/mm), or from the ratio of the FOV size to the sensor size, giving SR in micrometers (μm). This can be achieved using equations (10), and (11), respectively:

$$SR = FOV/CCD_{size} = FOV/d_i \; (lp/mm) \tag{10}$$

$$SR = d_i/FOV \; (\mu m) \tag{11}$$

3. Results and Discussion

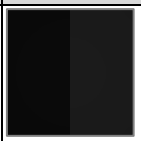
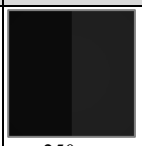
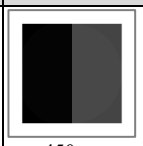
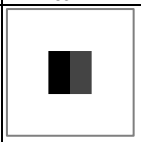
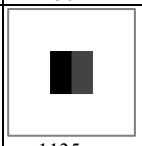
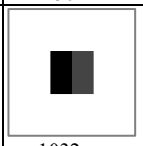
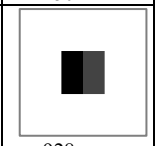
This section discusses the benchmarking and evaluation of the results from simulation of models developed. To begin with, the Radiography Simulator software which is designed to simulate a radiograph of particular shape or type of material of the sample was used. Firstly, the simulation of radiographs with components combination selected from section 2.1 was carried out with Radiography Simulator in [14] to generate the evaluation datasets. The generated datset is then analysed for correctness of parameters representing the selected simulation setup. The results are presented to contrast with the results from models simulation.

Secondly, the models are then simulated and the results are benchmarked against the Radiography Simulator results. The capability of the models to produce desirable results is then assessed, and lastly the results are used to select the best combination of components for a particular envisaged experimental requirement and constraint.

4.1 Models Benchmark Samples and Analysis Software

Secondary data were generated using Radiography Simulator to evaluate the success of the designed models. The dataset comprised of a simulated 12 cm × 1 cm × 12 cm ($l \times b \times h$) cube specimen made from two materials half each. The parameters used for simulation were from combination of the following detector system components specified in section 2.1: Andor iKon-L 936 CCD camera and MS in combination with the four lenses. Dataset of radiographs from MFD to the FOVmax were produced for each lens. Table 3 show a sample of data from each lens dataset. For each lens shown on the table, the first row consists of image at minimum OD with their respective OD position, second row consists of image at maximum OD also with position and the third row is the number of images generated per lens.

Table 3: Properties of images generated by Radiography Simulator

Lens	PC-E Micro NIKKOR f/2.8D ED 60mm	Nikon AF-S NIKKOR f/1.4G 55mm	Micro-NIKKOR f/2.8 50mm	AF-S Micro NIKKOR F2.8G ED 45mm
OD_{min}	 185 mm	 250 mm	 450 mm	 250 mm
OD_{max}	 1238 mm	 1135 mm	 1032 mm	 928 mm
Images	1053	885	582	678
Total Images				3198

ImageJ open source imaging software, which is freely available to the scientific community [15] was used to analyse the simulated images and measure the sample size, FOV_{min}, FOV_{max} and SR in each of the dataset in Table 3 as shown in Figure 4.

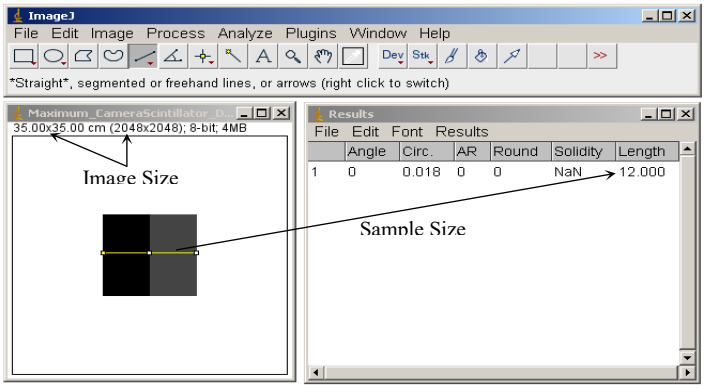


Figure 4: Radiograph measurement using ImageJ image analysis software

The results from ImageJ are tabled in Table 4, and they demonstrated that the simulated sample and components parameters selected are intrinsic to the simulated images.

Table 4: FOV range and SR range analyzed using ImageJ software

Lens	60mm PC-E Micro NIKKOR f/2.8D ED	55mm AF-S NIKKOR f/1.4G	50mm Micro-NIKKOR f/2.8	45mm AF-S Micro NIKKOR F2.8G ED
FOV range (mm)	73.9 – 495	109.0 – 495	216.1 – 495	133.5 – 495
SR range (µm)	25.5 – 170.9	37.6 – 70.9	74.6 – 170.9	46.1 – 170.9
SR range (lp/mm)	39.2 – 6.4	26.6 – 6.4	13.4 – 6.4	21.7 – 6.4

5.1 Simulation

Simulation of detector system components parameters for the design was undertaken using models developed in section 3.1. The results would aid in selecting the optimal combination of detector system components, constrained by experimental requirements. The data from physical detector system component listed in 2.1 were used in components combinations. The derived models were implemented in a spread sheet and the results are presented in the following subsections.

- *Field-of-View Area Saturation Determination*

To determine the FOV saturation (FOV_{max}) for a particular, the camera sensor and the scintillator aspect ratios were matched to determine the saturation size. This was realized using the aspect ratio matching for a camera sensor and scintillator. This was repeated the for combinations of the three Andor cameras and two scintillators. The aspect ratio matching was accomplished througt one of the fitting equation in (6) – (8) in each case. The results are given in Table 5. Column 1 shows the camera model; and column 2 and 4 indicate the match or type of mismatch for MS and LS respectively. The FOV_{max} with dimensions indicated in size by $W \times H$ for each combination is shown in column 3 and 5 for MS and LS respectively.

- *Object Distance Determination for a Variable Field-of-View*

The ODmax from which the FOV_{max} is derived is calculated from equation (5). The OD range data-points were simulated from zero (0) mm to OD_{max} with fixed predefined OD shift intervals of 1mm. In normalizing the data-points, the lens MFD was taken into account and all ODs below this point were discarded, staying on the useful OD range. The MFD and the resulting OD_{max}, are the OD range limits in each case shown in column 4 and 7 for MS and LS respectively.

Table 5: Aspect Ratio Matching for Sensor-Scintillator Combination with the Field-of-View Saturation

Camera name; Aspect Ratio	Medium Scintillator (LS) - (400 x 400 mm²)		Larger Scintillator (MS) - (350 x 350 mm²)	
	Sensor Scintillator Match	*FOV Size Saturation (W ×H)* mm²	*Sensor Scintillator Match*	*FOV Size Saturation (W×H)*mm²
Zyla 5.5 HF ;1.185	$x/y > W/H$	400 × 337	$x/y > W/H$	350 × 295
iKon-L 936 ;1.0	$x/y = W/H$	400 × 400	$x/y = W/H$	350 × 350
iKon-XL230 ;0.997	$x/y < W/H$	399 × 400	$x/y < W/H$	349 × 350

In evaluating the developed models, the parameters' values from benchmark dataset agree 100% with the Camera-Lens-Scintillator simulation using models in the result on Table 6 for Andor iKon-L 936 CCD camera and MS in combination with the four lenses.

The plots on Figure 5 show the OD range per camera using all lens-scintillator combinations. From the OD ranges determined, the variable FOV would then be realized by utilizing equation (4), and varying the *OD* within the OD range, resulting in a FOV ranging from the position of *MFD* where FOV_{min} is realized, to the position of OD_{max} were FOV_{max} is realized. Data points for FOV were simulated by ussuming a 1mm shift intervals for *OD*. The achieved FOV ranges for each combination are shown in Table 6 column 5 for MS and column 8 for LS. The evaluation dataset for Andor iKon-L 936 camera combined with the PC-E Micro NIKKOR f/2.8D ED 60mm lens match the result from models. The FOV change behavior was plotted over the OD range in Figure 5 for each lens-scintillator combination using different cameras shown in Table 6 column 1. MS was used in Figure 6 (a), (b), and (c) and LS in (d), (e) and (f).

Table 6: Simulation results for combinations of Camera-Lens-Scintillator

Andor Camera; Aspect Ratio(δ)	Lens focal length	Maximum Spatial Resolution (lp/mm; µm)	Medium scintillator			Larger scintillator		
			Object Distance Range (mm)	*Field-of-View Range (mm)*	*Minimum Spatial Resolution (lp/mm; µm)*	*Object Distance Range (mm)*	*Field-of-View Range (mm)*	*Minimum Spatial Resolution (lp/mm; µm)*
Zyla 5.5 HF; 1.185	45 mm	22.9;43.7	250-782	146.4-458	5.85; 136.71	250-894	146.4-523	5.12;156.29
	50 mm	14.1;70.7	450-870	236.9-458	5.85; 136.71	450-994	236.9-523	5.12;156.29
	55 mm	28.0;35.6	250-958	119.5-458	5.85; 136.71	250-1094	119.5-523	5.12;156.29
	60 mm	41.3;24.2	185-1045	81.1- 458	5.85; 136.71	185-1194	81.1-523	5.12;156.29
iKon-L 936; 1.0	45 mm	21.7;46.1	250-927	133.5-495	7.31; 170.92	250-1059	133.5-566	6.40;195.26
	50 mm	13.4;74.6	450-1031	216.1-495	7.31; 170.92	450-1178	216.1-566	6.40;195.26
	55 mm	26.6;37.6	250-1135	109.0-495	7.31; 170.92	250-1297	109.0-566	6.40;195.26
	60 mm	39.2;25.5	185-1238	73.9-495	7.31; 170.92	185-1415	73.9-566	6.40;195.26
iKon-XL 230; 0.997	45 mm	23.0;43.5	250-927	133.3-494	11.74; 85.21	250-1059	133.3-565	10.27;97.35
	50 mm	37.2;26.9	45 -1031	215.8-494	11.74; 85.21	450-1178	215.8-565	10.27;97.35
	55 mm	18.8;53.3	250-1135	108.9-494	11.74; 85.21	250-1296	108.9-565	10.27;97.35
	60 mm	12.7;78.6	185-1238	73.8-494	11.74; 85.21	18 -1416	73.8-565	10.27; 97.35

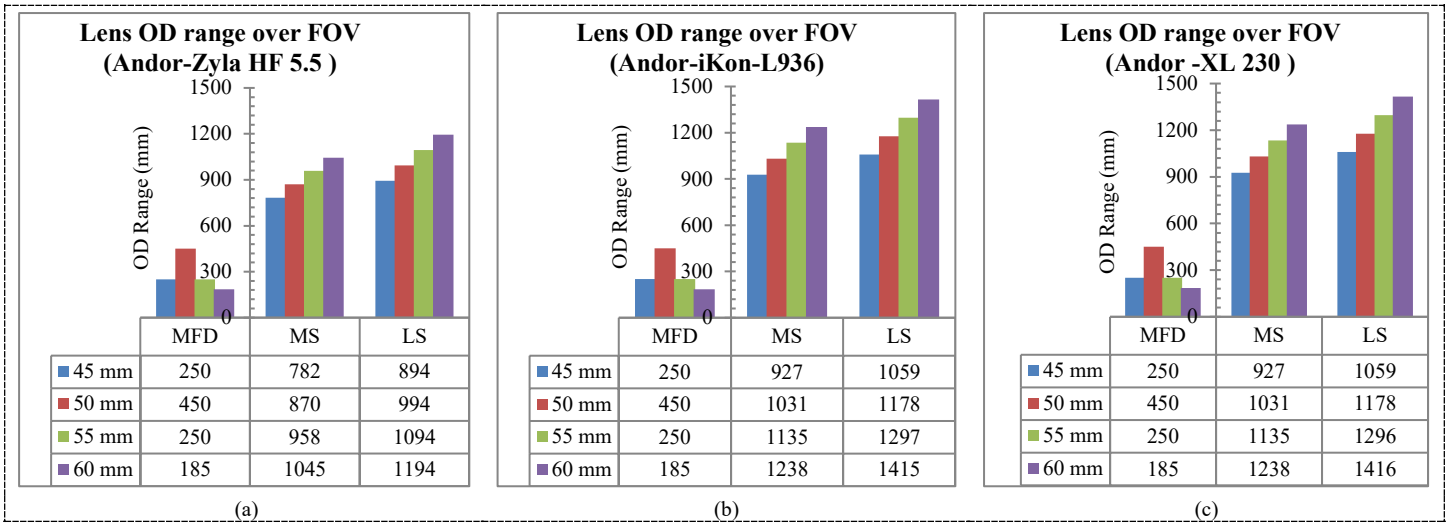


Figure 5: OD range per lens-scintillator combination for Andor cameras: (a) Zyla 5.5-HF, (b) and (c) iKon-L 936 and Andor iKon-XL 230

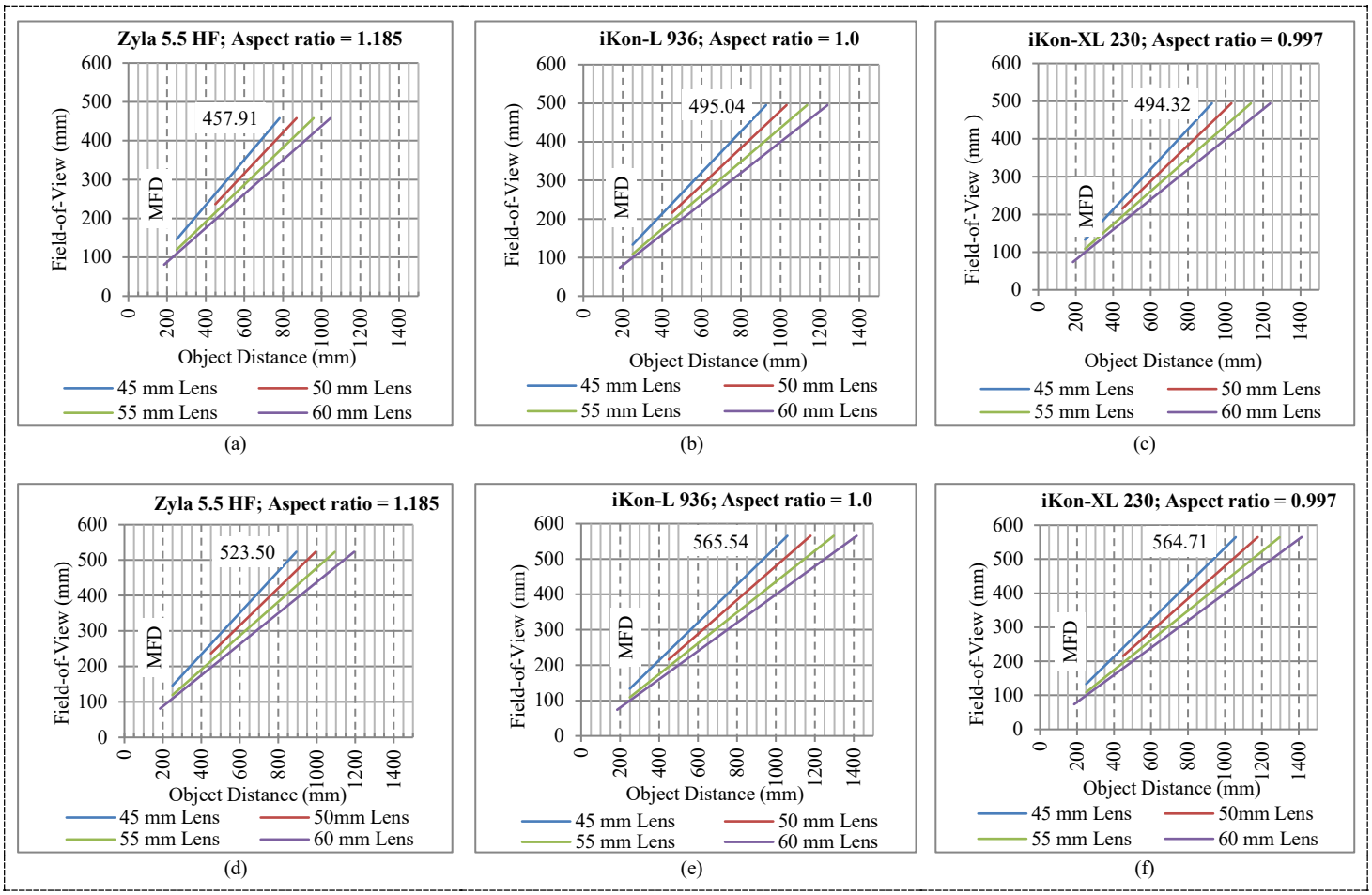


Figure 6: FOV range over OD for all camera-lens combination: First column (Zyla 5.5 HF), second column (iKon-L 936) and third column (iKon-XL 230)

- *Field-of-View Change for Variable Spatial Resolution*

As the FOV size changes (*IFOV*) with change in OD, SR also changes owing to a fixed sensor size. In simulating the SR each camera-lens-scintillator combination would produce at each FOV data point, equation (10), derived from the ratio of *FOV* to the d_i was used, giving results in *lp/mm* and vice versa, using equation (11), giving results in μm; both defines SR range. A 1mm shift interval for FOV data points was adopted over OD range. High SR range limit using both scintillators are shown in Table 6 column 3 and the low range limits in column 6 and 9 for MS and LS respectively. Figure 7 shows change in SR over the FOV size using various camera-lens combinations with the LS. Figure 7 (a), (b) and (c) shows the variation of SR in *lp/mm* whilst Figure 7 (d), (e) and (f) the same results shown in μm.

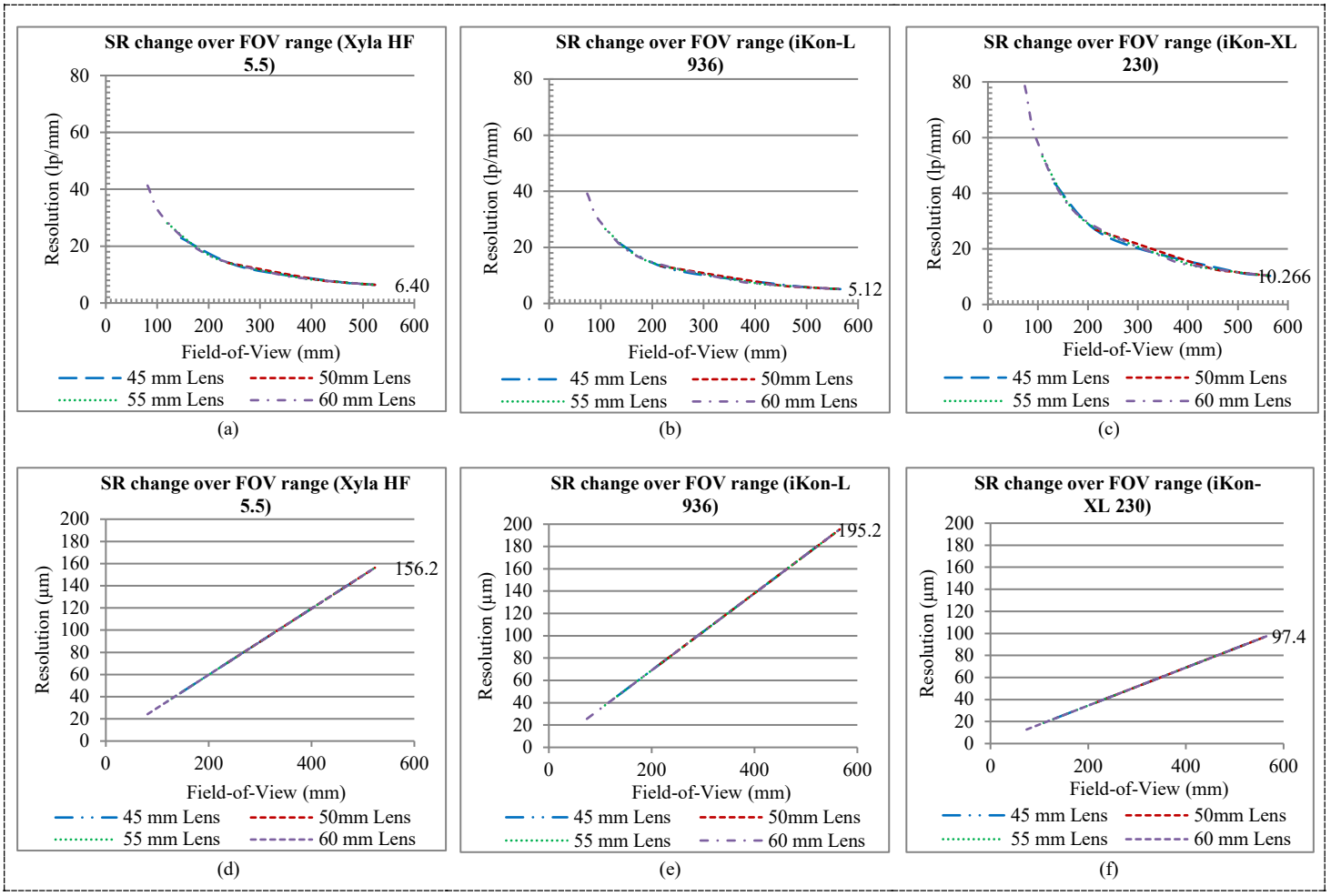
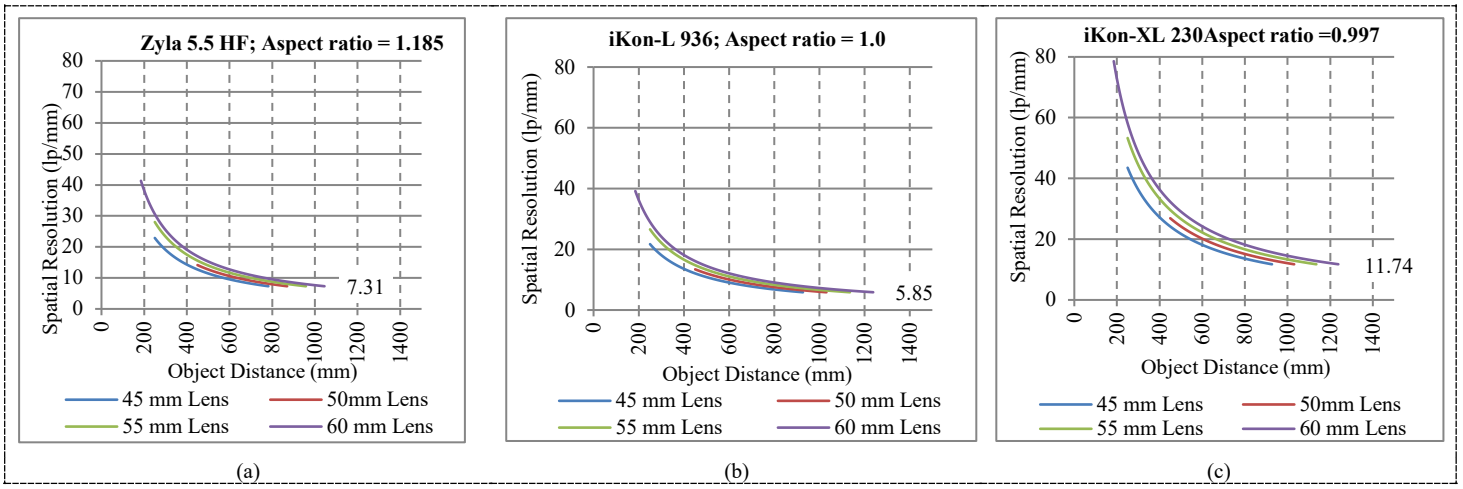


Figure 7: SR change over FOV range for all camera-lens combinations using LS: (a, b and c) in line pairs per mm and (d, e and f): in micrometers pixel size

- *Spatial Resolution and Object Distance*

As the length of OD is changed to vary the FOV, the SR also changes. The change in the SR as a function of OD is demonstrated for all camera-lens combinations and presented in Figure 8 (a), (b) and (c) for the MS and Fig. 8 (d), (e) and (f) for the LS.



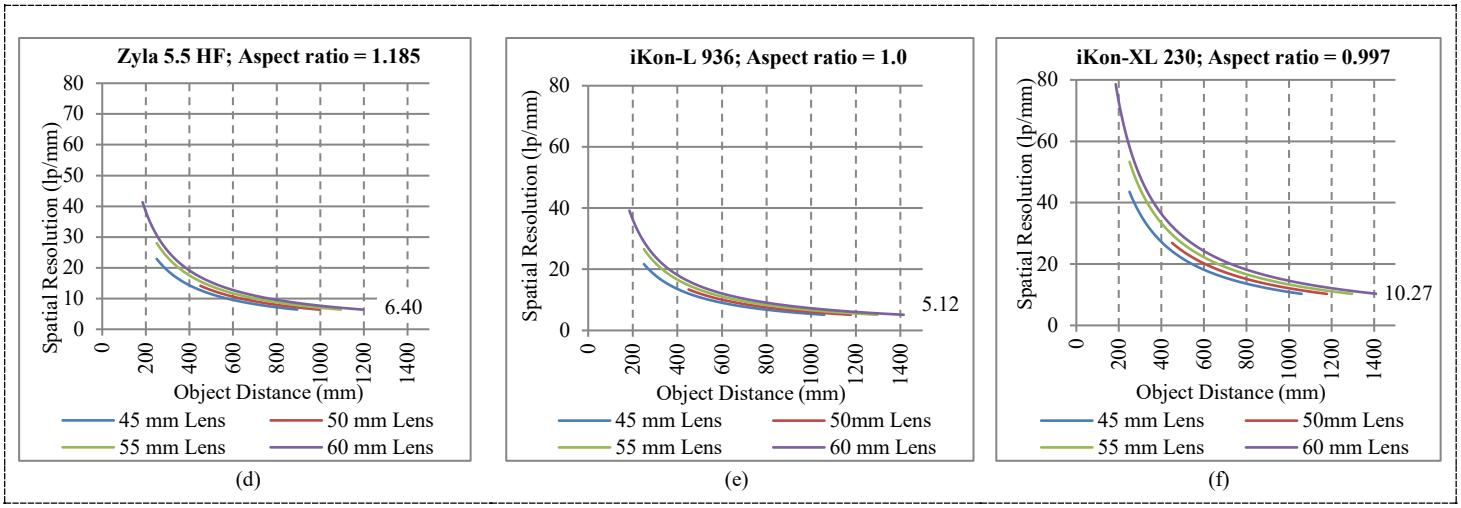


Figure 8: SR changes over OD range for all camera-lens combination using: MS in (a, b and c) and LS in (d, e and f)

The results in Figure 7 and Figure 8 show that the OD increases with increased FOV and OD decreases as SR increases. Figure 7 shows that the effective pixel size becomes bigger with larger FOV, indicating lower SR. The covariance due to the SR, FOV, and the correlation between OD and FOV, for the three cameras (each with a combination of four lenses over LS), were calculated using the covariance function (*Cov*) given in equation (12).

$$Cov(x,y) = \sum [(x_i - x_m) \times (y_i - y_m)]/(n-1) \quad (12)$$

where: x_i = a given x value in the dataset.
x_m = the mean, or average, of the x values.
y_i = the y value in the dataset that corresponds with x_i.
y_m = the mean, or average, of the y values.
n = the number of data points.

The strength of these relationships was calculated using Spearman's correlation coefficient (ρ), given by (13).

$$\rho = 1 - (6\sum d_i^2)/(n\,(n^2-1)) \quad (13)$$

where: ρ = Spearman rank correlation coefficient
d_i = difference between the ranks of corresponding variables
n = number of observations

The results of this relationship are tabled in Table 7.

Table 7: Field-of-View, Spatial Resolution Covariance and Correlation per Camera-Lens Combination over LS

Lens	Zyla 5.5 HF	iKon-L 936	iKon-XL 230
	Covariance; Correlation	*Covariance; Correlation*	*Covariance; Correlation*
PC-E Micro NIKKOR 45 mm f/2.8D ED	-443.8; -0.947	-488.56; -0.934	-978.6; -0.934
Nikon AF-S NIKKOR 50 mm f/1.4G	-174.5; -0.979	-221.14; -.970	-442.9; -0.970
Micro-NIKKOR 55 mm f/2.8	-589.8; -0.930	-628.8; -0.915	-1258; -0.915
AF-S Micro NIKKOR 60 mm F2.8G ED	-922.4; -0.893	-939.67; -0.874	-1883; -0.874

The graphs in Figure 7 demonstrate an existing negative monotonic relationship between the SR and the FOV. Results also show that there is strong negative correlation between SR and the FOV (-90 %) as confirmed by the results in Table 7. Overall results show that, despite the trade-offs that accompanies each choice of combination. The selection lies in the balance between SR and a larger FOV range to accommodate specimens of different sizes.

Larger scintillators offer a large maximum FOV useful in the examination of large specimens without having to examine them in parts. However, the results given in Figure 6 show that, despite the lens or camera in use, larger scintillators demand longer ODs to cover the entire FOV. This contributes to the overall size of the detector system.

Smaller scintillators offering a small maximum FOV are good for a compact design. They are ideal when investigating small specimens. However, they complicate experiments involving large specimens, owing to specimens being scanned in sections. This requires increased time for examination and expert knowledge of image stitching during post-processing. While the camera chip size and the required SR should be the primary consideration in selecting the size of the scintillator, the underlying rationale should be based on the expected size of specimens and trade-offs.

Larger AOV lenses are ideal for a compact design where examinations involving large specimens and large FOVs are necessary; however, these lenses have long MFD and are unable to achieve a very small FOV. They produce small FOV ranges due to immediate FOV saturation. While it is useful to employ a larger FOV, it must be observed that the anti-correlation between the FOV and SR means the trade-off between the two is unavoidable. In contrast to larger AOV lenses, smaller AOV lenses have short MFDs which allow them to be placed very close to the object plane, thereby increasing the reproduction rate. These lenses are ideal in a design intended for smaller specimens necessitating a smaller FOV to maximize the SR; however, they demand a longer displacement for larger FOV applications, increasing the size of the detector system.

By assuming the sample sizes between 10 cm × 10 cm to 30 cm × 30 cm as a target size of samples envisaged, this suggests a FOV range as an experimental requirement. An adequate minimum FOV in the given range of components would be achieved by a 60mm PC-E Micro NIKKOR f/2.8D ED lens. This combined with the iKon-L 936 CCD camera and the 35mm × 35mm FOV scintillator would allow a SR range of 25.5 – 170.92 µm pixel size. In other work CBDS have previously reported a SR at the region of ≈ 100 to 500 µm and recently less than 50 µm [2]. If the budget as a constraint allows, with 1.3m space available at the position where the detector must be located. This detector system would make a bests combination for requirements stated.

4. Conclusion

In conclusion to the component selection strategy, the CBDS design should optimize the SR by managing the FOV. In pursuit to achieve this, the parametric study for the design of a CBDS was presented. We developed models to determine the maximize performance of the CBDS based on experimental requirements and given constraints. In particular, we tried to attain a balance between the FOV and the SR in order to provide a versatile detector system with respect to both variables through simulation. A clear understanding of the different constraints and experimental requirements would aid in the choice of components using the developed models.

To evaluate our models, we benchmarked the results with the secondary data and 100% agreement between the results demonstrated the desirable functionality of the models. The chosen scenario of experimental requirements and constraint has shown that these models are effective in guiding the development of the best detector system with given limitations.

In addition, we proposed a use of linear translation system for the changing the OD to achieve a monotonic FOV and SR. As future work, we plan to develop software for components combination selection. This software will take experimental requirements and constraints as input and return the best possible components combination that can be achieved.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] A.P. Kaestner, Z. Kis, M.J. Radebe, D. Mannes, J. Hovind, C. Grünzweig, N. Kardjilov, E.H. Lehmann, "Samples to Determine the Resolution of Neutron Radiography and Tomography," Physics Procedia, **88**(September 2016), 258–265, 2017, doi:10.1016/j.phpro.2017.06.036.

[2] E.H. Lehmann, G. Frei, G. Kühne, P. Boillat, "The micro-setup for neutron imaging: A major step forward to improve the spatial resolution," Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **576**(2–3), 389–396, 2007, doi:10.1016/j.nima.2007.03.017.

[3] J. Rueckel, M. Stockmar, F. Pfeiffer, J. Herzen, "Spatial resolution characterization of a X-ray microCT system," Applied Radiation and Isotopes, **94**(September), 230–234, 2014, doi:10.1016/j.apradiso.2014.08.014.

[4] H. Pleinert, E. Lehmann, S. Körner, "Design of a new CCD-camera neutron radiography detector," Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **399**(2–3), 382–390, 1997, doi:10.1016/S0168-9002(97)00944-3.

[5] N. Kardjilov, M. Dawson, A. Hilger, I. Manke, M. Strobl, D. Penumadu, F.H. Kim, F. Garcia-Moreno, J. Banhart, "A highly adaptive detector system for high resolution neutron imaging," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **651**(1), 95–99, 2011, doi:10.1016/j.nima.2011.02.084.

[6] E. Lehmann, M. Raventos, R.P. Harti, P. Trtik, A. Kaestner, D. Mannes, C. Grünzweig, "Methodical Progress in Neutron Imaging at PSI," Physics Procedia, **88**(September 2016), 250–257, 2017, doi:10.1016/j.phpro.2017.06.035.

[7] D.S. Hussey, D.L. Jacobson, M. Arif, P.R. Huffman, R.E. Williams, J.C. Cook, "New neutron imaging facility at the NIST," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **542**(1–3), 9–15, 2005, doi:10.1016/j.nima.2005.01.004.

[8] C. Tötzke, I. Manke, A. Hilger, G. Choinka, N. Kardjilov, T. Arlt, H. Markötter, A. Schröder, K. Wippermann, D. Stolten, C. Hartnig, P. Krüger, R. Kuhn, J. Banhart, "Large area high resolution neutron imaging detector for fuel cell research," Journal of Power Sources, **196**(10), 4631–4637, 2011, doi:10.1016/j.jpowsour.2011.01.049.

[9] E.H. Lehmann, P. Vontobel, G. Frei, C. Brönnimann, "Neutron imaging—detector options and practical results," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **531**(1–2), 228–237, 2004, doi:10.1016/j.nima.2004.06.010.

[10] M. DINCA, "The implementation of a charge coupled device (CCD) camera in a neutron imaging system for real time and tomography investigations," Neutron Imaging:A Non-Destructive Tool for Materials Testing (IAEA-TECDOC Series), (September), 53–67, 2008, doi:978–92–0–110308–6.

[11] K. Eismann, S. Duggan, T. Grey, Real world digital photography, Pearson Education, 2010.

[12] N. Kardjilov, A. Hilger, I. Manke, M. Strobl, M. Dawson, S. Williams, J. Banhart, "Neutron tomography instrument CONRAD at HZB," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **651**(1), 47–52, 2011, doi:10.1016/j.nima.2011.01.067.

[13] E.H. Lehmann, P. Vontobel, G. Frei, G. Kuehne, A. Kaestner, "How to organize a neutron imaging user lab? 13 years of experience at PSI, CH," in Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2011, doi:10.1016/j.nima.2010.12.135.

[14] R. Nshimirimana, A. Abraham, G. Nothnagel, A. Engelbrecht, "X-Ray and Neutron Radiography System Optimization by Means of a Multiobjective Approach and a Simplified Ray-Tracing Method," Nuclear Technology, 1–20, 2020, doi:10.1080/00295450.2020.1740562.

[15] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," Nature Methods, **9**(7), 671–675, 2012, doi:10.1038/nmeth.2089.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 3, 257-262 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Microstrip Patch Antenna Designs with Quarter-Circular and Semi-Circular Slots in Patches for Wireless Communication Applications in Frequency Range of 1.2 GHz-4.6 GHz

Cihan Dogusgen Erbas*

Electrical and Electronics Engineering Department, Istanbul Yeni Yuzyil University, Istanbul, 34010, Turkey

ARTICLE INFO

Article history:
Received: 31 August, 2020
Accepted: 04 May, 2021
Online: 27 May, 2021

Keywords:
Microstrip patch antenna
Wireless communication
Patch truncation

ABSTRACT

In this study, 6 unique microstrip patch antennas with quarter-circular and semi-circular slots in patches are proposed. It is aimed to investigate the effect of these slots in the designed topologies. The antennas operate for wireless applications including Personal Communication Service (PCS), 3rd Generation (3G), The Standard for Wireless Fidelity (WiFi)/Wireless Local Area Network (WLAN)/Bluetooth, Long Term Evolution (LTE) and Global System for Mobile Communications at 1.8 GHz (GSM-1800). Simulation results for antenna performance parameters such as fractional bandwidth, gain, radiation pattern, radiation efficiency and total efficiency are presented.

1. Introduction

This paper is an extension of work originally presented at 11th International Conference of Electrical and Electronics Engineering [1]. Multi-band feature is often required in wireless communication systems because simultaneous operation of multiple technologies could be desirable. One antenna type that satisfies that feature is microstrip patch antennas. Microstrip patch antennas have been preferred in many wireless communication systems. [2, 3]. Operation in multiple frequency bands can be obtained by placing slots in patches and/or ground planes [4–6]. Various studies report on slotted microstrip patch antenna geometries: In one report [7], a microstrip patch antenna with a slotted rectangular radiating patch and a defected ground plane suitable for Bluetooth, LTE, Universal Mobile Telecommunications System Band II (UMTS-II), Industrial, Scientific and Medical band (ISM), WLAN, Worldwide Interoperability for Microwave Access (WiMAX), RFID, Wireless Broadband (WiBro) and WiFi is examined. In [8], the authors examine a microstrip patch antenna with an air-filled substrate and 3 open slots. The antenna operates between 1.77 GHz – 2.65 GHz. In [9], the authors focus on a dual-band microstrip patch antenna that consist of two layers separated by an air gap, and a triangular ground slot. The structure operates at 1.8 GHz and 2.4 GHz.

In this article, 6 microstrip patch antennas with quarter-circular and semi-circular slots in patches aimed for wireless operations are proposed. Geometry of the antennas are unique. It is intended to investigate the effect of patch truncations in these antenna topologies. The truncations are located symmetrically on the patches. Antenna performance parameters are simulated. Simulations as well as optimizations are carried out by utilizing ANSYS HFSS.

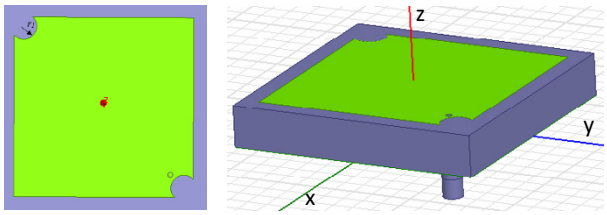
2. Antennas

2.1. Antenna 1

Antenna 1 has 2 identical quarter-circular slots in patch. These slots are located two diagonal patch corners. Radius of quarter-circles is given by r_1 (1.50 mm). The antenna is in square shape as all antennas in this article. Its operation range is 1.3905 GHz – 4.5739 GHz with a fractional bandwidth (FB) of 106.74%. Note that a simulation of the slotless configuration yields a resonance at 2.05 GHz using a coaxial probe as a feed located at (7, 9) mm. The frequency operation range is 1.365 GHz – 2.410 GHz with a fractional bandwidth of 55.36%. Operation frequency range is based on -10 dB impedance bandwidth ($|S_{11}|$<-10 dB). Height of the substrate is 4.30 mm. Its relative permittivity value is 15.5, and loss tangent value is 0.0001. Same substrate material and substrate height are used in all designed antennas. Length of the ground plane and patch without slots are 25 mm and 21.40 mm, respectively, which are the same as those of all designed antennas.

*Corresponding Author: Cihan Dogusgen Erbas, Email: cihan.dogusgen@yeniyuzyil.edu.tr

Antenna feed is a coaxial probe placed at (8, 8) mm in xy plane where center of the antenna is denoted as (0, 0) mm. All antennas in this article are fed by one coaxial probe. Antenna 1 is illustrated in Figure 1.

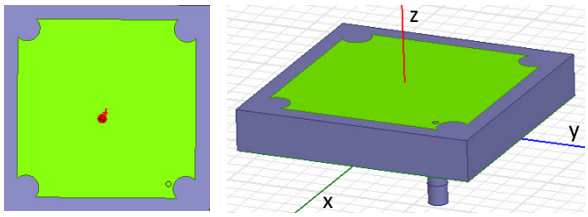


(a) (b)

Figure 1: Antenna 1 (a) patch, (b) side view

2.2. Antenna 2

Antenna 2 has 4 identical quarter-circular slots in patch (see Figure 2). These slots are located at all patch corners. Radius of quarter-circles is given by r_1. The structure' s operation range is 1.4101 GHz – 4.5791 GHz (FB=105.82%). Antenna feed is placed at (8, 9) mm.

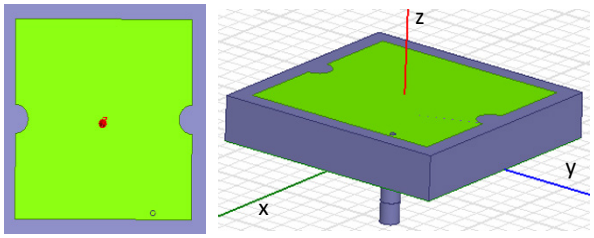


(a) (b)

Figure 2: Antenna 2 (a) patch, (b) side view

2.3. Antenna 3

Antenna 3 has 2 identical semi-circular slots in the patch. These slots are located in the middle of two patch edges. The antenna is shown in Figure 3. Semi-circle radius is given by r_1. The structure' s operation range is 1.3287 GHz – 4.0406 GHz (FB=101.02%). Antenna feed is placed at (10, 6) mm.

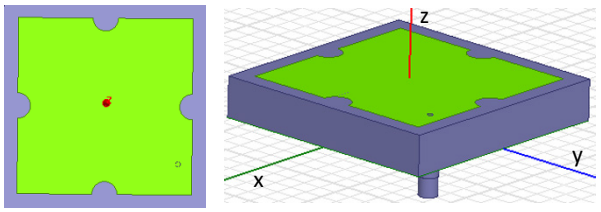


(a) (b)

Figure 3: Antenna 3 (a) patch, (b) side view

2.4. Antenna 4

Antenna 4 has 4 identical semi-circular slots in the patch. These slots are located in the middle of all patch edges. The antenna is shown in Figure 4. Semi-circle radius is given by r_1. The structure' s operation ranges are 1.3543 GHz – 2.3231 GHz (FB=52.69%), and 2.8465 GHz – 4.1123 GHz (FB=36.38%). Antenna feed is placed at (7, 9) mm.

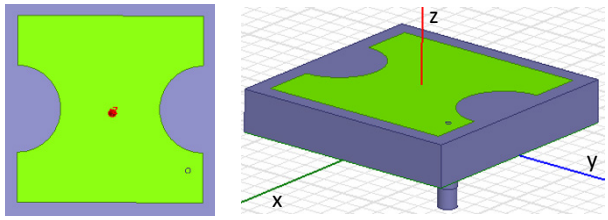


(a) (b)

Figure 4: Antenna 4 (a) patch, (b) side view

2.5. Antenna 5

Antenna 5 has 2 identical semi-circular slots in the patch. These slots are located in the middle of two patch edges. The antenna is depicted in Figure 5. Semi-circle radius is given by r_2 (5 mm). The structure' s operation ranges are 1.2659 GHz – 2.1436 GHz (FB=51.49%), and 2.7428 – 4.6657 GHz (FB=51.91%). Antenna feed is placed at (7, 9) mm.

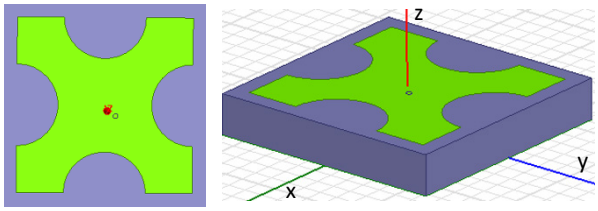


(a) (b)

Figure 5: Antenna 5 (a) patch, (b) side view

2.6. Antenna 6

Antenna 6 has 4 identical semi-circular slots in the patch. These slots are located in the middle of all patch edges. The antenna is illustrated in Fig. 6. Semi-circle radius is given by r_2. The structure' s operation range is 1.3465 GHz – 1.8263 GHz (FB=30.24%). Antenna feed is placed at (5.7, 5.3) mm.



(a) (b)

Figure 6: Antenna 6 (a) patch, (b) side view

3. Simulation Results

Simulation results regarding the antenna perormance parameters such as return loss, gain, and radiation pattern are presented for the designed antennas. The return loss and radiation pattern parameters are expressed in dB while the gain values are denoted in dBi. Radiation efficiency and total efficiency values are given in linear scale. Operation frequencies of interest are 1.9 GHz for PCS, 2.1 GHz for 3G, 2.45 GHz for WiFi/WLAN/Bluetooth, 2.6 GHz for LTE and 1.8 GHz for GSM-1800. Table 1 shows the the operations that the antennas are suitable for.

Table 1: Antennas and operations they are suitable for

Antenna	Operation
Antenna 1, 2, 3, 4	PCS, 3G, WiFi/WLAN/Bluetooth, LTE
Antenna 5	PCS, 3G
Antenna 6	GSM-1800

3.1. Results for Antenna 1

Return loss of antenna 1 is given in Figure 7. Maximum gains at 1.9 GHz, 2.1 GHz, 2.45 GHz and 2.6 GHz are 0.087 dBi, 0.055 dBi, 0.052 dBi and 0.058 dBi, respectively. Radiation efficieny values for the same frequencies (in order) are 0.982, 0.975, 0.964 and 0.967, respectively. Total efficiency values for the same frequencies (in order) are 0.923, 0.934, 0.882 and 0.871, respectively.

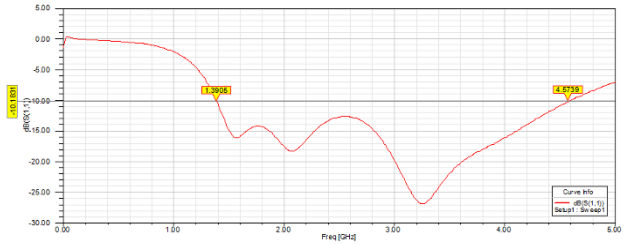


Figure 7: Return loss of antenna 1

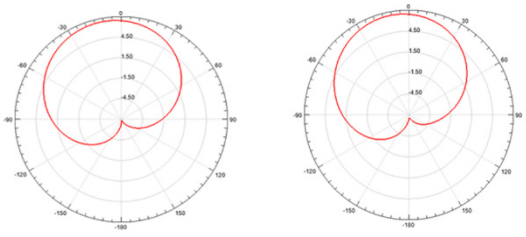
Simulated radiation patterns for antenna 1 are shown in Figure 8. Note that radiation pattern plots represent the radiation intensity for all antennas.

3.2. Results for Antenna 2

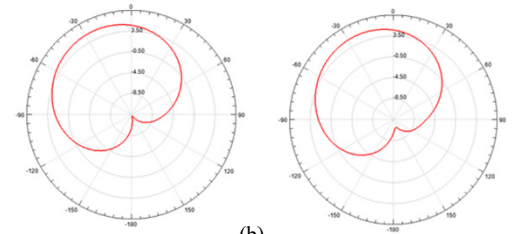
Return loss of antenna 2 is given in Figure 9. Maximum gains at 1.9 GHz, 2.1 GHz, 2.45 GHz and 2.6 GHz are 0.097 dBi, 0.059 dBi, 0.044 dBi and 0.045 dBi, respectively. Radiation efficieny values for the same frequencies (in order) are 0.985, 0.981, 0.973 and 0.970, respectively. Total efficiency for the same frequencies (in order) are 0.924, 0.885, 0.842 and 0.856, respectively. Simulated radiation pattern plots are depicted in Figure 10.

3.3. Results for Antenna 3

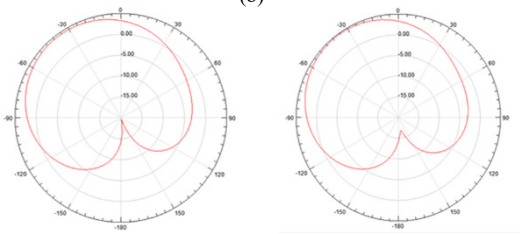
Return loss of antenna 3 is given in Figure 11. Maximum gains at 1.9 GHz, 2.1 GHz, 2.45 GHz and 2.6 GHz are 0.085, 0.042, 0.047 and 0.066 dBi, respectively. Radiation efficieny values for the same frequencies (in order) are 0.968, 0.989, 0.982 and 0.974, respectively. Total efficiency values for the same frequencies (in order) are 0.911, 0.940, 0.881 and 0.872, respectively. Simulated radiation pattern plots are shown in Figure 12.



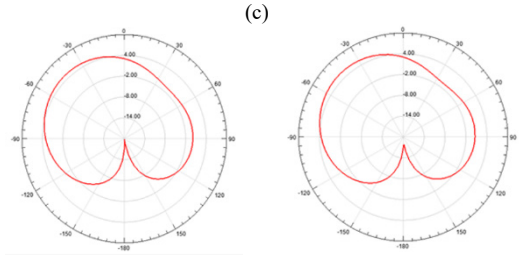
(a)



(b)



(c)



(d)

Figure 8: Antenna 1 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz, (c) f=2.45 GHz, (d) f=2.6 GHz

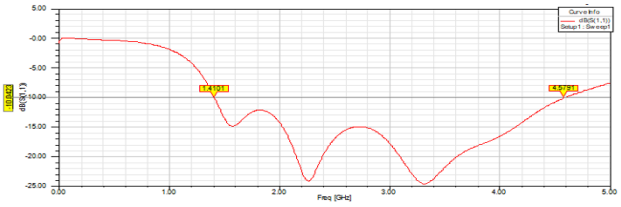
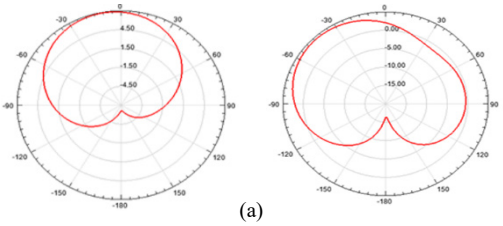
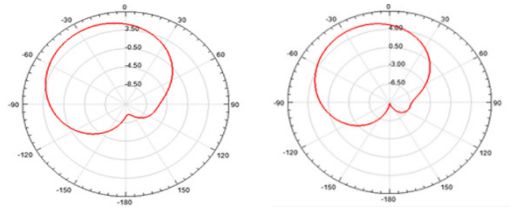


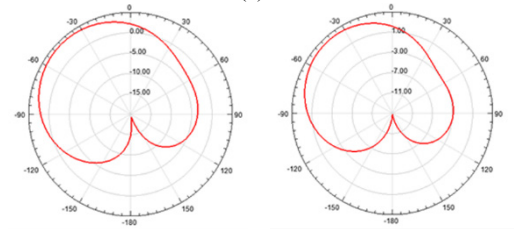
Figure 9: Return loss of antenna 2



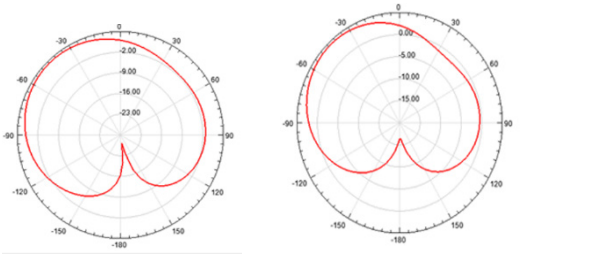
(a)



(b)



(c)



(d)

Figure 10: Antenna 2 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz, (c) f=2.45 GHz, (d) f=2.6 GHz

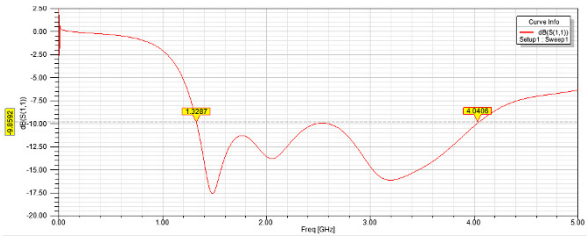
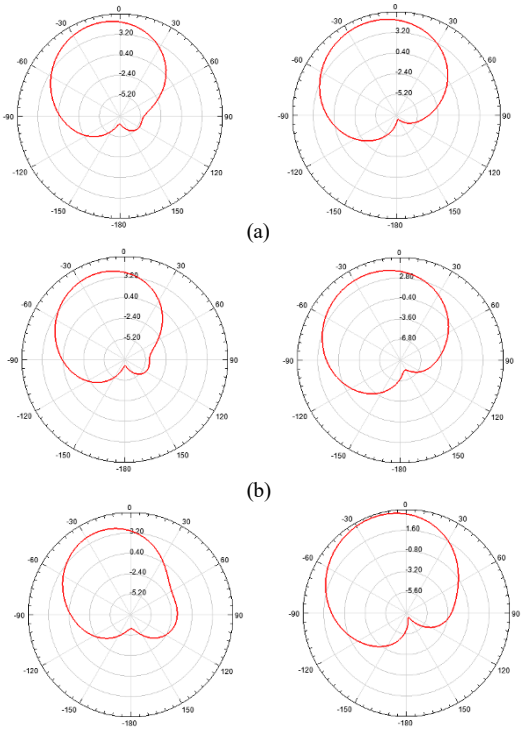


Figure 11: Return loss of antenna 3

3.4. Results for Antenna 4

Return loss of antenna 4 is given in Figure 13. Maximum gains at 1.9 GHz, 2.1 GHz, 2.45 GHz and 2.6 GHz are 0.074, 0.079, 0.082 and 0.087 dBi, respectively. Radiation efficiency values for the same frequencies (in order) are 0.978, 0.988, 0.974 and 0.982, respectively. Total efficiency values for the same frequencies (in order) are 0.935, 0.948, 0.847 and 0.841, respectively. Simulated radiation pattern plots are illustrated in Figure 14.

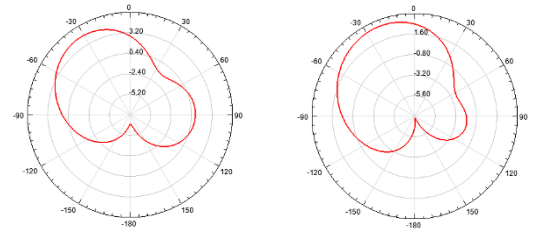


(a)

(b)

(c)

Figure 14: Antenna 4 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz, (c) f=2.45 GHz, (d) f=2.6 GHz.



(d)

Figure 12: Antenna 3 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz, (c) f=2.45 GHz, (d) f=2.6 GHz.

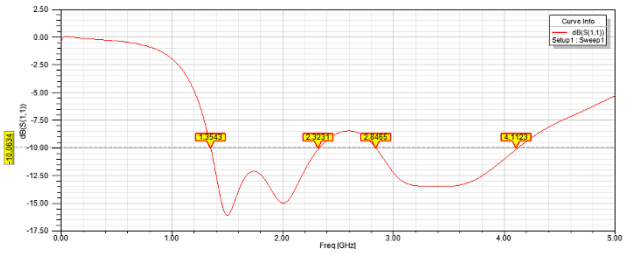
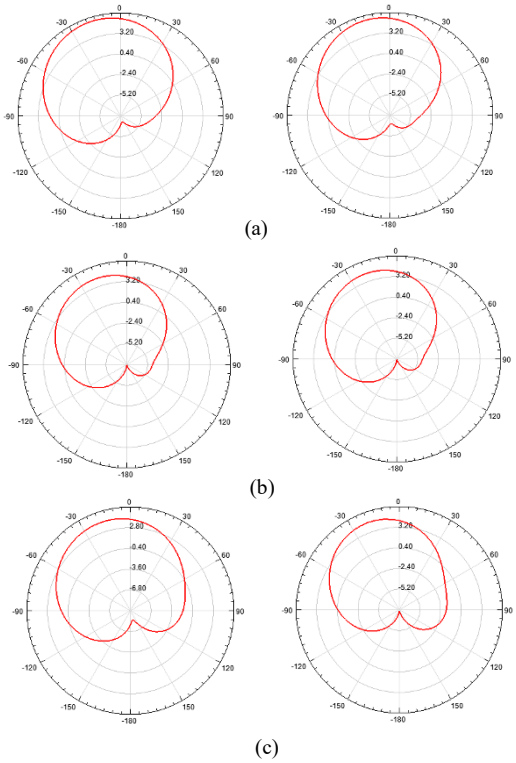


Figure 13: Return loss of antenna 4

3.5. Results for Antenna 5

Return loss of antenna 5 is given in Fig. 15. Maximum gains at 1.9 GHz and 2.1 GHz are 0.053 and 0.062, respectively. Radiation efficiency values for the same frequencies (in order) are 0.984 and 0.997, respectively. Total efficiency values for the same frequencies (in order) are 0.958 and 0.910, respectively. Simulated radiation pattern plots are given in Fig. 16.

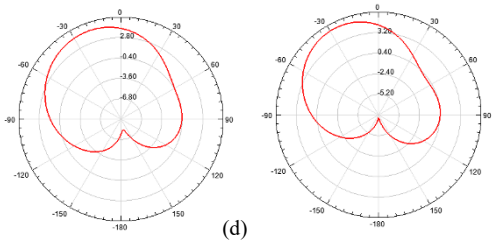


(a)

(b)

(c)

Figure 15: Antenna 5 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz, (c) f=2.45 GHz, (d) f=2.6 GHz.



(d)

Figure 14: Antenna 4 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz, (c) f=2.45 GHz, (d) f=2.6 GHz

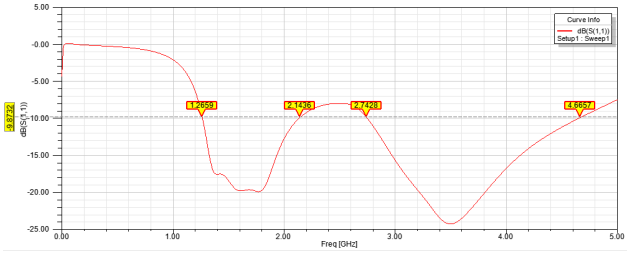
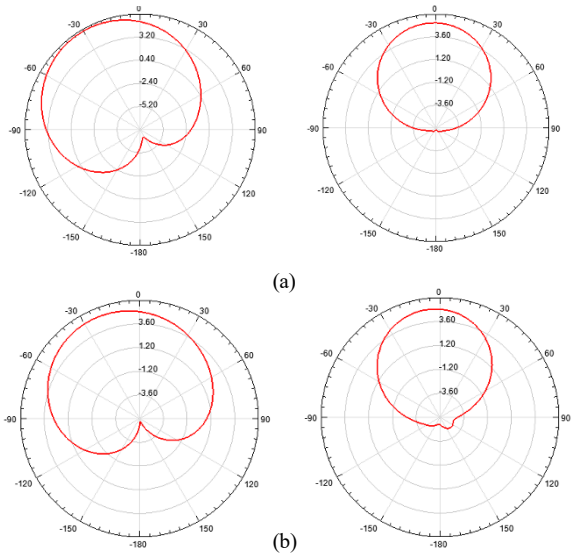


Figure 15. Return loss of antenna 5

3.6. Results for Antenna 6

Return loss of antenna 6 is given in Figure 17. Radiation efficiency and total efficiency values at 1.8 GHz are 0.985 and 0.912, respectively. Simulated radiation pattern plots are illustrated in Figure 18.



(a)

(b)

Figure 16: Antenna 5 radiation patterns of ϕ=0^{O} (left), and ϕ=90^{O} (right): (a) f=1.9 GHz, (b) f=2.1 GHz

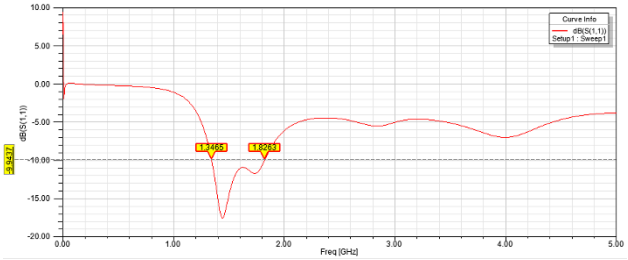


Figure 17: Return loss of antenna 6

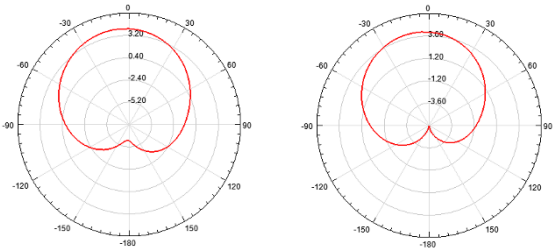


Figure 18: Antenna 6 radiation patterns at f=1.8 GHz: ϕ=0^{O} (left), and ϕ=90^{O} (right)

4. Discussion

Considering the designed antennas, antenna 1, antenna 2, antenna 3 and antenna 4 are suitable for PCS, 3G, WiFi/WLAN/Bluetooth and LTE. Antenna 5 cannot be used for WiFi/WLAN/Bluetooth and LTE operations due to the related frequency bands. Likewise, antenna 6 is only suitable for GSM-1800 operation. Highest fractional bandwidth (106.74%) is obtained for antenna 1. Operation frequency band of antenna 2 is shifted upward compared to that of antenna 1. Similarly, an upward band shift is clear by taking into account the frequency band of antenna 3 as well as the lower and upper frequency ends of antenna 4. Also, an increase in the lower end of frequency band is clear by comparing antenna 5 and antenna 6 frequency values. This occurs because the antenna capacitance decreases as the slot dimension increases, which yields larger resonant frequencies. Note that as the antenna capacitance decreases, the resonant frequency increases. Maximum gains for all cases are similar. The case for PCS operation of antenna 2 has the highest maximum gain (0.097 dBi). Directional radiation patterns are obtained for all cases. Maximum radiation occurs in boresight for PCS operation of antenna 6. We have the highest radiation levels for PCS operation of antenna 2. Radiation efficiency and total efficiency values are relatively high indicating that the usage of a substrate with a high dielectric constant does not significantly lower the efficiencies. Simulations of surface current distribution on the antenna patches for different phase angles at all application frequencies reveal multiple current paths, which yield resonances. Note that the resonance occurs when the path length of the current equals to half of the resonant wavelength in the dielectric material. More information on current path length and resonant frequency analysis can be found in [10]. In order to verify the reliability of the simulation tool HFSS, following procedure is performed: Based on the theoretical calculations of patch width and patch length by using a resonant frequency [11], a reverse computation is carried out for the slotless microstrip patch antenna: The resonant frequency is analytically computed as 2.088 GHz by using the known values of the width and relative dielectric constant. Then, the length of the patch is computed as 15.48 mm by utilizing the calculations of effective length and extension length. Analytically computed values of the resonant frequency and the patch length are consistent with the simulated values. Note that a simulated resonant frequency for the slotless microstrip patch antenna is found to be 2.05 GHz.

5. Conclusion

Six square microstrip patch antennas with slots in patch are designed. Impact of quarter-circular and semi-circular patch truncations on the antenna performance parameters is examined. Results regarding the simulations of return loss, maximum gain, radiation pattern, radiation efficiency and total efficiency are presented. Optimizations are carried out in order to improve impedance matching.

References

[1] C. Dogusgen Erbas, B. Demirbas, C. Yaman and A. Okatan, "Design and comparison of microstrip patch antennas with different patch shapes for wireless communications" in 11th International Conference of Electrical and Electronics Engineering, 646-650, 2019, doi: 10.23919/ELECO47770.2019.8990499.

[2] M. Kumar, V. Nath, "Introducing multiband a wideband microstrip patch antennas using fractal geometries: development in last dacade", Wireless Pers. Commun., **98**, 2079-2105, 2018, doi: 10.1007/s11277-017-4965-x.

[3] C. A. Balanis, "Antenna theory: analysis and design", Wiley, 2012.

[4] G. Sami, M. Mohanna, M. L. Rabeh, "Tri-band microstrip antenna design for wireless communication applications", NRIAG Journal of Astronomy and Geophysics, **2**, 39-44, 2013, doi: 10.1016/j.nrjag.2013.06.007.

[5] M. R. Ahsan, M. T. Islam, M. H. Ullah "A microstrip-fed reformed rectangular shape slotted patch for simultaneous operation in GPS and WLAN bands", Microwave and Optical Technology Letters, **57**, 2204-2207, 2015, doi: 10.1002/mop.29290.

[6] C. Y. D. Sim, B. H. Yang, "A single layer dual-band CP microstrip antenna for GPS and DSRC applications", Journal of Electromagnetic Waves and Applications, **22**, 529-539, 2008, 10.1163/156939308784150308.

[7] J. Kaur, Nitika and R. Panwar, "Design and optimization of a dual-band slotted microstrip patch antenna using Differential Evolution Algorithm with improved cross polarization characteristics for wireless applications", Journal of Electromagnetic Waves and Applications, **33**, 1427-1442, 2019, doi: 10.1080/09205071.2019.1612283.

[8] W. W. Li, Q. H. Li, Y. Meng, J. Y. Wang, et. al. , "A broadband microstrip patch antenna with multiple open slots", Microw. Opt. Technol. Lett., **61**, 626-632, 2018, doi: 10.1002/mop.31646.

[9] N. Hassan, Z. Zakaria, W. Y. Sam, I. N. M. Hanapiah, "Design of dual-band microstrip patch antenna with right-angle triangular aperture slot for energy transfer application", International Journal of RF and Microwave Computer-Aided Engineering, **29**, 1-11, 2018, doi: 10.1002/mmce.21666.

[10] R. Varma, J. Ghosh and R. Bhattacharya, "A compact dual frequency double U-slot rectangular microstrip patch antenna for WiFi/WiMAX", **59**, 2174-2179, 2017, doi: 10.1002/mop.30705.

[11] C. A. Balanis, "Antenna Theory: Analysis and Design", Wiley, 2005.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 261-268 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Fuzzy Controller Based SAPF for Power Quality Enhancement of Distribution System Integrated with Wind Energy Source

Vikas Kumar Sharma*, Lata Gidwani

Electrical Engineering Department, Rajasthan Technical University, Kota, 323021, India

ARTICLE INFO

Article history:
Received: 08 September, 2020
Accepted: 01 November, 2020
Online: 10 November, 2020

Keywords:
Wind energy
Distribution System
Power quality
Fuzzy controller
APCT
SAPF

ABSTRACT

Wind energy is emerged rapidly as the most important and viable sustainable power source due to the mature technology and wide availability. The integration of wind energy sources and utilization of nonlinear loads having different characteristics results in several challenges in a distribution system. One of which is power quality issue .Hence, in this work, the impact of power quality issues in a distribution system integrated with the wind generation system is studied in the uses of a shunt active power filter (SAPF) and presence of non-linear load conditions. The realization of SAPF is carried out using active power component theory (APCT) for reference current extraction, proportional-integral (PI) controller, and fuzzy logic controller (FLC) for dc-link voltage regulation and a basic hysteresis band current controller method for the extraction of switching pulse for operational the inverter. To enhance power quality in the proposed system, a modified fuzzy logic-controlled (MFLC) is developed to improve the transient performance of shunt APF. The proposed model is developed in MATLAB and results are given so as to show the presentation of the performance of this compensation technique in terms of mitigation of system harmonics, reactive power compensation, and enhancing the power factor.

1. Introduction

This Electrical energy has become an integral part of human life and its use is increasing globally by leaps and bounds to improve the living standard. Global concern of utilization of green and clean energy sources for sustainable power systems and shortage of fossil fuel reserves create opportunities for renewable energy sources (RES) of generation. In this manner, the expanding dangers to the earth from the vitality created by the broad utilization of fossil fuels have left us with the thought of renewable energy sources (RES) as the main conceivable arrangement. The demand for sustainable power sources has highly increased due to a lack of fossil fuel, environmental impact and global warming problems of the world. The classification and trends of wind energy chnology has been discussed [1-4] in detail. Wind power is the most developing, promising, and viable sustainable power source that has been dissipated over the word, and utilization of this energy source has been expanding vary large scale. The integration of distributed generation (DG) is directed to increments in vitality productivity and a decrease in emission. The ever-increasing load demand and scarcity of electric power have attracted the attention of power engineers to evolve the new techniques for the effective utilization of electrical energy. Such evolutions have increased the use of power electronics devices and controllers in manifold over the last two decades for more efficient operation of the electrical equipment and to transfer more power through the power networks, but they degrade the power quality in the distribution network connected with wind generation source.

The PQ problem of grid-connected wind energy conversion has been discussed [5-8] by many authors in past years. In [9], the author discussed the power quality issues related to flicker, reactive power, current harmonic distortion, stability of wind generator during the faulty conditions. Further, there are a lot of disturbances that create PQ problems such as harmonics distortion, voltage interruption, transient disturbances, frequency variations, etc. Out of these disturbances the main polluting impact will be created by harmonic distortion but at the same time, it does not mean that power conditioning is only confined to harmonic filtering. In [10], the author introduced the power electronics devices to operate on the sinusoidal voltage but they inject harmonics in the power system networks and increase reactive power demand. The injected harmonics are responsible for the distortion of voltage and current waveshapes. These nonlinear

*Corresponding author: Vikas Kumar Sharma., vikasvs1985@gmail.com

loads are a big reason of harmonics, and extra neutral currents in the power system network. In [11], the author reviewed the different sources of customer- generated harmonics include AC/DC converters adjustable speed drives, switching power supplies, AC phase control circuit, cyclo converters, magnetizing current of transformer, arc furnaces, welding machines, traction systems, switch-mode power supplies (SMPS), computers and programmable logic controllers (PLC's), etc. Moreover, the single-phase loads like personal computers (PCs), fluorescent lights, printers, Xerox machine, television, etc. are also responsible for excessive neutral current distribution system.

The capability and use of SAPF has been investigated in [12-14]. A SAPF could be used as the potential solution to mitigate the utility current disturbances and enhancing power quality in power distribution system integrated with wind generation system. The shunt connected custom power devices are especially useful to maintain power factor, balancing of load, limitation of harmonics and voltage regulation at distribution voltage level, etc. The SAPF operated in the current control mode is capable of injecting negative harmonic current into the grid. Hence, in today's scenario, active power is popular and preferred over the passive filter. Shunt APF has the ability to mitigate the current harmonics from the system but series has sued for voltage harmonics. In [15], the author discussed for fast and precise estimation of reference current, predictive and adaptive controllers based on the neural network have been proposed. The quick estimation of compensating currents a lot of artificial and soft computing techniques was used. Because of the added advantage of ANN to compute the Fourier series coefficient the authors have preferred this over others improved hysteresis current controller technique is overcome wide variation of switching frequency, hence, reduce the switching frequency. In [16], the author implemented a lot of on the hysteresis based current controller in literature like as adaptive frequency domain control of PWM strategy, three-dimensional PWM technique, etc.

In [17], the author developed a standalone wind-diesel hybrid system, in which an active shunt filter is introduced to mitigate the harmonics of load current. In [18-19], they are presented the recent work on the grid integration of renewal energy with SAPF for current harmonic elimination. The dc-link voltage control is achieved by either PI or fuzzy logic technique in SRF based SAPF for proposed system. In [20], the author discussed the purpose of analysis the harmonic reduction and SAPF control technique using harmonic extractor-based method to renewable energy source integrated with grid. In [21**]**, the author proposed the SRF reference current method based shunt active power filter with PI controller to mitigate the current harmonics from the grid connected wind at PCC simultaneously non-linear load conditions. In [22], the author implemented the ICC controller is proposed for shunt active filter for enhance power quality in renewable source interfaced grid system. In [23], the author proposed the system performance parameter like output voltage, frequency and THD level at variable wind speed and nonlinear load conditions. The validity of proposed SAPF is analyzed with fuzzy logic based PWM technique. In [24], the author implemented a technique for reduce of harmonics in hybrid SAPF connected distribution network to enhance the power quality. The optimal controller was made by the improved optimization algorithm.

The recent literatures on the grid integration of wind energy with SAPF mainly focus on the study of current harmonics reduction by the grid current THD analysis. However, a smaller number of publications had been reported for proposing distribution grid level implementation of SAPF for power quality enhancement with the integration of wind energy source. Hence, there is a need to investigate the methods which may be implemented at the grid level to mitigate the power quality problem due to wind energy into the distribution grid. The performance of APC- PI theory with direct and indirect current control methods are compared and found that in both conditions, the neutral current compensation is almost same but the remarkable improvement is observed in power factor, reactive power compensation and load balancing with indirect control method. It is proved through proposed simulation that the indirect current control method performance is better than direct current control for the distribution grid integrated with wind generation system. So, proposed Indirect APC-MFLC methods are suitable to compensate harmonics and reactive power, neutral current during unbalanced loading conditions. The system is capable of maintaining unity power factor as well as sinusoidal source currents. Hence, the power quality of distribution system is improved.

The major contributions of this paper are integration and design of APCT based SAPF in proposed system for enhances the power quality problem. The proposed MFLC controller strategy is validated through APC theory by extensive simulation results for enhancement in overall power quality problem of the proposed system. This paper is classified in five sections, first is the introduction in Section 1. Section 2 mathematical modeling and detail of the proposed model and proposed SAPF control method in Section 3, simulation result and discussion presented in Section 4. The conclusion is shown in Section 5.

2. Proposed Model

A PMSG based wind energy connected through AC-DC-AC converter to distribution grid with a non-linear load connected to load bus shown in Figure 1. A shunt APF connected with the system for enhancement of power quality in non-linear loading conditions and PMSG integrated with the grid. The proposed system is not only capable of transfer the wind energy to the grid, but will also act as a shunt active power filter (SAPF) to mitigate the current harmonics and regulate reactive power feed by the nonlinear loads. Table 1 provides parameter of wind turbine and PMSG.

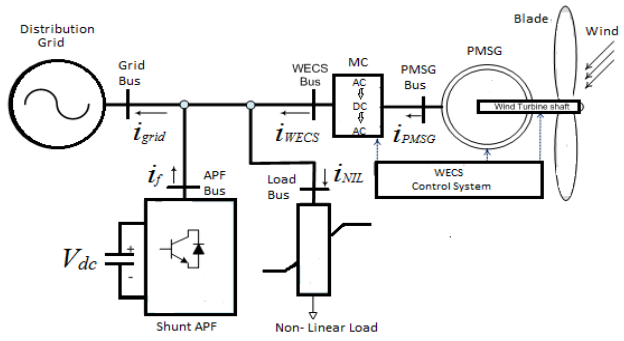


Figure 1: Block Diagram of Distribution Grid- Integrated Wind Generation System with Shunt APF and Non-linear Load.

A. Modeling of Wind Turbine System

The wind speed is sum of these components.

$$V_w(t) = V_B(t) + V_N(t) + V_G(t) + V_R(t) \quad (1)$$

where, V_w is the wind speed, all of them in "meter/ second" and time t is in "second Here this wind speed signal have four component, i.e. the basic speed of wind (V_B), a fluctuating wind (V_N) that describes a steady increase in wind speed, gust component (V_G) and a gradient wind component (V_R). The mechanical energy given wind turbine is determined as follow [3]

$$P_{\text{wind}} = \frac{1}{2} C_p(\lambda, \beta) \pi r^2 \rho_{\text{air}} V_w^2 \quad (2)$$

Power coefficient (C_p) is the ratio of mechanical power generated by the turbine to the power available in the wind. Which λ is the tip speed ratio, C_p is given as [3]

$$C_p = c_1\left(c_2 \frac{1}{\alpha} - c_3\beta - c_4\beta^x - c_5\right) \quad (3)$$

β is a function of the pitch angle of rotor blades, when β is equal to zero.

$$\frac{1}{\alpha} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{1+\beta^2} \quad (4)$$

Therefore, the tip speed ratio can be given as [3]

$$\lambda = \omega_a r / V_w \quad (5)$$

The mechanical torque turbine is defined as follow [3]

$$T_{mech} = \frac{0.5 C_p \pi r^2 \rho_{air} V_w^2}{\omega_a} \quad (6)$$

The driven train is treated as one concentrated mass model to time efficiency and precision. It is defined as:

$$\frac{d}{dt}\omega_a = \frac{1}{j}(T_{mech} - T_{elect} - B_x\omega_a) \quad (7)$$

where, B_x is damping coefficient, ω_g is angular speed (rad/sec), T_{elect} is the electromechanical torque (Nm),T_{mech} is the aerodynamic torque and J equivalent rotational inertia of the generator (kg-m^2)

B. Modeling of PMSG

The direct axis and quadrature axis voltage equation of generator is define as

$$V_d \equiv -R_s i_d - L_d \frac{di_d}{dt} + \omega_e L_q i_q \quad (8)$$

$$V_q \equiv -R_s i_q - L_q \frac{di_q}{dt} - \omega_e L_d i_d + \omega_e \mu_g \quad (9)$$

ω_e is the rotor speed in rad/s of the generator, describe by

$$\omega_e = \omega_a P \quad (10)$$

The electromagnetic torque of PMSG is given by

$$T_{elect} = \frac{3}{4} P\left[\mu_g + \left(L_d - L_q\right) i_d\right] i_q \quad (11)$$

Then, the angular speed of generator is defined as [4]:

$$\frac{d}{dt}\omega_a = \frac{1}{j}\left(T_{mech} - T_{elect} - B_x\omega_a\right) \quad (12)$$

Table 1: parameter of wind turbine & generator

Parameter	Value
Generated Power, (P)	2 MW
Stator Resistance (R_s)	0.7305 mΩ
Stator d-axis Inductance (L_d)	1.21 mH
Stator q-axis Inductance (L_q)	2.31 mH
Permanent Magnet Flux (μ_g)	6.61 Wb
Inertia (J_{eq})	10000 kg.m^2
Radius of Blades (r)	38 m
Air Density (ρ)	1.205m^3

3. Proposed Control Method for SAPF

The Voltage Source Inverter has seen widely utilized in shunt APF setup to minimize power quality issues. It is also utilized for energy storage during ideal system conditions, but the charging and discharging of this capacitor controlled utilized a control theory. Hence, only the active filter fit for producing counter harmonic current and feeding the reactive power to point of common coupling (PCC). Figure 2 presents the configuration of the proposed shunt APF with a distribution system integrated into the wind energy conversion system (WECS). The main aim of proposed method is to evaluate the control strategy for current harmonics reduction, compensation of reactive power and balancing of load current at PCC when the supply current is distorted.

3.1. Active Power Component (APC) or Unit Vector Template Theory

Shunt APF can effectively handle the harmonics and reactive power problems simultaneously. Presently shunt APFs are designed to encounter all the harmonics which is generated by nonlinear loads but the distribution system is highly unbalanced, so it is required to mitigate the neutral current burden and make the source current sinusoidal for the calculation of the desired compensation current. The unit vector templates are determined as

$$\begin{aligned} U_a &= \frac{v_{Sa}}{V_{SP}} = sin\theta \\ U_b &= \frac{v_{Sb}}{V_{SP}} = \sin(\theta - 120^0) \\ U_c &= \frac{v_{Sc}}{V_{SP}} = \sin(\theta + 120^0) \end{aligned} \quad (13)$$

In a three-phase system, it is represented as i_a^*, i_b^*, and i_c^*. Neutral current ought to be zero, along these references neutral current taken as zero in a balanced system. Capacitor voltage error goes

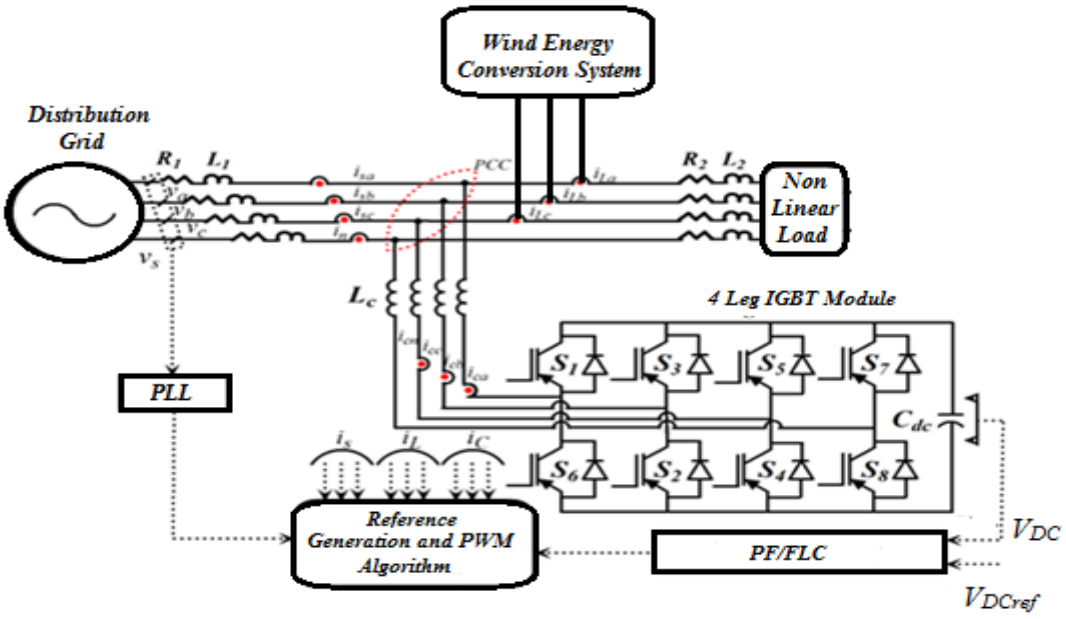


Figure 2: Proposed Shunt APF Configuration with Distribution System

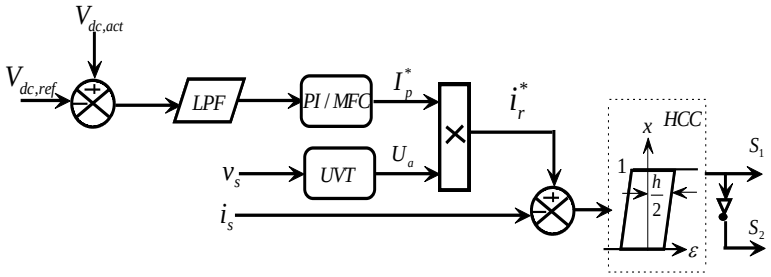


Figure 3: Block Diagram of UVT/APCT Algorithm with Indirect Current Control Method

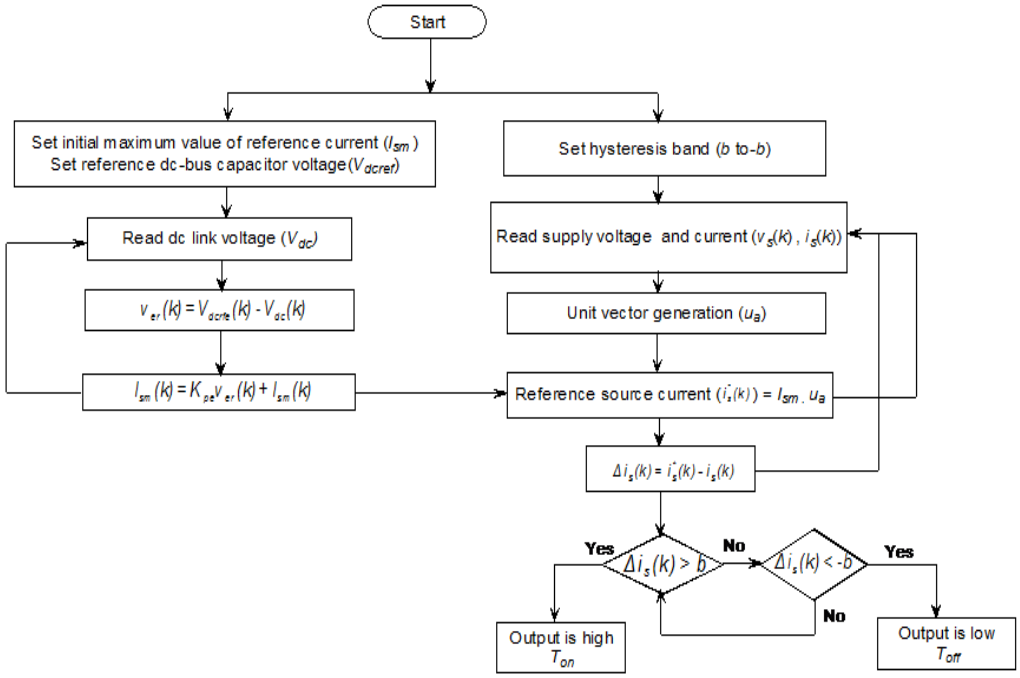


Figure 4: Flowchart of Active Power Component Theory (APCT) Algorithm

through the low power filter. The filtered dc voltage is contrasted with the reference value. Reference current can be composed as

$$i_a^* = I_m U_a,\ \ i_b^* = I_m U_b,\ \ i_c^* = I_m U_c,\ \ i_n^* = I_m U_n \qquad (14)$$

As present in the Figure 3 that capacitor voltage error goes through the LPF, which remove the high-frequency ripple from the voltage. This filtered dc voltage is compared with the reference value. Now error between the reference and grid current is processed by utilizing the following equations

$$\left.\begin{array}{c} i_a^* = I_m U_a \\ i_b^* = I_m U_b \\ i_c^* = I_m U_c \\ and \\ i_n^* = I_m U_n \end{array}\right\} \qquad (15)$$

Figure 4 presents the flowchart of the APCT algorithm with indirect current control method.

3.2. DC voltage Controller

DC voltage controller calculates response of the model and increasing proportional above a value can generate swaying. Integral control expels the error from the output. To actualize the PI controller, the dc link voltage is estimated the reference value. The transfer function of the PI-controller is given from [18] as follows

$$H(s) = (k_i / S) + k_p \qquad (16)$$

For the usage of the fuzzy rule over the PI controller in close loop filter, the dc voltage signal is compared with reference signal and reference error will be produced as follow [18]:

$$e = V_{\text{dc, reference}} - V_{\text{dc, actul}} \qquad (17)$$

The reference for nth sampling time is taken as a change of error in fuzzification process is determining as follows [18]-[19]

$$C_{e(n)} = e(n) - e(n-1) \qquad (18)$$

3.3. Proposed Fuzzy Ligic controller

The proposed modified fuzzy logic controller (MFLC) of the system is used to control the response of the dc voltage. They can be defined, as

$$x = \left(x_i, U_i, \left\{T_{x_i}^1, T_{x_i}^2, \cdots, T_{x_i}^{k_i}\right\}, \left\{\mu_{x_i}^1, \mu_{x_i}^2, \cdots, \mu_{x_i}^{k_i}\right\}\right) \quad \text{for } i = 1, 2, \cdots, n \qquad (19)$$

$$y = \left(y_i, V_i, \left\{T_{y_i}^1, T_{y_i}^2, \cdots, T_{y_i}^{l_i}\right\}, \left\{\mu_{y_i}^1, \mu_{y_i}^2, \cdots, \mu_{y_i}^{l_i}\right\}\right) \quad \text{for } i = 1, 2, \cdots, m \qquad (20)$$

where x & y is the input & output linguistic variable form a fuzzy output space. The size of a term set, $T_{(xi)} = K_i$ is called as the fuzzy partition number of X_i. In the case of a two – input-one-output proposed MFC system, If $T_{(x1)} = 9$ and $T_{(x2)} = 9$, than the number of fuzzy control rules is 9×9. A bell shaped relationship function has been used.

$$\mu_{xi}(x) = \exp\left(-\frac{(x - m_i)^2}{\sigma_l^{\ 2}}\right) \qquad (21)$$

This modified fuzzy controller rules are developed and approved dependent on AND Boolean logic operator. The new rules are characterized by fundamental stability position and moment in time response. To use of fuzzy sets, we restructure operators of the standard set method to fit the particular relationship functions of fuzzy logic for a value between zero & one. Here two variables are utilized as a contribution of MFLC, first is Error E and second change of error C_e. The moment in time response has been divided into four regions A_1, A_2, and A_3, A_4 with two arrangements of point, first is the peak value index (b_1, b_2), and second cross over-index (c_1, c_2). The fuzzy operator AND is formed from a grouping of fuzzy prepositions. Based on the above theory, rules are defined as ZE, PN NB, NS NM, PS, PB, and PB, PM for processing the dc-link response shown in Table 2. Figure 5 shows the moment in time response of a stable closed loop-system.

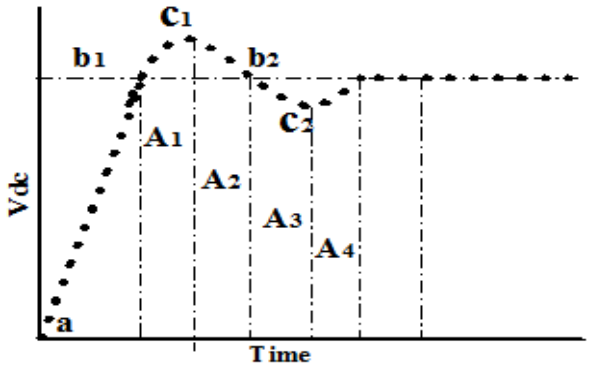


Figure 5: Moment in Time Response of a Stable Closed Loop System.

Table 2: Mechanism of Proposed modified fuzzy rule

Error/ Change error (E)/(Ce)	**NB**	**NM**	**NS**	**NN**	**ZE**	**PN**	**PS**	**PM**	**PB**
NB	NB	NB	NB	NB	NB	NB	NB	NM	ZE
NM	NB	NB	NM	NM	NM	NM	NS	ZE	PM
NS	NB	NM	NM	NS	NS	NS	ZE	PS	PM
NN	NB	NM	NS	NN	ZE	ZE	PS	PM	PB
ZE	NB	NS	NS	ZE	PN	PN	PS	PM	PB
PN	NS	NB	NS	PN	PS	PS	PS	PM	PB
PS	NM	NS	ZE	PS	PS	PS	PM	PM	PB
PM	NB	ZE	PS	PM	PM	PM	PM	PB	PB
PB	ZE	PM	PB	PB	PB	PB	PB	PB	PB

4. Result and discussion

The performance analysis of the distribution grid integrated with wind generation system is study with proposed SAPF under the use of different controllers for unbalanced loading condition. The impact of PQ issues in a distribution grid integrated with wind generation system is investigated in the presence of SAPF. The effectuation of SAPF is carried out using active power component theory (APCT) for reference current generation with direct and indirect current control method. Proportional integral (PI) controller is proposed with modified fuzzy logic controller (MFLC) for dc link voltage regulation and a basic hysteresis band current controller method. For study state analysis purposes, the load change with an increase at t=1.2 sec, and again it is changed with decreasing at t= 2.3 sec. The RMS load current at t=0.2 sec is in phase "a" 139.2 Amp, phase "b" 103 Amp, and in phase "c" 68.8 Amp. This result presents the unbalanced condition in the distribution system. Without compensation the neutral current flow in 73.3 Amp, the reactive power drawn is 20 kVAR, power factor degradation is below 0.8, and load current THD is 22.50%. A SAPF with a hysteresis-based APC technique is used to enhance the performance of the system according to certain power quality standards. In this technique, the reference signal is compared with actual utility load current.

4.1. Case 1: Direct current control with APCT-PI

In the direct current control method with APCT-PI, the switching pulses are generated through hysteresis controller by comparing the sensed three phase shunt APF current with their computerized generated reference current signal. The performance of the direct current control method with the APCT technique as shown in figure 6. The following observation has been made out after compensation reactive power flow between souse and load is 7 KVAR, natural current is around 7 Amp and system power factor is nearly 0.94.

4.2. Case 2: Indirect current control with APCT-PI

In the indirect current control method, the PWM switching pulses are obtained by comparing the sensed three phase current with their reference current which are generated by the reference current extraction methods. The performance of the in-direct current control method with the APCT technique is shown in figure 7. The following observation is made out after compensation reactive power flow between souse and load is 6 KVAR, natural current is around 6.8 Amp and system power factor is nearly 0.996. Hence the performance of the in-direct current control loop with unit vector template theory is found better than direct control so modified FLC implanted on indirect current method.

4.3. Case 3: Indirect current control with Modified Fuzzy Logic controller with APCT

It provides the batter result in a comparison of harmonics mitigation, reactive power compensation, and power factor improvement. From the figure 8, following observation, has been made out after compensation reactive power flow between souse and load is 3.7 KVAR, natural current is around 4.7 Amp and system power factor is nearly 0.999. Hence the performance of the in-direct current control loop with MFLC based APC theory is found suitable for SAPF. The comparisons of harmonics mitigation are shown in figure 9.

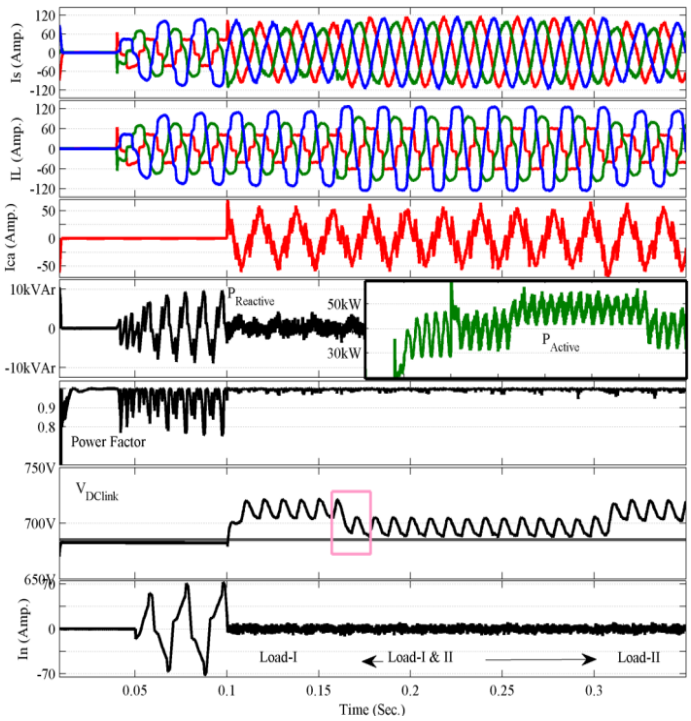


Figure 6: Indirect current control method shown as I_{sourse}, I_{load}, $I_{netural}$, $P_{react.}$, PF, V_{DC} and I_n with APC-MFLC theory.

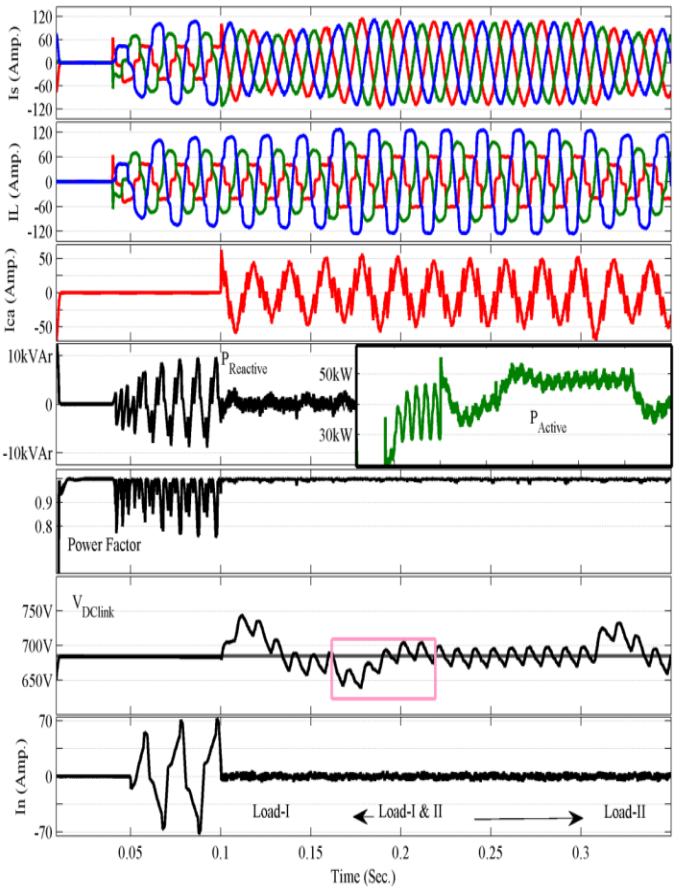


Figure 7: Direct current control method shown as I_{sourse}, I_{load}, $I_{netural}$, $P_{react.}$, PF, V_{DC} and I_n with APC-PI theory.

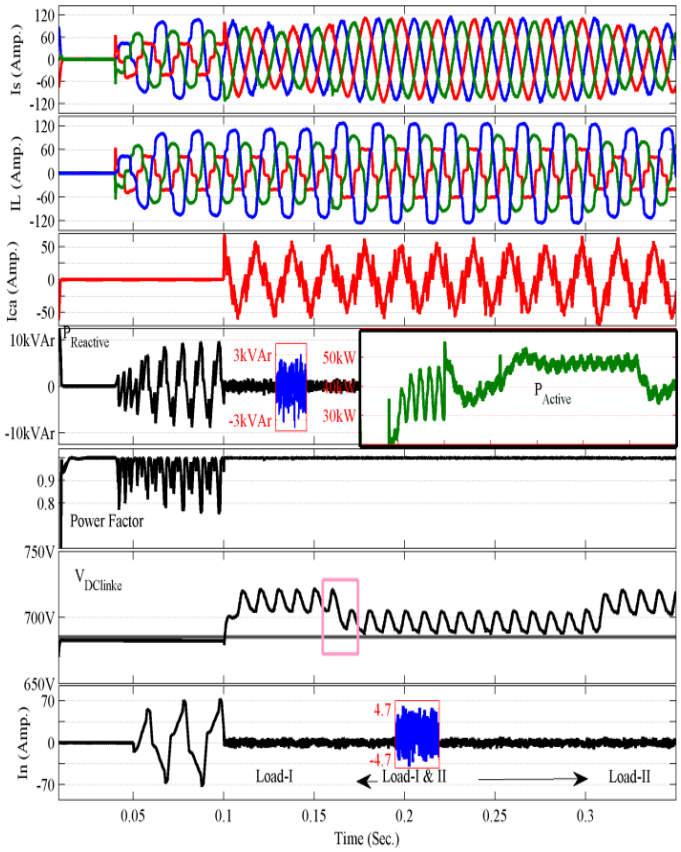


Figure 8: Indirect current control method shown as I_{sourse}, I_{load}, $I_{netural}$, $P_{react.}$, PF, V_{DC} and I_n with APC-MFLC theory.

4.4. Performance comparison of direct APC-PI, Indirect APC-PI and Proposed Indirect APC-MFLC

The comparison of overall performance of proposed system are discussed in Table 3.The performance of APC- PI theory with direct and indirect current control methods are compared and found that in both conditions, the neutral current compensation is almost same but the remarkable improvement is observed in power factor, reactive power compensation and load balancing with indirect control method. It is proved through proposed simulation that the indirect current control method performance is better than direct current control for the distribution grid integrated with wind generation system. So, proposed Indirect APC-MFLC methods are suitable to compensate harmonics and reactive power, neutral current during unbalanced loading conditions. The system is capable of maintaining unity power factor as well as sinusoidal source currents. Hence, the power quality of distribution system is improved.

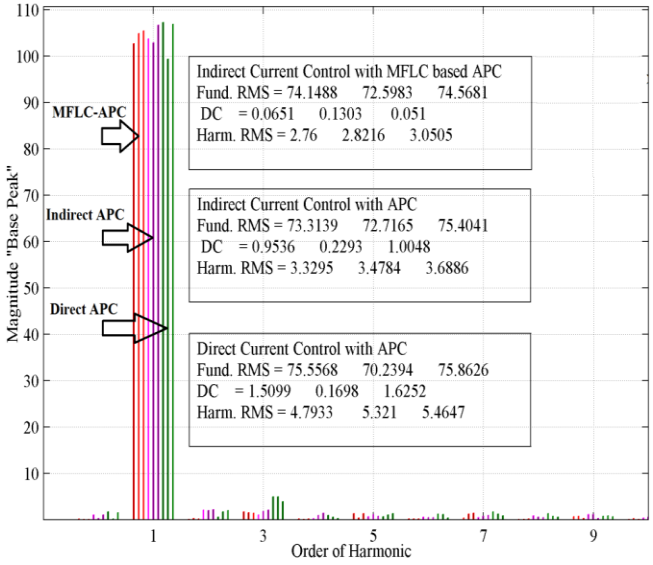


Figure 9: THD with Direct APCT-PI, Indirect APCT-PI & Indirect APCT-MFLC.

5. Conclusion

This paper presents a SAPF based on APC theory for reference current generation and proposed MFLC method for dc-link voltage control to enhance power quality of the distribution grid integrated with wind energy source under non-linear load. The performance of APC theory with direct and indirect current control methods is compared and found that in both conditions, the neutral current compensation is almost the same but the remarkable improvement shown in power factor, reactive power compensation, and load balancing with the indirect current control method.

Table 3: Comparison of direct APCT-PI, Indirec apct-pi and indirect Apct-MFLC

Parameter	**Direct control with APCT**			**Indirect direct control with APCT**			**Indirect direct control with MFLC- APCT**		
	I_{sa}	I_{sb}	I_{sc}	I_{sa}	I_{sb}	I_{sc}	I_{sa}	I_{sb}	I_{sc}
Balancing of source current I_s	108.4	98.06	105.1	103.6	102.8	106.6	110.8	1.8.1	108.1
Harmonic RMS after comp.	4.793	5.321	5.464	3.329	3.478	3.688	2.760	2.821	3.050
Neutral current I_n after amp.	7.0			6.8			4.7		
Reactive power after comp. (kVAr)	7			6			3.7		
Power Factor (PF)	0.94			0.996			0.999		

Hence, it has been proved through a simulation study that the indirect current control loop performance is better than the direct control loop for the unbalanced distribution system. The performance of the proposed method using APC-MFLC based realized SAPF is compared to that of SAPF designed based on reported strategies such as direct APC-PI controller and indirect APC-PI controller. The simulation results show the effectiveness of the proposed control algorithms to mitigate the harmonics, compensate the reactive power, and neutral current and to improve power factor with the balancing of source current using the MATLAB/Simulink.Finally, the proposed algorithms are capable of improving power quality problems in the distribution system with wind energy sources under non-linear load.

Conflict of Interest

All authors declare no conflicts of interests.

References

[1] N. F. Mohamed, "Wind energy conversion systems, classifications and trends in application," in 2019 5th International Renewable Energy Congress," 1-6, 2014, doi: 10.1109/IREC.2014.6826922.

[2] G. Shafiullah, M. Amanullah, A. Shawkat Ali, P. Wolfs, "Potential challenges of integrating large-scale wind energy into the power grid–A review," Renewable and Sustainable Energy Reviews, **20**, 306-321, 2013, doi: 10.1016/j.rser.2012.11.057.

[3] C. N. Bhende, S. Mishra, S. G. Malla, "Permanent magnet synchronous generator-based standalone wind energy supply system," IEEE Transastion Sustainable Energy, **2**(4), 361–373,2011, doi: 10.1109/TSTE.2011.2159253.

[4] W. Xueguang, W. Weisheng, D. Huizhu, C. Yunping, "Application of Models of the Wind Energy Conversion System to Wind Power Dynamic Analysis," in 1998 IEEE Power Systems Technology Conf., Beijing, 1406-1411, 1998, doi: 10.1109/ICPST.1998.729318.

[5] A. S. Bubshait, M. G. Sim, A. Mortezaei, T. D. C. Busarello, "Power quality achievement using grid connected converter of wind turbine system," in 2015 1st Annu. Meeting IEEE Ind. Appl. Soc.,1–8, 2015, doi: 10.1109/IAS.2015.7356817.

[6] M. Singh, V. Khadkikar, A. Chandra, "Grid synchronisation with harmonics and reactive power compensation capability of a permanent magnet synchronous generator-based variable speed wind energy conversion system," IET Power Electronics, **4**(1), 122-130, 2011, doi: 10.1049/iet-pel.2009.0132.

[7] Y. Oğuz, İ. Güney, H. Çalık, "Power quality control and design of power converter for variable-speed wind energy conversion system with permanent-magnet synchronous generator," The Scientific World Journal, **2013**(4), 122-132, 2013, doi: 10.1155/2013/783010.

[8] A. Bubshait, A. Mortezaei, M Simões, T. Busarello, "Power quality enhancement for a grid connected wind turbine energy system," IEEE Trans. on Industry Applications, **53**(3), 2495-2505, 2017, doi: 10.1109/TIA.2017.2657482.

[9] S. Kasa, P. Ramanathan, S. Amasamy, D. P. Kothari "Effective grid interfaced renewable sources with power quality improvement using dynamic active power filter," Electrical Power and Energy Systems, **82**, 150–160, 2016, doi:10.1016/j.ijepes.2016.03.002.

[10] H. Qazi, M. Wazi Mustafa "Review on active filters and its performance with grid connected fixed and variable speed wind turbine generator," Renewable and Sustainable Energy Reviews, 57, 420–438, 2016, doi: 10.1016/j.rser.2015.12.049.

[11] W. Ullah, M. Mekhilef, M. Seyedmahmoudian, "Active power filter for mitigation of power quality issues in grid integration of wind and photovoltaic energy conversion system," Renewable and Sustainable Energy Reviews, **70**, 635–655, 2017, doi: 10.1016/j.rser.2016.11.091.

[12] A. Hoseinpour, S. Masoud Barakati, R. Ghazi, "Harmonic reduction in wind turbine generators using a Shunt Active Filter based on the proposed modulation technique," International Journal of Electrical Power & Energy Systems, **43**(1), 1401-1412, 2012, doi: 10.1016/j.ijepes.2012.06.052.

[13] S. Kasa, P. Ramanathan, S. Ramasamy, "Effective grid interfaced renewable sources with power quality improvement using dynamic active power filter," International Journal of Electrical Power & Energy Systems, **82**(4), 150–160, 2016, doi:10.1016/j.ijepes.2016.03.002.

[14] D. Olivindo, J. A. Soa, I.R. Machado, "Shunt active power filter for energy quality improvement in distributed generation systems," in 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), 146-151, 2017, doi: 10.1109/ISIE.2017.8001238.

[15] H. Mojgan, Z. A. Abu, T. Arash, S. Mohammadsoroush "An Overview on Current Control Techniques for Grid Connected Renewable Energy Systems," 2012 2nd International Conference on Power and Energy Systems (ICPES 2012), Singapore, 119 – 126, 2012.

[16] O. F. Kececioglu, H. Acikgoz, C. Yildiz, A. Gani , M. Sekkeli, "Power Quality Improvement Using Hybrid Passive Filter Configuration for Wind Energy Systems," Journal of Electrical Engineering and Technology, **12**(1) , 207-216, 2017, doi: 10.5370/JEET.2017.12.1.207.

[17] M. Kouadria, T. Allaoui, M. Denai, "Power Quality improvement in isolated Wind-Diesel Power System," WEASE Trans. on Systems and Control, **12**(50), 490-498, 2018, doi: 10.5370/JEET.2017.12.1.207.

[18] J. Fei, K. Ma, S. Zhang, W. Yan, Z. Yuan "Adaptive current control with PI-fuzzy compound controller for shunt active power filter," Mathematical Problems in Engineering, **2013**(1), 150-158, 2013, doi.org/10.1155/2013/546842.

[19] P. Barik, G. Shankar, P. K. Sahoo "Power quality assessment of microgrid using fuzzy controller aided modified SRF based designed SAPF," International Transactions on Electrical Energy Systems, **29**(12), 1-24, 2019.

[20] V. N. Jayasankar**,** U Vinatha "Advanced Control Approach for Shunt Active Power Filter Interfacing Wind- Solar Hybrid Renewable System to Distribution Grid," Journal of Electrical System, **14**(2), 88-102, 2018, doi:10.1002/2050-7038.12289.

[21] S.H. Qazi, M.W. Mustafa, U. Sultana "Current harmonics mitigation from grid connected variable speed due to nonlinear loads using shunt active power filter," journal Teknnology, **79**(4), 2017, doi: 10.11113/jt.v79.4702.

[22] S. Ramasamy ,K. Sudheer "Enhancement of Power Quality in Renewable Source Integrated Three Phase System with ICC Controller," Journal of Green Engineering,, **8**(1), 71–88, 2018, doi: 10.13052/jge1904-4720.815.

[23] K. Ramesh, G.T. Arasu "Power Quality Improvement in Wind Energy System using Seven Level shunt active power filter," Asian Journal of Technology, **17**(3), 189-196, 2018, doi:10.36478/ajit.2018.189.196.

[24] G. Arunsankar, S. Srinath, "Optimal controller for mitigation of harmonics in hybrid shunt active power filter connected distribution system: An EGOANN technique," Journal of Renewable and Sustainable Energy, **11**(2), 1–16, 2019, doi: 10.1063/1.5085028.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 327-335 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Mathematical Modelling of Output Responses and Performance Variations of an Education System due to Changes in Input Parameters

Najat Messaoudi[*,1], Jaafar Khalid Naciri[2], Bahloul Bensassi[1]

[1]*Laboratory of Industrial Engineering, Information Processing and Logistics, Hassan II University, Faculty of Sciences Ain Chock, Casablanca, 20100, Morocco*

[2]*Laboratory of Mechanics, Hassan II University, Faculty of Sciences Ain Chock, Casablanca, 20100, Morocco*

ARTICLE INFO

Article history:
Received: 02 November, 2020
Accepted: 31 December, 2020
Online: 22 January, 2021

Keywords:
Modelling
Education system
Complex system
Performance
Enterprise modelling

ABSTRACT

"This paper is an extension of work originally presented in the 4th International Conference on Systems of Collaboration, Big Data, Internet of Things & Security -SysCoBIoTS'19". The use of complex and dynamic systems modelling to social systems is quite recent and its pertinence in the case of an educational system is continually increasing. For the concrete management of educational systems, a global approach is required. This approach must take into consideration the effects of many parameters that can act and interact together thus making, as a result, the system more or less efficient. Our aim is to develop a model that can capture the dominant dynamics of these systems while being at the same time simple enough to be useful for analyzing, simulating, and quantifying the impact of different parameters on the global performances of educational systems. By viewing education systems as skills production systems and by applying Business Processing modelling methods, a modelling of education systems is proposed in the present work which allows studying the effects of a set of parameters on the behavior of the system and its performance. The focus will be done on the study of the impact of learners' input competence on the performance of a training unit and on the performance of a training program. The obtained simulation results allow us to analyze the evolution of a training program's behavior as well as estimates its performance under the effect of the variation of simulation factors. These results enable to measure the performance variation according to the learners' input competence, their ability to acquire skills, and to the class size. This modelling enables us to test solutions for performance improvement.

1. Introduction

Education is a determining factor for growth and development and education systems and socio-economic development are closely related and interact with each other [1]. Indeed, an education system enables individuals to improve their productivity and increase their employability, and at the community level, it improves the competitiveness and attractiveness of the economy through the availability of skilled human capital. As a result, developing effective education systems whose performance can be quantified and ensured becomes an important issue at many levels for managers and all stakeholders.

But several questions arise: what defines an education system as a high-performing one? How can the performance of an education system be evaluated? What makes an education system efficient? What are the factors that impact its performance? How can we act on these factors to improve this performance?

Answers to these questions are not obvious as far as education systems are complex systems and having tools allowing the objective evaluation of their performance is still a goal to achieve due to the multidimensional and complex nature of these systems.

In [2]-[4], the authors have dealt with the reasons and some consequences of the complex nature of education systems and their nonlinear behavior.

This complexity is related to the existing interconnections between the different levels of educational systems starting from kindergarten to reach high schools and universities. Low performance at one level tends to affect the global performance at subsequent levels. Furthermore, individual's performance in

[*]Corresponding Author: Najat Messaoudi, Email: najatm2013@gmail.com

https://dx.doi.org/10.25046/aj060137

schools is influenced by several other systems. Education is in fact part of a larger system whose elements, such as economics, culture, society, and politics, interfere with it.

Education systems are also time-dependent and, changes in education often take a long time. Implemented policies may show results many years later, sometimes decades after their implementation. Another issue is that behaviors and actions of one part of the educational system tend to affect other parts of the system. This complexity can lead to unpredictable consequences of implemented actions. Effective management of education systems, therefore, requires a global approach that oversees the effect of all the parameters considered as a whole to ensure an efficient system operation.

Modelling can be one of the useful methods to evaluate performance and to identify the factors or areas of improvement of this performance. Modeling a system consists of designing a representation of this system for analysis and simulation purposes to answer questions. There can be several models for the same system with different levels of precision depending on the phenomena to be studied. Identification of the appropriate model will depend on the questioning under consideration [5].

The use of complex and dynamic systems modelling to social systems is quite recent and its pertinence in the case of an educational system is continually increasing [2]. The main challenge is to develop a model that can capture the dominant dynamics of these systems while being at the same time, simple enough to be useful for analyzing, simulating, and quantifying the impact of different parameters on the global performances of these systems

Several works have focused on the modelling of education systems. They concerned areas such as design, operations, analysis of performance, and management.

In the field of production engineering, studies on education systems modelling have focused on developing an analogy between education systems and systems of production of goods and services. Once this analogy was established, methods and methodologies for production systems modelling were applied to education systems while adapting them to the specificities of the latter. This works has enabled the transfer of business process modelling tools to the field of education and thus provided models for various purposes such as quality assurance, computerization of processes, or the design of new systems.

In this way, in [6], the author used the SADT method (Structured Analysis and Design Technique) for the analysis and the functional decomposition of a school using Le Moigne's systemic approach. Based on CIMOSA (Computer Integrated Manufacturing Open Systems Architecture) and IDEF0 (Integrated Computer-Aided Manufacturing Definition), he proposed modelling of the skills production system and engineering design of this system.

In [7], a model for resource specification for the planning of training activities using UML (Unified Modelling Language) and a model of the business processes of an education system using MECI (Modélisation d'Entreprise pour la Conception Intégrée) are proposed.

In higher education, in [8], the author proposed a model for quality improvement by identifying the skills that graduates must acquire and proposes the activities and means to achieve this objective by using business process modelling BPM.

In relation to the increasing use of digital technologies, the author in [9] proposed a personalized training path per learner according to his/her individual capacity to optimize academic performance using Petri nets.

In [4], the author investigated the application of dynamic systems modelling for educational policy analysis to a better understanding of the dynamics of the current system and the design of evolution scenarios for the future. He applied this approach during a case study during the school year 2007-2008 in the American state of Rhode Island.

Another class concerns the demographic models used for the simulation of educational policies and strategies. This is the case of the EPSSIM code (Education Policy and Strategy Simulation Model) [10] which is a generic model designed by UNESCO. It provides technical and methodological support for the elaboration of educational development plans. It is a demographic model in which schooling objectives are considered as decision variables and expenditures are calculated as a consequence of the achievement of these objectives. It essentially allows the students flow simulation and costs calculations.

These models, which remain partial, can simulate the behavior of the system for a reduced number of basic quantitative parameters and require a high degree of granularity to ensure their operationality and reliability of their results

This paper aims to present an educational system model for studying the impact of various parameters on the behavior of the system and to estimate its performance. The focus will be on the study of the impact of learners' competencies deficit on the performance of a training program.

2. Modelling of an education system

2.1. Modelling a training unit

A previous work [11], allowed us to propose a model of an education system that makes it possible to simulate its behavior to predict its long-term evolution, to evaluate its performance, and to simulate the impact of some factors on the evolution of this performance.

According to previous works which consider that an educational system can be defined as a competence production system [6] and also that any system which includes, even partially, in its missions the increase of learner skills can be considered as a part of an educational system [12], the approach conducted for the modelling of such system consists in proposing a decomposition of this system to introduce the notion of training unit which is the basis of the proposed modelling.

Despite the specificities of each country, educational systems at the global level are fairly homogenous and exhibit a quite similar central core. Therefore, most educational systems around the world have four educational levels: pre-primary, primary, secondary, and tertiary. The present simulation considers also that each training cycle, provided usually in an educational institution,

can be further divided into training programs, with each program composed of a limited number of courses consisting of a defined number of sessions.

Similarly, to the goods or services production systems, educational systems have organizational units, these are the training units. A training unit aims to transform business object inputs (can be physical or informational) into business object outputs within a framework of objectives and constraints fixed by the external conduct rules of the unit. A training unit can correspond to different parts; it can be a whole educational system, a training cycle, a training institution, or even a single classroom session. Thus, the reconstruction of the global system can be done by grouping a range of training units in parallel and/or in series according to well-defined rules.

To model the training unit, we relied on models based on BPM tools and mainly the work related in [13] where the author suggests using a generic processor to characterize the activities of both the physical and decision systems to evaluate its performance. A generic processor is an entity that performs one or more basic transformation activities, and that fits into the structure, thus modifying its behavior and performance. By organizing a set of generic processors in a network, it creates a particular state of the production system which is characterized by a performance level.

Thus, the training unit is an entity of the training system with the objective to transform the inputs competencies' level of the learners into a higher level of output competencies respecting the constraints and objectives of the educational system and thus to contributes to the achievement of the overall objectives of the training system. The reconstruction of the global system will then be reconstructed by grouping together a set of training units, either in parallel or in series according to well-defined rules.

Thus, the objective of a training unit is to increase the level of competencies of learners starting from an input reference competency C_r, to a level above an output reference competency C'_r on a fixed time h, by using resources and based on conduct and achievement data as shown in Figure 1.

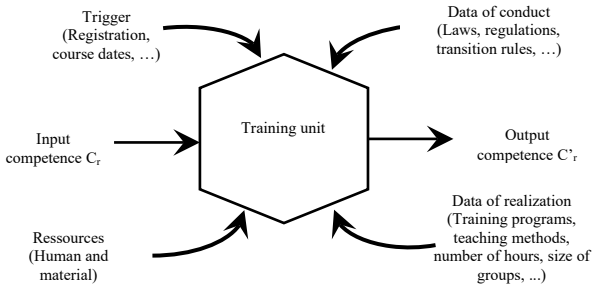


Figure 1: Model of a training unit

Educational systems present a particularity insofar as learners are co-producers of the service they use. In fact, each student has his own sensitivity to the process of skills acquisition. The objectives to be achieved by a training unit are defined during the design of this training unit. In this way, the reference input competencies needed to correctly follow this training activity, the reference output competencies to be reached, the resources required to carry out the activity, the conduct of data, and the achievement data are defined.

The interest of this modelling lies in the possibility of decomposing or aggregating the education system into networks of generic processors of different levels and by extension the possibility of evaluating the global or detailed performance of the training system on any level of abstraction.

The educational system is modeled as a series of training units interacting with each other to produce the targeted competencies.

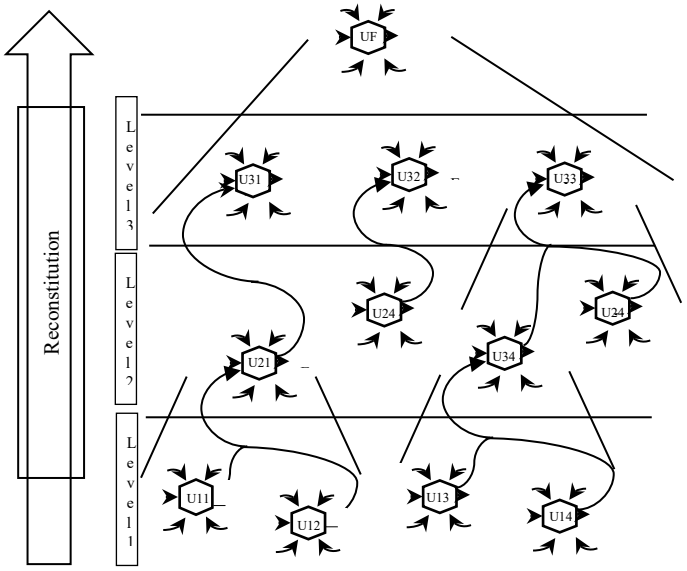


Figure 2: reconstitution of training units into a training program

Figure 2 illustrates an example of aggregating a set of training units by levels to reconstruct a training program. Each level groups together a defined set of training units and each training unit aims to achieve a defined increase in learner's competencies. Resources to be mobilized for a training unit to carry out its activity are defined, as well as related data of realization, data of conduct, and the launcher of the activity of the processor. The rules relating to the access of a learner from one level to the next level are also fixed, thus determining the time evolution of the training program.

This model allows the identification of what the training unit does, its function, as well as the determining factors for its functioning. Subsequently, the current modelling is associated with a simulation model which is built on the educational production function [14] and the learner's theories' results [15] in order to simulate the behavior of the system over time and the effect of some factors on its performance in a quantitative manner.

2.2. Mathematical model a training unit

The training unit is associated with a mathematical model based on the educational production function, that is written in a general form as follows [14]:

$$A = f^*(X1, \ldots, Xn) \quad (1)$$

where A is some measure of output, expressed as a function of a set of variables X1…..Xn.

In [11], the output A was the increase of competence level of a learner, noted as ΔC_i, while the parameters X_i that will be taken into consideration, at first approximation, are the duration of the training unit, the personal capacity of each learner μ_i, the input competence which represents the characteristics of the learners, the class size and a score for the pedagogical methods which represent some parameters of the resources and the data of conduct. The

other parameters remain constant throughout the simulation and therefore they will not impact the variation of ΔC_i.

Thus, expression (1) may be expressed as follows:

$$\Delta C_i = f(h, \mu_i, k, g) \quad (2)$$

where ΔC_i is the competence level's increase of the learner i, h the training unit's duration, μ_i the individual capacity of competence acquisition of the learner i, k the coefficient representing the gap between the input competence of learner i and the required input competence of the training unit; g the coefficient taking into account the effect of the class size and the pedagogical methods.

An approach to specifying the function f and the relation linking the parameters h, μ_i, and g to the output ΔC_i to carry simulations is presented in previous work [11].

Expression (2) becomes as follows:

$$\Delta C_i = \frac{(C_r' - C_r)}{2} \cdot \left[1 - \frac{erf\left(\sqrt{\alpha}\cdot\left(-h+\frac{h_r}{2}\right)\right)}{erf\left(\sqrt{\alpha}\cdot\frac{h_r}{2}\right)}\right] . \mu_i . g(n) . k\left(\frac{C_i}{C_r}\right) \quad (3)$$

Where C'_r is the output competence's reference, C_r is the input competence's reference, h the training unit's real duration, h_r the training unit's reference duration, α a parameter characterizing the progressivity of the evolution of the level of competence during the training.

The coefficient μ_i, which is specific to each learner, expresses his ability to achieve the expected competencies with more difficulty (μ_i<1) or more ease (μ_i> 1) than was projected in the design of training. The reference value for μ_r is 1.

For the coefficient g, we assume, in a first approximation, that a change in excess (resp. in default) of the class size is not beneficial (resp. beneficial) to the learner, so we can use a linear relation to model this coefficient as follows:

$$g(n) = b \text{ if } n > n_{max}$$

$$g(n) = b+((a-b)/2)*n \text{ if } n_{min} < n < n_{max} \quad (4)$$

$$g(n) = a \text{ if } n < n_{min}$$

where n is the ratio of the number of students in the group to the number recommended by the pedagogical method, a and b are the maximum and minimum values taken by the function g(n) when the group size is above a value n_{max} or below n_{min}.

Note that for the simulation results that will be presented here, we have used another relation instead of (4), which seems more adapted even if it presents similar characteristics, it is given by:

$$g(n) = b + \frac{(a-b)}{(1+n^4)} \quad (5)$$

A possible lack of the input competence C_i is taken into consideration through a factor $k(C_i/C_r)$ which reflects the difference between the learner's input competence C_i and the reference input competence C_r required by the training unit.

2.3. Performance of a training unit and a training program

Following [11], performance can be defined as the ratio of the actual result obtained as process output to an expected result under standard conditions to be defined.

The training unit's result is determined as all competence increases accomplished by the students and validated following the outbound validation rules.

The expected standard result can be considered as the sum of the increases in the levels of competencies that are assumed to be validated by the learners when all the parameters of the production function of the training unit defined by relation (3) are taken equal to the reference values of the parameters.

Therefore, the training unit performance can be expressed as follows:

$$P = \frac{\sum_{i=1}^{n}(C_i' - C_i)}{(C_r' - C_r)*n_r} \quad (6)$$

where P is the performance of the training unit, (C'_i - C_i) the competence level increase of learner i having successfully validated the training unit, n is the number of learners who validate the unit, and (C'_r-C_r)*n_r is the expected increase in the competence level of the class when all parameters are equal to their reference values.

A training program is considered as a network of interconnected training units distributed by the level of training with rules for transition from one level to another. The training units of a training level are connected in parallel and form a block and each training unit performs a defined competence increase activity. Thus, a training program that consists of n blocks, each block consisting of k training units, is schematized by a network of m processors such that m=n*k. The training units are denoted U_{ij} where i (1 ... n) denotes the number of the block to which it belongs and j (1 ... k) its classification within the block.

To define the performance of the training program, the performance of the training blocks making up the training program must first be defined.

We can then define the performance of block i of training units P_i as follows:

$$P_i = \frac{\sum_{j=1}^{k}\sum_{x=1}^{g_{ij}}(C'_{xij} - C_{xij})}{\sum_{j=1}^{k}(C'_{rij} - C_{rij})*g_{rij}} \quad (7)$$

where $\sum_{j=1}^{k}\sum_{x=1}^{g_{ij}}(C'_{xij} - C_{xij})$ is the sum of all the increases in competencies of the learners who have validated the training units that constitute the training block, $\sum_{j=1}^{k}(C'_{rij} - C_{rij}) * g_{rij}$ is the sum of all the increases in competencies expected by the training units when the other parameters values are those of the reference data.

The performance of the training program can be calculated as the average performance of the training blocks weighted by the respective numbers of learners and can be translated by the following expression:

$$P = \frac{\sum_{i=1}^{n} P_i * g_i}{\sum_{i=1}^{n} g_i} \quad (8)$$

where, P_i is the performance of training block i, g_i is the overall number of learners in block i, and n the number of blocks in the training program.

Thus, a model based on a mathematical formulation and a numerical simulation is obtained, it allows the calculation of the performance of a training program, based on the performance of the training units constituting this program. The effects on the overall performance of the training program of the variation of different parameters can thus be simulated and quantified.

This model was used in [11] to show the effects of the individual capacity of the learner's competence acquisition, as well as the pedagogical method's sensitivity in association with the group's size on the performance of a training unit.

The objective of this paper is to analyze the effect of the learner input competencies deficit factor on the behavior of a training unit and its performance and also to extend the approach to the case of a training program.

3. Impact of the input competence factor: results and discussion

For the chosen model, the education system is approached as an organized whole for skills production, composed of training units linked in parallel, and series according to rules of arrangement. Each unit has its own training objective, which is to increase specific skills of learners from a given level of competence at the input, assumed to be achieved by all entrants to the unit, to a higher level of competence at the output. The specificity of such a system is that the learners are co-producers of the service they use, which means that their characteristics will influence this increase in competencies. In fact, not all entrants to training have reached the required input level of competence and do not also have the same aptitudes for the acquisition of new skills.

Several studies have shown the important role of the learner in the process of competencies acquisition [16]-[18].

In a study using a multilevel model to identify the most important variables affecting students' performance [16], the authors have found evidence that both the students' and their families' characteristics play a role in the former's performance.

As stated in [19], precedent results and skills already acquired are strongly linked to the ease with which students will develop new skills, and thus have a real effect on their academic performance. In [20] and [21], the authors highlighted the importance of prior knowledge for success in studies.

In previous works [11], the personal characteristics of the learners are incorporated into the model through two factors. The first factor is the "personal capacity for competence acquisition μ_i" factor, which is specific to each learner and for each training unit and which is dedicated to taking into account the specific aptitudes of a learner allowing him to more easily assimilate new content. The impact of this factor on the increase in competencies and on the performance of a training unit was analyzed.

The second factor is the learner's prior knowledge through a "k" factor which quantifies the inadequacy between the student's input competence and the competencies entry level required by the training unit. In what follows, we will focus on the role of the latter, by establishing the effect of the factor k on the development of student's competence, on the training unit's performance, and on the training program's performance.

3.1. Evolution law of the coefficient $k(C_i/C_r)$

The goal sought is to establish an expression of the factor k which includes in the model defined by equation (3) the fact that a learner who would have a lack of competence at the entry of the training unit is less efficient in achieving the expected skills increase during the training. For that, we introduce the following expression:

$$k\left(\frac{C_i}{C_r}\right) = \left(\frac{C_i}{C_r}\right)^2 e^{\left[1-\left(\frac{C_i}{C_r}\right)^2\right]} \quad (9)$$

Equation (9) allows a variation of the coefficient k between the values 0 and 1, with k = 0 when $C_i = 0$ and k = 1 for $C_i = C_r$. This reflects the fact that a learner with a zero value of input competence, cannot progress through the training, and a learner who has required competence at the entry should achieve the expected level of competence at the output. While if a learner has achieved a level of competence C_i above the reference input level C_r, the benefit of training is less and allows only a small increase of competence.

The graph below shows the evolution of the factor k as a function of the ratio C_i/C_r.

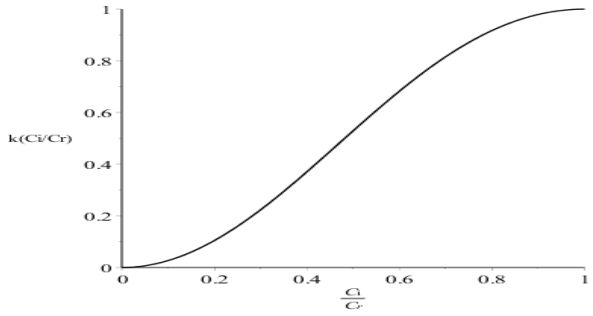


Figure 3: Evolution of the coefficient k as defined to the expression (9)

The mathematical model of the training unit (3), coupled to (9), is then given by the following expression:

$$\Delta C_i = \frac{(C_r' - C_r)}{2} \cdot \left[1 - \frac{erf\left(\sqrt{\alpha}\cdot\left(-h+\frac{h_r}{2}\right)\right)}{erf\left(\sqrt{\alpha}\cdot\frac{h_r}{2}\right)}\right] \cdot \mu_i \cdot \left(b + \frac{(a-b)}{(1+n^4)}\right) \cdot \left(\frac{C_i}{C_r}\right)^2 e^{\left[1-\left(\frac{C_i}{C_r}\right)^2\right]} \quad (10)$$

We will therefore consider three situations for the simulations. The first one shows the effect of the input competence on the competence increase's evolution, the second simulation concerns the effect of this factor on the training unit's performance and the third simulation reveals the effect of this factor associated with other factors on the training program's performance.

3.2. Simulation on the evolution of the increase of competence

The simulations to be carried out aim to determine the effect of the learners' skill level at the beginning of a training course on the progression of skills acquisition during the training process.

To illustrate this impact, we consider three groups of learners, the first group identified by a level of skills C_i below the required reference level C_r, the second having a skills level C_i equal to the required level, and the third having a level above the required level.

For the three groups, we consider all the other parameters of the simulation set to their reference values i.e. the training hours' number performed h is consistent regarding the reference duration of the training unit h_r ($h = h_r$), the students have a competence acquisition capacity $\mu_i = \mu_r$ and the coefficient g(n) is equal to 1.

Expression (10) above then allow simulation of the increase of competence in each of the three cases

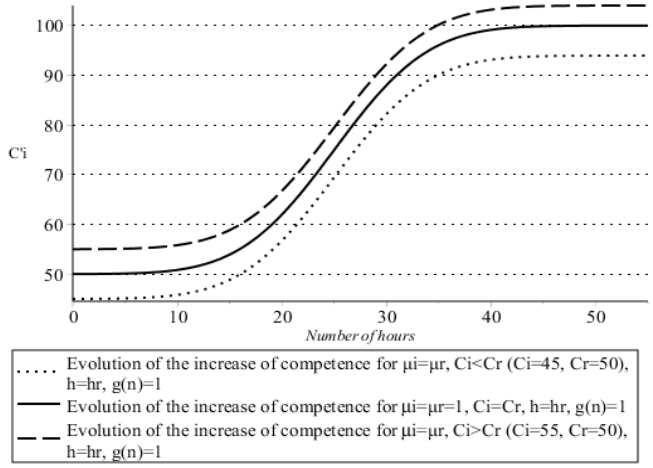


Figure 4: Input competence's effect on the competence increase

Figure 4 shows the input competence's effect on competence increase. The learner who starts the training with a deficit of competence's input (curve point), will not reach the expected competence's output. On the opposite, a learner with a competence's input greater than the required level accomplishes an output competence higher than the expected reference competence.

These results are in concordance with the studies [19], [20] and [21] that establish links between preceding results and skills previously acquired by the learners and the ease with which they acquire new skills

3.3. Simulation of the input competence's effect on the performance of a training unit

The goal set here is to simulate the effect of the variation in the input competence factor "k" on a training unit performance. Given the particularity of the educational systems in which every learner is a co-producer of his own competence increase according to his individual capacity of acquiring competencies μi, this factor is introduced in the following simulations.

Since the purpose is to focus on the input competence factor's variation, the other parameters are set to their reference values for the training unit, namely (C_r, C'r, h_r, n_r) as well as the training unit validation's rule that defines the minimum value of ΔC that the learner must reach to validate the training unit.

To realize the simulation, we consider, for a training unit, the following parameters reference values. The reference values are C_r=50 for the input competence, C'_r =100 for the output competence, h_r = 50 for the training unit duration, and n_r = 100 for the learner's number. We also consider that learner i validate the training unit if $(C'_i - C_i) \geq (C'_r - C_r)$. The simulation of the effect of the input competence on the training unit performance is executed according to the algorithm illustrated in Figure (5).

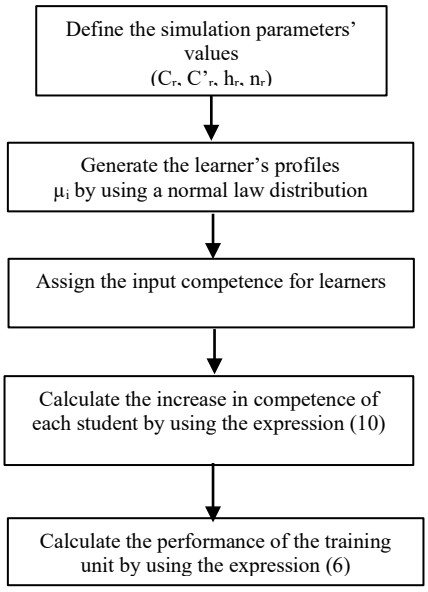


Figure 5: The simulation's algorithm

The first step for the simulation is to define the simulation parameter's values, the used values are n_r=100, C_r =50, C'_r =100 and h_r =50, and the training unit validation's rule that determine the minimum value required for ΔC_i that the learner has to reach in order to validate the unit. The second step is the generation of the student's profiles by using a normal law distribution for the parameter μ_i. The third step consists, for every simulation, in assigning the input competence C_i for the students. The fourth step is the calculation by relation (10) of the increase in competence for each learner. In the end, we calculate the performance of the training using relation (6).

Four simulations are carried out by varying the learners' input competence. In the first case, the input competence of learner's C_i is equal to the reference input competence of the training unit C_r, In the second case, the C_i deficit is of -3% compared to C_r, for the third case the C_i deficit is of -5% compared to C_r and the for the last case the C_i deficit of -7% compared to C_r.

The simulations' results are illustrated in the graph below:

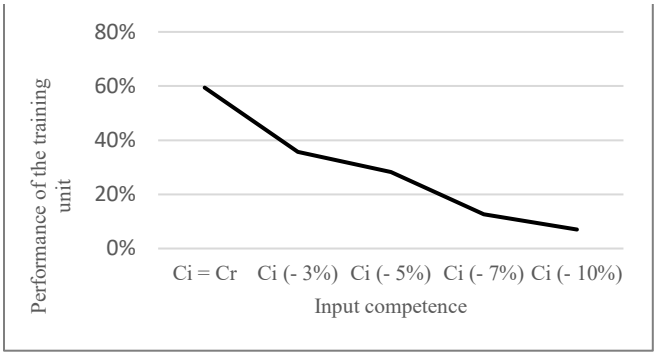


Figure 6: Impact of the input competence deficit on the performance of a training unit

These results highlight the effect of the deficit of learners' input competencies on the training unit performance.

We note that when all learners have the required input competence to follow the training unit, i.e. C_i=C_r, the performance of the training unit is around 59%. This percentage is consistent

and is due to the distribution in normal law retained for the personal capacity of competence acquisition μ_i, which is an essential factor of the learner. If these learners have an input competence deficit of -3% compared to the reference input competence, then this performance decreases and becomes of the order of 36%. Therefore, the improvement of a training unit performance, requires to check the input competence of a learner, so that the follow up of the competence acquisition process can be done properly.

3.4. Simulation of the impact of the input competence on the performance of a program training

The purpose is the stimulation of the input competence's impact and other parameters as the individual capacity of competence acquisition μ_i the class size on the performance of a training program and monitor its behavior over time.

The organization of the training program is illustrated by the following figure:

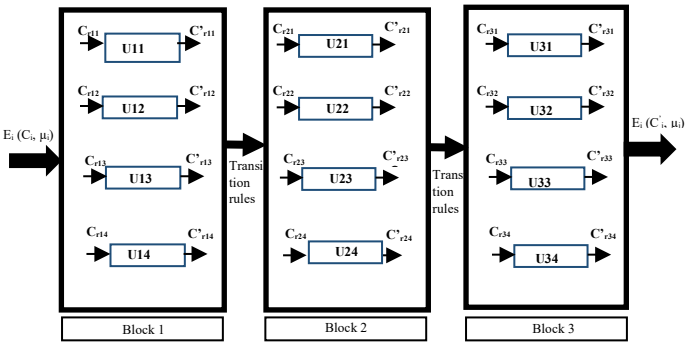


Figure 7: Organization of a training program

The training program for the simulation consists of three training blocks and each training block consists of four training units. The data of the realization of the training units are given in Table1.

Table 1: Training unit realization data and mobilized resources initialized to reference data

Level of training 1	Reference input competence	Reference output competence	Reference duration	Reference group size	Number of hours realized	Real group size	Number of groups
U 11	50	100	50	25	50	25	2
U 12	50	100	50	25	50	25	2
U 13	50	100	50	25	50	25	2
U 14	50	100	50	25	50	25	2
Level of training 2	**Reference input competence**	**Reference output competence**	**Reference duration**	**Reference group size**	**Number of hours realized**	**Real group size**	**Number of groups**
U 21	100	150	50	20	50	20	2
U 22	100	150	50	20	50	20	2
U 23	100	150	50	20	50	20	2
U 24	100	150	50	20	50	20	2
Level of training 3	**Reference input competence**	**Reference output competence**	**Reference duration**	**Reference group size**	**Number of hours realized**	**Real group size**	**Number of groups**
U 31	150	200	50	15	50	15	2
U 32	150	200	50	15	50	15	2
U 33	150	200	50	15	50	15	2
U 34	150	200	50	15	50	15	2

For the simulation, we consider that 100 learners are enrolled per year for level 1, 80 for level 2, and 60 for level 3.

The algorithm in Figure 8 illustrates how the simulation is performed

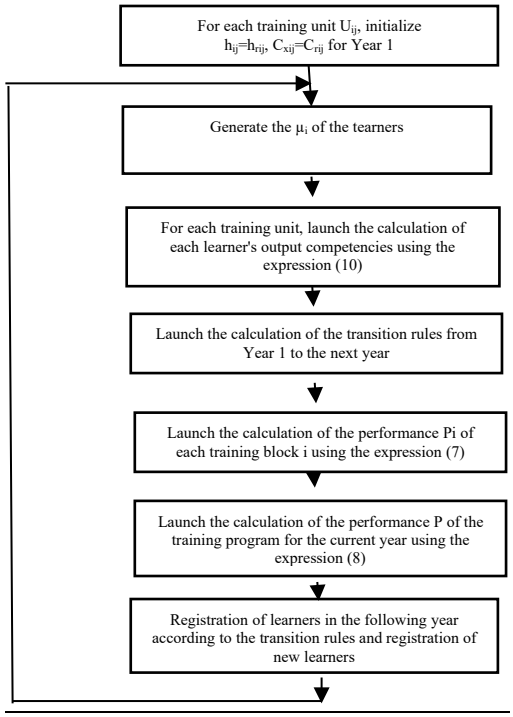


Figure 8: Algorithm for simulating the performance of a training program

The simulation is performed for seven iterations up to Year 7 of the training program. The objective is to track the behavior of the variation in performance over the seven years of training. The simulation is carried out with a constant number of new learners NI entering Level 1 for each year of study.

Figure 9 below is an illustration of the simulation's result.

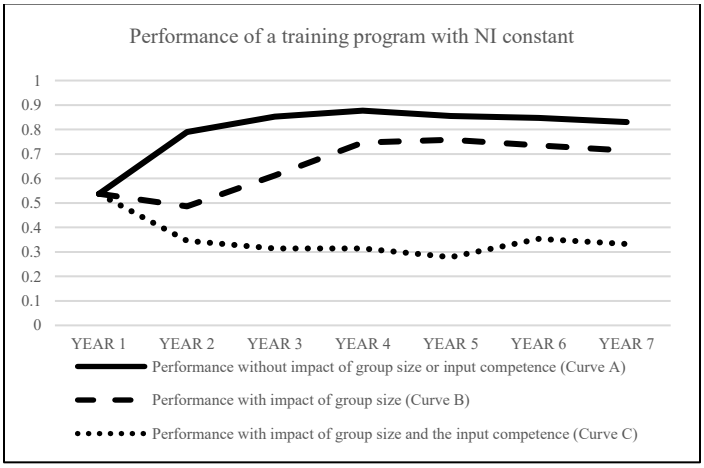


Figure 9: Simulation of the performance of a training program with NI constant

Figure 9 illustrates the variation in the performance of the training program on three different curves. The first curve (Curve A with continuous line) represents the impact of the personal capacity of competence acquisition μ_i. In this case, the coefficients of the group size g(n) and the input competencies $k(C_i/C_r)$ are kept equal to 1. The second curve (curve B with dashed line) represents the impact of the personal capacity of competence acquisition μ_i

and the coefficient of the group size g(n). The factor of the input competencies $k(C_i/C_r)$ is kept equal to 1. The third curve (Curve C with dotted line) reflects the impact of the personal capacity of competence acquisition μ_i, the coefficient of group size g(n), and the input skills factor $k(C_i/C_r)$.

For curve (A), in the first year of the training program, the performance of the training program is close to 50%. This is due only to the impact of the learners' μ_i factors, which are distributed according to a normal distribution whose average is 1. Thus, nearly 50% of the learners have a μ_i factor less than 1 and therefore cannot reach the required output competencies. For year 2 this performance increases to nearly 80%. This is explained by the fact that the learners who move up to the higher levels in year 2 satisfy the required input competencies and their μ_i are greater than or equal to 1, which allows them to reach the output competencies. Moreover, for learners at the different levels who fail and repeat some training units, they validate these units since they have re-enrolled with their output competencies from year 1 as input competencies for year 2 that are higher than the required competencies, thus making up for their μ_i deficit. The 20% performance deficit is mainly due to the new learners who enrolled in level 1, 50% of whom have a $\mu_i<1$ factor. Beyond this year, performance stabilizes at just over 80%. Thus, under the conditions of this simulation, i.e., without considering the effects of group size and the deficit in the input competence, the impact of the μ_i factors is offset by the increase in competencies of learners with a $\mu_i<1$ for the second enrolment in the same training unit.

For curve (B), in addition to the impact of μ_i, the impact of the group size is included. By comparing curve (A) and curve (B), the impact of group size is clearly apparent with a decrease in the performance of Year 2 of the training program. In fact, only μ_i has an impact on the performance of year 1. Thus, learners with a μ_i <1 in year 1 are re-enrolled in the same training units in year 2, thus increasing the group size of these training units and consequently this year's performance is impacted by this increase in the number of learners. From year 3 onwards, performance increases and then stabilizes at the end of year 5. This is due, on the one hand, to the increase in the input competencies of learners who repeat a given level, which partially compensates for the impact of μ_i and group size, and, on the other hand, to the stabilized number of learners in the higher levels.

For curve (C), in addition to the impact of μ_i and group size, the impact of learners' input competencies is included. Comparing this curve with the other two curves, we see that the factor of learner input competencies negatively affects the performance of the training program which stabilizes at the end of year 7 at a rate of 40%.

The results of this simulation allow us to compare and visualize the behavior and variation in the performance of a training program under the effect of the different simulation factors. These results also make it possible to quantify this variation in performance. Indeed, by acting on the learners' input competencies through specific training actions for their improvement, the performance of the training program can be improved and go from a rate of 40% to a rate of nearly 80%. If we also act on the size of the group, performance can be further improved. Another performance improvement lever lies in the improvement of the learners' personal capacity to acquire competencies. This simulation is carried out with a constant number of new learners who enroll in the training program each year.

Another simulation is carried out to study the effect of a change in the number of new registrants at the start of the program in year 1. We will consider three situations, the one for which the number of entrants is 100, then the one where this number is increased by 20% then the one where it is decreased by 20%

The results of this simulation are shown in the following graph:

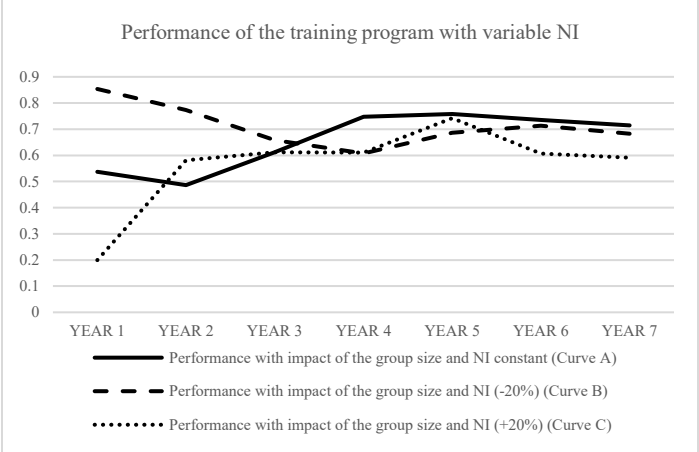


Figure 10: Simulation of the performance of a training program with variable NI

Figure 10 illustrates the variation in the performance of the training program along three different curves. The first curve (curve A with a continuous line) represents the impact of the personal capacity of competence acquisition μ_i and the group size with the new number of learners conforming to the reference number of the block of level 1. The factor of the input competence $k(C_i/C_r)$ is kept equal to 1. This curve corresponds to curve B of the previous simulation. This curve is considered as the reference curve for this simulation. The second curve (curve B with a dashed line) represents the impact of the μ_i and the coefficient of the group size g(n) with a number of new learners at the entrance 20% lower than the reference number. The factor of the input competence $k(C_i/C_r)$ is kept equal to 1 and, the third curve (curve C with dotted line) represents the impact of μ_i, the coefficient of the group size g(n) with a number of new entrants 20% higher than the reference number. The factor of the input competence $k(C_i/C_r)$ is kept equal to 1.

The performance of the training program in the start-up year "Year 1" varies according to the number of learners at the entrance. When the number of new entrants is reduced (curve B), the group size factor g(n) becomes greater than 1, which compensates for the effect of μ_i and therefore the performance of the training program increases, whereas when the number of new entrants is increased (curve C), the group size factor g(n) becomes less than 1, which contributes to the impact of μ_i in reducing the performance of the training program.

Whereas in the second "Year 2", the results are reversed. Indeed, in the case of curve B, given the number of learners who validated in year 1, the group size of the different blocks in year 2 increased and consequently g(n) becomes less than 1 and impacts the increase in learners' competencies and subsequently performance. The performance follows the same behavior until year 5, after which the performance increases, stabilizes, and approaches the performance of curve 1.

As for the behavior of the performance in the case of curve C, it follows the opposite behavior than that of curve B and in the stabilization phase, from year 5 onwards, the performance stabilizes at a rate of 60%, i.e. a decrease of 10% compared to the reference curve A.

These simulations on the performance of a training program illustrate the impact of the factors selected on this performance and make it possible to estimate, in a quantitative manner, the effects of variations in these factors on performance taken separately or in combination. The results of these simulations provide a better understanding of the interaction between the parameters and their impact on performance and can be used to test solutions for performance improvement.

4. Conclusion

By viewing education systems as skills production systems and by applying Business Processing Modelling methods, a modelling of education systems is proposed in the present work. The breakdown of the whole system allows us to introduce the training unit concept and then to organize these units in networks with units connected in series and parallel.

The objective of each training unit is to transform the learners' level of competencies into higher-level output competencies while respecting the objectives and constraints set by the educational system. Each training unit thus contributes to the achievement of the general objectives of the training system. This is achieved by mobilizing resources and according to rules of realization and rules of conduct.

Thus, the model proposed for the training unit has made it possible to determine the key factors influencing its performance. This model was coupled with a simulation model that allows to study the impact of the variation of these factors on the evolution of the system's behavior and performance and to simulate this impact.

The presented results of simulations show the effects of the learners' input competence deficit on the evolution of the process of competence increase and consequently on the performance of the training unit. Thereafter, the impact on the overall performance of a training program of this factor cumulated with those of the factors of the learners' personal capacity to acquire competencies, and the group size is studied.

This modelling allows us to evaluate the performance of the training unit and the training program in a quantitative way and to simulate the behavior of this performance by varying the different factors. This modelling enables us to test solutions for performance improvement.

Let us note that a validation of the model on case studies remains to be carried out and will require in particular the establishment of strategies for identifying the various parameters involved in the laws introduced to characterize the learner's aptitudes.

References

[1] T. McCowan, "The Role of Education in Development," Higher Education for and beyond the Sustainable Development Goals, Palgrave Studies in Global Higher Education, 2019, https://doi.org/10.1007/978-3-030-19597-7_2

[2] N. Ghaffarzadegan, R. Larson and J. Hawley, "Education as a Complex System," System Research and Behavioral Science, Wiley Online Library, 2016, DOI: 10.1002/sres.2405.

[3] M. Koopmans, D. Stamovlasis, Complex Dynamical Systems in Education: Concepts, Methods, and Application, Springer, 2016. DOI 10.1007/978-3-319-27577-2

[4] J. Groff, "Dynamic systems modeling in educational system: Design & Policy," New approaches in educational research, **2**(2), 72-81, 2013. DOI: 10.7821/naer.2.2.72-81

[5] F.B. Vernadat, "Techniques de modélisation en entreprise : Application aux processus opérationnels, " Economia, 1999.

[6] C. Clementz, Modélisation des systèmes de production de compétences: apports à l'ingénierie pédagogique, Ph. D Thesis, Université de Metz, 2000. https://hal.univ-lorraine.fr/tel-01775474

[7] O. Bistorin, Méthodes et outils d'aide à la conception des processus opérationnels des systèmes de formation, Ph. D Thesis, Université Paul Verlaine de METZ, 2007. https://tel.archives-ouvertes.fr/tel-01749013

[8] M. Dragna et D. Ivanaa, "Business Process Modeling in Higher Education Institutions. Developing a Framework for Total Quality Management at Institutional Level," Procedia Economics and Finance, **16**, 95-103, 2014. doi: 10.1016/S2212-5671(14)00779-5

[9] P. SeungSoon, "Applying Petri nets to model customized learning and cooperative learning with competency," International journal computer science and network security, **8**(2), 2008. DOI: 10.15693/ijaist/2016.v53i53.6-12

[10] UNESCO, "Education, Policy & Strategy: Simulation Model EPSSim Version 2.1." 2005.

[11] N. Messaoudi, J. Khalid Naciri and B. Bensassi, "Simulating effect of parameter variation on the performance of an educational system," International Journal of Simulation and Process Modeling, **13**(3), 189-199, 2018. DOI: 10.1504/IJSPM.2018.10014176

[12] O. Bistorin, C. Pourcel and P. Padilla, "Integrated model for operational process design of learning system," 8e Conférence Internationale de MOdélisation et SIMulation, Hammamet, 2010.

[13] H. Kromm, J. Deschamps, and G. Doumeingts, "Modélisation de processus pour une évaluation par niveaux de détail successifs, " in 3e Conférence Francophone de Modélisation et de Simulation MOSIM'01, 2001.

[14] S. Bowles, Towards an educational production function, National Bureau of Economic Research, 1970. http://www.nber.org/chapters/c3276.pdf

[15] M. Benko, "On the Status of the Model of Growth of Knowledge by Repetition on the Field of Didactics," Nitra, 2013.

[16] F. Mourji and A. Abbaia, "Les déterminants du rendement scolaire en mathématiques chez les élèves de l'enseignement secondaire collégial au Maroc : une analyse multiniveaux," Revue d'économie du développement, **21**(1), 127–158, 2013. DOI : 10.3917/edd.271.0127

[17] J.D. Williams and M.Somers, "Family, classroom, and school effects on children's educational outcomes in Latin America," School Effectiveness and School Improvement, **12**(4), 409–445, 2001. DOI: 10.1076/sesi.12.4.409.3445

[18] R.M. Felder and R. Brent, "Understanding student differences," Journal of Engineering Education, **94**(1), 57–72, 2005. DOI: 10.1002/j.2168-9830.2005.tb00829.x

[19] R. J. Marzano, Building background knowledge for academic achievement, Alexandria: Association for supervision and curriculum development, 2004. ISBN 0-87120-972

[20] D. Muijs and D. Reynolds, "Student background and teacher effects on achievement and attainment in mathematics: A longitudinal study," Educational research and evaluation, **9**(3), 289-314, 2003. https://doi.org/10.1076/edre.9.3.289.15571

[21] B. Hemmings, P. Grootenboer and R. Kay, "Predicting mathematics achievement: the influence of prior achievement and attitudes," International Journal of Science and Mathematics Education, **9**(3), 691-705, 2011. https://doi.org/10.1007/s10763-010-9224-5

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 336-341 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Recent Impediments in Deploying IPv6

Ala Hamarsheh[*,1], Yazan Abdalaziz[1], Shadi Nashwan[2]

[1]*Computer Science, Arab American University, Jenin, 240, Palestine*

[2]*Computer Science Department, College of Computer and Information Sciences, Jouf University, Sakaka, 2014, Saudi Arabia*

ARTICLE INFO

Article history:
Received: 20 November, 2020
Accepted: 06 January, 2021
Online: 22 January, 2021

Keywords:
Internet Protocol Version 6
Tunneling
Translation
IPv6 Adoption
IPv6 transition techniques

ABSTRACT

Internet Protocol version 6 is being adopted on slow pace and it is taking a long time. This paper intends to discuss the transition process between IPv4 and IPv6 and the major obstacles that prevent deploying IPv6 worldwide. It presents the IPv4 exhaustion reports results and where are the IPv4 address pool. Then it presents the methods that have been used to prolong the life expectancy of IPv4. After that it describes and discusses the mechanisms that have been used to deploy IPv6. Additionally, it describes the recently proposed mechanisms to overcome the problems encountered by the ISPs in migrating to IPv6. Furthermore, it shows the mechanisms that have been proposed to motivate the ISPs to start deploying IPv6 on their access networks. Finally, it presents a comparison between these mechanisms from the authors' point of view.

1. Introduction

'Internet Protocol Version 4" (IPv4) [1] is the current considered primary Internet protocol. It uses a 32-bit address. That provides more than 4 Billion addresses. Due to the frequent use of the internet and the abundance of devices that need an Internet connection, IPv4 was drained. As a temporary solution, many mechanisms have been proposed to prolong the life of IPv4. In the end, there must be a permanent solution. Internet Engineering Task Force (IETF) proposed a solution in which to convert IPv4 to a new protocol which is the "Internet Protocol Version 6" (IPv6) [2].

IPv6 was created by the IETF as a successor protocol to IPv4. It was proposed in 1994 and it became an internet standard on 14 July 2017 [3]. IPv6 uses a 128-bit address which provides approximately 3.4 x 1028. That will fulfill the huge need for IP addresses. The problem is that IPv6 can't be deployed all at once because of the large number of users and devices. Also, the lack of readiness of the infrastructure may delay the process knowing that the users can't be forced to update or change their devices so that they can use IPv6. IPv4 and IPv6 are not compatible and they can't be deployed both at the same time. IETF has proposed many solutions to start deploying IPv6 alongside IPv4. The plan is to start integrating IPv4 with IPv6 until IPv6 is fully deployed.

This Paper is an extension of work originally presented in [4]. In this research paper, we will present the methods that have been used to conserve the public IPv4 addresses. Then we will present some reasons that prevent the IPv6 deployment process. After that, we will present the mechanisms that have been proposed to integrate IPv6 alongside IPv4. Finally, we will compare and discuss these mechanisms from the authors' point of view.

2. IPv4 Exhaustion

The "Internet services providers" ISPs get the IP addresses from the Regional Internet Registry (RIR). There are five RIRs each one is responsible for providing IP addresses in a particular area of the world. which are: "American Registry for Internet Numbers (ARIN), Asia-Pacific Network Information Center (APNIC), Latin America and Caribbean Information Center (LACNIC), Reseaux IP Europeens Network Coordination Center (RIPE NCC), and the African Network Information Center (AFRINIC)". A statement has been made by APNIC that they have been operating under the last/8 framework since April 2011. At that time, they declared that they have 11,028,480 addresses assigned as available and have 15,728,128 unassigned addresses. For more information about the other RIRs reports check [5].

3. Methods to Preserve "Public IPv4 Addresses"

Many mechanisms that have been proposed to prolong the life expectancy of IPv4 the first one is the "Network Address Translation (NAT)" [6]. In short, NAT is a translation mechanism that translates IPv4 public addresses to a group of IPv4 private

[*]Corresponding Author: Ala Hamarsheh, Jenin, P.O Box 240, Palestine, E-mail: ala.hamarsheh@aaup.edu

addresses. rapid growth of internet users and IP addresses consumption has affected on NAT and it is not useful solution anymore. However, IETF has proposed some mechanisms which are:" Large Scale NAT (NAT444)" [7], and "Address plus Port (A+P)".

3.1. Large Scale NAT (NAT444)

Large Scale NAT or NAT444 is a technology that has been considered by the ISPs to extend the life of IPv4. Using NAT444 ISPs have been able to provide the users with IPv4 during the transition to the IPv6 process. It works by adding a Carrier Grade NAT (CGN) to the ISP's network. It translates one public IPv4 address among various CPEs and the private address in the CPE will again be translated into many private IPv4 addresses to arrive at the end-user with two NATs. NAT444 has many disadvantages, for instance, users may affect each other on bad behaviors because they share one public IPv4 address among a large number of users. Also, losing geolocation information because translation zones will cross traditional geographic boundaries see [8]. VoIP and video applications might be affected by NAT444 with latency, or packet loss, etc. see [9].

3.2. Address Plus Port (A+P)

A+P is a sharing schema that allows multiple users on the same CPE to use the same public IPv4 address. It extends the IPv4 address by using some of the port numbers as additional endpoints identifiers in the TCP/UDP header. Port Range Router (PRR) runs the process of assigning the range of IP addresses among the CPEs. A+P schema can provide 65536 ports for each public IPv4 address and each port can provide an extra 65536 ports. The idea behind A+P is that it divides the public IPv4 address without the need for translating it see [10] for more details.

4. Transition Approaches

Since IPv6 and IPv4 are not corresponding to each other and deploying IPv6 cannot be done all at once, the IETF has proposed techniques and mechanisms in order to achieve a smooth transition to IPv6. The smooth transition can be achieved by deploying IPv6 at the same time as IPv4. To allow communications between IPv4-Only networks and IPv6-Only networks IETF has proposed the transition mechanisms. These mechanisms are Dual-Stack, Tunneling, and Translation Approaches.

4.1. Dual-Stack Approach

The dual-stack approach is providing the user's connectivity for both IPv4 networks and IPv6 networks. A dual-stack device is any device that connects to the internet and configured with multiple IP addresses IPv4 and IPv6 at the same time. The idea behind the dual-stack approach is to configure all devices with both IPv4 and IPv6 addresses so that the devices can communicate over both networks [11]. It uses the perspective protocols (DHCPv4, DHCPv6) [12], [13] to assign addresses to the dual-stack devices. The network administrator is responsible for enabling perspective protocols. See Figure 1 which describes the Dual-Stack approach.

4.2. Tunneling Approach

The tunneling approach is a protocol that allows the movement of data from a network to another through a different type of network. IPv4/IPv6 Routers and Hosts encapsulate the IPv6 packet in an IPv4 packet and send it through an IPv4 network to another IPv6 network. Tunneling can be used in various ways for instance Router-to-Router, Host-to-Router, Host-to-Host, and Router-to-Host. Because the hosts and routers need to explicitly configure the tunneling endpoints Router-to-Router is probably used. Figure 2 illustrates the Tunneling approach.

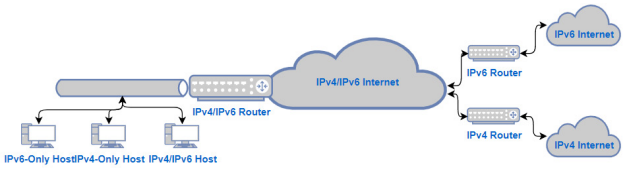


Figure 1: Dual-Stack Approach

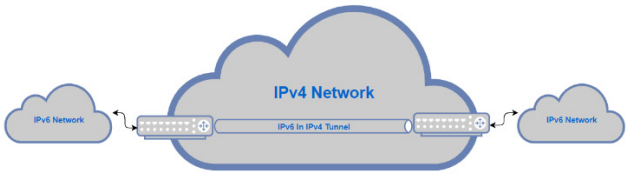


Figure 2: Tunneling Approach

Tunneling has two types which are: Static Tunneling and Automatic Tunneling.

- Static Tunnels: in static tunneling the address will be configured at the tunnel endpoint and the configuration in this technique is manual. The encapsulating node is the one responsible for storing the tunneling information. It stores the tunnel endpoint address which will be the destination address. in addition, it stores the routing information in order to determine which packet to tunnel. Also, it uses the match technique and the prefix mask to the destination to the packet [14]-[16].
- Automatic Tunnels: in automatic tunneling the IPv6/IPv4 nodes can automatically determine the tunnel endpoint and That can be extracted from the IPv6 address. definitely the IPv6 must be backward compatible with IPv4. Also, the nodes possess the ability to decide which packets are auto tunneled and which are not. The main difference between the dynamic and static tunneling is the automatic determination of the tunnel endpoint.

4.3. Translation Approach

The previous approaches could be useful when communication is needed between two isolated IPv4 networks or two isolated IPv6 networks, while they have no use when an IPv4 only network tries to communicate with an IPv6 only network. The translation approach is a mechanism that allows an IPvX-Only network to communicate directly with an IPvY-only network. It translates the IPvX packet to an IPvY packet to allow communication. Figure 3 describes the translation approach.

The translation mechanism can be classified into two approaches [11]: Host-Based Approach and Network-Based Approach.

- Host-Based Translation

  Host-based translation is needed when there is an incompatibility between the running application type and the current host connectivity. So that, the IPvX will be translated to communicate with the IPvY and vice versa. In host-based translators, the changeover is between the application layer and the IP communication layers. There are three host-based translators was proposed which are: Bump-In-the-Stack (BIS) [17], Bump-In-the-API (BIA) [18], Bump-In-the-Host (BIH) [19]. In the next section we will present the host-based techniques in addition to another technique the author was proposed which is Decupling Application IPv4/IPv6 Operation from the Underlying IPv4/IPv6 Communication (DAC) [20].

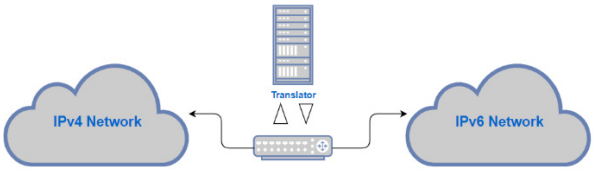


Figure 3: Translation Approach

- Network-Base Translation

  In network-based translation, the IP header will be translated to every packet. That to provide the connectivity between IPv4-only networks and IPv6-only networks and vice versa [11]. An example of network-based translators is Stateless IP/ICMP Translation (SIIT) [21], and Stateful Address and Protocol Translation from IPv6 Client to IPv4 Servers (NAT64) [22].

5. IPv6 is Not Fully Deployed Yet

Transition to IPv6 is indeed a big breaking change in the networks. It can't be done all at once for many reasons. This section intends to describe these reasons from the author's point of view.

5.1. End User

End-users are not directly benefiting from the transition process; however, it is considered as an effective part in achieving it. The end-user might pose an obstacle for deploying IPv6. As we discussed earlier the deployment process of IPv6 must be smooth. A smooth transition can be done without affecting the users taking into account the users' application and the host configurations should be upgraded. The end-user can delay the process knowing that not all users have the proper knowledge to update or implement the upgrades.

There are some upgrades that should be implemented on the IP stack, "Transmission Control Protocol (TCP)", and "User Datagram Protocol (UDP)" in addition to the users' old applications [11], [23], [24]. The host configuration process should be automatically performed without the user knowledge because the user can't be expected to be qualified to deal with the manual configurations or any other technical configurations.

Achieving the transition process smoothly could take a while till IPv6 is fully deployed. Within this time there will be IPv6 only networks and IPv4 only networks that need to be able to communicate with each other. The solution to that problem is the dual connectivity so that dual networks can communicate to both IPv4 and IPv6 networks [25]. The users' applications can be expected to be updated to support dual connectivity if it has good support but there are many applications that are not supported such as home-made applications and out-of-business applications and many other applications that will have the same problem.

There are many mechanisms that have been proposed as a solution to guarantee dual connectivity between incompatible hosts. IETF has proposed mechanisms such as "Bump In The Stack (BIS)" [17], "Bump In the API (BIA)" [18], [26], and "Bump In The Host (BIH)" [19]. Another mechanism that has been proposed by the author of this paper is "Decupling Application IPv4/IPv6 operation from the Underlying IPv4/IPv6 Communication (DAC)" [20], [27], [28].

- Bump in the Stack (BIS)

Bump-In-the-Stack is a technique that uses the SIIT [29] algorithm to translate the IPv4 packets (i.e. headers) to IPv6 and the IPv6 packets (i.e. headers) to IPv4. In BIS the translation module is inserted between the TCP/IPv4 module and the card driver module. The translation module takes the data that flow between the modules that it is inserted between and translates the packets from IPv4 to IPv6 and vice versa. The DNS server is responsible for assigning the IP addresses. So, the user does not know about the other types of IP addresses that it communicates with. The main idea behind BIS is to provide the communication between IPv4 applications and IPv6 hosts.

- Bump in the API (BIA)

Is a technique that used to translate the IPv4 socket API functions into IPv6 socket API functions and vice-versa. It detects the IPv4 socket API functions and invokes the equivalent IPv6 socket API functions so that there will be no need for translating the full IP header. BIA expects that TCP/UDP IPv4 and TCP/UDP IPv6 to existing on the local node. Unlike BIS, BIA inserts the API translator between the socket API module and the TCP/IP module in the dual-stack hosts.

- Bump in the Host (BIH)

BIH technique is a host-based mechanism. It integrates both BIS and BIA techniques together and translates from IPv4 to IPv6. BIH consists of two implementations which are a translator in the API socket which is inserted between the TCP/IP module and socket IP module the other implementation is the translation protocol which inserts it between the network card driver and the TCP/IP module. BIS intends to let the old applications which uses NAT to communicate with the IPv6 only applications.

- Decoupling Application IPv4/IPv6 Operation from the Underlying IPv4/IPv6 Communication (DAC)

DAC is a technique Allows applications that are compatible with IPv4-only stack and running on hosts with dual stack or with IPv6-only connectivity to communicate with IPv6 hosts. Additionally, it enables applications that are compatible with IPv6-only stack and running on hosts with dual stack and connectivity or with IPv4-only connectivity to communicate with IPv4 hosts. In theory, DAC is a layer that should be inserted between the

application layer and the IP communication stack on top of the API functions and it should perform the same functionalities that the API native functions perform. Translation from IPv4 and IPv6 and vice versa is only done on necessary basis. DAC will be effective between applications with an incompatible type and an incompatible networks connectivity. Furthermore, DAC can be installed at an IPv4 only host, IPv6 only host, and IPv4/IPv6 host.

5.2. Service Provider Network

The Internet Service Providers (ISPs) are the other effective part of the deployment process. as we discussed before the ISPs are cannot provide the users IPv4 addresses anymore and they need to deploy the IPv6 so that they can keep up with the internet users' swift increasing and provide the users with the required IP addresses. The transition process requires an upgrade or change of the infrastructure to one that can withstand the IPv6. That might be very expensive for the ISPs. IETF has proposed a technique that allows the ISPs to start deploying IPv6 over IPv4 networks. This technique is IPv6 Rapid Deployment on IPv4 Infrastructure (6rd) [30], [31].

6rd is an automatic tunneling technique was proposed to support rapidly deployment of IPv6 on existing IPv4 environment. In order to transmit the IPv6 traffic over IPv4 environment, 6rd uses encapsulation technique to transfer IPv6 packets through the IPv4 networks and it uses the ISPs' own IPv6 prefix other than a well-known prefix. Since it uses the network-specific prefix (NSP) each ISP can use its particular prefix that's mean the 6rd operational scope is limited to the ISP's domain. Moreover, the tunnels will be from the ISP's border relay (6rd gateway) to the customer edge (CE). see [30] and [32] for more details.

ISPs have faced determinations that hindered them from starting the deployment process of IPv6 through IPv4 infrastructure. One of these limitations is the need to upgrade or change the 6rd gateways (CE) which will affect them by raising the costs of deploying IPv6. In addition to the time that the process of Configuring the CEs would take after change and upgrade which might stall the process. Also, 6rd protocol upholds only one level of DHCPv4 between the CE and the border router. Never forget to mention the firewalls which might block the tunneled traffic that related to 6rd.

The authors of this research paper have suggested a few techniques that concurs the 6rd obstacles. These techniques are: "Deploying IPv6 Service Across Local IPv4 Access Network (D6across4)" [33], [34], "Configuring hosts to Auto-detect (IPv6, IPv6-in-IPv4, or IPv4) network connectivity (CHANC)" [35], and "Deploying IPv4-only Connectivity across Local IPv6-only Access Networks (D4across6)" [36]. The next section of this research will discuss these techniques and how they have overcome the limitations and obstacles that faced the ISPs before.

- Deploying IPv6 Service Across Local IPv4 Access Network (D6across4)

D6across4 is a mechanism that intends to motivate the ISPs to start the deployment process and start offering IPv6 services to their customers. It is a protocol mechanism to start deploying IPv6 service to host over the existing ISP's IPv4 network. D6across4 assigns an algorithmic mapping between the IPv4 and IPv6 addresses in the ISP network. It automatically resolves the tunnel server (TS) domain name to determine the tunnel endpoints. It consists of customers' hosts and one or more TSs. Same as in 6to4 it encapsulates the IPv4 packets then forwards them to follow the topology of the IPv4 in the ISP network.

While 6rd and other protocols need upgrade and change the 6rd gateways and the network devices in both ISPs' and end users' sides. D6across4 tries to reduce both cost and time of the deployment process of IPv6 which is one of the problems that faced the ISPs and that prevented them from starting. Even though, in order for D6across4 can encapsulate/de-capsulate traffic IPv6 packets through the local IPv4 networks, it needs to be installed at ISPs' side at the network components, and at the uses side particularly at the applications hosts. See for more details [33].

There are some limitations that could be found in D6across4 that could affect deploying IPv6. These limitations are: if D6across4 has been deployed on a large scall networks it could face some issues regarding the scalability and the performance issues. So, it only could be integrated on a relatively small network. Also, since the D6across4 protocol is a stateful operation, for each packet receival it needs to access the mapping tables that could be exhaustion and consider as an additional load and it would be higher than any other stateless protocol. In addition, using a well-known prefix might shorten IPSs' capability of managing and controlling their traffic. Just like 6rd, D6across4 could be affected by firewalls and its traffic might be blocked.

- Configuring hosts to Auto-detect (IPv6, IPv6-in-IPv4, or IPv4) network connectivity (CHANC)

Same as D6across4 CHANC is a mechanism to encourage the ISPs to start the deploying IPv6 through IPv4 environment. Unlike D6across4 CHANC protocol doesn't need any upgrading or changing in the CEs which represents the extra costs that cause the delay of the deployment process. CHANC protocol gives the ISPs the ability to offer IPv6 to their customers through IPv4 infrastructure with the minimum cost and automatically configure the end-user's host. In fact, CHANC automatically configure the hosts to give them the ability to automatically detect every connectivity with every type that their ISP provide and locating the relay server would also be automatic. Also, CHANC uses the HTTP-in-IPv4 mechanism for transmitting IPv6 traffic through IPv4 environment.

CHANC has many advantages that overcome the other proposed protocols which are:

- The configurations in the end-users' hosts are done automatically which ease the process instead of manual configurations.
- It reduces the costs because there is no need for changing or upgrading the IPv4-only networks on both ISPs' and end-users' sides.
- Because of the IPv4 only access network, the ISPs can immediately begin offering IPv6 without the need support IPv6.
- It prevents the firewalls from blocking the tunneled packets of CHANC. It encapsulates the packets in HTTP protocol which permit the access to the traffic through utmost all firewalls.
- It simplifies administrating IPv6 traffic by allowing every ISP to use the prefix of its own.

Table 1: Comparison between the previously presented techniques and mechanisms

Protocol	Category	Installed At	Functionality	Limitations
BIS	Translation-based	"between the TCP/IP module and the network card driver"	Permit communication between IPv4 only applications on dual stack machines and IPv6 hosts	- Does not work with multicast communication. - invalid for embedded addresses. - It cannot be combined with a secure DNS. - It cannot employ security overhead the network layer.
BIA	Translation-based	"between the socket API module and the TCP/IP module"	Permit communication between IPv4 only applications on dual stack machines and IPv6 hosts	- invalid embedded addresses. - Does not uphold multicast. - Difficulties in translating APIs because of IPv6 API's advance new features.
BIH	Translation-based	"between the TCP/IP module and the network card driver or between socket API module and the TCP/IP module"	Permit communications between IPv4 legacy application and IPv6 only hosts and dual stack host	- invalid for embedded addresses. - Does not uphold multicast. - Does not uphold all types of applications.
DAC	Translation-based	"between the communication stack and the application layer"	Permit communications for both IPv4 and IPv6 applications through IPv4 and IPv6 connections with both IPv4 and IPv6 type capable remote applications.	- invalid for embedded addresses.
6rd	Tunneling-based	between the ISP border relay and the customer edge	Permit IPv6 connectivity through ISPs' IPv4 environment	- Single point of failure. - requires upgrade and change in the CPEs. - Cannot cross firewalls.
D6across4	Tunneling-based	ISP side and end users' host	Permit IPv6 connectivity through ISPs' IPv4 environment	- Single point of failure. - cannot cross firewalls.
CHANC	Tunneling-based	ISP side and end users' host	Permit IPv6 connectivity through ISPs' IPv4 environment	- Single point of failure.

- CHANC is a stateless-based translation mechanism that uphold end to end address translucence while the communication is between IPv4 and IPv6.

The Table below presents a comparison between the techniques and protocols that have been presented in this research.

6. Summary

The internet development has led to exhaustion in the IPv4 addresses and most of the ISPs are running out of IPv4. That makes the deployment of IPv6 has become mandatory. IETF has offered various mechanisms to extend the usage of IPv4. The proposed mechanisms and protocols are considered as a temporary solution. Afterall, the deployment of the IPv6 is unavoidable. ISPs are the major factor that affect the deployment directly. They need to start upgrading and changing the infrastructure to support IPv6. Also, IETF has offered various techniques and protocols to encourage the ISPs to start offering IPv6 alongside with IPv4 or even IPv6 through IPv4 infrastructure. These techniques and protocols can be classified into three classifications: Dual stack, Tunneling, and translation. In addition, various techniques and mechanisms have been offered by the authors of this research and by IETF for instance BIS, BIA, BIH, DAC, 6rd, D6across4, and CHANC. These mechanisms intend to encourage the ISPs and facilitate the deployment process in order for them to start the smooth transition.

Conflict of Interest

Ala Hamarsheh, Yazan Abdalaziz, Shadi Nashwan declare that they have no conflict of interest.

References

[1] J. Postel, "Internet Protocol," California, USA: Internet Engineering Task Force RFC791, 1981.
[2] S. Deering, R., Hinden, "IP Version 6 Addressing Architecture," California, USA: Internet Engineering Task Force RFC 4291, 2006.
[3] S. Deering, R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," California, USA: Internet Engineering Task Force RFC 8200, 2017.
[4] A. Hamarsheh, Y. Abdalaziz, "Deploying IPv6 Service Across Local IPv4 Access Networks," in 2019 International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, April 3-4, 2019.
[5] G. Huston. IPv4 Address Report [Online]. 2018. Available: http://www.potaroo.net/tools/ipv4/index.html
[6] P. Srisuresh, M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations," California, USA: Internet Engineering Task Force RFC 2663, 1999.
[7] I. Yamagata, Y. Shirasaki, A. Nakagawa, J. Yamaguchi, H. Ashida, "NAT444," Internet Engineering Task Force Draft (work in progress), 2012.
[8] C. Donley, L. Howard, V. Kuarsingh, J. Berg, J. Doshi, "Assessing the Impact of Carrier-Grade NAT on Network Applications," California, USA: Internet Engineering Task Force RFC 7021, 2013.
[9] S. Bradner, J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices," California, USA: Internet Engineering Task Force RFC 2544, 1999.
[10] R. Bush, "The Address plus Port (A+P) Approach to the IPv4 Address Shortage," California, USA: Internet Engineering Task Force RFC 6346, 2011.
[11] A. Hamarsheh, M. Goossens, "A Review: Breaking the Deadlocks for

Transition to IPv6," IETE Technical Review, **31**(6), 405-421, 2014, doi: 10.1080/02564602.2014.950348

[12] E. Nordmark, R. Gilligan, "Basic Transition Mechanisms for IPv6 Hosts and Routers," California, USA: Internet Engineering Task Force RFC 4213, 2015.

[13] A. Durand, R. Droms, J. Woodyatt, Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion," California, USA: Internet Engineering Task Force RFC 6333, 2011.

[14] R. Gilligan, E. Nordmark, "Transition Mechanisms for IPv6 Hosts and Routers," California, USA: Internet Engineering Task Force RFC, 1996.

[15] Y. Abdalaziz A. Hamarsheh, "Analyzing The Ipv6 Deployment Process In Palestine", International Journal of Computer Network and Information Security(IJCNIS), **12**(5),31-45, 2020, DOI: 10.5815/ijcnis.2020.05.03.

[16] M. Al-Fayoumi S. Nashwan, "Performance analysis of SAP-NFC protocol," International Journal of Communication Networks and Information Security (IJCNIS), **10**(1), 125–130, 2018.

[17] K. Tsuchiya, H. Higuchi, Y. Atarashi, "Dual Stack Hosts using the "Bump-In-the-Stack" Technique (BIS)," California, USA: Internet Engineering Task Force RFC 2767, 2000.

[18] S. Lee, M-K. Shin, Y-J. Kim, E. Nordmark, A. Durand, "Dual Stack Hosts Using "Bump-in-the-API" (BIA)," California, USA: Internet Engineering Task Force RFC 3338, 2002.

[19] B. Huang, H. Deng, T. Savolainen, "Dual-Stack Hosts Using "Bump-in-the-Host" (BIH)," California, USA: Internet Engineering Task Force RFC 6535, 2012.

[20] A. Hamarsheh, M. Goossens, R. Alasem, "Decoupling Application IPv4/IPv6 Operation from the Underlying IPv4/IPv6 Communication (DAC)," American Journal of Scientific Research, Eurojournals Press, **14**, 101-121, 2011.

[21] X. Li, C. Bao, F. Baker, "IP/ICMP Translation Algorithm," California, USA: Internet Engineering Task Force RFC 6145, 2011.

[22] M. Bagnulo, P. Matthews, I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers," California, USA: Internet Engineering Task Force RFC 6146, 2011.

[23] Y. Abdalaziz, A. Hamarsheh, "Analyzing The IPv6 Deployment Process In Palestine," International Journal of Computer Network and Information Security(IJCNIS), **12**(5), 31-44, 2020, doi: 10.5815/ijcnis.2020.05.03

[24] S. Nashwan, "SAK-AKA: A secure anonymity key of authentication and key agreement protocol for LTE network," International Arab Journal of Information Technology (IAJIT), **14**(5), 790–801, 2017.

[25] A. Hamarsheh, M. Goossens, A. Al-Qerem, "Assuring Interoperability Between Heterogeneous (IPv4/IPv6) Networks Without using Protocol Translation," IETE Technical Review, **29**(2), 114-132, 2012, doi:10.4103/0256-4602.95384

[26] S. Nashwan, "SE-H: Secure and efficient hash protocol for RFID system," International Journal of Communication Networks and Information Security (IJCNIS), **9**(3), 358–366, 2017.

[27] A. Hamarsheh, M. Eleyat, "Performance Analysis Of Ain-Pt, Ain-Slt And Siit Network-Based Translators," in Advances on P2P, Parallel, Grid, Cloud and Internet Computing Proceedings of the 12th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2017), Lecture Notes on Data Engineering and Communications T. Springer, 10, Palau Macaya, Barcelona, Spain, 367-378, 2017, DOI: https://doi.org/10.1007/978-3-319-69835-9_35.

[28] S. Nashwan, "AAA-WSN: Anonymous access authentication scheme for wireless sensor networks in big data environment," Egyptian Informatics Journal, in press, https://doi.org/10.1016/j.eij.2020.02.005.

[29] Nordmark E, "Stateless IP/ICMP Translation Algorithm (SIIT)," Internet Engineering Task Force RFC 2765, 2000.

[30] W. Townsley, O. Troan, "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) Protocol Specification," California, USA: Internet Engineering Task Force RFC 5969, 2010.

[31] A. Al-Qerem, F. Kharbat, S. Nashwan, S. Ashraf and K. Blaou, "General model for best feature extraction of EEG using discrete wavelet transform wavelet family and differential evolution," International Journal of Distributed Sensor Networks, **16**(3), 1–21, 2020.

[32] R. Despres, "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)," California, USA: Internet Engineering Task Force RFC 5569, 2010.

[33] A. Hamarsheh, M. Goossens, R. Alasem, "Deploying IPv6 Service Across Local IPv4 Access Networks," in 10th WSEAS International Conference on TELECOMMUNICATIONS and INFORMATICS (TELE-INFO '11), pp. 94-100, Lanzarote, Canary Islands, Spain, 27-29, 2011.

[34] A. Hamarsheh, M. Goossens, "Exploiting Local IPv4-only Access Networks to Deliver IPv6 Service to End-users," International Journal of Computers and Communications, **5**(3), 2011.

[35] A. Hamarsheh, M. Goossens, R. Alasem, "Configuring Hosts to Autodetect (IPv6, IPv6-in-IPv4, or IPv4) Network Connectivity," KSII Transactions on Internet and Information Systems," **5**(7), 1230-1251, 2011, doi:10.3837/tiis.2011.07.002.

[36] A. Hamarsheh, "Deploying IPv4-only Connectivity across Local IPv6-only Access Networks," IETE Technical Review, **36**(4), 398-411, 2018, doi: 10.1080/02564602.2018.1498031.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 341-346 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Enhanced Power Utilization for Grid Resource Providers

Tariq Alwada'n[1,*], Thair Khdour[2], Abdulsalam Alarabeyyat[3], Ali Rodan[4,5]

[1]*School of Computing, Engineering & Digital Technologies, Teesside University, Middleborough, UK*

[2]*Computer Information Systems Department, Al-Balqa' Applied University, Amman, Jordan*

[3]*Software Engineering Department, Al-Balqa' Applied University, Amman, Jordan*

[4]*Computer Information System Department, the Higher Colleges of Technology, AL Ain, UAE*

[5]*King Abdullah II School for Information Technology, the University of Jordan, Amman, Jordan*

ARTICLE INFO

Article history:
Received: 26 August, 2020
Accepted: 31 October, 2020
Online: 20 November, 2020

Keywords:
Grid Computing
Multi-Agent system
FIPA
Grid Resource Broker
Jade

ABSTRACT

A grid is a system that can manage and organize services and resources that spread amongst different control domains, employ interfaces and protocols, and offer a high quality of services. The integration of Multi-Agent Systems (MAS) with a grid environment significantly affects grid performance. MAS is considered a suitable solution for open systems that modify frequently. Grid offers a wide range of resources for its users, and some of these resources might not be used or utilized for some time before any new jobs come to the grid for processing. Usually, these resources are accumulated in massive data centers to cater to the grid users' growing demand. This accumulated will lead to consuming a considerable amount of electricity for their operation. The ever-increasing usage of grid computing has led to an increase in electrical energy usage by massive servers in their data centers. In this paper, we have proposed an automated system composed of modern agents that can be used the power wastage resulted from inactive servers inside the data centers for a specific amount of time. The proposed technique depends on switching inactive virtual and physical machines to lower power positions (Sleep/Wakeup or switched off) while still preserving customers' performance requirements. The Automated system has been tested and evaluated using the Jade tools. The results show that the newly proposed method can reduce power wastage for inactive grid resources.

1. Introduction

This paper is an extension of work presented initially in the ITT2019 conference [1]. Grid computing is an improved technology that involves a pool of linked machines that be owned by numerous organizations in diverse locations to construct a distributed system. The grid system can be accustomed to working with complicated scientific and business problems. It is created to assist in the sharing of distributed and diverse resources and to simplify resolving a considerable volume of computing issues [2]. The essential purpose of grid computing is to improve reliability, cut the cost of computing, and enhance flexibility by changing computers from an object that we purchase and operate privately to an object managed by a third organization [3]. However, grid computing is encountering various challenges, such as finding the right resources and decreasing the number of discarded jobs [4], [5]. Figure 1 displays an example of a grid computing environment [6], where the grid may contain different types of resources which need a resource management system in order to coordinates them.

The grid structural design shown in Figure 2 contains the following components [7]: a grid portal, which is also named a grid interface. It is used by the authorized grid clients to access the grid providers (resources). The portal should hide the complexity of the grid from clients by employing an uncomplicated interface. In this manner, it enables the sorting of grid job requisites.

The second element is the grid broker, where this element is described as the core of grid computing. It performs critical jobs in creating an effective grid system by connecting the user demands to the grid providers to fulfill the clients' tasks. The broker is also responsible for reducing transmission delays and increasing the

*Corresponding Author: Tariq Alwada'n, Email: t.alwadan@tees.ac.ae

resource assigning jobs through resources by avoiding the depending on a certain one. The grid broker's primary role is to locate and reserve the most appropriate support for the users' tasks. The broker receives the requests from the clients and sends the job's description files to the clients' requirements is responsible for monitoring the running tasks and hands over the outputs to their related users. To process the grid users' requests, the broker asks the Replica Catalogue Information Service to locate the available resources and data in the grid.

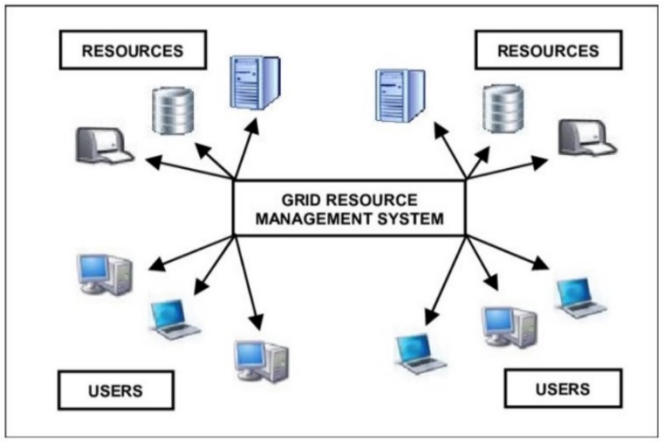


Figure 1: Grid Computing Environment

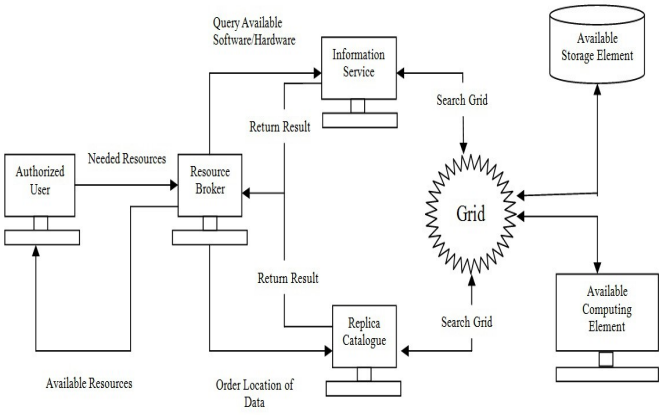


Figure 2: Grid Architecture. [8]

The third element is the Information Service. It is the directory service of the grid, where the grid keeps the information about all grid resources (nodes) and all the running jobs on those nodes. The grid resources have to register and update their information with the Information Services and their available resources. The grid broker retrieves all the information about the available and free resources from the Information Services. The information stored in the Information Services can be dynamic or static. The static information is employed to specify the operating system specifications and hardware requirements. In contrast, the dynamic information is associated with the current job running on the grid resources, the type of software specification, the available time in resources, disk space, and policies.

Replica Catalogue (RC) is the fourth element. RC is the directory that presents information, which helps in locating the saved data in the grid. To simplify employing the grid's data, the grid broker requests the RC for information about locations of the data in the grid and the way to identify the access mechanism needed to utilize this data.

To locate the best resources for the grid clients' requests, the grid broker gets the job specifications from the grid portal and looks for the suitable resources that can fit these specifications. For this purpose, the broker inquires for resources information and the data stored in the grid from the Information Service and the Replica Catalogue. Later it chooses the resource or resources that able to fit the job specifications.

Processing and modeling jobs have grown to be more complicated as the system scopes increased. Accordingly, it is difficult to carry out tasks with such an extensive system using centralized methods. Multi-Agent System (MAS) presents an excellent manner for delivering control. It also can be used to explore and change the local communications amongst entities [9]. Furthermore, artificial intelligence mechanisms can be combined with it. Intelligent agents are incorporated to operate autonomously in a changing environment. The intelligent agent also has the following main features: [10].

- Autonomy: the power to behave dynamically to a particular point in support of the clients and programs and adjust how agents accomplish their jobs.
- Cooperation and Communication: the power to cooperate and communicate with other agents to swap information, receives orders, and respond.
- Learning: the power to enhance the performance steadily while behaving with the external world.
- Re-activity: the ability to react to exterior calls and moderate the agents behave concerning these calls.
- Pro-activity: the ability to create choices as a consequence of internal evaluations.

In this introduction, we have introduced some background about Grid and MAS as a suitable solution for open systems, such as Grid, that modify frequently. This paper is organized as follows. Section 2 introduces related works to this research. Section 3 gives the system model. The experiment and results are presented in Section 4. Lastly, the conclusion of this research is given in Section 5.

2. Related Works

In recent times, combining the MAS with grid computing has taken the interest of lots of researchers. Most topics that have been concerned in the area of grid computing are how to achieve better use of grid data storage, service distribution, infrastructures, cost, and energy efficiency. Nevertheless, the worry in the agent environment's area is on the intelligent side of agents and the procedures that can be exploited to solve complicated problems.

The authors in [11] suggested a multi agent system that can support the growth of the IDAPS micro-grid where this proposed system includes a DER agent, database agent, and control and a user agent. The Agents transmit their requirements using a TCP/IP protocol embedded in the IEEE FIPA standard to assure the system capability. In addition, authors in [12] suggested a model for agent-based grid computing (AGEGC). Based on their model, AGEGC

was created by using the MAGE (Multi-Agent Environment) platform. This research included various levels, such as; information tier, knowledge tier, service-oriented tier, and operating system tier. AGEGC focuses on service-oriented tier provided that current processing system.

In recent times, there are many paradigms concerning the growth of agent systems, such as; MOLE [13], DMARS [14], RETSINA [15]. Lots of companies are contributing to building characteristics for standard agent models; for example, OMG [16] KSE [17] and FIPA [18].

Garimella in [19] employs a new preserve server that works in assistance with the Dynamic Soft Real-Time system [20]. The purpose of the Garimella system is to reserve the resource of the CPU beforehand. By using this system, the user request for some quality of service conditions. E.g. of that the percentage of CPU needed and the beginning time and the duration. Once the request for the preservation is begun, the reserved resources are ready to be reachable for the users. One more proposed model is the Resource Broker (RB) presented in [21] merged with a new preserve server shown in [19]. Both system have continual and quick response times and numerous negotiation probability for the users. Nevertheless, the volume of re-negotiation introduces an extra-large cost to the system, in the event of an unpredicted lack of resources.

The authors in [22] presented several algorithms for keeping an improved preservation approach for supercomputer scheduling systems. Those algorithms improve traditional organizing techniques by combing the reservation demand mechanism with scheduling regular tasks mechanism. Using this way of reservations, let clients ask for more than one resource in parallel from the scheduling systems at the same time. However, the mechanisms used in [22] preserve the whole time slot, which means the resources are not carried on in a distributed way by multiple users for the same amount of time.

3. System Model

Our suggested architecture is adopted from the grid architecture in [23], as shown in Figure 3. It employs Client/Server architecture since it is the most famous architecture in the heterogeneous system [24]. Our proposed architecture uses multi-tier architecture, considering that there is a necessity for the parted database to save the grid resources' location. Our grid architecture includes the following entities; the grid portal, the grid broker, and the grid resource providers.

A grid portal is an interface for an authorized client to send their jobs to the grid. A portal should hide the complexity of the grid from clients by employing an uncomplicated interface. In this manner, it enables the sorting of grid job requisites. The grid broker can connect with agents that operate, control, and support both the grid clients and grid resource providers. The broker is an agent handler who organizes and monitors the running jobs and maintains the grid clients updated with their running jobs.

The grid broker in this paper is moved from operating as a database warehouse that saves the available resources and conditions and matches the clients' jobs with these resources into an agent manager with the ability to control event creations and create and eliminate several agents in the system. The grid broker in our proposed architecture, is a MAS that can interconnect with the grid system on account of grid users to find the information regarding grid resource provider whom they can fit the users' requirements and choose the best one among them. Furthermore, our advanced system can provide and monitor the implementation of the jobs and support the grid users with live feedback concerning the job processes.

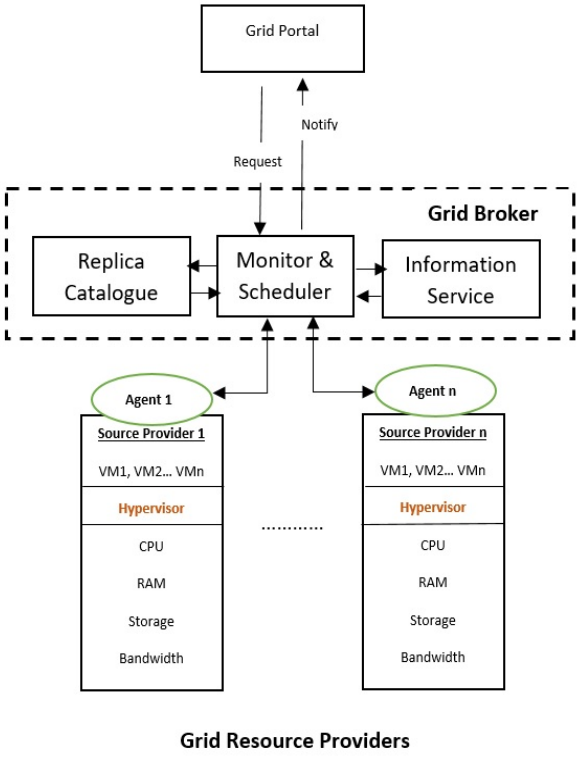


Figure 3: System Architecture

The grid broker is split into three primary entities. The first one is the monitor & scheduler. This entity is in charge of getting the grid clients' tasks and deciding the best available resource providers. The scheduler requests all data about the available resources from the Information Service entity and the data information saved in the Replica Catalogue (RC) entity. The monitor has the authority to monitor and provision the currently running jobs in the grid resource providers. The grid broker might have the improvements and ability to split clients' tasks between multiple resource providers to cut the cost and speed up the operations.

After the Monitor & Scheduler transmit the jobs to the available resource(s) for processing, it receives frequent feedbacks from the agents at the resource providers to verify the standing of the processing jobs and the percentage of the job completed to the grid broker. This point helps the grid broker to provision the running jobs' accomplishment, and by its turn, it will help determine an alternate resource provider(s) in case of failure at any point. The local agent at the resource providers is similarly in charge of updating the grid broker as the job is finished.

The resource provider can include several physicals and virtual machines. The Hypervisor in every machine is in charge of organizing and generating virtual machines on the upper of every physical machine. To react to the fast development of clients'

requests for data processing ability and storage, grid resource providers create a massive amount of data centers throughout the world. The grid datacenters typically include a large number of interconnected resources from physical machines and virtual machines, which will consume a considerable amount of electricity for their operations.

The proposed system depends on switching inactive virtual and physical machines to lower power positions (Sleep/Wakeup or switched off) while still preserving customers' performance requirements. The local agents at the resource providers are responsible for this operation. Each virtual machine (VM) has a local agent that can report the VM utilization to the data center's primary agent. The utilization here means the CPU utilization. Suppose the utilization for a specific VM has a zero amount for 300 seconds. In that case, the data center's primary agent will automatically shut down this VM to reduce the power consumption of the physical machine containing that turned off VM.

4. Implementation and Testing

The Jade platform was chosen to execute our proposed topology. Jade stands for the Java Agent Development framework, developed to help build out and enhance agent software as ruled by the FIPA conditions for smart MAS [25].

We introduced a simulation that deals with 200 physical hosts. The specifications of physical and virtual machines are shown in Table 1 and Table 2, respectively.

Table 1: Specification of Physical Machines

Type	Name	MIPS	RAM (MB)	Bandwidth (Gbps)	Number of Hosts
1	HP ProLiant ML110 G4 servers	2000	4096	2	100
2	HP ProLiant ML110 G5 servers	3000	4096	2	100

Table 2: Specification of Virtual Machines

Type	MIPS	RAM (MB)	Bandwidth (Gbps)	Number of VM
1	2000	1024	100	200
2	1500	2048	100	200
3	500	512	100	200

To create real power-consuming data, we took advantage of the actual data of power consumption offered by the SPECpower benchmark [26] and [27], as shown in Table 3. The power consumption by physical machines can be precisely expressed by a linear relationship between CPU utilization and power consumption. It can be noticed from Table 3 that even at small utilization, the machine exhausts a large volume of power. Therefore it is necessary to shut down such machines or reduce the number of VMs in it when not in use to reduce the power consumption of the physical machine. Figure 4 shows part of our simulation.

Table 3: Power Consumption (Watts per hour) at different load levels.

Machine type	0%	20%	50%	70%	100%
1	86	92.6	102	108	117
2	93.7	101	116	125	135

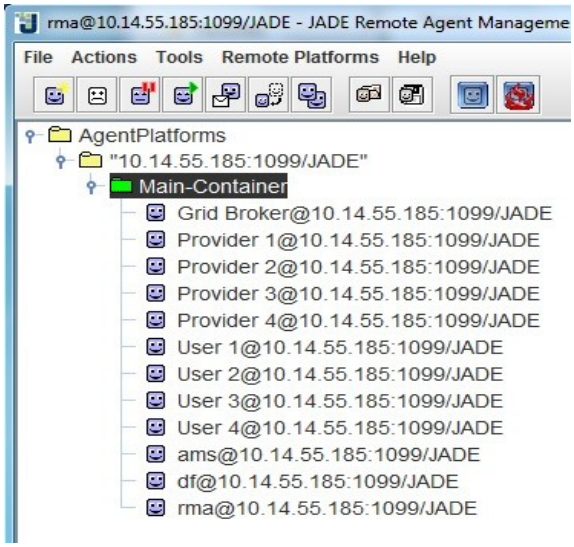


Figure 4: Part of the Summation Stage

The following stages explain the main steps of our simulation:

- Grid users state their tasks by choosing the required hardware and software specification to run their jobs through the grid portal.
- The grid broker receives the grid users' jobs and looks for the best provider(s) for those jobs. This step is done by helping the Information Services and the Replica Catalogue.
- Later, the grid broker sends the jobs to the suitable resource provider(s) and connects to the primary agent attached to that provider.
- The grid broker can divide the job between multiple resource providers to cut down the cost and speed up the operations.
- As the job is running at the resource provider, the provider's agent (primary agent) sends periodic feedbacks about the percentage of completion of the running job to the grid broker. This stage helps the grid broker to supervise the completion of the running job, and that will help find an alternative resource provider in case of failure.
- At the same time, the primary agent at the resource provider periodically checks the CPU utilization of its VMs at the physical machines.
- If the utilization reaches the zero point for 300 seconds, the primary agent will automatically shut down that VM.
- Finally, the grid broker sends the job outputs to grid users.

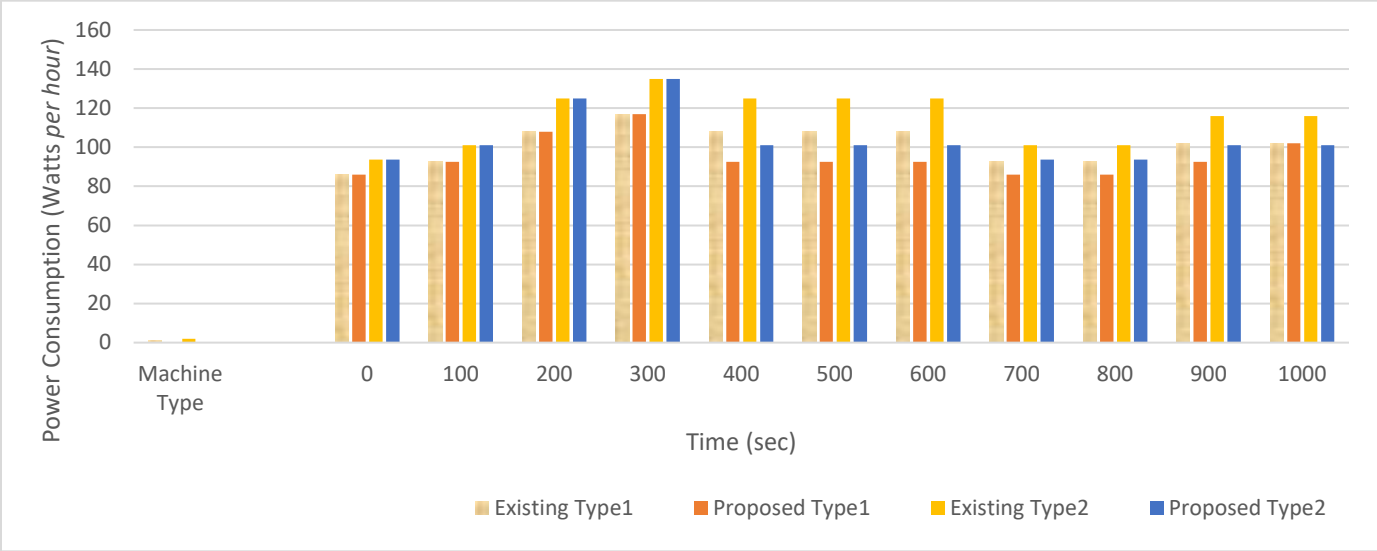


Figure 5: Simulation Results

From the simulation results in Figure 5. It can be seen that by using our proposed system, the power consumption has been slightly reduced for the two types of physical hosts (HP ProLiant ML110 G4 servers and HP ProLiant ML110 G5 servers), especially after the first 300s. This reduction is because the utilization of the physical machines has been reduced as the primary agents in both physical machines automatically shut down that VMs if the its utilization reaches the zero point for 300 seconds. As a result, the power consumption has been reduced in total for machines.

5. Conclusion

The purpose of the grid is to produce a broad scale, and heterogeneous systems intend to solve industrial or scientific problems. The integration of the Multi-Agent Systems (MAS) with the grid environment has a significant impact on grid performance. Grid offers a wide range of resources for its users, and some of these resources might not be used or utilized for some time before any new jobs coming to the grid for processing. As a result, the grid servers will consume a considerable amount of electricity for their operation. The growing usage of grid computing has led to an increase in electrical energy usage by the massive amount of servers in their data centers. In this paper, we have proposed an automated system composed of modern agents that can be used to reduce the power wastage resulted from inactive servers inside the data centers if the utilization of the VM reaches the zero point for 300 seconds, In this case the primary agent in the physical machine will automatically shut down that VM. The Automated system has been tested and evaluated, and the results show that the newly proposed method can reduce the amount of power wastage for the inactive grid resources.

References

[1] T. Alwada'n, A. Alarabeyyat, T. Khdour, and A. Rodan, "Utilizing Multi-Agent Systems in Grid Environments," Sixth Intentional Conference on HCT Information Technology Trends (ITT), Ras Al Khaimah, United Arab Emirates, 2019, doi: 10.1109/ITT48889.2019.9075099.

[2] T. Alwada'n, H. Aldabbas, H. Janicke, T. Khdour, and O. Aldabba. Dynamic "policy management in mobile grid environments", International Journal of Computer Networks & Communications (IJCNC), **4**(2), 2012.

[3] H. Alhakami, H. Aldabbas, and T. Alwadan. "Comparison between cloud and grid computing: Review paper", International Journal on Cloud Computing: Services and Architecture (IJCCSA), 2, 2012.

[4] T. Alwadan, H. Janicke, O. Aldabbas, and M. Alfawair. "New framework for policy support for mobile grid services", The 6th International Conference on Risks and Security of Internet and Systems (CRISIS2011), Romania 2011.

[5] T. Alwadan, H. Janicke, O. Aldabbas, and H. Aldabbas. "New framework for dynamic policy management in grid environments", In Recent Trends in Wireless and Mobile Networks, Third International Conferences, Springer, 2011.

[6] Ku. Ku-Mahamud and H. Abul Nasir. "Ant Colony Algorithm for Job Scheduling in Grid Computing", Asia International Conference on Modeling and Simulation, 2010, 10.1109/AMS.

[7] T. Alwadan, H. Janicke, A. Alarabeyyat, A. Alkharabsh, and T. Khdour. "Policy-based support for mobile grid services", International Journal of Computer Science Issues (IJCSI), 10, 2013.

[8] T. alwada'n, H. Aldabbas, and H. Janicke, editors. "Mobile Grid System: Policy and Privacy", LAP LAMBERT Academic Publishing, 2014.

[9] A. S. Grimshaw, M. A. Humphrey, and A. Natrajan. "A philosophical and technical comparison of legion and globus", IBM J. Res. Dev., March 2004.

[10] D. Talia. "Cloud computing and software agents: Towards cloud intelligent services", 2011.

[11] M. Pipattanasomporn, H. Feroze, and S. Rahman. "Multi- agent systems in a distributed smart grid: Design and implementation", 2009.

[12] Z. Shi, He Huang, J. Luo, Fen Lin, and H. Zhang. "Agent-based grid computing. Applied Mathematical Modelling", 2006, 30:629640, 07.

[13] J. Baumann, F. Hohl, K. Rothermel, and M. Straßer. "Mole – concepts of a mobile agent system", World Wide Web, Sep 1998.

[14] A. S Rao and M. P George. "Bdi agents: From theory to practice", In Proceedings of the First International Conference on Multi- Agent Systems (ICMAS-95), 1995.

[15] K. Sycara, A.e Pannu, M. Williamson, D. Zeng, and K. Decker. "Distributed intelligent agents", IEEE Expert: Intelligent Systems and Their Applications, December 1996.

[16] D. Milojicic, M. Breugst, I. Busse, J. Campbell, S. Covaci, B. Friedman, K. Kosaka, D. Lange, K. Ono, M. Oshima, C. Tham, S. Virdhagriswaran, and J. White. "Masif: The omg mobile agent system interoperability facility". Personal Technologies, Jun 1998.

[17] P. Patel-Schneider, D. P McKay, T. Finin Thomas, G. Ramesh Patil, R. Fikes, and R. Neches. "The DARPA Knowledge Sharing Effort: Progress Report". Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning. Morgan Kaufman, August 1992.

[18] fipa.org. "The foundation for intelligent physical agents", Website, 2019.

[19] G. Garimella, "Advance CPU Reservations with the Dynamic Soft Real-Time Scheduler", Master's Thesis, University of Illinois at Urbana-Champaign, 1999.

[20] H. Chu and K. Nahrstedt, "A Soft Real Time Scheduling Server in UNIX Operating System," European Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services (IDMS '97), Sep. 1997.

[21] K. Kim and K. Nahrstedt, "A Resource Broker Model with Integrated Reservation Scheme," IEEE International Conference on Multimedia and Expo 2000 (ICME '00), Aug. 2000.
[22] W. Smith, I. Foster, and V. Taylor, "Scheduling with Advanced Reservations," International Parallel and Distributed Processing Symposium (IPDPS '00), May 2000.
[23] T. Alwada'n, S. Alghyaline., A. Alazzam. "High Priority Requests in Grid Environment", International Journal of Advanced Computer Science and Applications, No. 10, 2020.
[24] C. Jean Dollimore, and T. Kindberg. "Distributed Systems: Concepts and Design", fourth Edition, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.
[25] F. Bellifemine, A Poggi, and G. Rimassa. "Jade - A FIPA- compliant agent framework", The Practical Application Company Ltd., 1999.
[26] A. Verma, P. Ahuja, A. Neogi. "pMapper: power and migration cost aware application placement in virtualized systems", Proceedings of ACM/IFIP/USENIX 9th International Middleware Conference, Leuven, Belgium, 2008.
[27] N. Patel, H. Patel. "Energy efficient strategy for placement of virtual machines selected from underloaded servers in compute Cloud", Journal of King Saud University - Computer and Information Sciences, 6, 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 342-348 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Study of latencies in ThingSpeak

Vítor Viegas[*,1,2], J. M. Dias Pereira[2,3], Pedro Girão[2,4], Octavian Postolache[2,5]

[1]*CINAV – Escola Naval, Base Naval de Lisboa, Almada, 2810-001, Portugal*

[2]*Instituto de Telecomunicações, Lisboa, 1049-001, Portugal*

[3]*ESTSetúbal/IPS, Instituto Politécnico de Setúbal, Setúbal, 2914-508, Portugal*

[4]*Instituto Superior Técnico, Universidade de Lisboa, Lisboa, 1049-001, Portugal*

[5]*ISCTE – Instituto Universitário de Lisboa, Lisboa, 1649-026, Portugal*

ARTICLE INFO

Article history:
Received: 12 November, 2020
Accepted: 07 January, 2021
Online: 22 January, 2021

Keywords:
IoT
ThingSpeak
Latency
Delay
Measurement

ABSTRACT

IoT platforms play an important role on modern measurement systems because they allow the ingestion and processing of huge amounts of data (big data). Given the increasing use of these platforms, it is important to characterize their performance and robustness in real application scenarios. The paper analyzes the ThingSpeak platform by measuring the latencies associated to data packets sent to cloud and replied back, and by checking the consistency of the returned data. Several experiments were done considering different ways to access the platform: REST API, MQTT API, and MQTT broker alone. For each experiment, the methodology is explained, results are presented, and conclusions are extracted. The REST and MQTT APIs have similar performances, with roundtrip times between 1 s and 3 s. The MQTT broker alone is more agile, with roundtrip times below 250 ms. In all cases, the up and down links are far from being symmetric, with the uplink delay showing higher variance than the downlink delay. The obtained results can serve as a reference for other IoT platforms and provide guidelines for application development.

1. Introduction

Platforms for IoT (Internet of Things) have become key components in measurement and control systems because they are able to ingest, store and analyze huge quantities of data, on a 24/7 basis, at reasonable prices. They are hosted on the "cloud", which is a fancy name for data centers spread all over the world, equipped with high bandwidth, large storage capacity, and heavy processing power.

The term "cloud" is interesting because IoT platforms have indeed a broader view of the physical processes, as they were somewhere above in the sky. They have a broader view in terms of space because they gather data from different locations, and a broader view in terms of time because they store data persistently. This new level of awareness has flattened the traditional five-level automation pyramid [1] because field devices can now communicate directly with the cloud. Intermediate levels are being bypassed leading to an horizontal structure that is the basis of "smart factory" and "connected manufacturing" [2], two core concepts of "industry 4.0" [3].

Today, the cloud concentrates huge amounts of data making it the ideal place to run large-scale data analytics. Deep leaning, based on artificial neural networks, has benefited a lot from this scenario because it needs lots of data (big data) to perform well. As long as the data are good (big and diverse), deep learning is able to find good computational models for complex processes, even the hardest ones. With a good model in hands, new things can be done (such as preventive maintenance, and just-in-time asset management), and old things can be improved (such as robust control algorithms, and automatic controller tuning). IoT platforms play a key role in this movement because they are the "stage" where things are happening.

IoT platforms began to be used at the top of the automation pyramid because these levels do not need accurate timing. Typical applications include monitoring and supervision [4]-[10], with the

[*]Corresponding Author: Vítor Viegas, CINAV – Escola Naval, +351210902000, vviegas2@gmail.com

goal of reducing or removing humans in the loop [11]. There are also some applications for closed-loop control [12], [13], but the constrains in terms of low-latency and real-time make harder the penetration of IoT platforms at lower automation levels.

Whether IoT platforms are used for monitoring, supervision, or real-time control, it is important to know how fast and reliable they are. For that purpose, we took a well-known IoT platform – the ThingSpeak platform [14] – and measured the time it takes to upload a data packet and receive a reply. This so-called "roundtrip time" is an indicator of *how fast* the platform is. By checking the resemblance of both packets, outgoing and incoming, we can also have an idea on *how reliable* the platform is. We chose ThingSpeak because it is very easy to use, is open source, is free (with some restrictions), has an active community, and provides a comprehensive set of features, including persistent data storage, data analytics based on MATLAB, easy access through ubiquitous protocols, and security over SSL. For these reasons, the ThingSpeak is one of the makers' favorite platforms [15].

There are similar works trying to characterize experimentally the behavior of cloud servers and IoT platforms. For example, in [16], the author gives a tutorial on network latency measurements using PlanetLab as testbench. He measured the roundtrip times of mobile and non-mobile devices, connected through wireless (WiFi and 3G) and wired (Ethernet) interfaces, while pinging five different AWS servers spread around the world. In [17], the author and his team studied the performance of a cloud database by measuring the time needed to complete a writing on the database and getting back a reaction. They used a Siemens PLC to generate data, the IBM Cloud to store data and fire events, and an industrial computer to catch those events, all connected through a MQTT broker. In [18], the author measured the roundtrip time associated to a MQTT broker when it was accessed from two different continents (Brescia in Europe and São Paulo in South America). In [19], the author evaluated the efficiency and roundtrip time of three protocols commonly used in IoT, namely CoAP (constrained application protocol), WebSockets, and MQTT (message queue telemetry transport). In [20], the author made a quantitative performance analysis of the CoAP and MQTT protocols over various conditions of network capacity, packet loss probability, and link delay. In all these works, a substantial effort was put in characterizing experimentally the behavior of cloud services and the protocols used to access them.

The remaining of this paper is organized as follows: section 2 gives an overview of the ThingSpeak platform; sections 3 and 4 analyze the latencies of the ThingSpeak platform when it is accessed through two different application programming interfaces (API); section 5 focus on the ThingSpeak MQTT broker alone; section 6 discusses the obtained results; and section 7 extracts conclusions.

2. ThingSpeak

The ThingSpeak platform provides resources to store and process data in the cloud. The data are accessed through two well documented APIs: a REST API [21] that communicates over HTTP and follows the request-response model; and a MQTT API [22] that communicates over TCP/IP and follows the publish-subscribe model. Both APIs support authentication through unique read/write keys, but only the REST API supports data encryption through HTTPS. The REST API works well for one-to-one communications, while the MQTT API is best suited for one-to-many communications. The ThingSpeak MQTT broker only supports QoS = 0 (equivalent to "deliver at most once" or "fire-and-forget").

The ThingSpeak platform organizes information in data channels. Each channel includes eight fields that can hold any data type, plus three fields for location, and one field for status. Each channel is also characterized by a unique ID, a name, and a free description. It is not possible to access the fields individually; all read/write operations are made at the channel level to optimize remote calls. All incoming data receive a sequential ID and a timestamp (with a 1 second resolution). Channels are private by default, but they can also be made public in which case no read key is required. Channels are provided at no charge for non-commercial projects as long as they require no more than 8200 messages/day (~5 messages/minute).

The ThingSpeak provides the following resources to control the dataflow:

- React: Executes an action when stored data meet a certain condition (e.g. when a given field of a given channel crosses a given threshold). The action can be as simple as the execution of a script or the issue of a remote message over HTTP.
- TimeControl: Orders the execution of an action once at a specific time, or periodically on a regular schedule, much like a software timer. The TimeControl supports the same actions as the React.
- ThingHTTP: Is a remote call over HTTP, useful to communicate with remote entities such as devices, websites, and web services.

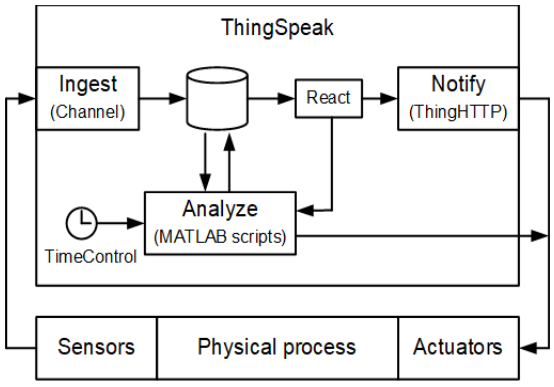


Figure 1: Dataflow inside the ThingSpeak platform

The ThingSpeak platform relies on MATLAB scripts to process stored data. Scripts can be associated to a TimeControl to run one-time or periodically, or to a React to run whenever a given condition is met. Scripts can use the MATLAB toolboxes listed in [23], as long as the user logs into ThingSpeak using its MathWorks account and is licensed to use them. This opens the door to powerful data analytics, supported by robust and well-known software libraries. The results can be visualized on the web, directly from the ThingSpeak site, through ready-to-use charts. The visualization experience can also be enriched with custom widgets and MATLAB plots.

Figure 1 shows the dataflow through the ThingSpeak platform. Data are ingested, stored in a database, and (optionally) analyzed by scripts that run periodically or when a given condition is met. Messages can be sent to third-party applications by a pre-configured ThingHTTP, or by a pre-programmed script using MATLAB functions. The present work focusses on measuring the time it takes to upload data and receive a reply, assuming no processing is made in the interim.

3. ThingSpeak accessed through the REST API

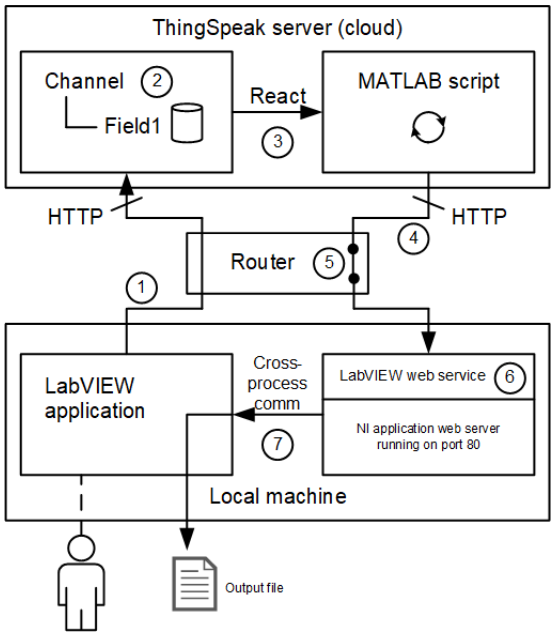
In this section we analyze the performance of the ThingSpeak platform when it is accessed through the REST API. We first explain the methodology used to measure latencies and then we present results.

3.1. Methodology

To measure REST API latencies, we built the closed data path shown in Figure 2, which includes the following stages (the numbers in the list correspond to the numbers in the drawing):

1. LabVIEW application: Is a custom application that runs on our local machine. It makes HTTP calls to the ThingSpeak server, collects the replies, and computes the time elapsed. Each call is a GET request that uploads an order number (n) that is incremented to identify the request. The timestamp of the request (t_{n0}) is registered to serve as reference for the roundtrip time.
2. ThingSpeak channel (name = TestChannel; ID = 515584; access = private): The channel contains a single field (Field1) to store the order number uploaded by the LabVIEW application.
3. React (condition type = numeric; test frequency = on data insertion; condition = TestChannel.Field1 ≥ 0; run = each time the condition is met): A reaction is fired each time a positive order number is received (which is always because n is an unsigned integer). The reaction instructs the MATLAB script to run (see below).
4. MATLAB script: Collects the timestamp at the ThingSpeak server and makes an HTTP call back to the local machine. As shown in Figure 3, the script is programmed to make a POST request to a web service running on the same machine as the LabVIEW application. The request sends back the order number and the timestamp at the ThingSpeak server (τ_n).
5. Router: Accesses from the local machine to an external server are inherently safe and the router forwards them transparently. However, connections in the opposite direction are potentially dangerous and are blocked by default. To overcome this problem, we had to forward port 80 on the local router, so that POST requests coming from the ThingSpeak server reach the LabVIEW web service. In other words, we had to expose the LabVIEW web service to the internet.
6. LabVIEW web service: Is a stateless routine that receives a POST request, extracts the attached order number (n) and timestamp (τ_n), and registers the timestamp of the reply (t_{n1}). The triplet (n, τ_n, t_{n1}) is then sent back to the LabVIEW application by means of a UDP socket. The web service is hosted by the NI Application Web Server running on port 80 of the local machine.

(n, τ_n, t_{n1}), adds the first timestamp (t_{n0}), and writes the quartet (n, t_{n0}, τ_n, t_{n1}) into the output file for further processing.




```
%get source
readChannelID=515584;
readApiKey='[redacted]';
n = thingSpeakRead(readChannelID,'ReadKey',readApiKey);
%get ThingSpeak's timestamp
tau = datetime('now','Format','HH:mm:ss.sss');
%call back
url='http://95.136.[redacted]:3582/MyWebService/MyCallBack';
%method = POST; media type = application/x-www-form-urlencoded
webwrite(url,'OrderNumber',n,'Tau',tau);
```

Figure 3: MATLAB script. The reading key of the channel and the IP address of the LabVIEW web service were erased for privacy

The LabVIEW application and the LabVIEW web service run both on our local machine (Intel i7-8550 CPU @ 1.80 GHz, RAM 16 GB, SSD 512 GB, NVIDIA GeForce MX150). The machine connects to an Ethernet port of a general-purpose router (model HS8247W from Huawei), which accesses the internet through a fiber optic link provided by Vodafone Portugal.

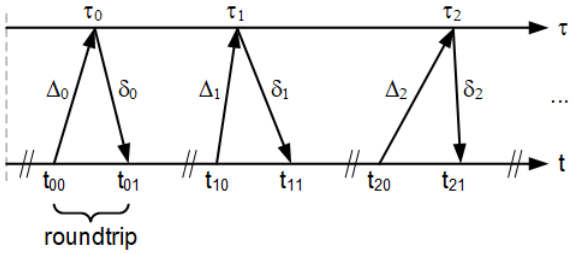
The LabVIEW application uploads order umbers (n) at multiples of 20 seconds to respect ThingSpeak free account limitations. On the nth upload, the quartet (n, t_{n0}, τ_n, t_{n1}) is saved on the nth line of the output file. Thus, it is possible to record the timelines shown in Figure 4, where the τ axis represents the timeline of the ThingSpeak server, and the t axis represents the timeline of the local machine. Of course, the two timelines are not

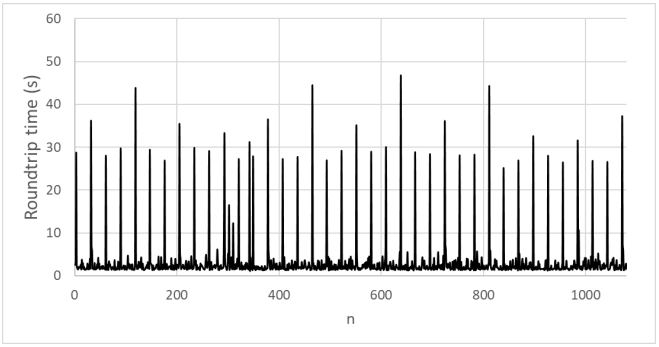
aligned because the clocks of the two systems are not synchronized. Yet, we can extract the following quantities from these timelines:

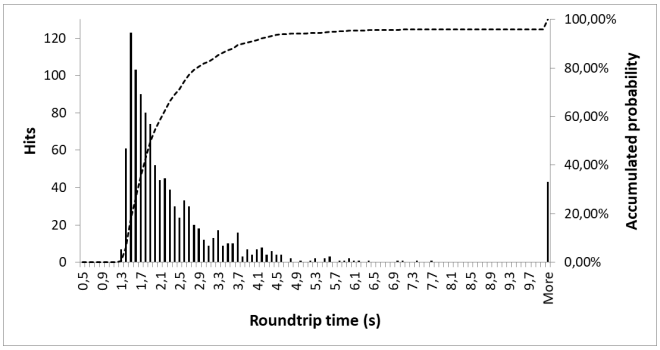
$$RT_n = t_{n1} - t_{n0} \quad (1)$$

$$d\Delta_n = \Delta_{n+1} - \Delta_n = (\tau_{n+1} - \tau_n) - (t_{(n+1)0} - t_{n0}) \quad (2)$$

$$d\delta_n = \delta_{n+1} - \delta_n = (t_{(n+1)1} - t_{n1}) - (\tau_{n+1} - \tau_n) \quad (3)$$







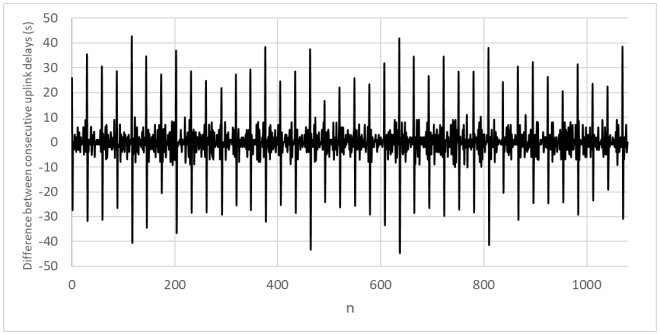
(b)

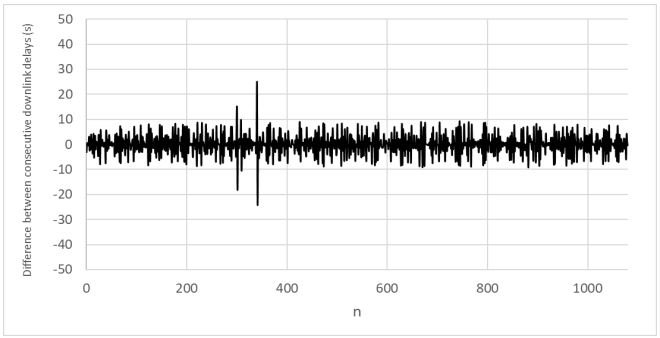
Figure 5: Roundtrip times of the REST API: a) plot; b) histogram

3.2. Results

We ran the closed loop illustrated in Figure 2 for 1080 times, from n=0 to n=1079, waiting approximately 20 seconds on each iteration. This number (1080) was a compromise between having statistically relevant results and limiting the test to a reasonable amount of time (almost six hours). Figure 5a shows the measured roundtrip times, and Figure 5b shows the corresponding histogram. The order number (n) was always replied correctly, which attests the robustness of the ThingSpeak platform.

Figure 5b we see that the roundtrip time has a heavy-tailed distribution, which is characteristic of multipath communication mediums as the Internet. The authors in [24] report similar results and suggest a lognormal distribution for the experimental data. In our case, we got mode = 1,66 s and a roundtrip time that is less than 2.9 s with a probability of 80%.





(b)

Figure 6: Difference between consecutive delays: a) uplink; b) downlink

Nevertheless, every 10 min (around 30 points) the roundtrip time increases very sharply up to tens of seconds, suggesting that the ThingSpeak platform stores data in temporary buffers, which, from time to time, are flushed and processed.

Figure 6 shows the difference between consecutive delays on the uplink and downlink directions. The differences are positive and negative because a higher delay on one iteration discounts on the next iteration. As expected, the uplink delay (Δ) is less stable than the downlink delay (δ) because the upload process is more complex than the reply (in terms of ThingSpeak internals).

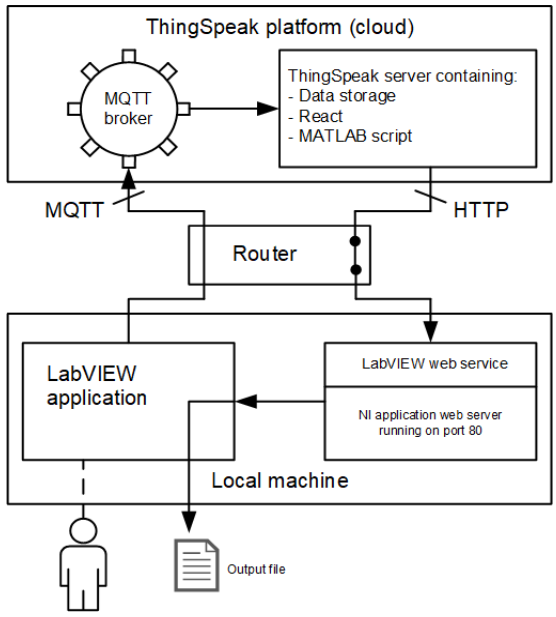
If we suppose that consecutive uplink delays are statistically independent, which makes sense because they correspond to different iterations, then the variance of the difference (dΔ or dδ) will be twice of the variance of the variable itself (Δ or δ). Therefore, if we compute the variances Var(dΔ) and Var(dδ) from the data of Figure 6, and divide the result by two, we get the variances of the delays: Var(Δ) = 37.52 s^2 and Var(δ) = 7.29 s^2. Applying the squared root, we get the standard deviations: σ_Δ = 6.13 s and σ_δ = 2.70 s.

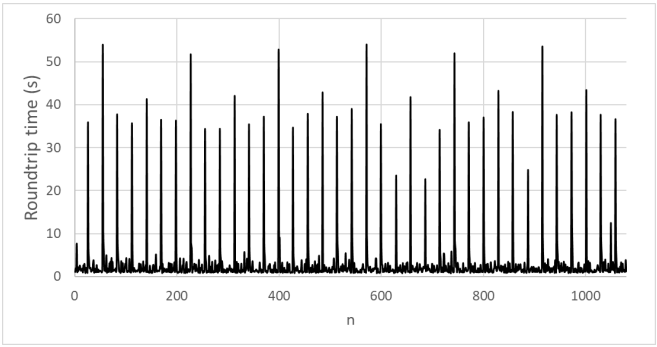
Finally, it is very difficult to infer about the mean value of the uplink and downlink delays because the clocks of the two systems

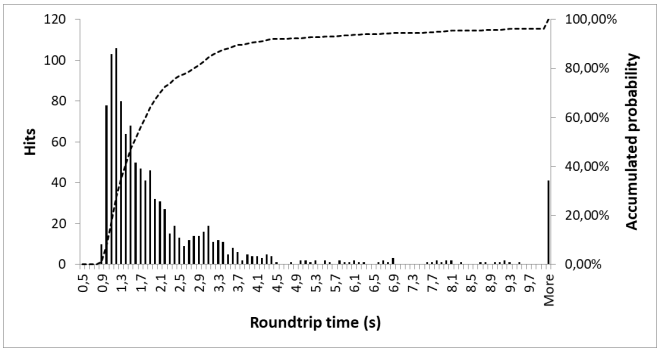
are not synchronized and, equally important, the two links are far from being symmetric.

4. ThingSpeak accessed through the MQTT API

built the closed data path shown in Figure 7, which is like that of Figure 2 with the difference that the order numbers are uploaded through the MQTT broker. The LabVIEW application was changed to publish order numbers (n) using a compatible MQTT driver [25]. The quartets (n, t_{n0}, τ_n, t_{n1}) were collected and saved into the output file as before.







(b)

Figure 8: Roundtrip times of the MQTT API: a) plot; b) histogram

We ran the closed loop 1080 times waiting approximately 20 seconds on each iteration. Figure 8a shows the measured roundtrip times, and Figure 8b shows the corresponding histogram. Again, the order number (n) was always replied correctly, with no mismatches or timeouts. From the graphs, we see that the roundtrip time follows a lognormal distribution with mode = 1,13 s and values below 2,8 s with a probability of 80%. We also see outliers coming every 10 min (as we saw in the REST API), suggesting that internal mechanisms of buffering and batch processing also apply to the MQTT API.

The random variables dΔ and dδ were also computed as before. The corresponding variances were extracted and divided by two, leading to Var (Δ) = 51.53 s^2 and Var (δ) = 5.99 s^2. Applying the squared root, we got the standard deviations of the uplink and downlink delays: σ_Δ = 7.18 s and σ_δ = 2.45 s.

5. MQTT broker alone

To analyze the performance of the MQTT broker alone we closed the loop without passing through the ThingSpeak server, as shown in Figure 9. The LabVIEW application was changed to publish order numbers (n) and subscribe the replies using the MQTT driver previously mentioned. In this case, only the triplets (n, t_{n0}, t_{n1}) were collected because the timeline of the ThingSpeak server (τ axis) is not available.

We ran the closed loop 1080 times waiting approximately 20 seconds on each iteration. Figure 10.a shows the measured roundtrip times, and Figure 10.b shows the corresponding histogram. As always, the order number (n) was always replied correctly, with no mismatches or timeouts. From the graphs, we see that the roundtrip time follows a lognormal distribution with mode = 0.189 s, and values below 0,250 s with a probability of 80%. We also see that the outliers observed in the previous tests have almost disappeared. This shows that the MQTT broker is more expeditious than the ThingSpeak server, probably because its buffering and processing needs are much less demanding.

6. Results and Discussions

Table 1 summarizes the latencies measured during all the experiments. From these results we can extract the following conclusions:

- The REST and MQTT APIs have similar performances with a slightly advantage for the MQTT API. In both cases, the roundtrip time was typically between 1 s and 3 s.
- The ThingSpeak server has internal mechanisms of buffering and batch processing that, periodically, introduce extraordinary delays. These mechanisms seem to be absent from the MQTT broker.
- The MQTT broker is very agile in distributing publications, with roundtrip times typically below 250 ms.
- The up and down links are far from being symmetric. The uplink is more complex since it has a higher variance.
- The tests lasted for several hours and were made in different days and in different times of the day, suggesting that the behavior of the ThingSpeak platform is time independent.

A final word about robustness: the order number (n) was never lost or corrupted, proving that the ThingSpeak is reliable, even for free accounts.

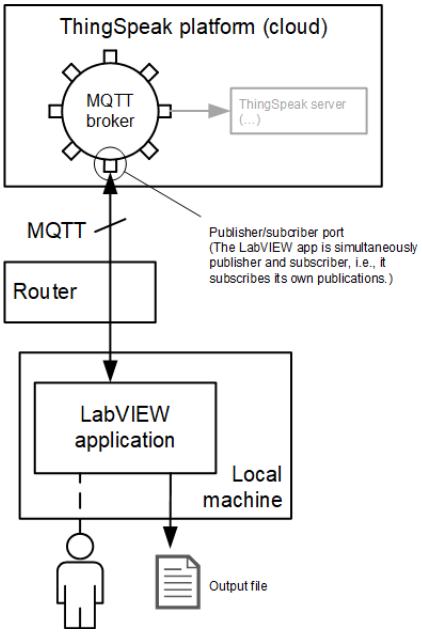
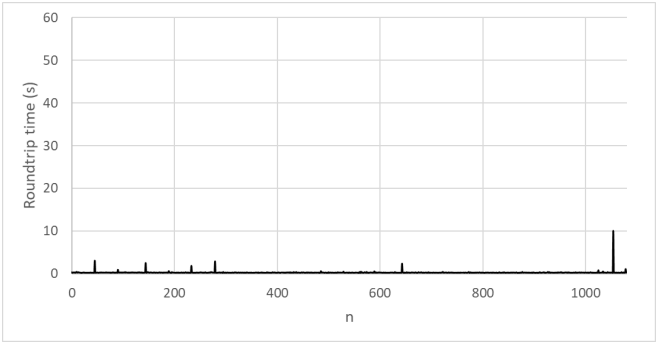
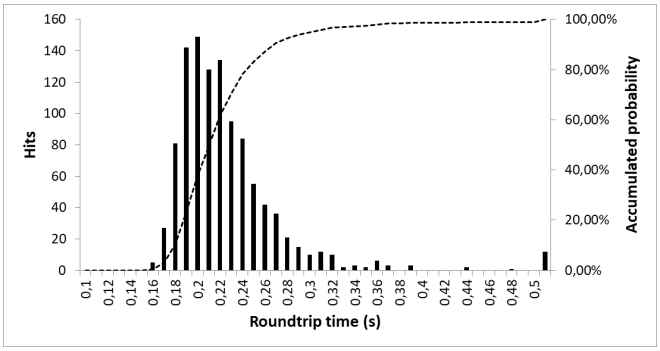


Figure 9: Data loop through the MQTT broker only



(a)



(b)

Figure 10: Roundtrip times of the MQTT broker alone (a) plot (b) histogram

Table 1: Experimental characterization of the ThingSpeak platform

		REST API	MQTT API	MQTT broker
Roundtrip time	**Min (s)**	1.25	0.86	0.158
	Mean (s)	3.28	3.27	0.243
	Mode (s)	1.66	1.13	0.189
	80th percentile (s)	2.90	2.80	0.250
Uplink delay	**Std. dev. (s)**	6.13	7.18	---
Downlink delay	**Std. dev (s)**	2.70	2.45	---

7. Conclusions

The paper reported the studies carried on the ThingSpeak platform to evaluate its responsiveness and reliability. We measured the time needed for a data packet to loop back through the platform, and we verified if its content has been corrupted during the trip. Tests were made for all access mediums (REST API, MQTT API and MQTT alone) covering periods of six hours.

We saw that the REST and MQTT APIs have similar performance with typical roundtrip times between 1s and 3s. We observed repetitive outliers that suggest that the ThingSpeak server has periodic mechanisms of buffering and batch processing. The MQTT broker alone did not show such outliers and performed significantly faster. In terms of reliability, no data was lost or corrupted.

We hope that the obtained results can serve as a reference for other IoT platforms and provide guidelines for application development.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] "ANSI/ISA-95.00.03-2013 Enterprise-Control System Integration - Part 3: Activity Models of Manufacturing Operations Management"

[2] R. Burke et al., "The smart factory - Responsive, adaptive, connected manufacturing," Deloitte Insights, https://www2.deloitte.com/us/en/insights/focus/industry-4-0/smart-factory-connected-manufacturing.html (accessed on 6 Oct. 2020).

[3] K. Schwab, "The fourth industrial revolution - What it means and how to respond," Foreign Affairs, 12 Dec. 2015, https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution (accessed on 6 Oct. 2020).

[4] J. Delsing, F. Rosenqvist, O. Carlsson, A. W. Colombo and T. Bangemann, "Migration of industrial process control systems into service oriented architecture," IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, Montreal, QC, 2012, 5786-5792, doi: 10.1109/IECON.2012.6389039.

[5] T. Hegazy and M. Hefeeda, "Industrial automation as a cloud service," in IEEE Transactions on Parallel and Distributed Systems, **26**(10), 2750-2763, 1 Oct. 2015, doi: 10.1109/TPDS.2014.2359894.

[6] R. Langmann and L. Meyer, "Automation services from the cloud," 2014 11th International Conference on Remote Engineering and Virtual Instrumentation (REV), Porto, 2014, 256-261, doi: 10.1109/REV.2014.6784271.

[7] H. Sequeira, P. Carreira, T. Goldschmidt and P. Vorst, "Energy cloud: Real-time cloud-native energy management system to monitor and analyze energy consumption in multiple industrial sites," 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing, London, 2014, 529-534, doi: 10.1109/UCC.2014.79.

[8] O. Givehchi, J. Jasperneite, "Industrial automation services as part of the cloud: first experiences," 2013 Jahreskolloquium Kommunikation in der Automation (KommA 2013), Magdeburg, Germany, 2013.

[9] O. Givehchi, H. Trsek and J. Jasperneite, "Cloud computing for industrial automation systems - A comprehensive overview," 2013 IEEE 18th Conference on Emerging Technologies & Factory Automation (ETFA), Cagliari, 2013, 1-4, doi: 10.1109/ETFA.2013.6648080.

[10] A. Ito, T. Kohiyama, K. Sato, F. Tamura, "IoT-ready industrial controller with enhanced data processing functions," in Hitachi Review, **67**(2), 208-209, Feb. 2018.

[11] J. Pretlove, C. Skourup, "Human in the loop," ABB Review 1/2007.

[12] L. Wang and A. Canedo, "Offloading industrial human-machine interaction

tasks to mobile devices and the cloud," Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), Barcelona, 2014, 1-4, doi: 10.1109/ETFA.2014.7005249.
[13] O. Givehchi, J. Imtiaz, H. Trsek and J. Jasperneite, "Control-as-a-service from the cloud: A case study for using virtualized PLCs," 2014 10th IEEE Workshop on Factory Communication Systems (WFCS 2014), Toulouse, 2014, 1-4, doi: 10.1109/WFCS.2014.6837587.
[14] "About ThingSpeak," https://thingspeak.com (accessed on 11 Oct. 2020).
[15] "Top IoT platforms for makers," https://ubidots.com/blog/top-iot-platforms/ (accessed on 30 Oct. 2020).
[16] Minseok Kwon, "A tutorial on network latency and its measurements," in Enabling Real-Time Mobile Cloud Computing through Emerging Technologies, IGI Global, 2015, 272-293, doi: 10.4018/978-1-4666-8662-5.ch009.
[17] Paolo Ferrari, Emiliano Sisinni, Alessandro Depari, Alessandra Flammini, Stefano Rinaldi, Paolo Bellagente, Marco Pasetti, "On the performance of cloud services and databases for industrial IoT scalable applications," in Electronics 2020, **9**(9), 1435, doi: 10.3390/electronics9091435.
[18] P. Ferrari, E. Sisinni, D. Brandão and M. Rocha, "Evaluation of communication latency in industrial IoT applications," 2017 IEEE International Workshop on Measurement and Networking (M&N), Naples, 2017, pp. 1-6, doi: 10.1109/IWMN.2017.8078359.
[19] S. Mijovic, E. Shehu and C. Buratti, "Comparing application layer protocols for the Internet of Things via experimentation," 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), Bologna, 2016, 1-5, doi: 10.1109/RTSI.2016.7740559.
[20] M. Collina, M. Bartolucci, A. Vanelli-Coralli and G. E. Corazza, "Internet of Things application layer protocol analysis over error and delay prone links," 2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC), Livorno, 2014, 398-404, doi: 10.1109/ASMS-SPSC.2014.6934573.
[21] "REST API," https://www.mathworks.com/help/thingspeak/rest-api.html (accessed on 6 Oct. 2020).
[22] "MQTT API," https://www.mathworks.com/help/thingspeak/mqtt-api.html (accessed on 6 Oct. 2020).
[23] "Access MATLAB Add-On Toolboxes," https://www.mathworks.com/help/thingspeak/matlab-toolbox-access.html (accessed on 6 Oct. 2020).
[24] I. Antoniou, V.V. Ivanov, Valery V. Ivanov, P.V. Zrelovc, "On the log-normal distribution of network traffic," in Physica D: Nonlinear Phenomena, **167**(1-2), 72-85, July 2002. doi: 10.1016/S0167-2789(02)00431-1.
[25] "mqtt-LabVIEW," https://github.com/cowen71/mqtt-LabVIEW (accessed on 8 Oct. 2020).

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 349-355 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Deep Learning based Models for Solar Energy Prediction

Imane Jebli[*,1], Fatima-Zahra Belouadha[1], Mohammed Issam Kabbaj[1], Amine Tilioua[2]

[1]*AMIPS Research Team, E3S Research Center, Computer Science Department, Ecole Mohammadia d'Ingénieurs, Mohammed V University in Rabat, Avenue Ibn Sina B.P. 765, Agdal Rabat 10090, Morocco*

[2]*Thermal and Applied Thermodynamics, Mechanics Energy Efficiency and Renewable Energies Laboratory, Department of Physics, Faculty of Sciences and Techniques Errachidia, Moulay Ismail University of Meknès, Boutalamine Errachidia, B.P-509, Morocco*

ARTICLE INFO

Article history:
Received: 10 November, 2020
Accepted: 06 January, 2021
Online: 22 January, 2021

Keywords:
Photovoltaic energy prediction
Deep Learning
Recurrent Neural Network
Long Short-Term Memory
Gated Recurrent Units

ABSTRACT

Solar energy becomes widely used in the global power grid. Therefore, enhancing the accuracy of solar energy predictions is essential for the efficient planning, managing and operating of power systems. To minimize the negatives impacts of photovoltaics on electricity and energy systems, an approach to highly accurate and advanced forecasting is urgently needed. In this paper, we studied the use of Deep Learning techniques for the solar energy prediction, in particular Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). The proposed prediction methods are based on real meteorological data series of Errachidia province, from 2016 to 2018. A set of error metrics were adopted to evaluate the efficiency of these models for real-time photovoltaic forecasting, to achieve more reliable grid management and safe operation, in addition to improve the cost-effectiveness of the photovoltaic system. The results reveal that RNN and LSTM outperform slightly GRU thanks to their capacity to maintain long-term dependencies in time series data.

1 Introduction

Currently, the global trend is to integrate renewable energies to generate electricity and rethink its energy mix. In this regard, the use and expansion of renewable energy have become an important element in maintaining energy security, as well as to establish an ecological and sustainable electrical system [1]. Nevertheless, the energy context is affected by price variations, changes in demand and the volatility of renewable energy production. In order, to confront these main challenges of the energy mix, it is essential to have a good forecasting model for renewable energies, which is particularly useful for optimizing and adapting supply to demand. This must reliably provide advance information on the availability of power, which helps to achieve the grid's stable operation and to allow for optimal unit engagement and economic dispatching [2]. In this work, we concentrate on solar energy forecasting. Photovoltaic (PV) forecasting methods fall into two groups: direct and indirect prediction models. As described in [3, 4], solar radiation at different time scales is predicted with different methods and then converted to power based on the panel characteristics in the case of indirect prediction, while direct predictions are performed directly from the output power of the plant. Besides, Statistical, physical, artificial intelligence (AI) and hybrid approaches were used in predicting PV solar energy. However, thanks to their great learning and regression capabilities, AI and especially, Deep Learning (DL) techniques [5], have been widely used in this area. They have the capability to extract in-depth features from PV power datasets and get more reliable predictive outputs. Authors in [6] have applied Recurrent Neural Network (RNN) as a good tool for predicting solar irradiation time-series using recorded meteorological data. Besides, Li et al. [7] have proposed an RNN-based forecasting method for short-term PV power prediction. Moreover, another research in [8] has proposed a hybrid deep learning model using Wavelet Packet Decomposition (WPD) and Long Short-Term Memory (LSTM) to forecast PV power for one hour ahead with an interval of five minutes. WPD is applied to split the PV power output time-series and next the LSTM is implemented to forecast high and low-frequency subseries. The results of this study indicate that the WPD-LSTM method outperforms in various seasons and weather conditions than LSTM, Gated Recurrent Units (GRU), RNN, and Multilayer Perceptron (MLP). Authors in [9] have presented a deep LSTM based on historical power data to predict the PV system output power for one hour ahead. The aim in [10] to suggest a hybrid model using LSTM and attention mechanism for short-term PV power forecasting, and employed LSTM to extract features from the historical data and learn in sequence the information on long-term dependence, to

[*]Corresponding Author: Imane JEBLI, Avenue Ibn Sina B.P. 765, Agdal Rabat 10090, Morocco, Email: jebli1imane@gmail.com

apply the attention mechanism trained at the LSTM to target the extracted relevant features, which has greatly enhanced the original predictive power of LSTM. An LSTM approach for short-term forecasts based on a timescale that includes global horizontal irradiance (GHI) one hour ahead and one day ahead have been applied in [11]. Furthermore, authors in [12] have implemented univariate and multivariate GRU models employing historical solar radiation, external meteorological variables, and cloud cover data to predict solar radiation. In [13], authors have presented multivariate GRU to predict Direct Normal Irradiance (DNI) hourly. And the suggested method is evaluated compared to LSTM using historical irradiance data. In addition, a hybrid deep learning model have introduced in [14] that combines a GRU neural network with an attention mechanism for solar radiation prediction. All these research contributions cited above are valuable. However, they are not able to identify the important parameters that would have an impact on the accuracy of predictions. and make real-time predictions for efficient and optimized management. Given the fluctuating nature of output power generation as a function of meteorological conditions such as the temperature, the wind speed, the cloud cover, the atmospheric aerosol levels and the humidity level, which leads to high uncertainties on the output power of PV [15]. Besides, according to the forecast horizon, PV forecasts range from very short to long term. In general, researchers focus on short-term, hourly and daily forecasts as opposed to long-term. The first are of major significance for the management of PVs and the related security constraints (planning, control of PV storage and the rapprochement of the electricity market, providing the secure operation of generation and distribution services, and reinforcing the security of grid operation), while the long-term forecasts are useful particularly, for maintenance [16]. In this context, we study the efficiency of three DL models. We also focus in this study, on real-time prediction of solar energy in order to help planner, decision-makers, power plant operators and grid operators to making responsive decisions as early as possible, and to manage smart grid PV systems more reliably and efficiently [17]. Moreover, the use of real-time prediction allows to adjust to changes in production and to react to complex events (exceptionally high or low load production or consumption). In addition, it decreases the amount of operating reserves needed by the system, thus reducing system balancing costs [18]. However, the models we have selected give good results and seem to be suitable for long-term forecasting thanks to their power as a deep learning model and their ability to perform complex processing on huge data sets. Moreover, we outline that our work has been experimented with Moroccan's case, particularly in the region of Errachidia. In this perspective, Moroccan decision-makers have launched a global plan to improve the percentage of renewable energy in the energy mix and substantially improve energy efficiency. In view of increasing the percentage of electricity generation capacity from renewable energy sources (42% by 2020 and 52% by 2030) [19]. Therefore, we believe that the case of Morocco remains an interesting case study that could lead to important findings since it is not only a promising future PV energy supplier, but also one of the leading countries in the global energy transition, particularly in Africa. The remainder of this paper is organized into four sections. Section 2 gives an overview of our comparative methodology. Subsequently, section 3 presents the approach applied to elaborate the forecast models. Section 4 presents and discusses the obtained results. Finally, section 5 summarizes the conclusions and perspectives of this study.

2 Methodology

Solar energy prediction is a key element in enhancing the competitiveness of solar power plants in the energy market, and decreasing reliance on fossil fuels in socio-economic development. Our work aims to accurately predict the solar energy. For this purpose, we explore architectures of the RNN, LSTM and GRU algorithm which is suitable to be used for forecasting such time-series data, and we experiment and evaluate them in Morocco's case, especially Errachidia area. This section presents at first the basis architecture of this recurrent neural network before explaining the important steps that we follow to build our models and perform our comparative study.

2.1 Recurrent Neural Network architecture

Recurrent neural network (RNN) is a category of neural network used in sequential data prediction where the output is dependent on the input [20]. The RNN [21] is capable to capture the dynamic of time-series data by storing information from previous computations in the internal memory. RNN has been applied in a context where past values of the output make a significant contribution to the future. It is mainly used in forecasting applications because of its ability to process sequential data of different length.

The basic principle of RNNs is to consider the input of a hidden neuron which takes input from neurons at the preceding time step [22]. To this purpose, they employ cells represented by gates that influence the output using historical data observations are given to generate the output. RNN is particularly efficient for learning the dynamic temporal behaviors exhibited in time-series data [23]. In RNN, the hidden neuron ht for a given input sequence x_t, it has information feedback from other neurons in the preceding time step multiplied by a W_{hh} which is the weight of the preceding hidden state h_{t-1} can be calculated sequence by Eq.1. x_t is the input at instant time t, W_{xh} is the weight of the actual input state, tanh is the activation function. The output state y_t is computed according to the Eq.2 where W_{hy} is the weight at the output state.

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \quad (1)$$

$$y_t = W_{hy}h_t \quad (2)$$

The LSTM is a special kind of RNN, designed to avoid and resolve the vanishing gradient problems that limit the efficiency of simple RNN [8]. LSTM network has memory blocks that are connected through a succession of layers. In the LSTM cells, there are three types of gates: the input gate, the forget gate and the output gate. This makes it possible to achieve good results on a variety of time-series learning tasks, particularly in the nonlinear of a given dataset [24]. Each block of LSTM handles the state of the block and the output, it operates at different time steps and transmits its output to the following block and then the final LSTM block generates the sequential output [9]. Besides, LSTM is a robust algorithm, which

allows the recurrent neural network to efficiently process time-series data. Its key component is memory blocks which have been released to address the vanishing gradient disadvantage by memorizing network parameters for long durations [25].an LSTM block gets an input sequence after that, the activation units are used by each gate to decide whether or not it is activated. This operation enables the change of state and the adding of information passing through the block conditional. In the training phase, the gates have weights that can be learned. In fact, the gates make the LSTM blocks smarter than conventional neurons and allow them to remember current sequences [10]. LSTM is flexible and estimates dependencies of different time scales thanks to its ability to perform long task sequences and to identify long-range features. Basically, LSTM starts with a forget gate layer (f_t) that uses a sigmoid function combined with the preceding hidden layer (h_{t-1}) and the current input (x_t) as described in the following equations:

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{3}$$

$$\check{c}_t = \tanh(W_c.[h_{t-1}, x_t] + b_c) \tag{4}$$

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{5}$$

$$o_t = \sigma(W_o.[f_{t-1}, x_t] + b_o) \tag{6}$$

$$c_t = f_t.c_{t-1} + i_t.\check{c}_t \tag{7}$$

$$h_t = o_t.\tanh(c_t) \tag{8}$$

where i_t, $\check{c}_t$, f_t, o_t, c_t and h_t represent the input gate, cell input activation, forget gate, output gate, cell state, and the hidden state respectively. W_i, W_c, W_f and W_o represent their weight matrices respectively. b_i, b_c, b_f, and b_o represent the biases. x_t is the input, h_{t-1} is the last hidden state, h_t is the internal state. σ is the sigmoid function.

2.2 Gated Recurrent Units architecture

The GRU [26] is a particular type of recurrent neural network proposed by Cho in 2014 which, is the same as the LSTM) in terms of use to resolve the issue of long-term memory network and back-propagation. The GRU network involves the design of several cells that store important information and forget those that are deemed irrelevant in the future [27]. The feedback loops of the GRU network can be regarded as a time loop, because the output of the cell from the past period is taken as an input in the following cell in addition to the actual input. This key property allows the model to remember the patterns of interest and to predict sequential time-series datasets over time. Unlike the LSTM, the GRU [28] consists of only two gates: update and reset gates. the GRU substitutes the forget gate and the input gate in the LSTM with an update gate that verifies the cell state of the previous state information that will be moved to the current cell, while, the reset gate defines whether new information will be included in the preceding state. For this reason, GRU has proven to be one of the most efficient RNN techniques, because of its capacity to learn and acquire dependencies over the long term and observations of varying length [29]. This characteristic is particularly useful for time-series data, and it is helpful in reducing the computational complexity [30]. GRU cells are described using the following equations:

$$z_t = \sigma(W_z.[h_{t-1}, x_t]) \tag{9}$$

$$r_t = \sigma(W_r.[h_{t-1}, x_t]) \tag{10}$$

$$\hat{h}_t = \tanh(W.[r_t - h_{t-1}, x_t]) \tag{11}$$

$$h_t = (1 - z_t).h_{t-1} + z_t.\hat{h}_t \tag{12}$$

$$y_t = \sigma(W_0.h_t) \tag{13}$$

where z_t and r_t respectively correspond to the output of the update gate and the reset gate, while W_z, W_r, W, and Wo respectively are the weights of each gate. σ and tanh respectively are the sigmoid and the hyperbolic tangent activation function, and x_t is the network input at time t. ht and h_{t-1} correspond to the hidden layer information of the actual and precedent time, and $\hat{h}_t$ is the candidate state of the input. Initially, the network input x_t at the time of the hidden state h_{t-1} and t at the last state computes the output of the reset gate and the update gate by Eq.9 and 10. Next, after the reset of the reset gate r_t calculates the amount of memory stored, the implicit layer $\hat{h}_t$ is computed by Eq.11, which is the new information at the actual time t. Then, by using Eq.12, the update gate z_z determines the amount of information was removed at the previous time, and how much information is stored in the candidate hidden layer ht at this time. The hidden layer information h_t is subsequently added. Finally, the output of the GRU is switched to the next GRU gated loop unit according to Eq.13: which is a sigmoid function.

3 Our approach for predicting solar energy

In this study, we apply three different DL models to predict the PV solar energy output, especially RNN, LSTM and GRU for the half-hour ahead. The aim of this study is to compares the relevance of the mentioned models, in order to identify the best algorithm to be used for predicting solar energy. The process of the method employed in this work is illustrated in Figure 1. It is organized into four main stages: data collection, data pre-processing, model training and parameterization, and model testing and validation. This section details how each of these processing/selection steps was performed, as well as key details on the architecture and parameters of the developed models.

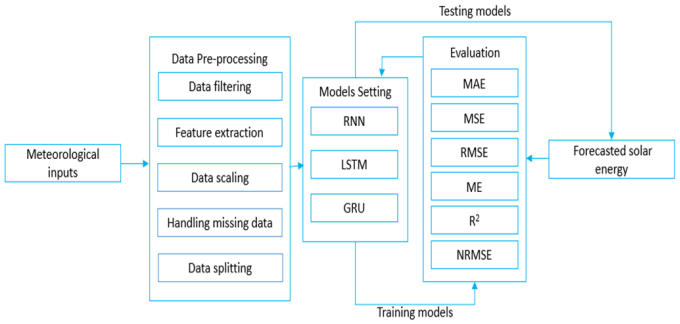


Figure 1: Our method for predicting solar energy

3.1 Dataset

The datasets used in this work, they represent meteorological data of Errachidia province, which is situated in the sunny region of Morocco, it benefits from an important solar energy potential during all the year. This data consists of measurements from three years (2016, 2017 and 2018), and they are used to predict the solar energy level based on a set of measurable weather conditions, in purpose of analyzing the performance of the studied models.

3.2 Preprocessing and Feature Selection

Pre-processing and feature selection are fundamental steps in DL approaches for making collected data in an appropriate form. Preprocessing allows numerous tasks and operations to convert the source data in a clean data, in a way that it can be easily integrated into DL models. It impacts the accuracy of the model and its results. Feature selection provides the most appropriate features that properly affect the learning process, and can minimize the number of variables to efficiently enhance the accuracy of the model and avoid costly computations. Therefore, in this work, we proceed with the following steps to prepare and select features from the source data: extraction of target data inputs, selection of relevant features, filling in the null values, normalization of the data, and adjustment of the steps to be taken into account for prediction.

The source data give multiple features but not all of them are important for use in the prediction. In our case, we have selected four important features: solar energy, temperature, humidity and pressure. These features are the most and highly correlated variables to the targeted output (solar energy) according to Pearson correlations [31]. The objective of this step is to help developing most accurate models that learn from the most correlated data in order to give the most accurate forecasts. We note that we have selected the Median in order to fill missing data.

The features of a given dataset are usually presented at different scales. To ensure that all these values are at the same scale and to add uniformity to our dataset, we employ the Min-Max scale that transforms all values between 0 and 1, which removes the noise from our data and simplifying the learning process of our models.

Besides, the meteorological data are time-series with a time step of 30 min, and therefore, the actual data values are needed as inputs for making predictions. Time-series data cannot use future values as input features, whereas the inputs of a time-series model are the values of past features. In this work, we adjust our model to learn from the past to predict PV solar energy for the next half hour. This choice has been adopted because of the size of the data and to perform real-time predictions for an efficient and optimized management. Finally, based on best practices in the area of data analytics, the dataset is split into 80% and 20% were respectively used to train and test the models developed.

3.3 Training and models parametrization

Each of the models developed must be parameterized during the training phase in order to be able to provide the most accurate forecasts while taking into account the size of the batches. After setting different parameters for our DL models during the training and evaluating the results obtained, the Adam optimizer has proved to be the best optimizer and was then selected as the common parameter for all proposed DL models. ADAM, which is widely used and works better than other stochastic optimizers in empirical results [32], would allow the models to learn quickly. In addition, the tanh activation function has shown a good fitting for all DL models. On the other hand, we notice that our DL models have different architectures and that the number of layers and neurons for example has been fixed after tuning several values and selecting the ones that have given the best precisions. Figure 2 illustrates the neural network architecture, and Table 1 shows details on the architectures and the key parameters of each model suggested for the dataset studied.

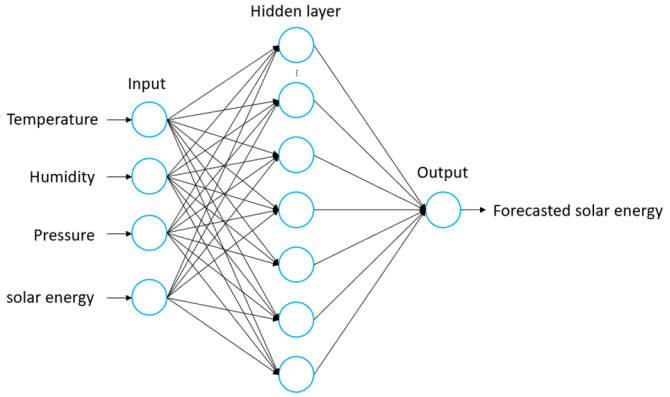


Figure 2: ANN architecture for predicting solar energy

Table 1: Architecture and parametrization of the models.

Models	Parameters				
	Layers	***Epocs***	***Activation Function***	***Optimizer***	***Batch size***
RNN	RNN cell Of 100 units	20	Tanh	Adam	12719
LSTM	LSTM cell Of 100 units	20	Tanh	Adam	12719
GRU	GRU cell Of 100 units	20	Tanh	Adam	12719

3.4 Performance metrics

To evaluate the performance of the proposed DL models for PV solar energy prediction. A set of evaluation metrics widely-used are employed to evaluate the results accuracy forecasting. In this step, we have applied the most suitable to the context of DL and regression problems. For this purpose, five performance metrics, whose equations are presented below, have been employed for models testing as well as training: MAE (Mean Absolute Error), MSE (Mean Square Error), RMSE (Root Mean Square Error), Max Error (ME) and R squared (R^2).

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \tag{14}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2 \tag{15}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \tag{16}$$

$$ME = \max_{1\leq j\leq n}|y_j - \hat{y}_j| \tag{17}$$

$$R^2 = 1 - \frac{SS_{reg}}{SS_{tot}} \tag{18}$$

where y is the actual output, $\hat{y}$ is the predicted output and n is the number of samples. Eq. 14 computes MAE as the average of the absolute errors (absolute values indicating the differences between the actual and the predicted values). Eq. 15 shows the MSE, which is the average squared errors (difference between the real values and what is estimated. Eq. 16 calculates RMSE, which is the square root of the MSE, and is applied in cases with small errors [33], whereas Eq. 17 measures the maximum residual error ME and reveals the worst error between the actual and predicted value. In addition, to calculate the predictive accuracy of proposed models, we also used R-squared, which is a statistical metric in a regression model, to identify the proportion of variance of the dependent variable that can be made explicit by the independent variable. Expressly, Eq. 18 indicates the R-squared measure to which the data correspond in the regression model (the goodness of fit) where SS_{reg} is the sum of the squares due to the regression and SS_{tot} is the total sum of the squares [34]. Eventually, in the different steps outlined in this section, we employed a set of technical tools for the implementation of the studied algorithms based on the Python libraries, including Pandas, NumPy, SciPy and Matplotlib in a Jupiter Notebook, in addition to Scikit-learn (sklearn) and TensorFlow.

4 Results and discussion

In this section, we present and discuss the results of our three studied models tested for PV solar energy forecasting. It should be noted that although R-squared (R^2) is a commonly used metric, it is not regarded as an effective indicator of the models fit with the data [35]. A low or high R^2 value does not always indicate that the model is wrong or that it is automatically right [36]. Due to the controversy concerning the efficiency of R^2 as an appropriate metric for determining the best regression model [35, 36], we additionally evaluate our results using other statistical metrics that are the most widely used: MAE, MSE, Max Error and RMSE. The RMSE remains the most frequently employed in the regression area [37]. Table 2 show the metrics values corresponding to solar energy prediction in real time of studied models. Considering the parameters presented in Table 2, we can see the results for one-half hour ahead prediction of our experiments for the recurrent neural networks: RNN, LSTM and GRU. In spite of their architectural differences, RNN and LSTM have demonstrated similar performance, showing that the better model is strongly dependent on the task. Thanks to their ability to handle long-term dependencies in sequential information for time-series predicting applications, both RNN and LSTM perform better with very high accuracy as shown by their related R^2 values which are respectively equal to 94.68% and 94.26%. GRU gives slightly less performance than the RNN and LSTM with 90.32 as value. It also causes errors unlike RNN and LSTM. We observe that for GRU, 2.74, 15.53, 3.94, 37.30 are given as MAE, MSE, RMSE and ME values, respectively. Furthermore, these measures are nearly similar for RNN and LSTM with MAE=1.83, MSE=8.53, RMSE=2.92and ME=30.42 for RNN, and MAE=1.90, MSE=9.20, RMSE=3.03 and ME=29.73 for LSTM. Figure 3, Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8 show their training and testing curves. In reality, more accurate prediction results mean less uncertainty and fluctuations in PV solar energy generation. Accuracy is a crucial factor in the efficient planning and use of electrical energy and energy systems with high PV power penetration. In short, as shown in Table 2, the prediction capability of RNN and LSTM seems very reliable compared to GRU under meteorological conditions, and they are better adapted to meet real-time PV solar energy prediction needs with high accuracy. Due to their ability to minimize errors during learning and test processes, they can reduce uncertainties related to the operation and planning of PV energy integrated management systems.

Table 2: Metrics related to PV solar energy dataset.

Models	Data	MAE	MSE	RMSE	ME	ME
RNN	Training	1.73	7.22	2.68	34.06	95.62
	Test	1.83	8.53	2.92	30.42	94.68
LSTM	Training	1.77	7.65	2.76	35.42	95.36
	Test	1.90	9.20	3.03	35.42	94.26
GRU	Training	2.63	13.69	3.70	34.26	91.70
	Test	2.74	15.53	3.94	37.30	90.32

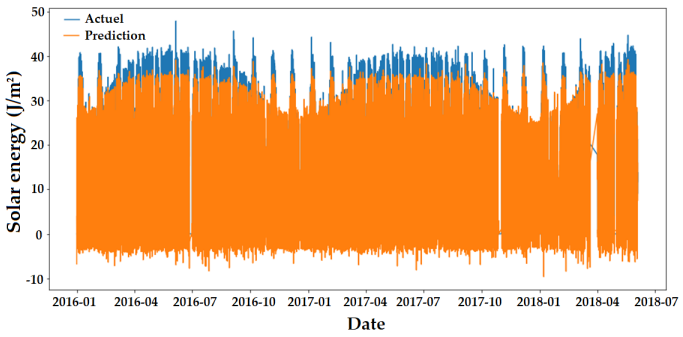


Figure 3: RNN training curves of solar energy

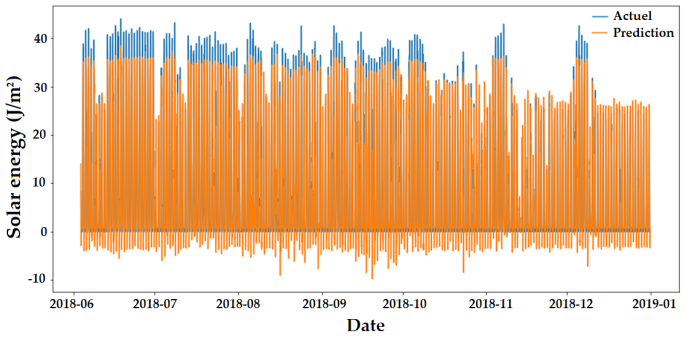


Figure 4: RNN testing curves of solar energy

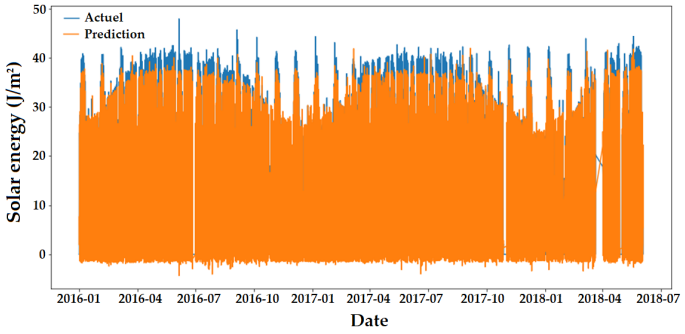


Figure 5: LSTM training curves of solar energy

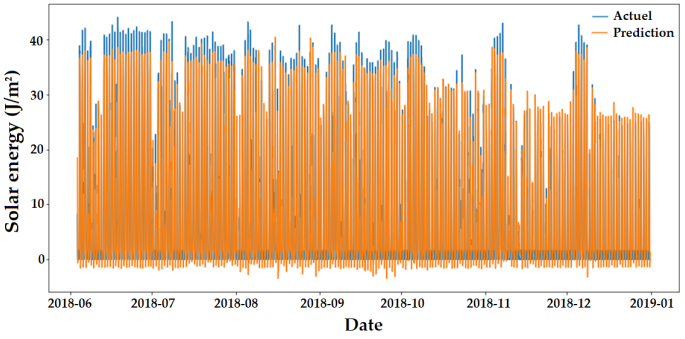


Figure 6: LSTM testing curves of solar energy

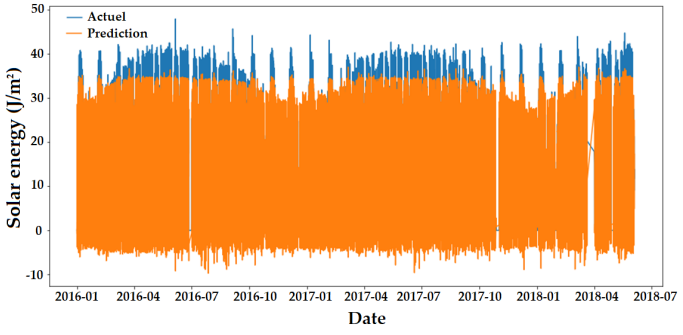


Figure 7: GRU training curves of solar energy

Besides, while RMSE is regarded in the literature as a good metric for evaluating and comparing regression models, it is still difficult to interpret properly. Therefore, we also normalize the RMSE values to provide a more meaningful representation of our results, which would allow more efficient conclusions. The normalized RMSE (NRMSE) represents the rate of the RMSE value and the range (the maximum value minus the minimum value) of the actual values. Table 3 shows a results synthesis based on this metric for real-time solar energy prediction. We can observe that RNN and LSTM provide good and similar NRMSE values with 0.055 and 0.058 respectively.

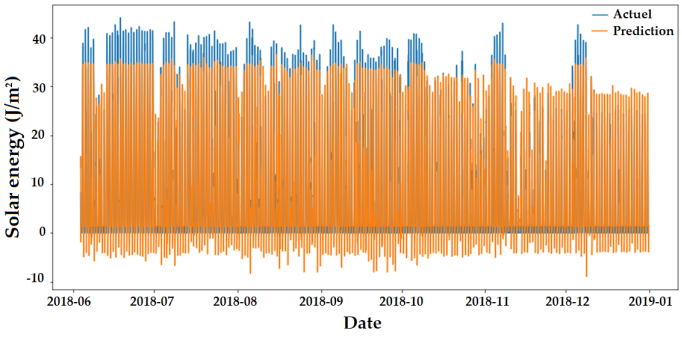


Figure 8: GRU testing curves of solar energy

Table 3: Results synthesis for real time prediction.

Models	Data	NRMSE
RNN	Training	0.055
	Test	0.066
LSTM	Training	0.058
	Test	0.068
GRU	Training	0.077
	Test	0.089

5 Conclusion and Perspectives

With the growing deployment of solar energy into modern grids, PV solar energy prediction has become increasingly important to deal with the volatility and uncertainty associated with solar power in these systems. In the literature, different models have been proposed for the prediction of PVs by using the AI techniques capacity. However, the majority of the research contributions in this area offer different techniques for making separate short-term and long-term forecasts, and do not focus on real-time forecasts. For this reason, we are working to find a model capable of producing continuous real-time forecasts using meteorological data, which is primordial for providing important decision support for power system operators to ensure more efficient management and secure operation of the grid and enhance the cost-effectiveness of the PV system. In this study, we investigated the efficiency of three different DL models: RNN, LSTM and GRU. The models were tested with data from a Moroccan region to make real-time PV forecasts. Further, the analysis of the results was based on six metrics MAE, MSE, RMSE, ME, R^2 and NRMSE to determine both forecast accuracy and margin of error. The results prove the efficiency of RNN and LSTM for real-time PV prediction thanks to their ability to address long-term dependencies in sequential information for time-series regression problems, compared to GRU that gives slightly less performance

than the RNN and LSTM. From these findings, it can be concluded that RNN and LSTM showed very high accuracy and low errors, they are reliable models that can minimize errors in the learning and testing process, in addition to reduce the uncertainties associated with the operation and planning of integrated PV energy management systems. Moreover, they are promising techniques that also appears to be suitable for long-term PV forecasting and should be studied and recommended as a possible unified and standard tool for real-time, short-term and long-term PV prediction. As future work of this study, we plan to enhance these models by employing other approaches, extend them to longer term PV forecasting, as well as to studying other promising deep learning methods in view of providing and deploying a powerful model for PV energy prediction.

Conflict of Interest The authors declare no conflict of interest.

References

[1] E. Hache, "Do renewable energies improve energy security in the long run? Septembre 2016 Les cahiers de l'économie - n^{o}109 Série Recherche," 2016.

[2] M.G. De Giorgi, P.M. Congedo, M. Malvoni, "Photovoltaic power forecasting using statistical methods: Impact of weather data," IET Science, Measurement and Technology, **8**(3), 90–97, 2014. DOI:10.1049/iet-smt.2013.0135.

[3] P. Li, K. Zhou, S. Yang, "Photovoltaic Power Forecasting: Models and Methods," 2nd IEEE Conference on Energy Internet and Energy System Integration, EI2 2018 - Proceedings, 1–6, 2018. DOI:10.1109/EI2.2018.8582674.

[4] U.K. Das, K.S. Tey, M. Seyedmahmoudian, S. Mekhilef, M.Y.I. Idris, W. Van Deventer, B. Horan, A. Stojcevski, "Forecasting of photovoltaic power generation and model optimization: A review," Renewable and Sustainable Energy Reviews, 81(August 2017), 912–928, 2018 DOI:10.1016/j.rser.2017.08.017.

[5] A. Youssef, M. El-telbany, A. Zekry, "The role of artificial intelligence in photovoltaic systems design and control: A review," Renewable and Sustainable Energy Reviews, 78(November), 72–79, 2017. DOI:10.1016/j.rser.2017.04.046.

[6] A.P. Yadav, A. Kumar, L. Behera, "RNN Based Solar Radiation Forecasting Using Adaptive Learning Rate," 442–452, 2013.

[7] G. Li , H. Wang, S. Zhang, J. Xin, H. Liu,"Recurrent Neural Networks Based Photovoltaic Power Forecasting Approach," 1–17, 2019.

[8] P. Li, K. Zhou, X. Lu, S. Yang, "A hybrid deep learning model for short-term PV power forecasting," Applied Energy, (November), 114216, 2019. DOI:10.1016/j.apenergy.2019.114216.

[9] M. Abdel-Nasser, K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," Neural Computing and Applications, **31**(7), 2727–2740, 2019. DOI:10.1007/s00521-017-3225-z.

[10] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, Y. Du, "Short-Term Photovoltaic Power Forecasting Based on Long Short Term Memory Neural Network and Attention Mechanism," IEEE Access, **7**, 78063–78074, 2019. DOI:10.1109/ACCESS.2019.2923006.

[11] Y. Yu, J. Cao, J. Zhu, "An LSTM Short-Term Solar Irradiance Forecasting Under Complicated Weather Conditions," IEEE Access, **7**, 145651–145666, 2019. DOI:10.1109/ACCESS.2019.2946057.

[12] J. Wojtkiewicz, M. Hosseini, R. Gottumukkala, T.L. Chambers, "Hour-Ahead Solar Irradiance Forecasting Using Multivariate Gated Recurrent Units," 1–13, 2019. DOI:10.3390/en12214055.

[13] M. Hosseini, S. Katragadda, J. Wojtkiewicz, R. Gottumukkala, "Direct Normal Irradiance Forecasting Using Multivariate Gated Recurrent Units," 1–15, 2020. DOI:10.3390/en13153914.

[14] K. Yan, H. Shen, L. Wang, H. Zhou, M. Xu, Y. Mo, "Short-Term Solar Irradiance Forecasting Based on a Hybrid Deep Learning Methodology," 1–13.

[15] M.Q. Raza, M. Nadarajah, C. Ekanayake, "On recent advances in PV output power forecast," Solar Energy, **136**(September 2019), 125–144, 2016. DOI:10.1016/j.solener.2016.06.073.

[16] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin, Z. Hu, "Photovoltaic and solar power forecasting for smart grid energy management," CSEE Journal of Power and Energy Systems, **1**(4), 38–46, 2016. DOI:10.17775/cseejpes.2015.00046.

[17] A. Tuohy, J. Zack, S. E. Haupt, J. Sharp, M. Ahlstrom, S. Dise, E. Grimit, C. Mohrlen, M. Lange, M. G. Casado, J. Black, M. Marquis, C. Collier, "Solar forecasting: Methods, challenges, and performance," IEEE Power and Energy Magazine, **13**(6), 50-59, 2015.DOI:10.1016/j.111799.

[18] G. Notton, M.L. Nivet, C. Voyant, C. Paoli, C. Darras, F. Motte, A. Fouilloy, "Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting," Renewable and Sustainable Energy Reviews, 87(December 2016), 96–105, 2018. DOI:10.1016/j.rser.2018.02.007.

[19] A.A. Merrouni, F.E. Elalaoui, A. Mezrhab, A. Mezrhab, A. Ghennioui, "Large scale PV sites selection by combining GIS and Analytical Hierarchy Process. Case study: Eastern Morocco," Renewable Energy, 2017. DOI:10.1016/j.renene.2017.10.044.

[20] A. Alzahrani, P. Shamsi, C. Dagli, M. Ferdowsi, "Solar Irradiance Forecasting Using Deep Neural Networks," Procedia Computer Science, **114**, 304–313, 2017. DOI:10.1016/j.procs.2017.09.045.

[21] A. Alzahrani, P. Shamsi, M. Ferdowsi, "Solar Irradiance Forecasting Using Deep Recurrent Neural Networks," In 2017 IEEE 6th international conference on renewable energy research and applications (ICRERA), 988-994, 2017.

[22] Y. LeCun, Y. Bengio, G. Hinton. "Deep learning" Nature **521**(7553), 436-444, 2015.

[23] H. Wang, Z. Lei, X. Zhang, B. Zhou, J. Peng, "A review of deep learning for renewable energy forecasting," Energy Conversion and Management, **198**(July), 111799, 2019. DOI:10.1016/j.enconman.2019.111799.

[24] X. Qing, Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," Energy, **148**, 461–468, 2018. DOI:10.1016/j.energy.2018.01.177.

[25] H. Liu, X. Mi, Y. Li, "Comparison of two new intelligent wind speed forecasting approaches based on Wavelet Packet Decomposition , Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Arti fi cial Neural Networks," Energy Conversion and Management, 155(July 2017), 188–200, 2018. DOI:10.1016/j.enconman.2017.10.085.

[26] Y. Wang, W. Liao, Y. Chang, "Gated Recurrent Unit Network-Based Short-Term Photovoltaic Forecasting," Energies, **11**(8), 2163, 2018. DOI:10.3390/en11082163.

[27] K. Cho, B. van Merrienboer, C. Gulcehre, F. Bougares, H. Schwenk, Y. Bengio, "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," arXiv preprint arXiv:1406.1078, 2014.

[28] M. C. Sorkun, C. Paoli, Ö. D. Incel, "Time series forecasting on solar irradiation using deep learning," 2017 10th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, 151-155, 2017.

[29] H.M. Lynn, S.B.U.M. Pan, "A Deep Bidirectional GRU Network Model for Biometric Electrocardiogram Classification Based on Recurrent Neural Networks," IEEE Access, **7**, 145395–145405, 2019. DOI:10.1109/ACCESS.2019.2939947.

[30] Z. Che, S. Purushotham, K. Cho, D. Sontag, Y. Liu, "Recurrent Neural Networks for Multivariate Time Series with Missing Values," Scientific Reports, 1–12, 2018. DOI:10.1038/s41598-018-24271-9.

[31] H. Zhou, Z. Deng, Y. Xia, M. Fu, "A new Sampling Method in Particle Filter Based on Pearson Correlation Coefficient," Neurocomputing, 2016. DOI:10.1016/j.neucom.2016.07.036.

[32] D.P. Kingma, J. Ba:, "Adam: A Method for Stochastic Optimization," Proceedings of the 3rd International Conference on Learning Representations (ICLR), p. No abs/1412.6980, 2014.

[33] F. Wang, Z. Mi, S. Su, H. Zhao, "Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters," Energies, **5**(5), 1355–1370, 2012. DOI:10.3390/en5051355.

[34] B.J. Miles, "R Squared , Adjusted R Squared," Wiley, 2014.

[35] C. Onyutha, "From R-squared to coefficient of model accuracy for assessing " goodness-of-fits "," Geoscientific Model Development Discussions, 1-25, 2020. doi: 10.5194/gmd-2020-51

[36] D.B. Figueiredo Filho, J.A.S. Júnior, E.C. Rocha, "What is R2 all about?," Leviathan (São Paulo), **3**, 60–68, 2011.

[37] I. Jebli, F. Belouadha, M.I. Kabbaj, "The forecasting of solar energy based on Machine Learning," (c), 2020 International Conference on Electrical and Information Technologies (ICEIT). IEEE, 1-8, 2020. DOI: 10.1109/ICEIT48248.2020.9113168

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 356-368 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Novel Infrastructure Platform for a Flexible and Convertible Manufacturing

Javier Stillig*, Nejila Parspour

University of Stuttgart, Institute of Electrical Energy Conversion, Stuttgart, 70569, Germany

ARTICLE INFO

Article history:
Received: 20 November, 2020
Accepted: 10 January, 2021
Online: 22 January, 2021

Keywords:
Industry 4.0
Factory of the Future
Intralogistics
Wireless Power Transfer

ABSTRACT

Sales behavior and the technical development of products influence their fabrication. As market influences become increasingly volatile and unpredictable, factories will have to adapt their manufacturing to market trends even more in the future. Adaptation is referred to as convertibility and can be achieved, among other things, by mobile and intercompatible machines. Enabling machines to be mobile, its power supply must be wireless and it should be possible to locate it on the shop floor at any time. With the help of the infrastructure platform Intelligent Floor, machines in future factories can be made more mobile than they are today. In combination with the novel autonomous guided vehicle BoxAGV, the platform offers a cost-efficient and highly flexible solution for in-house transport tasks. The transport of goods can be performed based on lot size one and thereby opens a wide field in logistics automation.
This paper is an extension of work originally presented in MELECON 2020 and describes the concept and the functionality of the open platform that is implemented so far. It shows how to apply the platform to a real industrial manufacturing environment and highlights the resulting manufacturing benefits. Finally, the next development steps on the platform and machine side are presented.

1 Introduction

The rigid linking of specialized individual machines by conveyor belts is the characteristic of today's production lines. Future production lines must consist of mobile, intercompatible machines that can be easily combined to form new production lines. This universality can be achieved with a reduction in functionality of the individual machine and by supplying it wirelessly with energy and data. The compatibility of the individual machines allows a higher-level control system to perform interconnected operation in order to follow in real-time market requirements.

These statements are verified in section 2 by using the automobile production as an example. Adapting today's productions to the new requirements, the Intelligent Floor (IF) poses as a universal platform for the building infrastructure and helps to achieve a flexible and convertible production. The IF is a raised floor that is currently under development. It is designed as a modular floor with individual panel elements that have sensor and actuator functionalities. The V0.5.x prototype can already supply production equipment by wireless power and measures the weight of objects on the floor panel. Furthermore, the IF has an Light Emitting Diodes (LED) display system in each panel included that is built-up with four LED stripes and located in a cross shape on the top of one panel.

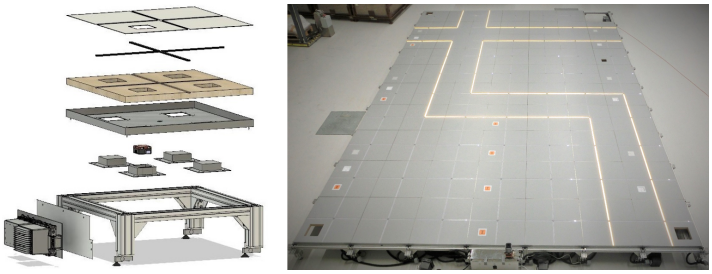


Figure 1: Exploded view of the IF and the prototype V0.5.2 test set-up, which has a size of 10 m × 7 m (L × W). When the picture was taken, as an example a walking path in the production was reproduced by the IF.

The IF consists of the four main components substructure, control and power unit, module inserts and panel (Figure 1). It receives its sensor and actuator capabilities from the module inserts or individually configurable panels. The IF basic version is already equipped with a detection and display functionality. Due to the modular design as a raised floor, the flexible positioning of machines on the IF is made possible through an omnipresent media supply. The IF can generate advantages for users and operators of production facilities that would not be achievable with conventional solutions. Two prototypes of the IF are currently being tested

*Corresponding Author: Javier Stillig, Pfaffenwaldring 47, 70569 Stuttgart, Germany, +49 711 68567819 & javier.stillig@iew.uni-stuttgart.de

in the ARENA 2036[1] using specific use cases of the convertible production.

A first prototype V0.1.1 of the novel Autonomous Guided Vehicle (AGV), hereafter named *BoxAGV* and shown in Figure 2. It is based on a simple frame construction made of aluminum profiles, to which four meccanum wheels are attached. The four drive motors are connected each to an inverter that receive their commands from the controller in real-time. The BoxAGV has installed a camera for identifying the route, which is mounted centrally in the frame. By using the Wireless Power Transfer (WPT) the battery of the vehicle can be charged dynamically. The entire vehicle technique is embedded in a standard box [1] made of polypropylene with the outer dimension 600 mm × 400 mm × 213 mm (L × W × H). It offers after its integration still 50% of the origin transport volume.

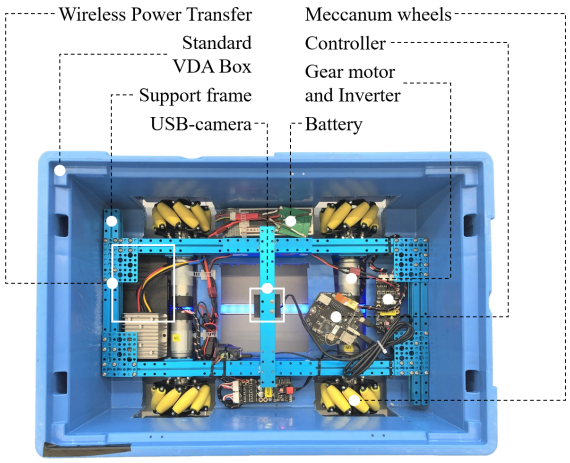


Figure 2: Design of the first BoxAGV prototype V0.1.1, where frame, driving and control components are built-in a standard VDA box.

The combination of IF and BoxAGV meets the requirements of convertible production in terms of logistics flexibility and usability better than the solutions known so far. To verify this hypothesis, a study has been started in the ARENA 2036. Using a 70 m^2 large-scale demonstrator of the IF that interacts with two BoxAGV prototypes, highlights the expected gains for future logistics systems. To initiate a technical discourse, the novel concept is therefore introduced in this paper and evaluated based on a real field application. Additional to the logistic applications, further use cases of the IF are outlined and the importance of WPT for mobile applications is analyzed in more detail.

2 Convertible Production

To understand the purpose of the IF and its benefits for the users, it is essential to get to know more about convertible production. Scientists, in particular the production technology institutes of the University of Stuttgart, have been researching the subject of convertible production since more than 25 years, e.g. [2]–[7]. One of the central findings of their research work, carried out so far, is that in response to the fast changing requirements of the global market, production equipment must be designed to be more modular, mobile and compatible with each other. This is a fundamental of the future factories production concept. To implement this concept is necessary in order to be able of continuing producing economically with a high value-added share. Taking the automotive industry as an example, it will be outlined which factors drive the change in requirements and what kind of solutions can be helpful.

2.1 Driving Factors

The requirement to transform existing production systems to convertible production systems will be illustrated based on two general market developments. These market developments are described using the automotive industry as an example. Based on the research study [8], the number of registered automobile models has increased by 349% between 1990 and 2014 in Germany. In addition, the number of possible equipment options increased with the model variants, so that today almost every new car allows several hundred million of configuration options and is consequently built as a one-off. The second market trend is the shortening of a model's life cycle. While the life cycle of a Volkswagen Golf 1 was nine years in 1983, it has been reduced to just four years in 2012 [9]. These two market movements are critical not only for the automaker, but also for all their suppliers, who must adapt their products and production facilities to market developments.

The driving force behind the increase in the number of variants and the reduction in product lifecycle times is the interaction between the market and suppliers. The customer's desire for choice and product individualization is matched by a large number of competing manufacturers who expect to gain sales advantages by offering variants ever faster. The ongoing connection of people through the Internet means that the market and suppliers are better connected than before. This means that a reduction in the number of variants and an extension of the life cycle time cannot be expected in future. Therefore, it is necessary for product makers to adapt their production concepts to the actual situation and to move to convertible production concepts [10, 11]. Here it is important to couple the product life cycle with the life cycle of the production machine. Thus, the rigid coupling of product and machine is not critical if the period of the product life cycle is equal to or longer than the life cycle of the machine. However, the case in which the life cycle of a product is much shorter than that of a machine is critical for automated manufacturing. This causes production difficulties: Machines can no longer be planned, built and commissioned as quickly as the product life cycle requires. As shown in Figure 3, it is therefore necessary to implement a system that synchronizes both life cycles. This synchronization is the main task of the convertible production, as shown by the example in subsection 2.3.

2.2 Solution

There are many possible solutions for synchronizing product and machine life cycles. Two solutions will be highlighted in this paper: Transparency in the production process and faster product changeovers. Transparency is understood as the knowledge of the current position and status of all objects in the production process. Especially with knowledge of the material positions, the flow of goods can be optimized and non-production-related idle times can

[1] ARENA 2036 is a flexible research platform for the mobility and production of the future and is located in Stuttgart.

be reduced or eliminated. By shortening production time, more products can be produced in the same time period. As a result, the proportion of production costs allocated to the product becomes smaller, enabling the product to be produced cheaper. A solution to manage the increase in the speed of product changeovers is the use of modular, manufacturer independent and compatible production equipment. In some cases, solutions for defined manufacturing steps already exist, as it can be shown in the next subsection 2.3.

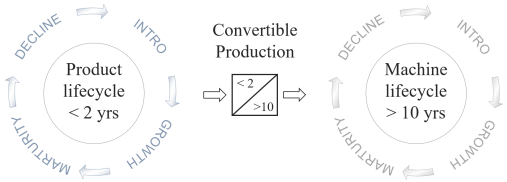


Figure 3: Product and machine lifecycle coupled by the convertible production.

2.3 Example

The relationship between product and machine life cycle is to be demonstrated using the example of wire harness production. In this case, the product is a wire harness (Figure 4). A wire harness is composed of different wires of various lengths and with a range of wire terminations. The use of wire harnesses is particularly widespread in the automotive industry. For example, an average equipped Volkswagen Golf 7 has a wiring harness with almost 1,000 different wires of a total length around 1.6 km and weighing up to 60 kg [12]. Each individual wire must be prepared individually for installation in the car. Due to the high demand quantities in the automotive industry, this requires a fully automatic machine that cuts wires to length, marks, strips and terminates them with a connector. Figure 5 shows a wire processing machine, type Zeta 640 / 650 that is manufactured by Komax. The machine is able to prepare up to 36 different wires with a cross-section from 0.22 to 6 mm^2 [13].

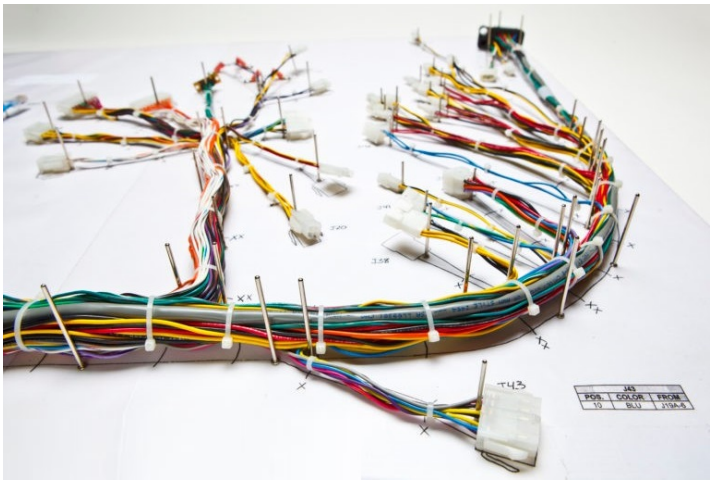


Figure 4: Wire harnesses are used in various industries, like in the consumer industry for wiring household appliances. One of its main advantages is the significantly reduced assembly time compared to single wire installation. Photo: TES.

Today, wiring harnesses take into account the vehicle equipment features ordered by the customer and they are built-to-order. This is done in order to reduce the overall vehicle weight and thus also reduce fuel consumption and exhaust emissions. Such a production strategy leads to a large number of theoretical combinations, which [12] puts at 425 billion. Due to the high variance of the wire harness design as a result of changing vehicle equipment options and cyclical model changes, the wire processing machine is modular in itself. For example, there are up to 13 process modules for the model shown, which take over the feeding of different sized ferrules or the wire labeling. This enables the machine operator to react quickly to product changes. With the help of the modular machine design strategy, the fast-moving product cycle is coupled with the long-lasting machine cycle.



Figure 5: Wire processing machine for the fully automatic production of prepared wires for the use in wire harnesses. Photo: Komax.



Figure 6: The photo shows a wire harness production facility where workers assemble prefabricated wires into a wire harness. Photo: OWC.

Looking at the subsequent production process of the wire harness manufacturing, it is obvious that the fully automatic wire preparation is followed by a completely manual production process. As shown in Figure 6, the prepared wires are now assembled by hand on a pin board to form a complete wire harness. So far, it has not

been possible to build a machine for the automatic production of wire harnesses with a development and construction time that is shorter than the product life of a wire harness. Alternatively, it would also be possible to build a modular machine or a combination of individual system modules that are capable of producing complex wire harness layouts automatically. However, this has not yet been achieved either. Thus, product and machine life cycles for wire harness manufacturing in the automotive industry are not synchronized by machines, but manually by workers as shown in Figure 6.

The example of wire harness production can also be applied to other production processes in and outside of the automotive sector. It highlights the demands made on the machine life cycle in terms of flexibility, modularity and degree of automation.

At the same time, fully automatic machines should also be durable and robust investment assets that must guarantee a high output over their life cycle. Therefore, it is necessary to readjust the balance between product and machine life cycle by using new processes and elements from the field of convertible production. The IF described below is a new element from the field of convertible production. It uses an example from intralogistics to show how product and machine life cycles can be synchronized with each other based on machines.

3 Intelligent Floor

The Figure 7 shows a test surface of the IF prototype V0.5.1, which consists of 12 identically constructed panels, their control and power units and its substructure. In addition to the IF basic functionality of visualization and detection, a WPT unit is built into each panel. One panel has a surface area of 1 m × 1 m and a thickness of 46 mm.The IF is constructed as a modular raised floor, which is adjustable in its height. In this specific case, the IF has a total height of 40 cm. The prototype V0.5.x is designed to carry static loads of min. 1,500 $^{kg}/_{m^2}$. A special feature of the IF is the interconnection of the energy units to form a DC 24 V grid. This enables higher power to be provided at single points within the grid than could normally be provided by just the power supply unit of each panel. Furthermore, each panel has a control unit for processing its sensors and actuators. The Panel Controller (P-CTRL) is the central element of the control unit and communicates with the higher-level control system, the Section Controller (S-CTRL), over the 100 Mbps Ethernet network. The S-CTRL itself manages by definition up to 256 panels. It can be implemented as a physically present controller or a cloud-based solution. Neighboring sections can be combined to a cluster, where the Cluster Controller (C-CTRL) is software-based controller. The IF prototype V0.5.x panels are already hot-plug capable and have the following sensor / actuator capabilities:

- The sensor system for object detection, which is based on load cells, is integrated into the substructure of the IF. Each load cell can measure up to 1,000 kg. Since each stand of the floor is equipped with a load cell, a network of weight values is built up over the area, which enables object detection with a positional accuracy of approx. 0.5 m.
- The panel's surface is divided into four quadrants by LED stripes consisting each of up to 144 $^{LED}/_{m}$. The LEDs are controlled individually, in groups or all together in color and brightness by the P-CTRL.
- The IF contains up to four module slots, where Figure 7 shows one with a DC 24 V / 240 W WPT unit equipped module slot. In the IF prototype V0.5.2, the module slots contain further functionalities. For example, there are installed an e-paper display modul, a radio-based object localisation module and a power socket to provide three-phase current AC 400 V / 16 A instead of the WPT unit.
- Depending on the application, it may be useful to combine other functions in addition to the basic functions. This can be done in one panel or over a certain panel area. As the example in Figure 11 shows, the visualization functionality and the WPT unit were used here to charge a BoxAGV.

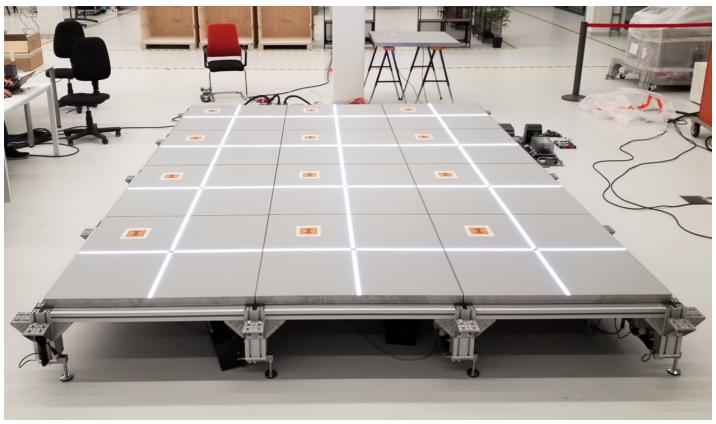


Figure 7: Prototype V0.5.1 of the IF, which was setup for test purposes in the ARENA 2036. It is assembled from 12 equal panels, each using one module slot for functionality expansion. The module is a 240 W WPT unit. Further, all stands are equipped with load cells, measuring the loads applied to the floor continuously.

3.1 Localization

The object localization in prototypes V0.2 and V0.5 is realized with the help of bending beam load cells, which are mounted under each stand of the substructure. During a measurement process, the weight force acts on the bending beam and generates an elastic deformation. This strain is converted into an electrical signal by the four strain gauges glued to the bending beam. In order to convert the signal into a weight value, the strain gauges are connected in a Wheatstone bridge and are made available to the P-CTRL by a subsequent measuring amplifier with an analog-to-digital converter. The controller forwards the weight values to a database located in the S-CTRL over the Ethernet network at intervals of 1 second. The database is designed as a ring buffer with a storage capacity of 256 panels, equal to one section, and a buffer time of 60 min. This allows the access to a total of 921,600 weight values.

In order to detect static or moving objects on the IF, the weight values of a time point that are above a defined threshold are extracted from the database. Since the position of the load cells is fixed and each weight value can be uniquely assigned to a load cell, it is possible to detect moving objects. The weight value itself can serve as an indicator of which type of object is being detected. For example, a measured value of 50 to 120 kg is an indication of a

human being. Object detection can be further refined by comparing the change in weight of several neighboring panels. If the weight sensors of panel Ⓐ measure the mass m at time t_0 and the weight sensors of panel Ⓑ, which lies next to panel Ⓐ, measure the mass m at time $t_1 > t_0$, it can be assumed that an object with mass m moves in time $t_1 - t_0$. Since the distance s between the panels Ⓐ and Ⓑ is known due to the uniform structure, the velocity v of the object is known as well. This is useful as another indicator of object types. For example, people in production areas typically move at speeds between 0.8 and 1.6 m/s. The object detection gains reliability if weight and velocity values are now combined.

Another possibility of object localization is currently being implemented in a panel of the IF prototype V0.5.2. With the help of a film-based sensor technology that is installed directly under the floor surface, the measuring accuracy and robustness is evaluated. Thus, the substructure of the IF's next prototype will be simplified and made ready for further applications outside industrial use.

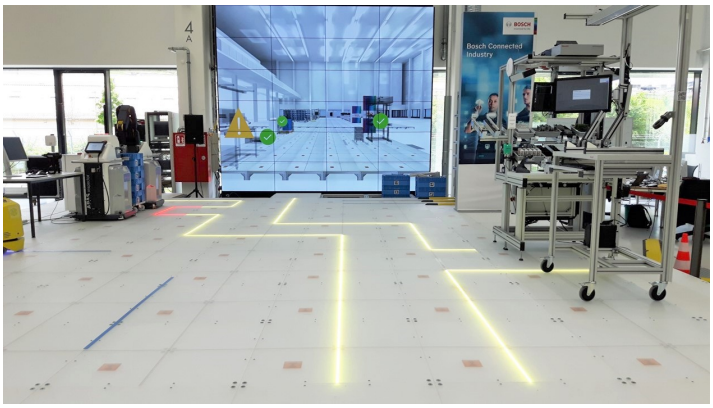


Figure 8: The prototype V0.2.1 uses the LED visualization system to show a walkway on the production area and serves as orientation for visitors. In the top left corner of the photo, an area with mobile production equipment can be seen. A red no-entry zone visually warns visitors that they are not allowed to enter this area.

3.2 Visualization

In the prototype V0.2.1, LED stripes with an LED density of 144 $^{\text{LED}}/_{\text{m}}$ are installed. Thus, a pixel density of 284 $^{\text{px}}/_{\text{m}^2}$ is achieved. In the panels of the prototypes V0.5.x, the pixel density has been reduced to 116 $^{\text{px}}/_{\text{m}^2}$, which is sufficient for information purposes and for steering the BoxAGV. Each LED is from type WS2812B and can be controlled in its color and brightness by the P-CTRL. One LED stripe is embedded in an aluminum profile, on which a translucent cover protects the LEDs. They are wired to a connection box for the data and power supply from the P-CTRL and the DC 5 V source. In version V0.5.x, the panel is connected via cable to the control and power supply. This connection will no longer be necessary in the following prototype generation to ensure a fast and ergonomic change of the panels.

Since the LED visualization system is mounted on the panel in a cross shape, a chessboard-like grid is created on the surface when several panels are installed. Static dot and line patterns, but also dynamic patterns can be displayed in this grid. Thereby the synchronous visualization is performed by the S-CTRL, which updates all P-CTRLs connected to the section at the same time. As shown in Figure 8, the visualization system enables man / machine and inter-machine communication easily. For example, dynamic LED markings can replace static adhesive tape markings for indicating temporary storage locations for materials. As well it can be used as a guidance system for workers to find other persons or specific objects. Instead of separate ANDON displays, production states can be visualized to the worker on the IF. Moreover, it can also be used for zone identification and in combination with a secure object detection, for safeguarding hazardous areas in production environments. This application example is described in more detail in the subsection 3.4.

3.3 Wireless Power Transfer

Convertible production demands high standards for the mobility of production equipment. To ensure mobility, machines must be freely positionable on the shop floor. Today's usual cable- and tube-based media supply of machines is contrary to the mobility requirements of convertible production. Therefore, the use of WPT systems is promising. The WPT primary units shown in Figure 9 transmit up to 240 W unidirectionally to a static, machine-side secondary unit.

The voltage level is DC 24 V on both WPT sides. The efficiency achievable during energy transfer is 91% over an air gap of 5 mm. As the technical data indicate, the WPT system shown is provided for low-power applications. For example, a workshop cart with electrical label printer was equipped with a secondary side, in order to operate the printer on each charging point of the IF. In addition to the low-power system, a mid-power system is available, whose transfer power is 1,000 W with an efficiency of 90% and a transfer distance of up to 10 mm. Also a high-power system is already in the prototype V0.5.2 in operation. It is able to transfer a power of 3 kW at a distance of 15 to 40 mm. Its max. lateral offset of 40 mm from the centers of both coils helps to charge machines easily. The IF panel's design allows to operate up to four WPT units, so that in theory an energy density of max. 12 $^{\text{kW}}/_{\text{m}^2}$ can be achieved.

However, since mobile and thus lightweight production equipment is required for convertible production in particular, the required electrical power is usually rather low. As described in [14], power densities of less than 4 $^{\text{kW}}/_{\text{m}^2}$ are enough for most applications. Fixed and semi-fixed production equipment with a higher power demand, can be supplied using conventional socket units in the raised floor system. One of the most significant challenges of convertible production is the location-independent and dynamically adaptable energy supply of the production equipment. Unfortunately, this challenge can only be solved to a limited extent with the industrial WPT systems on the market today. Therefore, the Institute of Electrical Energy Conversion (IEW) of the University of Stuttgart is researching on the following three fields of WPT:

Dynamic Charging Looking at the operation of industrial WPT systems, it is remarkable that most of them are designed for static coupling of primary (Tx) and secondary (Rx) units – so defined 0D systems. Even small deviations in $x-$ or $y-$direction from the ideal position of the Rx unit in relation to the Tx unit will reduce the transfer power and efficiency. Therefore, only WPT systems capable to couple Rx and Tx units at least along a route (1D systems) are suitable for the use in the convertible production environment. Even more appropriate would be a 1.5 or 2D system. The Table 1 summarizes the coupling of Tx and Rx units in $x-$ or $y-$direction. A

solution further developed by [15] is the introduction of separately switchable Tx modules in a grid-like structure that are able to operate as 1.5D system or better, but without having the disadvantages of today's solutions.

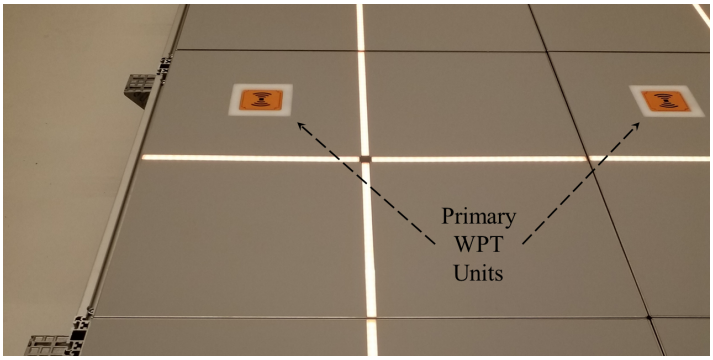
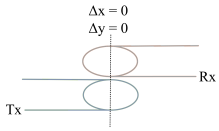
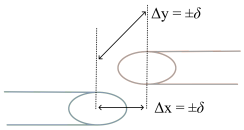
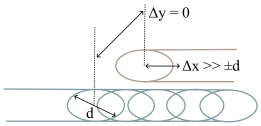
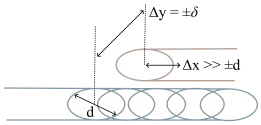
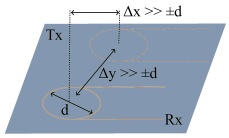


Figure 9: Two low-power Tx units each installed in an IF prototype V0.5.1 panel. The distance between the units is 1 m. Both Tx and Rx units have a uniform size of 100 mm × 100 mm × 47 mm (L × W × H) and can supply on Rx side 10 A at DC 24 V.

Table 1: Five categories of coupling between Tx and Rx coils in x and y planar movements without angular misalignment of the coil planes.

Cat.	Figure	Description
0D		The Tx and Rx are statically coupled, so that the two coils must be exactly aligned with each other.
0.5D		The Tx and Rx are statically coupled, but with a tolerance δ in two dimensions.
1D		The Rx side has one degree of freedom and is static in the second in relationship to the Tx side.
1.5D		The Rx is movable in one dimension and static with a tolerance δ in the second dimension.
2D		The Rx is completely dynamic in both dimensions, so that the Rx can be placed anywhere in relationship to the Tx device.

Bidirectional Power Transfer Moreover, the industrial WPT systems available on the market are characterized by their unidirectional power transfer. Bidirectional systems are not yet available on the market. However, bidirectional energy transfer would have many advantages in the environment of convertible production. For example, battery-buffered machines could help to compensate peak loads in production by feeding energy into the IF's DC grid via the bidirectional WPT system. At low-load times, the batteries can be recharged, which results in cost advantages for the operator in terms of a homogenized energy demand of his production [16, 17].

Power Density A so far unknown key figure is the power density

$$\rho_p = \frac{P_{inst}}{A}\text{ , where } P_{inst}\text{: installed power [W], A: area [m}^2\text{]}$$

needed to operate mobile machines in the environment of convertible production. This figure depends on the machines used and cannot be given as a general rule. However, with the operation of the prototypes V0.2.1 and V0.5.x, first empirical values are emerging which indicate a certain range. Measurements on lightweight mobile assembly and transport machines, showed a power requirement of less than 3.7 kVA per feeding point. For all equipment used in the ARENA 2036 test environment, a power input of 11 kVA was sufficient. In [18], the author calculates that the power requirement of an AGV with 500 kg payload is 347 W during steady driving. According to [14], a continuous power output density of 500 $^{W}/_{m^2}$ is sufficient to start the supply of mobile production equipment.

3.4 *Use Case 'Safety Zone'*

Beside the intralogistics use case of the BoxAGV, which is described in more detail in subsection 3.5, another IF application shall be presented briefly. With the help of the safety panels a safety zone can be marked by the LED system and monitored for unauthorized entry of persons or mobile machines. To protect humans from injuries caused by dangerous movements of machines, such movements must be encapsulated according to ISO 13857 or limited in force and torque to non-hazardous values according to ISO/TS 15066. For this reason, protective fences as enclosures, sensors for initiating emergency stops or robots suitable for man-machine operation are used today, among other things. Especially the enclosures are static in their nature and thus in conflict with the goals of the convertible production. The safety zone application described below was implemented on the IF prototype V0.2.1 and shows the potentials as well as the added value for the user and operator of the IF. The application classifies two types of safety zones:

Immobile machinery The safety zone around a production facility is normally realized with a fence and its access is monitored by safety devices triggering an emergency stop. To enable the access of persons and machines into hazardous areas, these entrances, e.g. doors and rolling gates, are secured with safety locks or safety light curtains. As well as for the safety enclosure of the production equipment and for access supervision in the safety zone, engineering time and capital expenditure must be considered. In the end, the solution remains inflexible with regard to the convertible production, because a relocation of the production facility always means a fully relocation of the safety enclosure. Considering the extensive mechanical and electrical installation of the safety equipment as well as the necessary acceptance test of the safety system, production managers usually decide against relocating machines despite more advantageous production logistics.

Mobile machinery Since the risk analysis of machines must also consider their surrounding space, the common installation of safety zones is much more complex than for stationary machines. This is due to the changing of the surrounding space. Based on the application example of AGVs, one of the central hazards of mobile machines for humans is the risk of becoming crushed [19]. Therefore, minimum safety requirements are placed on an AGV, which demand the presence of a braking system, a speed limiter and a device for object detection. By the interaction of AGV and IF, dynamic safety zones can be generated, which minimize the risk of an accident and at the same time reduce the expensive safety technology on board of the AGV (Figure 10).

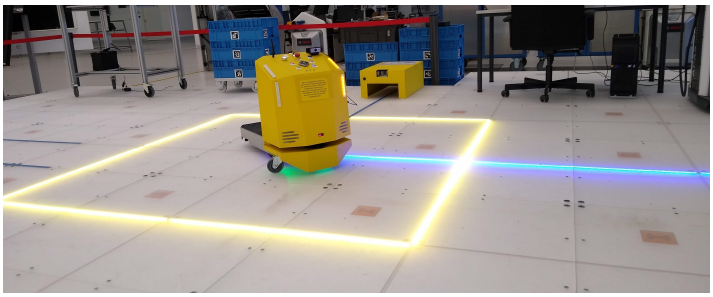


Figure 10: AGV and IF interact so that the yellow safety zone moves with the AGV. The IF detects an zone entry and can trigger the AGV to change direction or to stop.

3.5 Use Case 'BoxAGV'

Another application of the IF comes from the field of intralogistics and shows the use of the BoxAGV, which was adapted to the real production scenario described in [20]. The presented example relates to the production of interior parts, like headrests and side bolsters, for various vehicles. It is suitable, because the transition of a Manually Operated Intralogistics (MOL) to the new material supply concept of the In-plant Milkrun System (IMS) is already being discussed there. Furthermore, the paper includes a production layout and data on which the application example of the Novel BoxAGV System (NBS) can be based on. The comparison of the three material supply concepts does not claim to be complete and / or be generally valid. Further, it does not take into account the influencing factors that occur in practice, such as the robustness of the concepts with regard to the failure of critical components, fluctuations in demands or similar. The comparison rather serves as an indication for a possible operational capability of the NBS.

With the beginning of the 1990s, stationary storage capacities were gradually replaced by trucks moving in public roads. Responsible for this is the Just-In-Time (JIT) production concept [21, 22] which is widely used today. This relocation is not only to be criticized from an ecological point of view, but if the development is viewed from a business point of view, the JIT concept is extraordinarily successful: Warehousing and thus its capital tie-up can be significantly reduced. JIT can be scaled and transferred to the logistics of convertible production. This generates the following advantages:

- Liquidation of supermarkets (cf. central warehouse).
- Reduction of material handling (storage and retrieval).
- Minimum quantity of goods in the production process.
- Shorter throughput times and reduced capital tie-up.

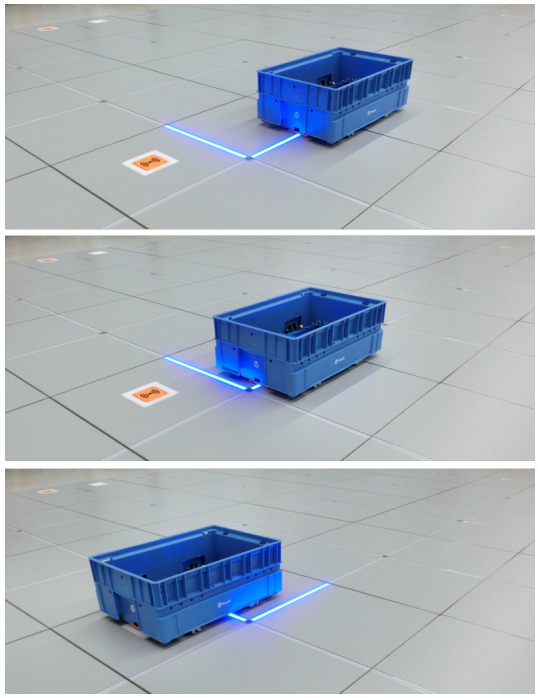


Figure 11: The three pictures show a time sequence in which the BoxAGV follows the dynamically controllable blue line to a WPT unit in order to get recharged.

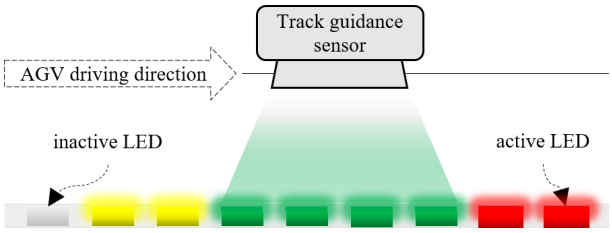


Figure 12: Based on the author's patent [23] an easy to use optical guidance system for mobile robot applications in combination with the IF is proposed. Conventional guiding lines that are painted on the floor, can be replaced by dynamically controllable LED guiding lines.

The IF is suitable in particular for use in internal logistics. Resources (goods, machines and employees) can be organized dynamically with the help of the IF visualization function. This means that it is no longer necessary to depend on rigid production layouts, but can, for example, redefine and reroute goods flows depending on the order situation. If this dynamic route modification is taken as a base for an AGV's track guidance, it is possible to react ad-hoc to unforeseen events in the production. Figure 12 shows that a single LED segment, for example consisting of eight single LEDs, is sufficient to realize track guidance. The segment not only sets the direction of the AGV, but also its speed. With only three colors the speed control can be built up:

- The BoxAGV remains at its current speed, while detecting a green light only.
- While detecting a yellow light, the BoxAGV controller will interpret this as an acceleration command.

- A red light decelerates the BoxAGV from its current speed to zero.

Today's AGVs used in industrial environments are expensive since, as autonomous transport machines, they must have cost-intensive components like battery, on-board computer, localization system and safety devices. As a result, the investment costs for a fully automated logistics are also high. The use of the simply designed BoxAGV, which is centrally navigated and supplied by WPT systems, is a promising approach for the automated logistics of the convertible production. Due to the BoxAGVs slow driving speed of max. 1 km/h and its max. payload of 50 kg, complexity and costs can be reduced. For example, costs can be saved using simple safety edges instead of laser scanners for mobile safety applications. Additional collision protection measures should also be considered. Here, the IF provides predictive safety through its built-in weight and object detection on the total production area. In order to illustrate the potential savings and technical possibilities of the NBS concept in comparison with the IMS solutions and the MOL, a real production scenario from the operational practice will be examined.

3.5.1 *Production Layout*

As shown in Figure 13, the predominantly manual production of car interior parts is divided into a logistics and a production sections with a total area of 1,350 m^2. One third of the available area is reserved exclusively for logistics applications and two third is available for production. A closer look at both sections reveal a further subdivision into manufacturing, storage, traffic, administration and infrastructure areas.

The division of areas is partly determined by the building's structure itself (stairwells, escape routes), organizational requirements (offices, social rooms) and the type of manufacturing (workplaces, storage areas, logistics routes). The organizational and production-specific areas can be partially or completely adapted by changes in the production technology. In the following, the areas specified by the construction are considered to be non-adaptable. Table 2 applies the area allocation to the manufacturing plant and determines the area that is required for manually operated production logistics. It is obvious that only one third of the total area can be used for value-added assembly activities, two third of the area is occupied by traffic routes (marked by red lines) and storage areas. Therefore, [20] also highlights that one of the most urgent problems is the lack of space in production. Furthermore, the turnaround time of incoming and / or outgoing materials is criticized too. Materials sometimes lie on the racks for longer than one shift, which means that the logistics process is not optimally synchronized with the production process.

The production logistics process is considered and the three applicable logistics concepts (MOL, IMS and NBS) are compared with each other in terms of their capital and operational expenditure. The comparison is based on the data published in [20] and the available practical and calculated values for the IF and BoxAGV. It serves to identify the market and technology position of the novel logistics concept NBS compared to existing systems, but is not to be understood as an extensive and generally valid research study. This is part of subsequent studies that foresee an implementation of the NBS concept in a real manufacturing environment.

The production is run in two shifts, with one shift lasting eight hours. Of these, 7.5 hours are net working hours. The working and assembly stations are marked as gray areas in Figure 13. The material buffer in the workstations is divided into input (blue) and output (green) materials and has a capacity for two hours of production time. In the current scenario, logistics is provided in each shift by two qualified workers who travel between the supermarket and the workstations with pallet trucks and handcarts. The data basis [20] is accompanied by a component list of a center console as shown in Table 3. It includes the main components only. Small parts such as screws or washers are not listed. The bill of materials also contains information on the packaging unit (PU) of the components and determines their turnaround time based on the target production quantity of 296 pieces per shift.

Table 2: The layout of the manufacturing area is shown in Figure 13, where traffic routes, production and storage areas are the main uses of the total area.

Type of area	**Share**	
Traffic	463 m^2	34%
Manufacturing	436 m^2	32%
– Area for the workers	*388 m^2*	*28%*
– Workbenches/Machines	*48 m^2*	*4%*
Storage	363 m^2	27%
Administration	74 m^2	6%
Infrastructure	14 m^2	1%
Total	1,350 m^2	100%

Table 3: Bill of materials containing the main components of a center console including the numbers of components per PU, the consumption per shift and the coverage of one PU. The components are provided in a standard VDA box with a capacity of 0.072 m^3, type SLC 8280, 6280 and 6410. `Original data taken from [20].`

Component	**Quantity**	**Consumption**	**Coverage**
① Support arm	192 pcs/PU	1.54 PU/shift	292 min/PU
② Retainer	180 pcs/PU	1.64 PU/shift	274 min/PU
③ Storage	24 pcs/PU	12.33 PU/shift	36 min/PU
④ Handle	1,250 pcs/PU	0.24 PU/shift	1,900 min/PU
⑤ Foam part 1	60 pcs/PU	4.93 PU/shift	91 min/PU
⑥ Foam part 2	16 pcs/PU	18.50 PU/shift	24 min/PU
⑦ Cover lid 1	40 pcs/PU	7.40 PU/shift	61 min/PU
⑧ Cover lid 2	54 pcs/PU	5.48 PU/shift	82 min/PU
⑨ Cup holder 1	21 pcs/PU	14.10 PU/shift	32 min/PU
⑩ Cup holder 2	432 pcs/PU	0.69 PU/shift	657 min/PU

Since [20] has not synchronized the products assembly location with the layout and the paper does not provide any additional bill of materials, the author has made the following definitions to determine the benchmark figures:

- Similarly, assembly products are produced in the manufacturing, where the center console is taken as example to all assembly stations.

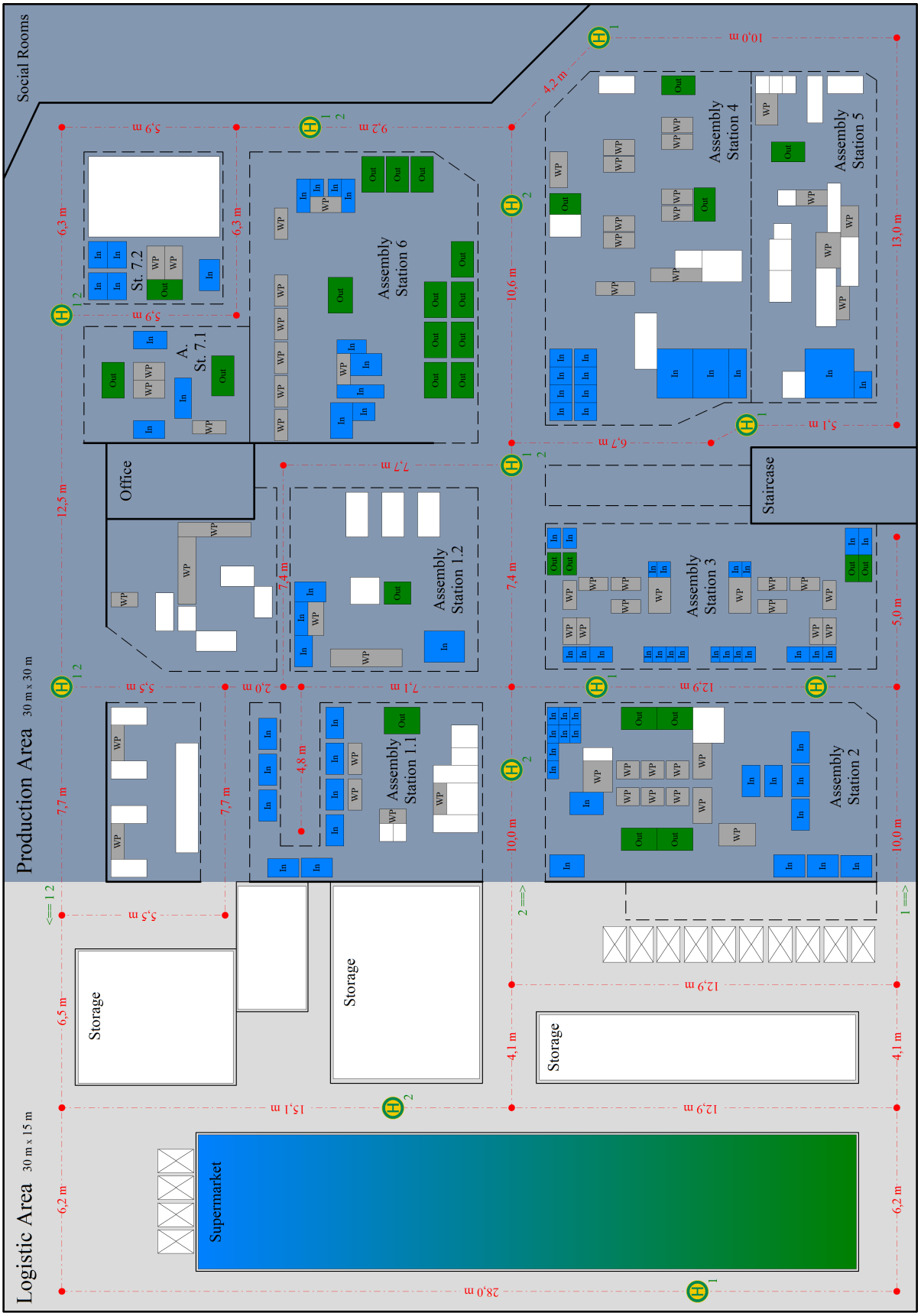


Figure 13: Layout of the logistics and production area for car interior parts, consisting of seven assembly zones, in which the blue rectangles represent goods entry areas, the green ones the goods exit and the grays the workstations. The walking and logistics routes are marked in red and their lengths are listed. The Ⓗ–symbols indicate the routes of two milkruns described in [20].

- Ten consoles will be packed into one PU in order to form one outgoing box.
- Milkrun stops are defined according to Figure 13, where the start and end point of a milkrun is located at the supermarket.

With the help of Table's 3 data and the assumptions made, the time-related box throughput in production can now be determined. The Figure 14 shows the time points and quantities of the incoming and outgoing boxes to be handled by the internal logistics over one station and shift. Based on the assumption that similarly complex products are also produced in the other six assembly stations, the handling process can be transferred to them as well. The process does not consider any disturbing effects in the production or logistics chain. It is therefore easy to see that a lack of synchronization between production and logistics, as well as disturbances acting on the logistic chain, can significantly increase the simultaneous material requirements, even though there are buffer at each station.

Furthermore, the number of box storage spaces and the total number of boxes in production can also be calculated from Table 3 and the assumptions made. For this purpose, an assembly station is divided into input, storage and output spaces. In the production start-up configuration, it is planned that all components are available in one box each and that there is also a buffer for each material to bridge a two-hour supply breakdown. The empties storage spaces at the assembly station are dimensioned accordingly. Table 4 shows that the throughput of boxes per shift is 94. This means that in average every five minutes, one box must be fed to or removed from the assembly station. In this case, a usage of the material buffer is not necessary. The buffer capacity is based on the requirement to guarantee a stand-alone manufacturing over two hours.

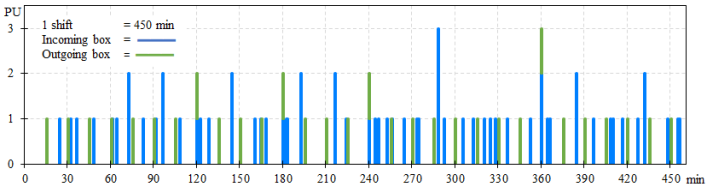


Figure 14: Throughput of input and output boxes at an assembly station over one shift. The incoming boxes with the ten main components of the center console are marked blue and the outgoing boxes with the finished product are marked green. In total 94 boxes are handled over per shift.

To dimension the storage capacity of an assembly station with components correctly, the autonomous production time of two hours is taken as a basis first. Table 3 shows for each component how much production time can be represented by a full box. This time per box is defined as the coverage. Dividing the buffer time of two hours by the coverage gives the number of boxes required. Since this result is usually not a natural number, the value is rounded up to full boxes. As there should always be a full box of each component at the workplace at the beginning of a shift, the result of the storage space calculation must be reduced by this box. The same number of full boxes in the material buffer must be matched by the number of empty box storage spaces. So, the assembly station has in total 21 storage spaces for empty boxes and 21+10 spaces for full boxes. In order to determine the minimum number of boxes in production, it is assumed that the number of boxes handled per shift as well as in the storage areas of the assembly stations must be provided. Applying this to the present production, a minimum number of 875 boxes is obtained. The following comparison of the three logistics concepts MOL, IMS and NBS bases on the figures of this subsection.

Table 4: Throughput and number of boxes of one assembly station, considering the start-up phase of production and the buffer for at least two hours continuous production.

Component	Start-up	Buffer for 2 h production		Throughput
	full boxes	full boxes	storage space	
Incoming boxes				
① Support arm	1 PU	0 PU	0 PU	1 PU/shift
② Retainer	1 PU	0 PU	0 PU	1 PU/shift
③ Storage	1 PU	3 PU	3 PU	12 PU/shift
④ Handle	1 PU	0 PU	0 PU	1 PU/shift
⑤ Foam part 1	1 PU	1 PU	1 PU	4 PU/shift
⑥ Foam part 2	1 PU	4 PU	4 PU	18 PU/shift
⑦ Cover lid 1	1 PU	1 PU	1 PU	7 PU/shift
⑧ Cover lid 2	1 PU	1 PU	1 PU	5 PU/shift
⑨ Cup holder 1	1 PU	3 PU	3 PU	14 PU/shift
⑩ Cup holder 2	1 PU	0 PU	0 PU	1 PU/shift
	10 PU	13 PU	13 PU	64 PU/shift
Outgoing boxes				
Ⓐ Center console	0 PU	8 PU	8 PU	30 PU/shift

3.5.2 *Comparison of the Logistics Concepts*

Manually Operated Logistics The MOL is characterized by using manual transport equipment and a low degree of digital information and processing systems. MOL is limited in its complexity and can easily be applied to low variant and continuously running productions. Since it is operated manually, high personnel expenditure is necessary to ensure the material supply – corresponding to the variety of parts, their throughput and the logistic routes length.

In-plant Milkrun System The IMS is an extension of MOL. Logistics employees are still in charge of driving tugger trains, loading and unloading boxes and checking the material quantities by visual inspection. In comparison to MOL, the operation of IMS requires more use of digital information and processing systems as well as the extension of organizational processes. For example, a production with IMS requires defined stops in the shop floor and a schedule with accurate arrival times. This requires a good synchronization of production and logistics that can be implemented with an appropriate production automation technology. In return, productivity gains and personnel cost savings are possible.

Novel BoxAGV System The NBS bases on the IF combinated with the BoxAGV. It represents a fully automated logistics concept. Each box can be controlled autonomously, without the investment

cost disadvantage of today's AGV solutions. In contrast to the logistics concepts MOL and IMS, no persons are required to perform the transport task. But the extensive use of networked information and processing systems is necessary or have to be implemented initially. In addition to the automation of the transport task and the associated reduction in personnel costs, the NBS offers a significant saving of space in production by transferring the material storage to the traffic routes. Due to the demand-driven material supply, NBS reduces the risk of production breakdowns and is therefore more robust than MOL and IMS.

Table 5: Investment and operating expenditures of the logistics concepts MOL, IMS and NBS in a quick comparison.

Cost category	MOL	IMS	NAS
Building	44,388 EUR	44,388 EUR	39,604 EUR
– Space costs/month/m² [24]		*2.74 EUR*	
– Total production area	*1.350 m²*	*1.350 m²*	*1.205 m² (1)*
Logistics equipment		– are not considered –	
Transport equipment	1,270 EUR	2,392 EUR	83,813 EUR
– 4 x Electric pallet truck [25]	*10,156 EUR*		
– 2 x Platform truck [26]		*5,580 EUR*	
– 6 x Trailer [27]		*3,558 EUR*	
– 351 x IF panel acc. Figure 15			*280,800 EUR*
– 854 x IF cover panel			*170,703 EUR*
– 438 x BoxAGV (2)			*219,000 EUR*
Staff costs	159,360 EUR	79,680 EUR	0 EUR
– Workers per shift	*2*	*1*	*0*
Annual expenditures	205,018 EUR	126,460 EUR	123,417 EUR

(1) *Elimination of the supermarket and its traffic area → 145 m².*
(2) *Each BoxAGV can transport up to 2 PU.*

Since the current state of development of NBS and the data basis of the described production do not allow a detailed cost comparison, the Table 5 is intended to provide a rough cost evaluation of the three logistics concepts shown. It shall be proven that NBS is principally interesting from its capital and operational expenditures compared to MOL and IMS. The table bases on a general operating lifetime of the transport equipment of eight years [28]. This period was also applied for the IF, so that the investment sum in the individual logistics concepts is depreciated on a linear method over eight years. Furthermore, the annual income of a logistics employee is set at 33,200 EUR for the comparison, to which an employer's contribution of 20% to social insurances is added to determine the staff costs [29].

The logistics concepts are compared in three cost categories. For the purpose of completeness, immobile logistics equipment, such as racks, signage and similar equipment, is also listed. But this cost category is not used to evaluate the concepts. Typically, transport equipment is included in the investment costs and staff in the operating costs. The cost of the building may fit into one category or the other, depending on whether the company owns the building or not. If the building is rented, the costs of rent, electricity, water and heating should be allocated to the operating costs. When comparing the annual expenditure, it is remarkable that the lower the degree of automation and rationalization in the logistics concept, the higher the costs. In the MOL concept the personnel costs are dominant and represent the largest part with three quarters of the total costs. The NBS concept is particularly capital-intensive, but it is cheaper over the eight years of operation compared to MOL. NBS and IMS are roughly equal in price. In addition to investment and operating expenditure considerations, the logistics concepts can also be evaluated according to other criteria. For example, the examination of adaptability to changing market and production conditions is an interesting evaluation criterion that should be considered before deciding on a specific logistics concept. Due to the focus of this paper on the NBS introduction, comparing and delimiting studies are described in subsequent papers. Therefore, a preliminary qualitative list of NBS logistic concept advantages is given here:

- The IF can detect obstacles along the traffic routes with its load cells and redirects the flow of materials automatically by changing the displayed route of the BoxAGV.
- Batch tracking can be carried out by using the built-in LED stripes or an e-paper display unit. The visualization makes it easier to find required batches or to implement space-saving chaotic storage on the shop floor.
- Only a limited number of traffic routes in production are suitable for tugger trains. This limits the automation gain of the IMS concept. As shown in the layout Figure 13, tugger trains are used on the main routes, because curves are not large enough or necessary safety distances [30] cannot be guaranteed. Logistics employees are therefore still forced to carry out material transports manually.
- The MOL is basically flexible, but sensitive to disruptions due to the low degree of interconnection between production and logistics. Bottlenecks that arise here, affect the entire manufacturing process. Due to its high degree of interconnection, the NBS is more robust and less error prone.
- Compared to standard AGV solutions, the distances traveled by the BoxAGV are rather short, so that a BoxAGV can generally be built more simply and therefore more economically. Standard AGVs are designed for a daily mileage of 15…20 km [31]. In contrast, the BoxAGV has to run an average travel distance of less than 200 m (warehouse → assembly station → warehouse) in one shift.

4 Conclusion

In the factory of the future, the floor, walls and ceiling will be the only elements that are installed fixed. Production equipment is mobile and can be flexibly combined to form new production lines. Particularly the transport of goods is performed automatically and made more flexible in the sense of the convertible production. To ensure the competitiveness and future viability of a production company according to the customer demands, the convertible production is becoming more and more essential. Storage space inside the factory like supermarkets is reduced, since the material is in a continuous flow on the BoxAGV. Communication takes place in real-time and humans are the decision makers, supported by assistance systems that operate on an artificial intelligent core.



Figure 15: Implementation proposal of the IF into the described production, in which the traffic routes are equipped with standard panels. On this area, the basic functionality of the IF *visualization*, *object detection* and *WPT* for BoxAGV charging is provided. The floor in the area of the assembly stations is built up with regular cover panels.

The IF supports convertibility best and is suitable as an universal and open infrastructure platform for the FOF. Using the IF in such an industrial environment, opens a wide range of applications and makes our current manufacturing processes faster, more flexible and more transparent. A contribution to the convertibility is made by the NBS concept, which prospectively offers cost and space advantages over current logistics solutions. It reduces the material stock at the assembly stations by providing a demand-driven supply. Supermarkets can be downsized by chaotic warehousing or eliminated at all by shifting logistics to the traffic routes. This enables rapid changes in company intralogistics and thus a quick response to market-relevant developments. The close integration of production and logistics also makes manufacturing processes faster and more efficient than today. It is not necessary to search for specific batches (boxes) because each BoxAGV can navigate to a defined point on command or the IF can show the user the way to the BoxAGV via its LED system. The BoxAGV can be used to save on personnel costs or to assign logistics staff to higher-value tasks too.

Overall, the combination of IF and BoxAGV seems to be promising for production companies, especially for companies with a high level of manual and multi-variant assembly. For this reason, a further development of the BoxAGV and the IF is planned. It includes the WPT of transportation and production equipment not only at specific spots, but also along a track and in an open area as described in [15].

References

[1] German Association of the Automotive Industry (VDA), "Small Load Carrier (SLC) System," Technical Report VDA 4500-1.

[2] H. Kühnle, "Paradigmenwechsel im Produktionsbetrieb – Die Fraktale Fabrik," in H.-J. Warnecke, H.-J. Bullinger, editors, Produktionsstrategie für das 21. Jahrhundert, 9–43, Springer Berlin Heidelberg, Berlin, Heidelberg, 1994.

[3] E. Westkämper, "Marktorientiertes Produzieren in dynamischen Strukturen," in H.-J. Warnecke, editor, Fabrikstrukturen im Zeitalter des Wandels – Welcher Weg führt zum Erfolg?, 9–22, Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.

[4] E. Westkämper, "Produktion in Netzwerken," in G. Schuh, H.-P. Wiendahl, editors, Komplexität und Agilität – Steckt die Produktion in der Sackgasse?, 275–291, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997, doi: 10.1007/978-3-642-60841-4_19.

[5] E. Westkämper, Technical report, Institut für Industrielle Fertigung und Fabrikbetrieb (IFF), Stuttgart, Stuttgart, 2006.

[6] W. Günthner, K. Furmans, K.-H. Wehking, K.-P. Rahn, M. Ten Hompel, Intralogistik im Dialog mit Forschung und Lehre, 239–276, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, doi:10.1007/978-3-540-29658-4_6.

[7] M. Hofmann, D. Korte, Neuartiges Logistikkonzept für die automobile Endmontage ohne Band und Takt, 85–100, Springer Berlin Heidelberg, Berlin, Heidelberg, 2020, doi:10.1007/978-3-662-60491-5_8.

[8] M. Mandat, "Produktportfolios der fünf großen deutschen Autohersteller in Deutschland: 1990 bis 2014: Das Dilemma mit der Vielfalt," .

[9] Volkswagen, "Absatz des VW Golf im Zeitraum der Jahre 1974 bis 2012 nach Modell," 2012.

[10] H. Meier, N. Hanenkamp, Komplexitätsmanagement im Lebenszyklus individualisierter Produkte im Maschinen- und Anlagenbau, 623–639, Gabler Verlag, Wiesbaden, 2005, doi:10.1007/978-3-663-01577-2_38.

[11] T. Stehle, U. Heisel, Konfiguration und Rekonfiguration von Produktionssystemen, 333–367, Springer Berlin Heidelberg, Berlin, Heidelberg, 2017, doi: 10.1007/978-3-662-55426-5_39.

[12] J. Stratmann, "Blick unter die VW Golf 7 Karosserie: Über die Nervenbahnen der modernen Automobile," .

[13] Komax, "Zeta 640/650, [Online]. Avail- able: https: // www.komaxgroup.com / en / Products-and-Solutions / Products / Harness-Manufacturing / Zeta-640-650 /" .

[14] J. Stillig, N. Parspour, "Advanced Manufacturing based on the Intelligent Floor," in 20th IEEE Mediterranean Electrotechnical Conference (MELECON), 248–253, 2020, doi:10.1109/MELECON48756.2020.9140660.

[15] J. Stillig, N. Parspour, "Novel Concept for Wireless Power Transfer Modules," in 2020 International Conference on Broadband Communication, Wireless Sensors and Powering (BCWSP), 167–172, 2020, doi:10.1109/BCWSP50066.2020.9249465.

[16] T. Kuhlmann, P. Spanier, M. Ehlich, Potenziale einer industriellen Gleichstromversorgung, 9–35, Carl Hanser Verlag GmbH & Co. KG, München, 1997, doi:10.3139/9783446466128.002.

[17] F. Regnery. [Online]. Available: https: // www.vde.com / resource / blob / 1888372 / 2fab7ab57295bebd7fbd3769314e9efa / faktencheck-bidirektionale-energiefluesse—download-data.pdf

[18] J. Kesselring, Prozessbegleitende Planung und Konfiguration von Fördertechnikanlagen unter Zuhilfenahme von virtuellen Konfigurationsmustern und Konfigurationsmodellen, Ph.D. thesis, Karlsruher Institut für Technologie (KIT), 2017, doi:10.5445/IR/1000075521.

[19] "Flurförderfahrzeuge – Sicherheitstechnische Anforderungen und Verifizierung, Teil 4: Fahrerlose Flurförderfahrzeuge und ihre Systeme," 2020.

[20] M. Knez, B. Gajšek, "Implementation of In-plant Milkrun System for Material Supply in Lean Automotive Parts Manufacturing," in 12th International Conference on Logistics and Sustainable Transport 2015 (ICLST), 121–126, 2015.

[21] M. Ellrich, "Vor- und Nachteile des Produktions- und Logistiksystems," Geographie Infothek, 2.

[22] L. Ickert, U. Matthes, S. Rommerskirchen, E. Weyand, M. Schlesinger, J. Limbers, Technical report, Bundesministerium für Verkehr und digitale Infrastruktur, Basel/Berlin, 2007.

[23] J. Stillig, "Verfahren zum Betrieb eines Spurführungssystems," 2019.

[24] Haufe Online Redaktion, "Logistik: Nebenkosten für Lager und Cross Docks leicht gestiegen," .

[25] STILL GmbH, n.d. [Online]. Available: https: // www.still.shop / ecu-15-c.html

[26] STILL GmbH, n.d. [Online]. Available: https: // www.still.shop / ltx-20.html

[27] Carl Beutlhauser Kommunal- und Fördertechnik GmbH & Co. KG, "Routenzug Katalog," Booklet.

[28] Bundesministerium der Finanzen [Online]. Available: https: // www.bundesfinanzministerium.de / Content / DE / Standardartikel / Themen / Steuern / Weitere Steuerthemen / Betriebspruefung / AfA-Tabellen / Ergaenzende-AfA-Tabellen / AfA-Tabelle AV.html.

[29] StepStone Deutschland GmbH, "Welches Gehalt verdient man im Bereich Logistik-Mitarbeiter/in?" n.d.

[30] DGUV Fachbereich Handel und Logistik, "Routenzüge – Einsatz von Schleppern und Anhängern als Routenzüge," Technical Report FBHL-008, 2018.

[31] Dahl Automation GmbH, n.d. [Online]. Available: https: // www.mobile-robots.de / 70 / robotize

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 369-377 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Data Aggregation, Gathering and Gossiping in Duty-Cycled Multihop Wireless Sensor Networks subject to Physical Interference

Lixin Wang[*,1], Jianhua Yang[1], Sean Gill[1], Xiaohua Xu[2]

[1]*TSYS School of Computer Science, Columbus State University, Columbus, GA 31907, USA*

[2]*College of Computing and Software Engineering, Kennesaw State University, Marietta, GA 30060, USA*

ARTICLE INFO

Article history:
Received: 07 October, 2020
Accepted: 15 January, 2021
Online: 22 January, 2021

Keywords:
Data aggregation
Data gathering
Data gossiping
Duty-cycled scenarios
Physical interference model
Wireless sensor networks

ABSTRACT

Data aggregation, gathering and gossiping are all vital communication tasks in wireless sensor networks (WSNs). When all networking devices are always active, scheduling algorithms for these communication tasks have been extensively investigated under both the protocol and physical interference models. However, wireless devices usually switch between the sleep state and the active state for the purpose of energy saving. A networking device with duty-cycled scenarios having sleep/awake cycles may need to transmit the message to all neighbors more than once. Taking the duty-cycled scenarios into consideration, communication scheduling algorithms for these tasks have been extensively investigated under the protocol interference model. As far as we know, scheduling algorithms for these communication tasks have not yet been investigated in WSNs with duty-cycled scenarios under the physical interference model. In this paper, we propose minimum latency scheduling algorithms for these communication tasks in duty-cycled WSNs under the physical interference model. Our innovative scheduling algorithms for both data gathering and gossiping achieve approximation ratios at most a constant time of $|T|$, where $|T|$ is the length of a scheduling period. The approximation ratio of our proposed data aggregation scheduling algorithm is less than or equal to a constant times $|T|$ with bounded maximum degree of the network.

1 Introduction

This paper is an extension of our work originally presented on the 15th IEEE International Conference on Mobile Ad-hoc and Sensor Networks (IEEE MSN 2019) [1].

Wireless sensor networks (WSNs) have been widely employed in automated battlefields, emergency disaster relief, environmental monitoring, real-time pollution monitoring, military surveillance, etc. Most of these WSN applications are time critical. The definitions of data gossiping, data gathering, and data aggregation can be found in [2]. The concepts of a link schedule, data aggregation schedule, data gathering schedule and data gossiping schedule can also be found in [2].

The latency of a data aggregation schedule is the total number of time-slots needed for the sink node to receive the aggregated message from all other networking nodes. Minimum-Latency Aggregation Schedule (**MLAS**) is the research problem for developing a scheduling algorithm for data aggregation in a multi-hop WSNs with the smallest latency. Similarly, the latency of a data gathering schedule is the total number of time-slots required for the sink node to receive a packet from every other networking node; the latency of a gossiping schedule is the total number of time-slots needed for every device to receive the packets from all other networking devices. Minimum-Latency Gathering Schedule (**MLGS**) is the research problem for developing a scheduling algorithm for data gathering in a multi-hop WSNs with the smallest latency. Minimum-Latency Gossiping Schedule (**MLGoS**) is the research problem for developing a scheduling algorithm for data gossiping in a multi-hop WSNs with the smallest latency. A great number of communication scheduling algorithms have been developed for MLAS, MLGS and MLGoS in multi-hop WSNs under either the physical interference model [2], or the protocol interference model [3]–[13] and the references therein.

However, wireless devices usually switch between the sleep state and the active state for the purpose of energy saving. A networking device with duty-cycled scenarios having sleep/active cycles may need several times to transmit the message to all neighbors. The duty-cycled scenarios is a popular energy-saving method for multi-

[*]Corresponding Author: Lixin Wang, TSYS School of Computer Science, Columbus State University, Columbus, GA 31907, USA. Email: wang_lixin@columbusstate.edu

hop WSNs. The GreenOrbs projects [14] and the VigilNet projects [15] are two actual applications of duty-cycled WSNs (DC-WSNs). Thus, all the existing approximation solutions for MLAS, MLGS, or MLGoS without duty-cycled scenarios are not working any more on DC-WSNs.

We use the same duty cycle model in this paper as the one we used in our short conference version [1]. The problem MLAS on duty-cycled multi-hop WSNs is represented by **MLAS-DC**. Similarly, MLGS and MLGoS on duty-cycled multi-hop WSNs are represented by **MLGS-DC** and **MLGoS-DC**, respectively. With duty-cycled scenarios, all three problems MLAS-DC, MLGS-DC and MLGoS-DC have been extensively investigated subject to the protocol interference and its simple variations [5], [16]–[24]. To the best of our knowledge, all three duty-cycled scheduling problems have not yet been investigated on duty-cycled multi-hop WSNs subject to the physical interference.

In this paper, we develop short communication schedules for the above three duty-cycled scheduling problems on multi-hop WSNs under the physical interference model which is a realistic model that handles wireless interferences more accurately. Under such a realistic model, a receiver can successfully receive the message if and only if the *Signal-to-Interference-plus-Noise-Ratio* (*SINR*) of the receiver is greater than a threshold value. According to the way the SINR ratio is computed, the SINR value is based on those transmissions are that will be scheduled concurrently in every time slot. Therefore, constructing a conflict graph to model the wireless interference is challenging. Such a nature of the physical interference makes it very hard to analyze the proposed communication scheduling algorithms subject to physical interference. Let $\|uv\|$ represent the Euclidean distance between the two devices u and v. Let σ represent the threshold value for a receiving device can successfully receive the wanted signal subject to physical interference, and ξ power strength of the background noise.Suppose that all the devices use a uniform and fixed power P for transmission. The path loss is represented by a positive reference loss parameter η, and a constant path-loss exponent κ (a value between 2 and 6). Assume that the communication radius of each device is normalized to one. If a device u sends a signal using a transmission power P, the power strength of this signal received by the receiving device v is $\eta P \|uv\|^{-\kappa}$.

By far, [25] is the only work that studied **MLAS-DC** for multi-hop WSNs subject to physical interference. When the SINR ratio of a node u is calculated in EQ(1) of [25], the received power $P_r(u)$ at u is incorrectly calculated in EQ(2) of [25] as it will be greater than the transmission power when the distance $\|uv\| < 1$. When $\|uv\| \geq 1$, the received power $P_r(u)$ is simply set to be the transmission power in [25]. Similarly, the total interference power at u is also incorrectly calculated in EQ(3) of [25]. Therefore, radio signal attenuation is not taken into consideration and the algorithms proposed in [25] for **MLAS-DC** are invalid under the physical interference model. As far as we know, there are no existing algorithms proposed for **MLGS-DC** or **MLGoS-DC** for multi-hop WSNs subject to the physical interference constraint.

Let $|T|$ denote the total number of time-slots in a scheduling period. Below are this paper's main contributions:

1. The approximation ratio of the algorithm we proposed for MLGS-DC is less than or equal to a constant times $|T|$;
2. The approximation ratio of the algorithm proposed for MLGoS-DC is less than or equal to a constant times $|T|$;
3. The approximation ratio of the algorithm proposed for MLAS-DC is less than or equal to a constant times $|T|$ when the network has bounded maximum degree.

All the notations used in this paper are listed in Table 1 for easy referencing.

The remaining of this paper is organized as follows. The related work for existing approximation scheduling algorithms for data aggregation, gathering and gossiping are given in Section 2. Section 3 gives some basic concepts and preliminary knowledge that are needed for the design of the scheduling algorithms. In Section 4, we create a tree rooted at the sink that will be used for routing. In Section 5, we present an approximation solution to MLAS-DC. In Section 2, we propose an approximation solution to MLGS-DC. An approximation solution to MLGoS-DC is developed in Section 7. Finally in Section 8, a conclusion for this paper and some future research are presented.

2 Related Works

First, let us review the related work for known algorithms that were proposed for MLAS in multi-hop WSNs when all the networking nodes are implicitly assumed to be always active [2]–[4], [7], [9]. If all the networking devices have uniform transmission radii normalized to one on a multi-hop WSN, the network topology forms a UDG (unit-disk graph). For convenience, the number of networking nodes is denoted by n, the radius of the network graph w.r.t. the sink node of data aggregation or gathering is denoted by R. For arbitrary interference radius, any optimal solution to MLAS uses at least R time-slots. So is $\log n$. [3] and [7] proposed two scheduling algorithms for data aggregation with the total latency less than or equal to $(\Delta-1)R$ and $23R+\Delta-18$, respectively, where Δ is the maximum degree when the transmission radius is the same as the interference radius. The single-level aggregation was investigated in [4] that developed aggregation scheduling algorithm. When the transmission radius is the same as the interference radius, three approximation algorithms were proposed in [9] with total latency less than or equal to $15R+\Delta-4$, $2R+\Delta+O(\log R)$ and $\left(1+O\left(\log R/\sqrt[3]{R}\right)\right)R+\Delta$, respectively. These existing scheduling algorithms for MLAS mentioned above were proposed subject to the protocol interference, and the networking devices are assumed to be always awake.

Subject to the physical interference or the protocol interference, MLGS in multi-hop WSNs has also been extensively investigated when all the networking devices are always awake [2], [10], [11], [13], [26]. Let us review the approximatino solutions developed subject to the protocol interference first. The NP-hardness of MLGS was verified by in [10] that developed a scheduling solution for MLGS that has approximation ratio at most four with the assumption that the communication topology could be reduced to a UDG. In [11], the author proposed a greedy approach for MLGS and proved it to achieve approximation ratio at most 4 in a general setting and achieve approximation ratio at most 3 when

Table 1: All notations used in this paper:

P	the uniform transmission power of every node
η	the parameter of the reference loss
κ	the path-loss exponent
$\|\|uv\|\|$	the Euclidean distance of the two nodes u and v
σ	the threshold ratio of the SINR model
ξ	the background noise in the SINR model
T	a scheduling period
$\|T\|$	the count of the time-slots in a scheduling period T
V	the set of all networking nodes
Δ	the maximum degree of the communication topology of the network
G_r	r-disk graph on V
SPT	a shortest path tree over V
$SPAN$	a spanning tree over V
$\zeta(x)$	the Riemann zeta function
$A(v)$	the active time slot of a node v
s	the sink node of aggregation and gathering
M	the largest latency of the shortest paths from all other nodes to node s in SPT
U_j	the set of the awake nodes in time-slot j
L_i	the subset of the nodes in V from which the latency of the shortest path to node s is i
I_j	the maximum independent set in U_j
I_{ji}	is defined to be $I_j \cap L_i$
$Lat(u, v)$	the Latency of the link (u, v)

the communication topology could be reduced to a UDG. When the transmission radius is the same as the interference radius, a scheduling solution for MLGS was proposed in [26] that has approximation ratio at most $(1 + 1/(k + 1))$ with k being a whole number. In [13], the author developed a scheduling solution for MLGS with the shortest gathering schedule in which several channels were used to speed up communications and make the scheduling solu-tions more efficient. For MLGS, when the interference radii r_i are uniform and no less than the transmission radius, an efficient scheduling solution was proposed in [13] having approximation ratio less than or equal to $2\beta_{r_i}/\lambda + 7$ with λ being the count of the channels that can be used and β_{r_i} the largest number of points in a half-disk of radius $r_i + 1$ with the distance between any two of them is greater than 1.

For MLGoS on multi-hop WSNs, under the UDG model, the authors of [8] designed a gossiping scheme that achieved approximation ratio 27, improved previously known algorithms which have approximation ratio 1000+. In [6], the author considered the problem MLGoS and presented respectively two algorithms with approximation guarantees of 20 and 34, thereby improving the approximation ratio of 27 presented in [8].

When all the devices are always active, [2] developed the shortest communication schedules for MLAS, MLGS and MLGoS subject to the physical interference. [2] developed scheduling algorithms for the four important group communication tasks under the physical interference model. Both the gathering schedule and gossiping schedule developed in [2] achieved constant approximation ratios. For MLGS, [2] proposed an efficient scheduling algorithm with approximation bound less than or equal to $2\beta_\rho$, where ρ and β_ρ are defined in EQ(1) and EQ(3), respectively. For MLGoS, [2] developed a gossiping solution with approximation bound less than or equal to $5\beta_\rho$. Under mild assumptions, the broadcast schedule and data aggregation schedule developed in [2] could also achieve constant approximation ratios.

However, none of the known research papers described above have considered the situation that some networking devices may be in an inactive state. Taking duty-cycled scenarios into consideration under a graph-based interference model, the problem MLAS-DC has been extensively investigated [5], [17], [20]–[23], [27], [28].

Both [5] and [17] developed short data aggregation schedules for MLAS-DC. Both [20] and [21] studied the MLAS-DC problem in order to minimize the communication delay and maximize sensor nodes' lifetimes, and propsed another distributed algorithm for MLAS-DC. All the above mentioned scheduling algorithms for MLAS-DC were proposed under the simple protocol interference model. A data aggregation scheduling algorithm was developed in [22] for MLAS-DC with minimum latency for the purpose of energy efficiency. [23] studied the MLAS-DC problem with no structures taken into consideration. A distributed aggregation algorithm was proposed for MLAS-DC under the protocol interference model, in which the routing tree used for aggregation and a data aggregation schedule without collision were produced at the same time by the algorithm. In [27], the author investigated the MLAS-DC and developed two distributed aggregation algorithms for MLAS-DC, where the routing tree used for data aggregation and an aggregation schedule were produced at the same time. In [27], the author developed an approximation algorithm for MLAS-DC with better approximation bound. When the networking nodes have changable radii, a aggregation schedule with less number of time slots was developed for MLAS-DC in [28] by constructing a nonlinear mathematical formula.

For MLGS-DC subject to protocol interference, [24] proposed

a data gathering scheduling algorithm that achieves an approximation bound of $10|T|$. For the problem MLGoS-DC, [18] proposed a solution called the UTB algorithm for the all-to-all broadcast in which every message must be sent separately without any combination or aggregation and a more efficient version UTB-P of this algorithm using the Prune method. The approximation bound of these two algorithms for MLGoS-DC are less than or equal to $17|T| + 20$. With the unit-size message model, [24] investigated MLGoS-DC and proposed a scheduling algorithm for gossiping that achieved approximation bound less than or equal to $20|T|$. The paper [19] studied the problem MLGoS-DC in uncoordinated duty-cycled multi-hop wireless networks with unbounded-size messages. The NP-hardness of MLGoS-DC was verfified in [19]. This paper also developed two approximation algorithms for MLGoS-DC with approximation bounds less than or equal to $3\Phi^2(\Delta + 6)|T|$, with $\Phi = \lceil 2(\Psi + 2)/3 \rceil$, and Δ the maximum degree of the network. The value of Ψ is equal to the ratio obtained via dividing the interference radius by the transmission radius of each device.

3 Preliminaries

In this section, we introduce some notations, parameters and concepts that are needed for the design of our scheduling algorithms to be proposed in this paper. Set

$$R = \left(\frac{\eta P}{\sigma \xi}\right)^{1/\kappa}.$$

Given a pair of devices u and v. Clearly, there is a link between u and v in the communication topology when there is no interference from other nodes that are transmitting simultaneously if and only if $\|uv\| \leq R$. We call R the *maximum transmission radius*.

Suppose that r is a positive parameter. Denote by R' be the largest edge length of a Euclidean minimum spanning tree over V, the set of all the networking nodes. Clearly, R' equals the minimum value of r satisfying that the *r-disk* graph over V is connected. The relation $R' \leq R$ is needed for the network to be connected. An additional assumption we need is that $R' \leq (1-\varepsilon)R$ which holds for certain small positive constant ε.

Denote by Γ the collection of all Euclidean distances of the networking nodes in V less than R, and greater than or equal to R'. That is, we have

$$\Gamma = \{\|uv\| : u, v \in V \text{ and } R' \leq \|uv\| < R\}.$$

Through a straightforward calculation, we have $1 \leq |\Gamma| \leq n(n-1)/2$ with n being the number of nodes.

Under the physical interference model, a collect I of networking nodes is referred to as *r-independent* when the following two conditions hold

(1) if every node in I sends a message at the same time, each such a transmission could be correctly received by every node whose distance from the transmitter is no more than r;

(2) the distance between any two nodes in I is larger than r.

Now we introduce the well-known Riemann zeta function denoted by $\zeta(x)$:

$$\zeta(x) = \sum_{j=1}^{\infty} \frac{1}{j^x}.$$

For each $r \in [R', R)$, define the following parameter ρ that is needed for the design of our scheduling algorithms to be proposed in later sections and its value is clearly explained in [2].

$$\rho = 1 + \left(\frac{\sigma(16\zeta(\kappa-1) + 8\zeta(\kappa) - 6)}{1 - (r/R)^{\kappa}}\right)^{1/\kappa}. \quad (1)$$

Now we introduce the concept of independence for a given collection of links under the physical interference model. A collection C of links is called *independent* subject to the physical interference if the following two conditions hold:

(1) when the senders of all the communication links in C transmit at the same time, the receivers of every communication link in C can correctly receive the message; (2) all the communication links in C are disjoint.

4 Construction of A Routing Tree

A spanning tree $SPAN$ with the root node $s \in V$ is created in this section. This spanning tree will be employed in routing of the group communication tasks studied in this paper. Let $r > 0$. Denote by G_r a connected and undirected r-disk graph over V.

To begin with, a directed graph G'_r is created based on G_r as follows: replaced every edge (u, v) of G_r by two directed links (u, v) and (v, u) on G'_r, and then remove all links leaving s from G'_r. For a vertex $u \in V$, let $A(u) \in T$ be the awake time slot of u. For a link $(u, v) \in E(G'_r)$, it is assumed that u is the sender and v is the receiver. The latency of (u, v) is defined as follows:

$$Lat(u,v) = \begin{cases} A(v) - A(u), & if\ A(u) < A(v); \\ A(v) - A(u) + |T|, & if\ A(u) \geq A(v). \end{cases} \quad (2)$$

We associate each link (u, v) of G'_r with the weight equal to its latency $Lat(u, v)$, then G'_r is a directed weighted graph with all link weights non-negative. Therefore, a shortest path tree SPT of G'_r rooted at s can be computed by applying the Dijkstra's algorithm on G'_r. The tree SPT consists of all the minimum-weight paths from all nodes other than s to s. The largest latency among all the shortest paths in SPT from all other nodes to sink s is represented by M. Clearly, any optimal solution to either MLAS-DC or MLGS-DC is greater than or equal to M.

Let us take the undirected graph G_r with eight nodes in Fig. 1 as an example to explain the constructions of the directed graph G'_r and the shortest path tree SPT rooted at s. In this example, a scheduling period $T = \{0, 1, 2, 3\}$ with $|T| = 4$. Each node is represented by a circle with its active time slot contained inside it.

Based on our discussion above, the directed graph G'_r is constructed below

We compute a shortest path tree SPT of G'_r with the root s via employing the Dijkstra's shortest path algorithm on G'_r. This SPT is shown in Fig. 3.

This tree SPT has three leaf nodes v_3, v_6 and v_7. Let $\langle v_6, v_5, v_2, s\rangle$ denote the path from v_6 to s through v_5 and v_2.

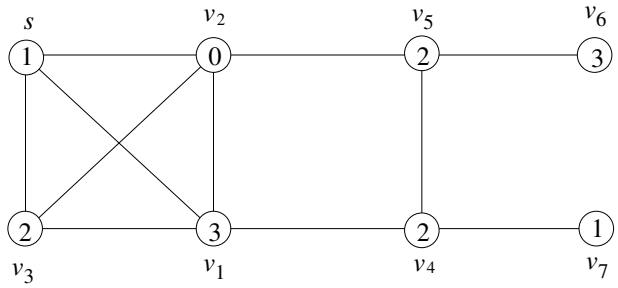


Figure 1: This network has eight vertices. We assume that a scheduling period T has four time slots, that is, $\{0, 1, 2, 3\}$. Each vertex contains a number that is the time-slot at which the vertex is awake.

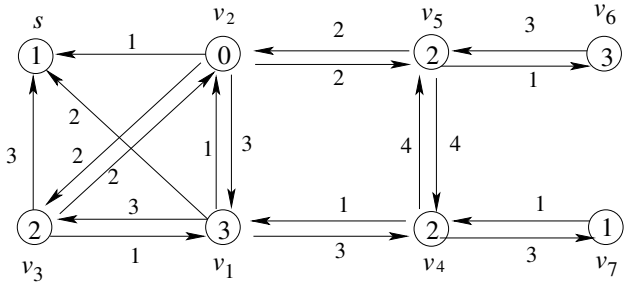


Figure 2: A directed graph with each link associated with the weight equal to its latency defined in EQ(2).

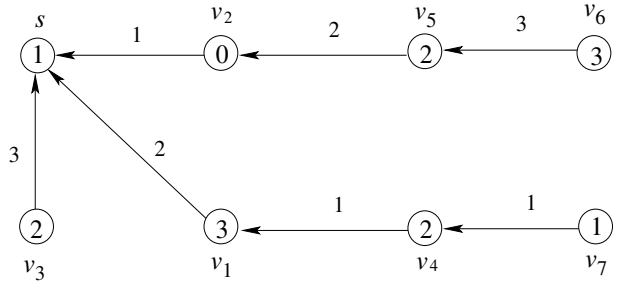


Figure 3: A shortest path tree T_{SPT} rooted at s of the above directed graph in Fig. 2.

Clearly, $Lat(\langle v_6, v_5, v_2, s\rangle) = 6$. Similarly, $Lat(\langle v_7, v_4, v_1, s\rangle) = 4$ and $Lat(v_3, s) = 3$. Therefore, the maximum latency $M = 6$ in this example.

Due to the likely interference with each other caused by the concurrent transmissions, however, the shortest path tree SPT constructed above cannot directly be used for routing of either data aggregation or data gathering. Therefore, a spanning tree $SPAN$ with root s must be created for the use of routing in data aggregation and gathering operations. Such a spanning tree $SPAN$ is constructed as follows:

The nodes in V are put into distinct layers L_0, L_1, L_2, $\cdots$, L_M according to the latency of the shortest path from each $v \in V$ with $v \neq s$ to s. Let's use the same example above to explain how to divide the nodes of V into different layers and construct the sets U_j with $j = 0, 1, 2, 3$. According to the shortest path tree SPT shown in Fig. 3, we have $L_0 = \{s\}$, $L_1 = \{v_2\}$, $L_2 = \{v_1\}$, $L_3 = \{v_3, v_4, v_5\}$, $L_4 = \{v_7\}$, $L_5 = \emptyset$, and $L_6 = \{v_6\}$. Based on EQ(**??**), it is easy to see that $U_0 = L_1 \cup L_5 = L_1$; $U_1 = L_0 \cup L_4$; $U_2 = L_3$; and $U_3 = L_2 \cup L_6$.

Note that likely interferences only happen at those active nodes that can receive messages at the time-slot that is under consideration. For each time slot $0 \leq j \leq |T| - 1$, we construct an MIS $I_j \subset U_j$ layer by layer by following the steps below, based on the disjoint union of the nodes in distinct layers of U_j given in EQ(**??**):

- Initially, $I_j = \emptyset$.
- Sort all the nodes in the layer $L_{A(s)-j+K_0|T|}$ (with smallest index) in the ascending order of their identifications if $L_{A(s)-j+K_0|T|} \neq \emptyset$, say $L_{A(s)-j+K_0|T|} = \langle v_1, v_2, \cdots, v_z\rangle$ with z being the size of $L_{A(s)-j+K_0|T|}$, and add the first node v_1 to I_j.
- If $I_j \cup \{v_2\}$ is independent in G_r, then $I_j \leftarrow I_j \cup \{v_2\}$. If $I_j \cup \{v_2\}$ is dependent in G_r, then the node v_2 is skipped and then the next node in the list is taken into consideration.
- Do the same thing for v_3 until the last node v_z in $L_{A(s)-j+K_0|T|}$ is either added to I_j or skipped.
- Then do the same thing for all the nodes in $L_{A(s)-j+k|T|}$ with $k = 1 + K_0, \cdots, K_1$ until all the nodes in the last layer $L_{A(s)-j+K_1|T|}$ are processed.

Clearly, the MIS I_j constructed above is an independent subset of U_j. For simplicity, let $I_{ji} = I_j \cap L_i$.For each layer $0 < i \leq M$,

every node $v \in L_i \setminus I_{ji}$ was skipped by the above algorithm for the construction of the MIS I_j. Thus, the union $I_{j1} \cup I_{j2} \cup \cdots \cup I_{ji}$ contains one or more neighbors of $v \in L_i \setminus I_{ji}$.

The algorithm for constructing the spanning tree $SPAN$ over V rooted at s was described in the conference version [1] and omitted here due to space limitation. Let's use the same example above to explain how to construct the MIS $I_j \subset U_j$ for each time slot $0 \le j \le |T| - 1$, and how to construct the tree $SPAN$ rooted at s. For $U_0 = L_1 = \{v_2\}$, $I_0 = \{v_2\}$. For $U_1 = L_0 \cup L_4 = \{s, v_7\}$, $I_1 = \{s, v_7\}$. For $U_2 = L_3 = \{v_3, v_4, v_5\}$, $I_2 = \{v_3, v_4\}$. For $U_3 = L_2 \cup L_6 = \{v_1, v_6\}$, $I_3 = \{v_1, v_6\}$.

Now the tree $SPAN$ with root s for this example is created. For $L_1 = \{v_2\}$, v_2's parent in $SPAN$ is s (the same parent as in SPT) as $v_2 \in I_0 \cap L_1$. For $L_2 = \{v_1\}$, v_1's parent in $SPAN$ is also s as $v_1 \in I_3 \cap L_2$. For $L_3 = \{v_3, v_4, v_5\}$, v_3's parent in $SPAN$ is s as $v_3 \in I_2 \cap L_3$; v_4's parent in $SPAN$ is v_1 as $v_4 \in I_2 \cap L_3$; v_5's parent in $SPAN$ is v_4 (a different parent from SPT) as $v_5 \notin I_2 \cap L_3$. For $L_4 = \{v_7\}$, v_7's parent in $SPAN$ is v_4 as $v_7 \in I_1 \cap L_4$. For $L_6 = \{v_6\}$, v_6's parent in $SPAN$ is v_5 as $v_6 \in I_3 \cap L_6$. The constructed tree $SPAN$ rooted at s is shown in Fig. 4 below:

Finally in this section, the concept of the first-fit distance-d coloring for a subset of V in the lexicographic order is presented and adopted from our prior work [2]. Denote by U a subset of V. We consider the lexicographic order of the nodes in U, then we sort all nodes in U from the left to right, $u_1, u_2, \cdots, u_k$ with $k = |U|$. *The first-fit distance-d coloring of* U in the lexicographic order is described below:

- positive integers 1, 2, 3, $\cdots$ represent colors;
- the first node u_1 in the sequence is assigned color 1;
- For $i = 2$ up to k, the node u_i is assigned the smallest possible number that is not utilized by any prior node u_j in the sequence with $j < i$ and $\|u_i u_j\| \le d$.

[2] proved that for a given independent set I of G_r, the first-fit distance-ρr coloring of I with the lexicographic ordering needs at most β_ρ colors. The value of β_ρ is given below (see Corollary 4 of [2]):

$$\beta_\rho = \left\lfloor \frac{\pi\rho^2}{\sqrt{3}} + \left(\frac{\pi}{2} + 1\right)\rho \right\rfloor + 1. \tag{3}$$

5 Aggregation Schedule under the Physical Interference Model Taking Duty Cycles into Consideration

In this section, we consider MLAS-DC subject to the physical interference and develop a short aggregation schedule for this problem. The sink node is represented by s. For a real number r satisfying $R' \le r < R$, we use EQ(1) to calculate ρ. Then follow the steps stated in Section 4, the spanning tree $SPAN$ rooted at s is constructed. The transmission follows a bottom-up approach.

Next we present the first-fit scheduling algorithm for MLAS-DC subject to physical interference. The first layer to transmit is the bottom layer L_M, and then the nodes in layer L_{M-1} transmit. The scheduling is based on a bottom-up approach. For each layer L_i $(0 < i \le M)$, we know that the nodes in L_i are active at time-slot j $(0 \le j \le |T| - 1)$. At this layer, the nodes in $L_i \setminus I_j$ will be scheduled to transmit first, followed by the tranmissions of the nodes in $L_i \cap I_j$. Then the nodes in layer L_{M-1} will be scheduled to transmit. For each layer among L_{M-1}, $L_{M-2}, \cdots$, L_0, the nodes in these layers will receive the aggregated data from their children nodes in the decreasing order of the indices of the layers $M-1, M-2, \cdots$.

For the details of our first-fit data aggregation scheduling algorithm proposed for the problem MLAS-DC subject to the physical interference, please refer to the conference version [1]. Assume the graph in Figure 1 represents an r-disk graph G_r for some $r \in \Gamma$. We use this example in Figure 1 to illustrate the data aggregation scheduling algorithm we proposed for MLAS-DC.

The tranmission schedules for aggregation outputted by the above algorithm are listed in Table 2. Totally, 6 scheduling periods are required for s to receive eventually the aggregated message in this example.

The algorithm for this example works as follows: At $L_6 = \{v_6\}$, $v_6 \in I_3 \cap L_6$ and its parent in $SPAN$ is v_5 which is active at time-slot 2. Thus, link (v_6, v_5) is scheduled at time-slot 2 of SP 1. We skip $L_5 = \emptyset$. For $L_4 = \{v_7\}$, $v_7 \in I_1 \cap L_4$ and its parent in $SPAN$ is v_4 which is also active at time-slot 2. Thus, link (v_7, v_4) is scheduled at time-slot 2 of SP 2. For $L_3 = \{v_3, v_4, v_5\}$, node $v_5 \in L_3 \setminus I_2$ and its parent in $SPAN$ is v_4 which is active at time-slot 2. Thus, link (v_5, v_4) is scheduled at time-slot 2 of SP 3. The nodes $v_3, v_4 \in I_2 \cap L_3$; v_3's parent in $SPAN$ is s which is active at time-slot 1. Thus, link (v_3, s) is scheduled at time-slot 1 of SP 4. Node v_4's parent in $SPAN$ is v_1 which is active at time-slot 3. Thus, link (v_4, v_1) is scheduled at time-slot 3 of SP 4. For $L_2 = \{v_1\}$, $v_1 \in I_3 \cap L_2$ and its parent in $SPAN$ is s which is active at time-slot 1. Thus, link (v_1, s) is scheduled at time-slot 1 of SP 5. Finally for $L_1 = \{v_2\}$, $v_2 \in I_0 \cap L_1$ and its parent in $SPAN$ is s which is active at time-slot 1. Thus, link (v_2, s) is scheduled at time-slot 1 of SP 6.

We summarize our above discussion in the following theorem. The proof of this theorem can be found in the short conference version [1].

Theorem 1 *The data aggregation scheduling algorithm for MLAS-DC is technically correct and produces a data aggregation schedule without any collision. The latency of the transmission schedule for aggregation outputted via our proposed algorithm is less than or equal to* $2\beta_\rho |T| (\Delta - 1)M$. *The proposed algorithm for MLAS-DC has approximation ratio upper bounded by* $2\beta_\rho |T| (\Delta - 1)$.

Note that both ρ and β_ρ are constants defined in EQ(1) and EQ(3), respectively. Therefore, the approximation bound for our scheduling algorithm proposed for MLAS-DC is less than or equal to a constant times $|T|$ when the maximum degree Δ is bounded.

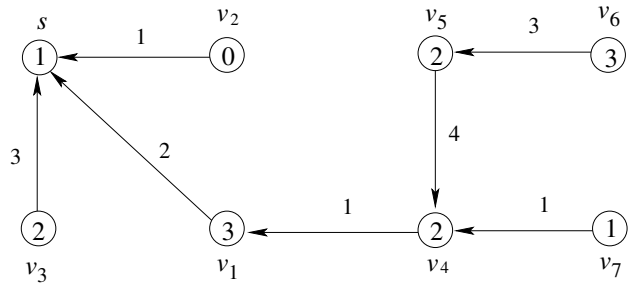


Figure 4: The construction of a spanning tree T_{span} rooted at s.

Table 2: The transmission schedule for the aggregation operation outputted by the algorithm proposed for MLAS-DC. In each row, "SP i" means the i-th scheduling period.

	timeslot0	timeslot1	timeslot2	timeslot3
$SP1$			(v_6, v_5)	
$SP2$			(v_7, v_4)	
$SP3$			(v_5, v_4)	
$SP4$		(v_3, s)		(v_4, v_1)
$SP5$		(v_1, s)		
$SP6$		(v_2, s)		

6 Gathering Schedule under the Physical Interference Model Taking Duty Cycles into Consideration

In this section, we develop a scheduling algorithm for MLGS-DC subject to physical interference. The sink node is represented by s. For a real number r satisfying $R' \leq r < R$ (see Section 3 for the definitions of R' and R), we use EQ(1) to calculate ρ. Then follow the steps stated in Section 4, the spanning tree $SPAN$ rooted at s is constructed.

For any time-slot $0 \leq j \leq |T| - 1$, each node in U_j is active and ready to receive packets from one of its children nodes in $SPAN$. Let

$$Children(U_j) = \{u : u \text{ is a child in } SPAN \text{ of a node in } U_j\},$$

which is the collection of children devices of the devices in U_j. The devices in $Children(U_j)$ will send their packets at time-slot j.

For each $0 \leq j \leq |T| - 1$, we adopt the data gathering heuristic obtained in our previous work [2] for WSNs with all nodes always active for the time-slot j. Our propsed data gathering scheduling algorithm will have $2n - 3$ rounds in total. For each round k, represent by E_k the collection of tree links of $SPAN$ whose label is k. Suppose that A_k represents the collection of the links in the inward s-arborescence made by the tree edges in E_k. For each round k, a dominator connects no more than one link in A_k (see [2]). For each round k and time-slot j, represent by $A_k^{(j)}$ the sub-collection of links of A_k whose sender belongs to the set $Children(U_j)$.

Next the data gathering scheduling algorithm for MLGS-DC subject to physical interference is presented. In each time slot $0 \leq j \leq |T| - 1$, the algorithm is divided into $2n - 3$ rounds sequentially, each of which is dedicated to a subset of links in the list:

$$A_{2n-3}^{(j)}, A_{2n-4}^{(j)}, \cdots, A_2^{(j)}, A_1^{(j)},$$

respectively.

For each round k and time-slot j, the links in $A_k^{(j)}$ are scheduled for transmission: Let S denote the set of the dominator ends of the links in $A_k^{(j)}$. A first-fit distance-ρr coloring of the set S is first calculated. At the i-th time-slot of the k-th round, according to this coloring, the link with a dominator end that is assigned the i-th color will send a packet. By Corollary 4 of [2], the maximum number of time-slots each of these rounds uses is β_ρ. The total number of rounds is $2n - 3$. Therefore, for each time slot j, the maximum time-slots used by the solution outputted by our proposed gathering scheduling algorithm is $\beta_\rho(2n - 3)$. Each scheduling period has $|T|$ time-slots in total. Thus, the entire gathering operation uses no more than $\beta_\rho(2n - 3)|T|$ time-slots.

We summarize our above discussion in the following theorem for MLGS-DC under the physical interference model.

Theorem 2 *The gathering scheduling algorithm described above for MLGS-DC can correctly produce a interference-free schedule of data gathering. The latency of the produced data gathering schedule is less than or equal to* $\beta_\rho(2n - 3)|T|$.

Clearly, any optimal solution to the problem MLGS-DC uses at least $n - 1$ time-slots, the approximation bound for the gathering scheduling algorithm presented in this section is less than or equal to $2\beta_\rho|T|$, which is a multiple of $|T|$ as β_ρ is a constant.

At the end, a data gathering schedule is computed by applying the above scheduling algorithm for each $r \in \Gamma$. The one with the smallest latency among all these resulting data gathering schedules will be chosen. Note that the number of element in Γ is a finite and we have its upper bound $|\Gamma| \leq n(n-1)/2$.

7 Gossiping Schedule under the Physical Interference Model Taking Duty Cycles into Consideration

In this section, we develop a scheduling algorithm that produces a short gossiping schedule for MLGoS-DC under the SINR model. Let $r = R'$ (see Section 3 for the definition of R'), we use EQ(1) to calculate ρ. The r-disk graph on V is denoted by G_r. Then we compute the graph center s for the graph G_r.

Our proposed gossiping scheduling algorithm for MLGoS-DC consists of two phases:

Phase 1: s collects all the packets from all other nodes.

Phase 2: s broadcasts all the n packets (including its own) to all other nodes.

For Phase 1, the data gathering scheduling algorithm for MLGS-DC presented in Section 6 will be adopted. After that, the sink node s sends all the n packets to all other $n-1$ nodes in Phase 2. In this phase, for each time slot $0 \leq j \leq |T|-1$, we adopt the algorithm that was used in the second phase of the gossiping algorithm from our prior work [2] we developed for MLGoS without duty-cycled scenarios. This algorithm works as follows: To begin with, a spanning tree $SPAN$ over V on the graph G_r is created according to the steps stated in Section 4. The routing used for the broadcast of Phase 2 is the outward spanning s-aborescence oriented from $SPAN$. Then we perform the following three steps:

- For all the dominating nodes, we calculate the first-fit distance-ρr coloring. Assume that this distance-ρr coloring uses Z colors. By Corollary 4 of [2], we have $Z \leq \beta_\rho$.
- For simplicity, suppose that the color 1 is assigned by the graph center s. We divide the time-slots into $2Z$ frames. In each frame of the time-slots, the first Z time-slots make a subframe of dominating nodes and the second Z time-slots make a subframe of connecting nodes. The center s sends one packet in each subframe.
- For each $1 \leq i \leq Z$, the connecting node with the color i receiving a packet in a subframe of dominating nodes sends the received packet in the i-th time-slot.
- For each $1 \leq i \leq Z$, the dominating node with the color i receiving a packet in a subframe of connecting nodes sends the received packet in the i-th time-slot.

According to Theorem 9 of [2] and the discussions thereafter, for each time slot $0 \leq j \leq |T|-1$, the latency of both phase 1 and phase 2 is at most $\beta_\rho(5n-5)$. Thus, our proposed gossiping scheduling algorithm generates a solution whose latency is less than or equal to $\beta_\rho(5n-5)|T|$.

The theorem below summarizes our discussions and analyses in this section

Theorem 3 *The gossiping scheduling algorithm presented in this section works and it generates a gossiping schedule without collision. This gossiping scheduling algorithm produces a solution whose latency is less than or equal to* $\beta_\rho(5n-5)|T|$.

Clearly, any optimal solution to the problem MLGoS-DC uses at least n time-slots, the approximation ratio of the gossiping scheduling algorithm presented in this section for MLGoS-DC is less than or equal to $5\beta_\rho|T|$, which is a multiple of $|T|$ as β_ρ is a constant.

8 Conclusion

In this paper, approximation algorithms for MLAS-DC, MLGS-DC and MLGoS-DC on duty-cycled multi-hop WSNs under the SINR model were proposed in order to minimized the communication latency. To the best of our knowledge, the paper is the first work that proposed efficient communication scheduling algorithms for these three problems subject to the physical interference. Denote by $|T|$ the count of time-slots contained in a scheduling period. Our proposed approximation algorithms for MLGS-DC has approx. ratios upper bounded by a constant time of $|T|$. So is the proposed approximation algorithms for MLGoS-DC. Our proposed approximation algorithms for MLAS-DC has an approximation bound no more than a constant times $|T|$ when the network maximum degree Δ is bounded. It has been well documented in the literature that analysis for scheduling algorithms developed for MLAS, MLGS and MLGoS under the SINR model is very challenging due to the impossibility of constructing related conflict graphs.

For future work in the design of scheduling algorithms for MLAS-DC, MLGS-DC and MLGoS-DC on duty-cycled multi-hop WSNs under the SINR model, the problems of proposing constant-approximation algorithms are still open, under either the SINR model or the protocol inference model. Any constant-approximation algorithms for MLAS-DC, MLGS-DC and MLGoS-DC will be significant.

9 Declarations

9.1 *Availability of data and material*

The data generated during the investigation of this research are included in this paper.

9.2 *Conflicts of interest*

The authors declare that there is no conflict of interest about the publication of this article.

9.3 *Funding*

This work is supported by National Security Agency NCAE-C grant H98230-20-1-0293 with Columbus State University, Georgia, USA.

References

[1] L. Wang, H. Liang, and P.-J. Wan, "Data Aggregation Scheduling in Duty-Cycled Multi-hop Wireless Networks Subject to Physical Interference, " the 15th IEEE International Conference on Mobile Ad-hoc and Sensor Networks (IEEE MSN), Dec. 11-13, 2019, DOI: 10.1109/MSN48538.2019.00041

[2] P.-J. Wan, L. Wang, O. Frieder, "Fast Group Communications in Multihop Wireless Networks Subject to Physical Interference," the 6th IEEE International Conference on Mobile Ad hoc and Sensor Systems (IEEE MASS), 526-533, 2009 DOI: 10.1109/MOBHOC.2009.5336958

[3] X.J. Chen, X.D. Hu, J.M. Zhu, "Minimum data aggregation time problem in wireless sensor networks," Lecture Notes in Computer Science 3794, 133-142, 2005.

[4] Y. P. Chen, A. L. Liestman, J. Liu, "A hierarchical energy-efficient framework for data aggregation in wireless sensor networks," IEEE Transactions on Vehicular Technology, **55**(3), 789-796, 2006, DOI: 10.1109/TVT.2006.873841

[5] Z. Chen, G. Yang, L. Chen, J. Xu, H. Wang, "A load-balanced data aggregation scheduling for duty-cycled wireless sensor networks," In Cloud Computing Technology and Science (CloudCom), 4th IEEE International Conference on, 888-893, December, 2012, DOI: 10.1109/CloudCom.2012.6427533

[6] R. Gandhi, Y.-A. Kim, S. Lee, J. Ryu, P.-J. Wan, "Approximation Algorithms for Data Broadcast in Wireless Networks," IEEE INFOCOM Mini-conference, 2009, DOI: 10.1109/TMC.2011.162

[7] S.C.-H. Huang, P.-J. Wan, C. T. Vu, Y. Li, F. Yao: "Nearly Constant Approximation for Data aggregation Scheduling in Wireless Sensor Networks," IEEE INFOCOM, 2007, DOI: 10.1109/INFCOM.2007.50

[8] S.C.-H. Huang, H. Du, E.-K. Park, "Minimum-latency gossiping in multi-hop wireless networks," ACM Mobihoc, 2008, DOI: 10.1145/1374618.1374662

[9] P.-J. Wan, C.-H. Huang, L. Wang, Z.-Y. Wan, X. Jia: "Minimum Latency aggregation Scheduling in Multihop Wireless Networks," ACM MOBIHOC, 2009, DOI: 10.1145/1530748.1530773

[10] J.-C. Bermond, J. Galtier, R. Klasing, N. Morales, S. Perennes, "Hardness and approx. of gathering in static radio networks," Proceedings FAWN06, 2006, DOI: DOI: 10.1109/PERCOMW.2006.62

[11] V. Bonifaci, P. Korteweg, A. Marchetti-Spaccamela, L. Stougie, "An approx. Algorithm for the Wireless Gathering Problem," Proceedings of SWAT, 328-338, 2006, DOI: 10.1007/11785293_31

[12] R. Gandhi, S. Parthasarathy, A. Mishra, "Minimizing broadcast latency and redundancy in ad hoc networks," in Proceedings of the 4th ACM international symposium on Mobile Ad hoc networking and computing (MobiHoc), 222-232, 2003, DOI: 10.1145/778415.778442

[13] P.-J. Wan, Z. Wang, Z. Wan, S. C.-H. Huang, H. Liu, "Minimum-Latency Schedulings for Group Communications in Multi-channel Multihop Wireless Networks," WASA, 2009, DOI: 10.1007/978-3-642-03417-6_46

[14] The GreenOrb Wireless Sensor Network Projects available at http://www.greenorbs.org/all/projects.htm, 2011

[15] P. Vicaire, T. He, Q. Cao, T. Yan, G. Zhou, L. Gu, T. F. Abdelzaher, "Achieving long-term surveillance in vigilnet," ACM Transactions on Sensor Networks (TOSN), **5**(1), 9, 2009, DOI: 10.1145/1464420.1464429

[16] K. Han, J. Luo, Y. Liu, A. Vasilakos, "Algorithm Design for Data Communications in Duty-Cycled Wireless Sensor Networks: A Survey," Communications Magazine, IEEE, **51**(7), 2013, DOI: 10.1109/MCOM.2013.6553686

[17] X. Jiao, W. Lou, X. Feng, X. Wang, L. Yang, G. Chen, "Delay efficient data aggregation scheduling in multi-channel duty-cycled WSNs," In 15th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 326-334. IEEE, 2018, DOI: 10.1109/MASS.2018.00055

[18] X. Jiao, W. Lou, J. Ma, J. Cao, X. Wang, X. Zhou, "Minimum latency broadcast scheduling in duty-cycled multihop wireless networks," IEEE Transactions on Parallel and Distributed Systems, **23**(1), 110-117, 2011, DOI: 10.1109/TPDS.2011.106

[19] X. Jiao, W. Lou, X. Wang, J. Ma, J. Cao, X. Zhou, "On interference-aware gossiping in uncoordinated duty-cycled multi-hop wireless networks," Ad Hoc Networks, **11**(4), 1319-1330, 2013, DOI: 10.1007/978-3-642-14654-1_24

[20] B. Kang, P. K. H. Nguyen, V. Zalyubovskiy, H. Choo, "A distributed delay-efficient data aggregation scheduling for duty-cycled WSNs," IEEE Sensors Journal, **17**(11), 3422-3437, 2017, DOI: 10.1109/JSEN.2017.2692246

[21] Z. Li, Y. Peng, D. Qiao, W. Zhang, "LBA: Lifetime balanced data aggregation in low duty cycle sensor networks," In 2012 Proceedings IEEE INFOCOM, 1844-1852, IEEE, 2012, DOI: 10.1109/INFCOM.2012.6195559

[22] F. Y.-S. Lin, Y.-F. Wen, "Multi-sink data aggregation routing and schedul-ing with dynamic radii in WSNs," Communications Letters, IEEE, **10**(10),
692-694, 2006, DOI: 10.1109/LCOMM.2006.06058

[23] Q. Chen, H. Gao, S. Cheng, J. Li, Z. Cai, "Distributed non-structure based data aggregation for duty-cycle wireless sensor networks," In IEEE INFOCOM 2017-IEEE Conference on Computer Communications, 1-9, IEEE, 2017, DOI: 10.1109/INFOCOM.2017.8056960

[24] X. Xu, J. Cao, and P.-J. Wan, "Fast Group Communication Scheduling in Duty-Cycled Multihop Wireless Sensor Networks," WASA, 2012, DOI: 10.1007/978-3-642-31869-6 17

[25] J. Tang, X. Jiao, W. Xiao, "Minimum-latency data aggregation in duty-cycled wireless sensor networks under physical interference model," In the 22nd Wireless and Optical Communication Conference, 309-314, *IEEE,* 2013, DOI: 10.1109/WOCC.2013.6676328

[26] X. Zhu, B. Tang, and H. Gupta, "Delay efficient data gathering in sensor networks," International Conference on Mobile Ad-Hoc and Sensor Networks, Springer, Berlin, Heidelberg, 380-389, 2005, DOI: 10.1007/11599463 38

[27] Q. Chen, H. Gao, Z. Cai, L. Cheng, J. Li, "Distributed low-latency data aggregation for duty-cycle wireless sensor networks," IEEE/ACM Transactions on Networking, **26**(5), 2347-2360, 2018, DOI: 10.1109/TNET.2018.2868943

[28] Y.-F. Wen, F. Y.-S. Lin, "Cross-layer duty cycle scheduling with data aggrega-tion routing in wireless sensor networks," Embedded and Ubiquitous Comput-ing. Springer Berlin Heidelberg, 894-903, 2006, DOI: 10.1007/11802167 90

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 378-385 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Decentralized Management System for Solid-State Voltage Regulators in Nodes of Distribution Power Networks

Igor Polozov[1], Elena Sosnina[2], Vladimir Kombarov[1], Ivan Lipuzhin[2,*]

[1]*Department of Computer technologies in design and production, Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod, 603950, Russia*

[2]*Department of Electric power supply, power engineering and power supply, Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod, 603950, Russia*

ARTICLE INFO

Article history:
Received: 07 October, 2020
Accepted: 06 January, 2021
Online: 22 January, 2021

Keywords:
Voltage Vector Regulation
Solid-State Voltage Regulator
Consensus Task

ABSTRACT

The article describes the concept and architecture of a decentralized control system for a solid-state voltage regulator (SSVR). The SSVR is a universal device for controlling the mode and operation parameters of medium voltage electrical networks. SSVR manage the amount of current in line using the vector voltage control method. The SSVR control system consists of two levels: the SSVR semiconductor converter control level (the technological control system), SSVR cluster management level (intelligent control system). The objective of the study is to develop an algorithm for managing the SSVR cluster located in the nodes of the distribution electric network. The Raft consensus algorithm for managing the computing cluster is applied to ensure reliable decentralized network management. The algorithm is iterative. Ethernet and PLC architectures are proposed for constructing a data transmission network between SSVR nodes. A simulation model of the SSVR cluster and its control system is developed to study the operation of the control system (node shutdown, loss of communication). The criteria for the normal operation of the intelligent control system are formulated and an algorithm for its operation in emergency situations is presented. The studies of the SSVR control system confirmed the operability of the developed control system in normal and emergency operation of the cluster on a simulation. The dependence of the control system response time on the number of cluster nodes is investigated. The maximum number of nodes in the SSVR cluster depending on the tasks being solved by the SSVR will be limited by the speed of the control system. If the number of cluster nodes increases, it is necessary to enlarge the minimum time interval between requests for control commands.

1. Introduction

The concept of the Internet of Energy (IoE) is the logical development of smart grid technology [1, 2]. IoE is a new architecture for building a power electricity system with decentralized management. A key role in this system is assigned to intellectualization and economic interaction between all network users (energy transactions).

Distribution networks of the future will have a complex configuration and a large number of types of distributed generation sources. Thus, power flows should be bi-directional unlike traditional networks. To implement this principle, electrical networks of all voltage classes, especially low and medium voltage networks, must become "controlled". From a technical point of view, this is hindered by the lack of Flexible AC Transmission System (FACTS) devices in medium voltage networks.

The design of solid-state voltage regulator (SSVR) is known [3]. The SSVR is a power flow control device built on the basis of two transistor converters (Figure 1). The SSVR control method is based on the voltage vector regulation principle – the SSVR manage the voltage drop vector across network rector L [4], and the converter itself acts as a controlled current source. The SSVR is installed in the nodes of the medium voltage distribution network and allows to control the network operation modes.

* Corresponding Author: Ivan Lipuzhin, 24 Minin St., Nizhny Novgorod, Russia, lipuzhin@nntu.ru

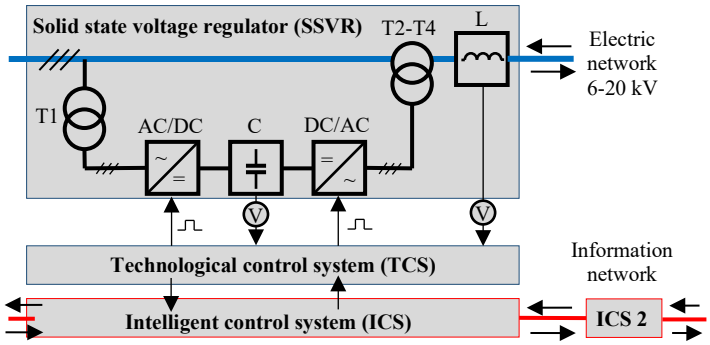


Figure 1: SSVR power unit and control system: T1 - shunt transformer; AC/DC - rectifier; C - capacitive storage; DC/AC - inverter; T2-T4 - serial transformers; L - network choke; V – voltmeter; ICS – intelligent control system of a neighbor node with SSVR

It is necessary to install several SSVRs in the distribution network to perform effectively all the specified functions. The required number of devices depends on the network configuration and the tasks being solved.

The SSVR has a two-level control system (Figure 1). The lower level system is a technological control system (TSC) [5]. The upper level control system is an intelligent control system (ICS). The TCS task is to form the required voltage vector and control the power elements of the converter. The ICS task is to manage the SSVR cluster, communication of SSVRs within a cluster and the command generation for the TCS.

The novelty of this work is that the RAFT consensus algorithm is applied in ICS to provide smooth voltage regulation and control of the electrical network with SSVR.

The article presents the results of the development and research of the SSVR decentralized control system of the upper level. The section 2 provides an overview of decentralized control systems used in the electricity industry. The section 3 describes the architecture of the SSVR ICS. The algorithms for the functioning of the ICS (the algorithms for managing a cluster and a network) are given in 4 and 5 sections. The section 6 describes the simulation model of SSVR ICS and the section 7 presents the simulation results of emergency operation modes.

2. Literature Review

Traditional power grid management systems require a central control room. Each control element must communicate with the control center and often not only the main, but also backup communication channels are required. Situations when communication with the control center is impossible may occur during the operation of networks and the need for autonomous operation arises. Decentralized management systems are the solution to this problem.

There are some tools for decentralized control of emergency switching based on reclosers for 6-10 kV distribution networks [6]. Various algorithms [7] of reclosers operation are possible to control emergency modes. The proposed methods of decentralized management are used only for emergency shutdown operations in a medium voltage network. This functionality is not sufficient for smart distribution networks. A smooth voltage regulation is necessary, and switching should be performed in both emergency and normal modes in order to optimize the power flows in electricity network of up to 20 kV.

The problem of creating a decentralized intelligent management system for networks with distributed generation is most fully considered in [8], where a method of multi-agent service restoration and the digital model of a power system with multi-agent control are proposed. The principles of building a multi-agent system, as well as the stages of its operation are determined.

The most common control system architecture is SCADA (Supervisory Control and Data Acquisition). As shown in [9], to solve similar problems the traditional for SCADA centralized control model requires either improving the existing communication infrastructure to provide higher data transfer rate or storing measured data at a local substation and then organizing data exchange between substations to ensure the operation of complex applications in real time. At the same time, the amount of saved data turns out to be less than the amount of data saved in the SCADA database because each substation is responsible only for its own network section. This allows data to be saved more frequently, such as once per second or millisecond depending on the application. Since most of the data is stored locally, the requirements for communication links between substations and control centers are reduced.

This approach requires the use of decentralized algorithms that can bypass reduced SCADA functionality to provide optimal local control. Local controllers are intelligent enough to coordinate their work and thus provide reliable global control.

Some of these features are found in modern Energy Management Systems (EMS), for example, power flow analysis and generation prediction, that now will have to be used locally. At the same time, the EMS system, together with the Distribution Management System (DMS) and the Outage Management System (OMS), are superstructures of SCADA systems, and do not allow to fully implement the necessary functionality within a distributed control system [10].

In modern smart grid it is necessary to predict possible events (based on monitoring and trend analysis) and take proactive measures based on this data instead of a passive reaction to events in the distribution network. This forecast applies to both generation and consumption. Power flows in such network model will not be strictly deterministic. It is obvious that such a complex unstructured network must have a powerful control system that coordinates the operation of all network components. To do this, all components of the network must "communicate" with each other and with the control center through special communication networks. One of the ways to solve this problem is to use the consensus algorithm [11].

In [12], a distributed control strategy is proposed to share unbalanced currents in three-phase three-wire isolated AC microgrids. The control algorithm is based on an approach where the microgrid is analyzed not as a three-phase system, but as three single-phase subsystems. The consensus algorithm is used in distributed control system [13] to share the imbalance and harmonics between different converters in three-phase four-wire droop-controlled microgrids.

An asynchronous averaging consensus protocol [14] is proposed to estimate the values of inseparable global information. The consensus protocol is then combined with a fully distributed primal dual optimization utilizing an augmented Lagrange function to ensure convergence to a feasible solution with respect to power flow and power mismatch constraints. This algorithm uses only local and neighbourhood communication.

The consensus algorithm can be used to reach an agreement for resources dispatch among multiple controllable components in distribution grids, taking into account the individual voltage support capability of sources [15]. This algorithm reduces the time required to reach a consensus by solving the fastest distributed linear averaging (FDLA) problem.

An algorithm based on the RAFT consensus is used in [16], but in order to ensure the security of the switching infrastructure of the network against a specific type of attack, called a timing attack. In contrast to these studies, the authors use the RAFT consensus algorithm to ensure smooth voltage regulation and control of the electrical network parameters.

3. SSVR Control System

The concept of the IoE implies that the network becomes active and manageable, and the efficiency of management will only increase with an increase in the number of control devices in the network. At each point of the electrical distribution network where the SSVR installation is required, a node consisting of three connected elements is formed:

- solid state voltage regulator (SSVR);
- technological control system (TSC) – realizes the physical execution of SSVR control commands, measurement of the parameters of its functioning and transfer them to the upper level control system;
- intelligent control system (ICS) – provides the functions of SSVR operational control, diagnostics of its state, collecting monitoring data and implementation of algorithms for intelligent control of the SSVR network in order to control effectively power flows, ensure the required power quality and reduce losses during its transmission.

The SSVR control system has a peer-to-peer network structure and the SSVR ICS units are the nodes of this network. The ICS is an industrial computer, geographically located near the SSVR (the maximum distance is determined by the specification of the interface for TCS connecting). Also, the ICS is equipped with an interface for communication with other SSVRs ICS to create a data exchange network on the operating modes of the devices and to implement the algorithms for active-adaptive power transmission network control. At the same time, the results of the work of these algorithms should be propagated to neighboring nodes with minimal delays and with high reliability. This sets high requirements for an information network built in parallel with an electrical one.

The physical implementation of this network is possible using various technologies, including power line communication (PLC) technologies. A scheme of a 6-20 kV distribution network section with transformer substations and jumpers for controlling power flows with SSVR nodes and corresponding TCS and ICS are shown in Figure 2.

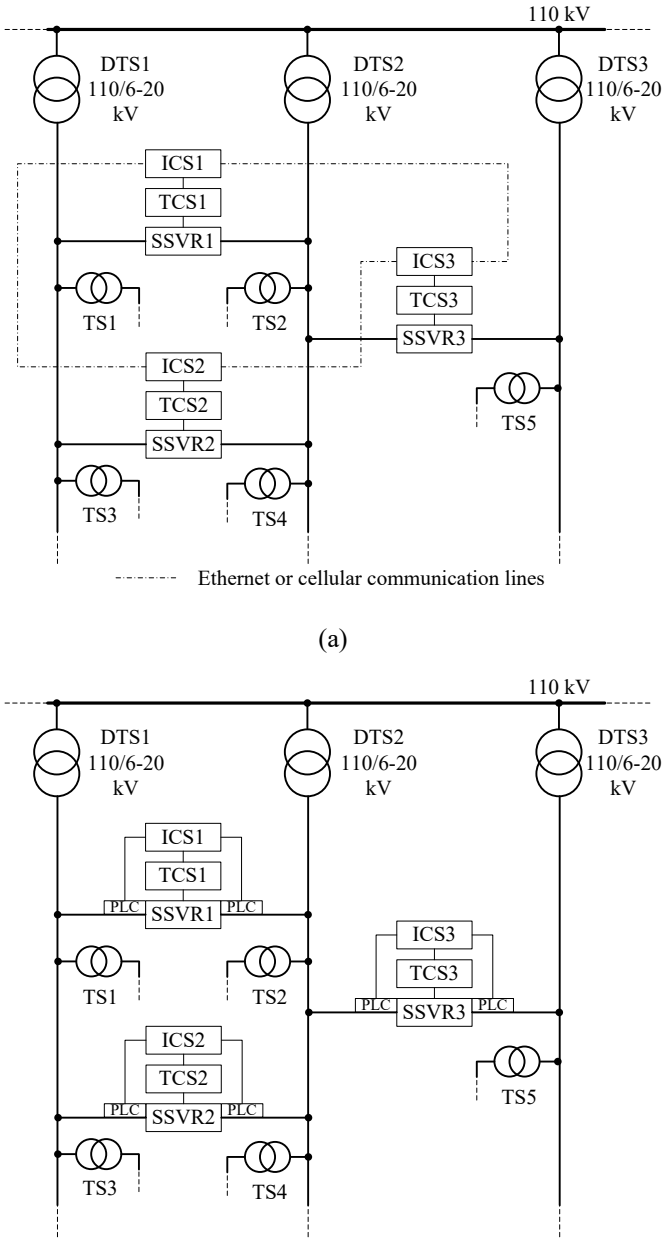


Figure 2: Structure of SSVR ICS: DTS1-3 – district transformer substations 110/6-20 kV; TS1-5 – transformer substations 6-20/0.4 kV; SSVR1-3 – solid-state voltage regulators; TCS1-3 – technological control system; ICS1-3 – intelligent control system: for Ethernet or cellular communications technology (a); for PLC technology (b)

In Figure 2a the ICS perform the function of routing information flows. The structure of the SSVR ICS for the PLC technology is shown in Figure 2b.

In addition to communication interfaces with TCS and other ICSs connected to SSVR located in the same electrical distribution network, there is an operational control client interface. This interface allows the operator to connect to the operating ICS in order to view the current parameters of its operation and, if necessary, correct the parameters of the operation of control algorithms for a specific SSVR or the SSVR network as a whole.

As indicated above, the automatic control in this system is decentralized and the regulation functions are performed in each SSVR ICS by local software implemented on the basis of the electrical network control algorithm (Figure 3). The developed control algorithm of the ICS is necessary to match the SSVR nodes in the distribution network. The function of distributing information about the state of network nodes is performed by the software for managing a group of industrial computers of the ICS. This group of ICS operates as a single system, called a cluster. The cluster control algorithm is used as the algorithm of this software.

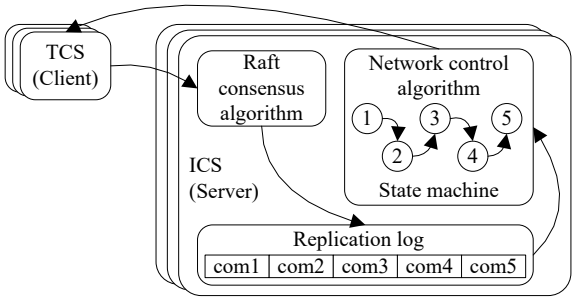


Figure 3: ICS functional diagram

As indicated above, the automatic control in this system is decentralized and the regulation functions are performed in each SSVR ICS by local software implemented on the basis of the electrical network control algorithm (Figure 3). The developed control algorithm of the ICS is necessary to match the SSVR nodes in the distribution network. The function of distributing information about the state of network nodes is performed by the software for managing a group of industrial computers of the ICS. This group of ICS operates as a single system, called a cluster. The cluster control algorithm is used as the algorithm of this software.

4. Network Control Algorithm

4.1. Management Task

Let's imagine a distribution electric network in the form of a graph, at the vertices of which there are transformer substations (TS) with their electricity consumers, and power lines act as arcs, and some neighboring TCs are connected by SSVR (Figure 4).

If the mode parameters in the network node deviate from the permissible (or specified by the operator), the SSVR ICS must send a signal to the TCS about the need for regulation. The effectiveness of the control system commands depends on the quality and completeness of the initial information about the network operation mode. The SSVR is capable of measuring the following parameters [5]:

- the RMS value of phase voltage before and after SSVR installation point;
- the effective value of the phase current in the line with SSVR;
- the phase angle of the line current relative to the voltage.

When information about deviations of the parameters and corresponding node number is obtained, it is possible to make decisions about control actions in order to bring these parameters to the expected values. The SSVR can regulate the amount of current in the line by controlling the magnitude of the phase voltage and its phase angle.

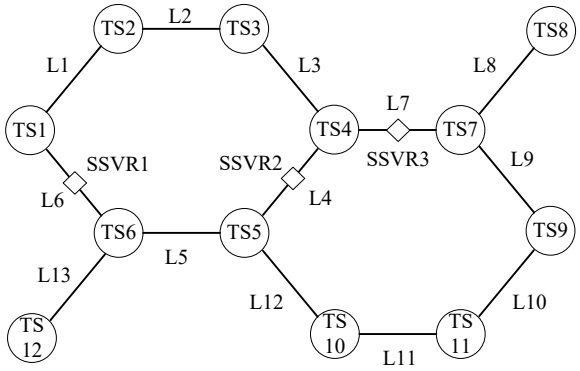


Figure 4: Graph of distribution electric network: L - power line; TS - transformer substation 6-20/0.4 kV; SSVR - solid state voltage regulator

4.2. Cluster Management Algorithm

To implement the adaptive network control algorithm, a logical computing cluster is formed on top of the electrical distribution network. This cluster formed of industrial computers, each of which controls its own separate SSVR, and is a control network node. It is assumed that all nodes in the cluster can communicate with each other without loss and without delay.

Cluster parameters are controlled by the cluster leader, which task is to control the power flows between the parts of the electrical network by changing the parameters of the cluster nodes. The leader is an arbitrary cluster node, chosen by voting between all cluster nodes, using the Raft consensus algorithm (consensus problem) [17].

The advantages of Raft are reliable leader election and guaranteed delivery of broadcast messages from the leader to the nodes. Despite its relative simplicity, Raft provides a safe and efficient implementation of a state machine on top of a clustered computing system. Raft ensures reliable delivery of signals to all nodes in a given order [18]. This ensures the transition of all state machines through the same sequences of states and each node is guaranteed to come into agreement with the other nodes.

Raft has the following features:

- assumes that there is always only one clearly distinguished leader on the cluster. Only this leader sends new records to other nodes in the cluster. Thus, the rest of the nodes follow the leader and do not interact with each other (except for the voting phase). If an external client connects to the cluster through a regular node, then all his requests are redirected to the leader and only from there come to the nodes.
- involves the decomposition of the cluster management task into several subtasks, the main of which are: leader election (voting) and replication of logs. Each of these tasks allows a more detailed separation. This makes the algorithm easier to understand and reduces the risk of errors in its implementation.
- logs cannot contain gaps. Records are added strictly sequentially. This simplifies the algorithm very much. In addition, the specifics of applied tasks, most often, do not allow correct work with logs containing gaps.

- allows to change easily the configuration of the cluster without interrupting operation: add or remove nodes.

An important circumstance is that Raft numbers strictly all records in the log. The records must go strictly sequentially. These records numbers play an important role in the operation of the algorithm. They determine the degree of relevance of the state of the node. The leader is always the most relevant node. The same numbers are used to number voting sessions. A node can vote only once for each vote request.

4.3. Network Control Algorithm

The network control algorithm is iterative. Each iteration of the algorithm consists of several steps:

- Each node reads its own voltage, phase and current values. Then, the parameter values are analyzed according to the specified criteria. The criteria include the maximum power flow through the node and the ability to share the load with an adjacent part of the network. The criteria are formed separately for each cluster node.
- The desired correction of the parameter values of each node is formed according to the calculation results, and then the desired parameters are transmitted to the cluster leader.
- After receiving the current values from each cluster node, the leader decides to change values on one of the nodes.
- It is assumed that the parameters should be changed iteratively and sequentially, by a small value and only at one node per iteration. The leader selects the node with the maximum required parameter correction among the nodes and sends a request to this node to change the parameter value by a certain step. The size of the step is set by the cluster settings.
- One of the cluster nodes receives a request from the leader to change his operating parameters, makes these changes and sends a message to the leader that the changes were successfully made. This concludes the iteration of the algorithm.

The indication that the node parameter values have been corrected is the change in the iteration number field. If the value of the iteration number is greater than the current number (on the leader's side), then the leader concludes that the node has recalculated the corrections of its parameters and sent updated values. Then the leader collects analysis data from all cluster nodes again. If changes in the correction of parameter values from the nodes are received, then the leader increments its current iteration number and repeats the steps to analyze and change the data.

Thus, at each iteration of the algorithm, small changes are made in the operation of the network at one of its nodes. Due to the fact that changes are made in the operation of only one node, and by a small value, the operation of the algorithm is stable and predictable and does not lead to an imbalance in the network.

Fault tolerance of the algorithm is achieved due to the fact that all network elements are combined into a cluster and each member of cluster is equal. The cluster leader which makes decisions about network management is unknown in advance and is chosen randomly from all the cluster members. That is, each cluster element contains all the necessary logic in order to become a leader, if the other members of the cluster elect it. The leader election procedure is finite and fault-tolerant, and is supported by the proven algorithm in cluster computing – Raft.

If the current cluster leader becomes unavailable due to problems in network connectivity and/or software or hardware failures, the cluster elects a new leader and the algorithm continues to work. If too many members drop out of the cluster due to possible failures, the algorithm determines that it does not have enough quorum to make decisions about network management and stops. In this case, a failure notification is sent to the operator.

The software simulation model of the SSVR ICS has been developed to study the operation of SSVR control algorithms in normal and emergency modes of network operation.

5. SSVR ICS Simulation Model

The simulation model allows to study the operation of the SSVR ICS in isolation without a real connection to the SSVR and without real communication channels. The software simulation model emulates the following objects and parameters:

- SSVR as ModBus slave [19];
- communication channels between SSVR and ICS;
- the topology of a network consisting of a set of SSVR ICS, with a set of communication channels connecting them.

The structural diagram of the simulation model is shown in Figure 5.

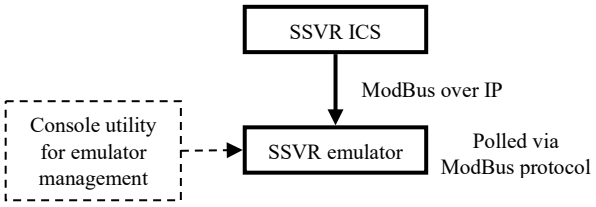


Figure 5: The structural diagram of the simulation model

The SSVR emulator has been written in C++ and is a Linux daemon (a background program launched by the operating system). The main task of the SSVR emulator is to respond to the requests of the ICS via the ModBus-over-TCP/IP protocol, just like a real SSVR does. The used implementation of the ModBus protocol defines a set of integer registers. For each of these registers in the SSVR emulator, a range of acceptable values and a period of values change are determined.

During its operation the SSVR emulator determines randomly the value of each register based on its range of changes, and gives it to the ICS in response to its requests. Also, the SSVR emulator changes the value of each register once per period. It is possible to change dynamically the settings for each ModBus register of the SSVR emulator without restarting the emulator itself using the console utility for the dynamic configuration of the emulator.

A configuration file with the network parameters of the available nodes in the topology is created for an instance of the

SSVR ICS software module to organize communication channels with other ICSs in the topology.

The console utility is used to change dynamically the settings for each ModBus register in the SSVR emulator without restarting this emulator, as well as to simulate changes in the network parameters and the occurrence of various emergency situations. So, for example, using this utility, it is possible to simulate the connection of an additional load to the SSVR by changing the values of the voltages and currents on the side after the controller.

6. SSVR ICS Simulation Model

Operating modes (scenarios) of the control network were simulated using the developed simulation model of the ICS SSVR. The cluster nodes were switched off one by one and the operation of the algorithm was monitored along with messages sent to the network operator.

The algorithm of actions of the control system when starting the system and the occurrence of emergency situations is shown in Figure 6.

Voting between the cluster nodes and the leader elections occurs when the ICS is turned on. Normal operation of ICS SSVR is ensured when two groups of conditions are met:

1. the cluster is working correctly;

2. the control algorithm works correctly.

The cluster works correctly when the following conditions are met:

- there is a cluster quorum n_q – the number of nodes in the cluster is sufficient to elect the leader (the minimum required number $n_{q.min} = 3$);
- the chosen leader is functioning normally;
- the set of nodes that have elected a leader is stable.

Algorithm quorum – there is a specified minimum number of controlled nodes which is necessary for the algorithm to work correctly.

Each time an alarm message is received indicating the nature of the problem, the operator must perform the required recovery and restart the system.

The simulation results are presented in Table.

The control system coped with the assigned tasks both in normal and emergency modes of operation. It was found that depending on the number and quantity of disconnected nodes two scenarios are possible:

- in scenarios B1 and C1, the remaining number of nodes was enough to select a leader, the system continued to work and perform the declared functions until the accident was eliminated;
- in scenarios B2, C2, D and E, the election of a new leader is impossible and the cluster stops its work until the accident is eliminated and the operator restarts the control algorithm.

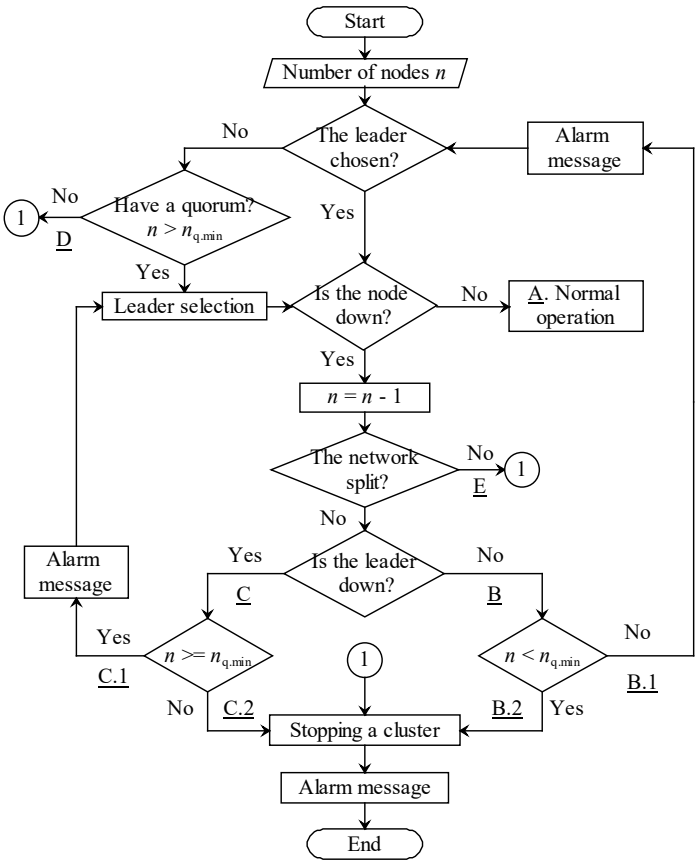


Figure 6: The control system algorithm: the letters A, B, C, D, E indicate the scenarios presented in the Table 1

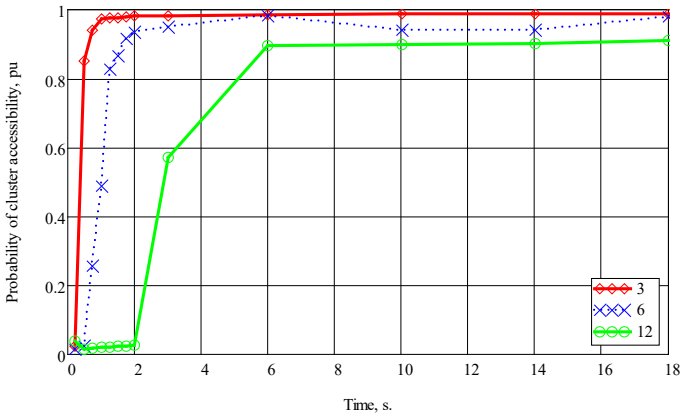


Figure 7: The dependence of the probability of cluster accessibility on the duration of the time interval between processing requests for a different number of network nodes

7. Determining the Cluster Speed Limit

The iterative control algorithm involves sending requests for data changes to the cluster. Sending commands takes some time, determined by the speed of the industrial computer of the node.

An experimental study on simulation models including 3, 6 and 12 nodes was carried out to determine the minimum cluster response time. Figure 7 shows the dependence of the probability of cluster accessibility on the duration of the time interval between processing requests.

Table 1: Modeling Results of Normal and Emergency Operation of the Control Network

Scenario	Result	Messages to operator
Scenario A. Normal network operation. No accidents.	The network topology does not change. The leader is chosen and does not change. There is a periodic measurement and iterative correction of the parameters of the nodes.	No messages
Scenario B.1. One of the nodes is down. The remaining number of nodes is sufficient for the operation of the cluster or for a quorum for a leader election.	The cluster continues its work according to the algorithm.	Alarm message indicates the ID of the node requiring recovery.
Scenario B.2. One of the nodes is down. One of the nodes is down. The remaining number of nodes is not sufficient for the operation of the cluster or for a quorum for a leader election.	The cluster stops its work.	Alarm message indicates the ID of the node requiring recovery.
Scenario C.1. The leader node is not working. The remaining number of nodes is sufficient for a quorum for voting	Election of a new leader.	Alarm message indicates the ID of the node requiring recovery.
Scenario C.2. The leader node is not working. The remaining number of nodes is not sufficient for a quorum for voting	The cluster stops its work.	Alarm message indicates the ID of the node requiring recovery.
Scenario D. No cluster quorum	The cluster stops its work.	Alarm message indicates the number of nodes.
Scenario E. The failure of one or more nodes led to the division of the network into two areas (partitions)	The cluster stops its work.	Alarm message indicates the ID of the node requiring recovery.

One can conclude that with an increase in the number of nodes in a cluster, the minimum allowable time interval between requests increases. From the given dependencies, it can be seen that the minimum interval between requests is 2.5 seconds for networks with a number of nodes up to 6 and 7 seconds for networks of 12 nodes. Such an interval is acceptable when regulating power flows in the electrical distribution network.

8. Conclusion

Decentralized management system architecture and an adaptive network control algorithm have been developed for solid-state voltage regulators (SSVR) in medium voltage distribution networks.

All SSVRs are combined into a single logical computing cluster using industrial computers, which are controlled by the cluster leader. The Raft consensus algorithm is used to find a leader and manage the entire cluster.

In the course of researching the control system on the simulation model, all cluster nodes were switched off one by one. In all simulated scenarios, the control system acted according to the laid down algorithm and followed the given commands. The research results confirmed the efficiency of the developed algorithm and the stability of its functioning.

Modeling clusters with different numbers of nodes showed that the minimum allowed time between requests for cluster nodes depends on its size. Thus, it is possible to determine the maximum allowable number of nodes in a cluster which will depend on the network topology and the tasks solved by the SSVR. The study of the performance of the algorithm proves the possibility of its application to solve the problem of power flow control in medium voltage distribution networks.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The study was carried out with a grant from the Russian Science Foundation (project No. 20-19-00541).

References

[1] A. Huang, “FREEDM system - a vision for the future grid,” in IEEE PES General Meeting, 1–4, 2010, doi:10.1109/PES.2010.5590201.2.

[2] Y.R. Kafle, K. Mahmud, S. Morsalin, G.E. Town, “Towards an internet of energy,” in 2016 IEEE International Conference on Power System Technology (POWERCON), 1–6, 2016, doi:10.1109/POWERCON.2016.7754036.

[3] E.N. Sosnina, A.I. Chivenkov, I.A. Lipuzhin, “Solid-State voltage regulator for a 6–10 kV distribution network,” in 2018 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), 1–7, 2019, doi:10.1109/FarEastCon.2018.8602948.

[4] Y.P. Kubarkov, A.V. Chivenkov, N.N. Vikhorev, A.V. Shalukho, I.A. Lipuzhin, “Algorithm of power line current formation at load node,” Vestnik of Samara State Technical University. Technical Sciences Series, **1**(61), 128–138, 2019, doi:10.14498/tech.201901128-138 (article in Russian with an abstract in English).

[5] E.N. Sosnina, A.I. Chivenkov, I.A. Lipuzhin, N.N. Vikhorev, “Control system for vector regulation of power flows in medium voltage network,” IOP Conference Series: Materials Science and Engineering, **643**, 012048, 2019, doi:10.1088/1757-899X/643/1/012048.

[6] N. Ji, S. Geiger, “Reducing outages in distribution by testing recloser controls,” in 2014 China International Conference on Electricity Distribution (CICED), 1748–1788, 2014, doi: 10.1109/CICED.2014.6992261.

[7] V. Sharipov, A. Minin, I. Mokhov, “A protection strategy for power grids equipped with Siemens vacuum reclosers,” in 2010 Innovative Smart Grid Technologies (ISGT), 1–5, 2010, doi: 10.1109/ISGT.2010.5434741.

[8] W. Li, Y. Li, C. Chen, Y. Tan, Y. Cao, M. Zhang, Y. Peng, S. Chen, “A full decentralized multi-agent service restoration for distribution network with DGs,” IEEE Transactions on Smart Grid, **11**(2), 1100–1111, 2020, doi: 10.1109/TSG.2019.2932009.

[9] C. Ien, D. Botting, A.D.B. Paice, J. Finney, O. Preiss, “When grids get smart,” ABB Review, **1**, 44–47, 2008.

[10] L. Strezoski, I. Stefani, B. Brbaklic, “Active management of distribution systems with high penetration of distributed energy resources,” in IEEE EUROCON 2019 -18th International Conference on Smart Technologies, 1–5, 2019, doi:10.1109/EUROCON.2019.8861748.

[11] D. Ongaro, J. Ousterhout, "In search of an understandable consensus algorithm," in 2014 USENIX Annual Technical Conference, 305–319, 2014.
[12] C. Burgos-Mellado, J. Llanos, E. Espina, D. Saez, R. Cardenas, M. Sumner, A. Watson, "Single-phase consensus-based control for regulating voltage and sharing unbalanced currents in 3-wire isolated AC microgrids," IEEE Access, **8**, 164882–164898, 2020, doi: 10.1109/ACCESS.2020.3022488.
[13] C. Burgos-Mellado, J. J. Llanos, R. Cardenas, D. Saez, D. E. Olivares, M. Summer, A. Costabeber, "Distributed control strategy based on a consensus algorithm and on the conservative power theory for imbalance and harmonic sharing in 4-wire microgrids," IEEE Transactions on Smart Grid, **11**(2), 1604-1619, 2020, doi: 10.1109/TSG.2019.2941117.
[14] B. S. Millar, D. Jiang, "Asynchronous consensus for optimal power flow control in smart grid with zero power mismatch," J. Mod. Power Syst. Clean Energy, **6**, 412–422, 2018, doi.org/10.1007/s40565-018-0378-4.
[15] K. Zhang, D. Recalde, T. Massier, T. Hamacher, "Fast online distributed voltage support in distribution grids using consensus algorithm," in 2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), 350–355, doi: 10.1109/ISGT-Asia.2018.8467821.
[16] A. Sheoran, V. Singh, "Consensus based Anomaly detection in Smart Grid," Fault-tolerant computer system design. Final project report. https://engineering.purdue.edu/ee695b/public-web/handouts/Projects/Reports/amit_vishwanath_smartgridsecurity.pdf.
[17] The Raft Consensus Algorithm. [Online]. URL: https://raft.github.io.
[18] C. Fluri, D. Melnyk, R. Wattenhofer, "Improving raft when there are failures," in 2018 Eighth Latin-American Symposium on Dependable Computing (LADC), 167–170, 2010, doi:10.1109/LADC.2018.00028.
[19] libmodbus. A Modbus library for Linux, Mac OS X, FreeBSD, QNX and Win32. [Online]. URL: http://libmodbus.org.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 386-392 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Smart Collar and Chest Strap Design for Rescue Dog through Multidisciplinary Approach

Fang-Lin Chao[1], Wei Zhong Feng[2,*], Kaiquan Shi[1]

[1]*Department of Industrial Design, Chaoyang University of Technology, Taichung, 436, Taiwan*

[2]*Department of Visual Communication, Chaoyang University of Technology, Taichung, 436, Taiwan*

ARTICLE INFO

Article history:
Received: 09 November, 2020
Accepted: 13 January, 2021
Online: 22 January, 2021

Keywords:
Rescue dog
IoT device
Concept design
Remote interaction
Design education

ABSTRACT

Rescuers escorted search dogs into the disaster area, using their unique sense of smell to find the injured. First, researchers summarize the design requirements in the search process from interviews with rescuers, and construct a conceptual prototype to confirm the interaction mode between the user and the dog. User central design invited people melt into the situation to identify product features. The ideas were selected based on the viability which increases efficiency. The main design proposal includes a strap and a smart collar. Smart sensing (heartbeat, speed, temperature, and GPS) can improve communication and increases the efficiency of rescue. The search area is large in many cases; therefore, we selected the WiFi or Ultra-wideband module as the wireless transmission medium when the rescue team enters this domain. The pre-deploy nodes connect and position with the smart collars. The instructor sends voice commands remotely to prompt the dog to return when the temperature is high. The smart collar design includes an elastic O-ring waterproof shell. Rescuers click the recall button, and the remote device sends a signal of dog returning. This proposed work looks more at user's needs through multi-disciplinary aspects of view, which enhance usability. The case consists of customer interviews, observation, concepts, evaluation (science/ device/ electronic packaging/ and App software); the design process also demonstrated a possible teamwork perspective in the industry. This scenario encourages cross-field extension for design education.

1. Introduction

Rescue dogs are sensitive to hearing and smell [1], and can be used as rubble rescue dogs and mountain rescue dogs after personalized training. This research proposed suitable devices for rescue task. During task, dog will present a stabilize bark to mark the sufferer position and special events [2]. The rescue team fights for gold rescue time to locate suffered people in the mountains or landslides space. Inappropriate equipment reduce efficiency.

The design goal is improving communication efficiency of the training with the help of smart collar through the design process (session 3). The design team collected requirements (session 3.1) through user interviews (session 2.1); after the midterm meeting, we built the prototype (session 4.1, 4.2) to support the interaction between the user and the dog. From the observation of the rescue member's situation, we confirm the rationality of the design (session 4.4, 4.5).

*Corresponding Author: W.Z. Feng, 79 Futong St., Taichung 436, Taiwan R.O.C., 886-918276088, feng.jimmy@msa.hinet.net

2. Literature survey

2.1. Rescue dog training situation and facilities

Figure 1a shows the rescue dog training cite in central Taiwan, the site is suitable for air scent search of the surrounding or trace the odors. Figure 1b is layout and facilities in Taichung training area. Trainer evaluate dogs and determine working potential to give the dog of their skills to fulfilling its purpose.

The training ground provides olfactory practice and obstacle crossing in a large area, and then enters the building of the injured person through overcome obstacles, such as suspended bridge structure. Simulation provides robustness under various conditions of the dog. Dogs in rocky areas need to be careful not to hurt their paws when moving (Figure 2). The training included the communication skill with the rescue personnel to notify the location of the victim [2]. When the road is in poor condition, the foot is often injured. Dogs often stick out their tongues to strengthen sweat; long-term search can lead to body failure [3]. Passing on the current physical condition allows calling for retreat.



(a)



(b)

Figure 1: Rescue dog training in Taiwan: (a) field survey site; (b) Taichung training area.



Figure 2: Training situation and facilities in rocky area.

2.2. Prior art of Strap

There are several ideas of dog lash as follows: Double D-ring rope buckle, LED warning; H-shaped chest for quick-fitting; X-shaped design using breathable fabric, and Y-shaped structure on the front chest. Double leash keep safe of manage reactive dog with a front and back ring. A leash shifts the center of gravity back towards trainer which gives trainer a pivot point [4] without restricting dog's shoulder for freely searching. Dog is dependent on trainer to explore the sights, trainer have to teach "how to follow lead" or teaching commands through strap.

Sokolowski explored 3D hand grip shape for dog walking which captured from civilian's [5] to fulfill product design perspective. Both cognitive and physical abilities help a successful handler design [6]. Figure 3 reveals an ergonomic X-strap dog harness. The traction belt is smooth and evenly bears multi-directional forces. Another design is an enhanced retractable leash [7] which provides multiple handles for controlling the movement. Plastic buckles are utilized with Velcro straps which eliminate the leash pulling forces at the dog's center of mass (middle chest) area.

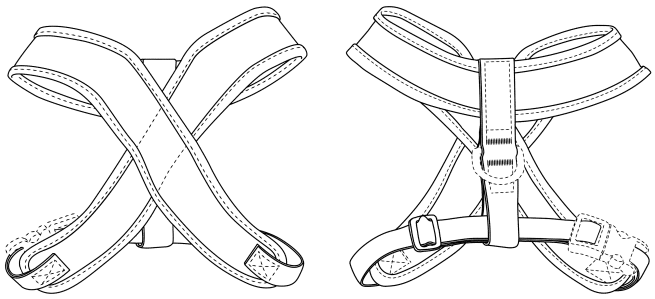


Figure 3: Ergonomic design X strap for dog, US 29/395,500 [6].

2.3. Scent and airflow

The skin disseminates cells with special odors in the metabolism, which form diffusion due to air convection. When dogs approach, they can feel the odor; after being trained in specific odors, it can increase sensitivity [2]. During the training process, the dog tracked across the sidewalk, turf, and forest path; the notification after sniffing helps people confirm the rescue path.

Researchers [3] developed a two-part dog wearable computer and interface. In the concept design process, the team first interviewed experts to correctly explain the dog's behavior. When the dog finds impure odor sources, timely encouragement by the trainer can improve its ability [8]. The experiment determined the distance and concentration that the dog can recognize, and was surprised to find that the dog can detect the smell at a distance of more than 62 meters.

2.4. IoT device in dog/ human interaction

The elderly-animal friendship bond surveyed elf-perceived criteria by the old people regarding their intimate association with their dogs [9]. A survey of 60 people [10,11] asked people classify the behavior of the nine dogs in the video as "friendly" or "radical." Researchers found that the dog's tail's movement is the main clue to explain the dog's sensation [12,13].

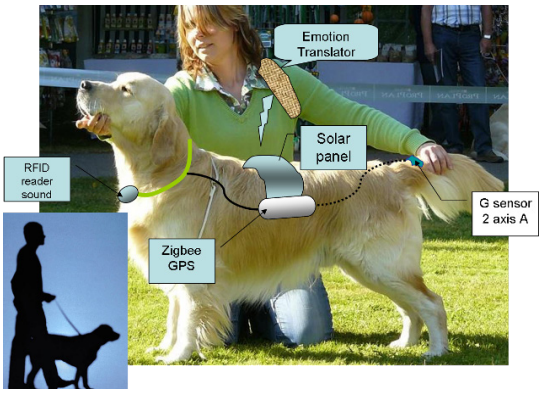


Figure 4: Placement of the major components and main controller [15]

Sensor and actuator help the communication between the blind and the dog; the power spectral density detection method based on breathing sound discovers the physiological state related to different odor detection tasks [14]. Previous study [15] proposed a remote interaction device implemented with Zigbee wireless technology between pet and elderly. The signal provided outdoors activities monitoring and communication, the module placed on

the back of the dog, and issue commands using a voice pager. The control chassis consist of battery, charger circuit and Zigbee circuit (Figure 4). Owner use vibrating actuator to command the dog remotely. Other study developed [16] an algorithm for locating victims in the nearly collapsed building with indoors positioning and extra IR and sound sensors.

2.5. Possibility of WiFi nodes

Many rescue areas do not present a WiFi connection. If WiFi is not available, the trainer depends on the sound to communicate. The mobile station could install, which provides the possibility of access in an emergency. Moreover, Ultra-wideband (UWB) and Long Range (LoRa) were utilized for a high-precision UAV positioning [17] with altitude error reduced to 1.74m. Simultaneous localization could be obtained through WiFi signal strength measurements [18]. Received signal strength data provide a convent way of practical implementation with a set of nodes.

The above technology is accurate, but this study focuses on the interaction between the trainer and the dog. To simplify the complexity, we use the WiFi module for the usability evaluation.

3. Design Method

The successful design of a product takes into account the user response, the alternative designs can amplify designers' understanding of the intended purpose [19]. The user central design reduced error and accelerating the decision-making process [20].

3.1. Design process

We collected user desire through interviews with rescue team. Organize their feelings into group records and use the data as a source of concept design. A scenario invited people melt into the pre-settled situation and identified needs. The conceptual ideas were evaluated based on the viable characteristics. The design flowchart shows in Figure 5.

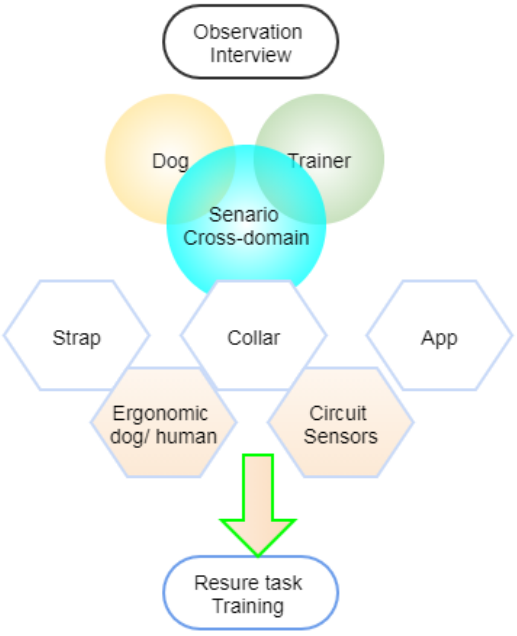


Figure 5: The design process and flowchart

Inquiring with professionals from the Taichung rescue team, the researcher also participated volunteer activity to observe the training details. Interview data shows: dog looks down and scans left and right for evidence. When the trainer pulls back, the neck will be red and swollen. An inappropriate chest strap will cause friction and injury to the fur. Don't hinder neck movement; prevent objects from hooking; use breathable materials when hot and humid.

The design begins with the scenario which deduced from insight of the task and following directions were issued;

1) Include small flashlights and emergency relief kits, but do not hinder neck movement.
2) The disaster may face rainy environmental and need waterproof capability.
3) Taiwan is hot and humid, and dogs do not dissipate heat well. Breathable and thin materials is preferred.
4) GPS-path records located the trapped; the App is needed to improve the efficiency of search.
5) Health datas (dog's heartbeat and body condition) are essential for physical condition control.

Figure 6 shows the selected design concepts of a tracking device and straps which reduces the obstacles interference during the mission. The design features overcome the current product are following:

1) Does not hinder the search activities. When the dog looks down, the branches may get caught in the leash or traction loop and be injured.
2) The product frame is breathable, avoiding heat stroke during dog wearing, which satisfied design priority of: comfort, safety and dog's acceptance. The strap disperses pulling forces on the muscles.
3) Dogs and rescuers can support each other through complementarity, but the configuration of equipment cannot interfere with the task.

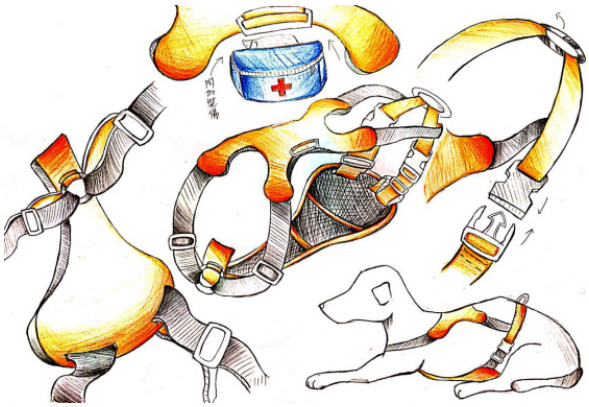


Figure 6: The idea sketch of design concept

3.2. Multidisciplinary approach

The implementation requires a cross-domain integration by running a multidisciplinary collaborative project. The ability to lead a team through the different work stages of a project is a fundamental contributor to its success [21]. Framework for the conceptual design of unmanned aerial vehicles (UAVs) proposed

in developing efficient winged design [22] for industrial designer. Teachers used cross-domain design and developed " airflow play with music" to realize the concept of flow guiding.

Comfort rescue dogs' products needs structural mechanics, material applications, and most importantly, software and hardware integration. Cross-domain presentation requires communication and interaction between departments. The conceptual designer proposes idea sketch and functions supported in the context of use, which discussed with the engineer in advance. Finally, it can be realized with appropriate technology. This process includes the assembly of electronic circuit boards, and software and hardware test to present the design concepts. Then through the deduction of the concept the design verified with the cross-domain team.

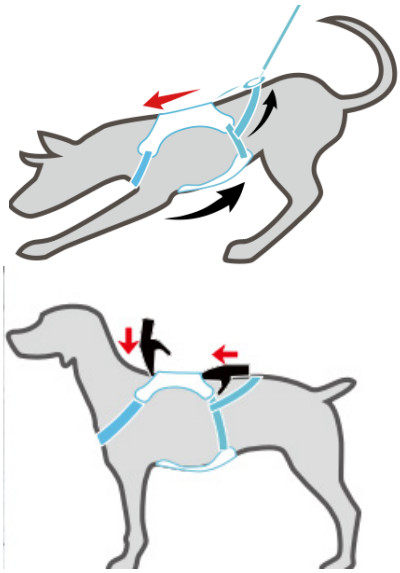
4. Concept design

4.1. Chest strap

Comfortable shoulder straps need to be worn quickly and fit the body without friction. The chest H-shaped backrest equipped with an adjustable double belt buckle. The Y-shaped structure made of breathable fabrics reduces tensile strength (nylon, plastic and fluorescent reflective tape). The front stainless steel O-ring can release and reduce pressure.



(a)



(b)

Figure 7: (a) the upper chest pad and storage, (b) the force applied to the body during exercise, the lower pad is made of insulating cotton and wear-resistant washed canvas (sandwich mesh).

Figure 7 illustrates the design features of the chest harness: (1) equipped the upper body of the dog with a space for light devices (batteries for smart devices, simple first aid supplies, and communication equipment) for survivor; (2) as shown in the figure 7a, the side rings inserted from both sides and placed on the upper back. During the search, the trainer followed the dog in front, and strap is with tension; the traction point placed in the upper region to avoid neck strain. The illustration (Figure 8) shows that the force is directed to both sides through the outer strap to distribute it. The D-ring connects both sides with the lower abdominal support belt, and control belt of trainer. The force in three directions joined at chest. The overall design of the chest harness shows in the diagram (Figure 9). The abdomen are soft materials to avoid gaps and friction. When the chest harness is properly fitted to the body, the dog will not hook on the protrusions or tree trench when moving.

The flexible adjustment structure avoids gaps. The harness is soft to coordinate with the body shape. The designed strap presents lots of curvature elements. Based on the ergonomic and strap's curvature, the software using for CAD drawing is the Catia version-6.

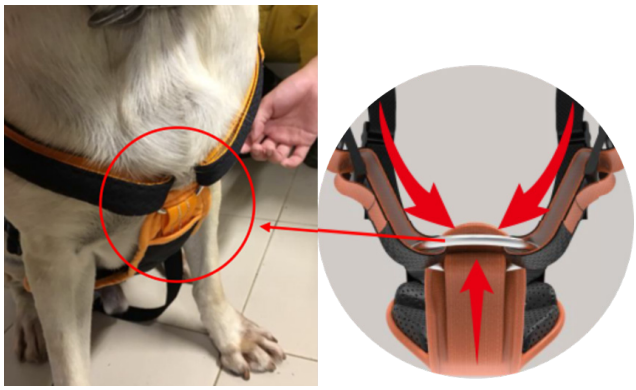


Figure 8: Combination of straps on chest



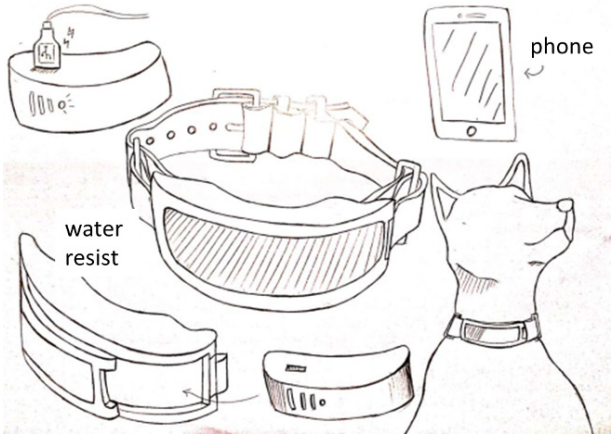
Figure 9: The 3D rendering of the chest strap

4.2. Smart Collars

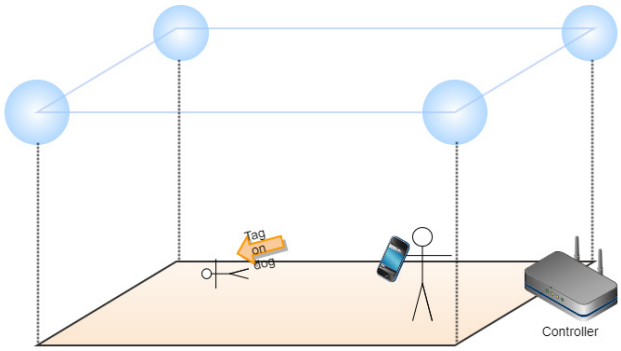
Combined the existing wireless transmission and positioning technology, the communication between the dog and the rescuer is better in a large range search. Traditionally, dog barks to hint trainer, this touching way extended nearby areas. With wireless device, dog barks trigger the signal of the transmission position

data and notify the trainer of the location of the victim (Figure 10a), wireless touching way extended more than 150 m diameter areas. The injury of the dog is a great loss of team, so their physical conditions are monitoring by trainer [23]. The sensor in the collar transmit the physical condition of dog, which allowing the trainer to adjust schedule. This function prevents the dog from being exhaustion.

The search mission is larger in the open field. The indoor Wi-Fi router operating on the 2.4 GHz frequency band has a radio wave of up to 50 m, and it can be greater than 92 m outdoors. Based on IEEE 802.15. 4a standard, above 200 meters electromagnetic wave range can be reach with UWB [24]. User selects suitable module based on desired range. Figure 10b shows the concept of Balloon floating positioning. The UAVs generate a lot of electromagnetic waves and noise, which affect the concentration of dogs. We chose movable balloon set to place the RF module along with the balloon; hovering over the disaster area, it can quickly move to the nearby search areas. The floating balloon will move slightly with the ambient wind but does not affect the position's accuracy much.



(a)



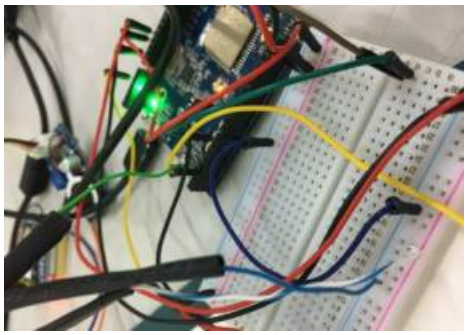
(b)

Figure 10 The proposed design: (a) smart ring module, (b) Balloon floating positioning. The UWB constructed with Semtech SX1276 Dragino LoRa Shield (with Arduino via SPI).

4.3. Implementation

A wireless remote connection collar was realized with Arduino platform, the system included a trainer side App to receive signal and arrange current task. The interaction between the smart collar and the rescuer's mobile phone are using Wifi signal. Since the Arduino system is an open platform, many Shield modules are available.

Sometimes, the extended search area cannot cover by Bluetooth signal; therefore, we selected the WiFi or Utra-wideband module as the wireless transmission medium. When the rescue team enters this domain, they will deploy nodes which connects with the tag on smart collars. The mobile phones also share connection and able to communicate with other devices.



(a)



(b)

Figure 11: Components used in the circuit: (a) wireless remote connection realized with Arduino Shield (Ameba RTL8195), WiFi antenna, temperature and humidity sensor (DHT22), and sound sensor (NEO-7M-CAPP), (b) 3D printed chassis structure.

The Arduino Shield [25] used include GPS, speakers, Grove main control version. The breadboard wiring implementation shows in Figure 11a. The appearance of the smart collar (3D printed) contains miniaturized circuit board and sensors (Figure 11b); the collar placed on the neck area of the rescue dog with an adjustable belt. During the task, the collar might contact water, to maintain the stability of the circuit, so we installed a rubber O-ring placed in chassis gap for waterproofing of the circuit board.

4.4. Software and Trainer's APP

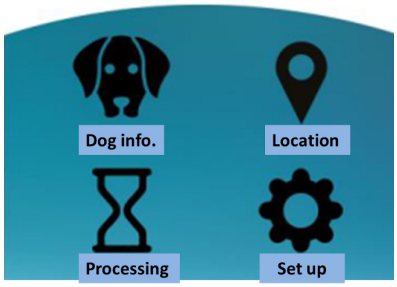
Heartbeat/ temperature/ and GPS status recorded by Smart collar. The squeak signal is identified with the trainer's App. The user interface can be implemented using Android APK. The dog's track record and the graphical display enhance recognition (Figure 12a). Collected dog's physical condition will display on phone screen so that the trainer can immediately prevent the dog from being exhaustion (Figure 12b).

The major Information are:

- Positioning: The dog's barks activates and informs the trainer which reveal the location of the target (Fig 12b) of the target in the mission.
- Call back: When the body temperature exceeds the set value or the task exceeds the time, the coach issues a warning. Use the withdrawal button to call the module attached to the dog and make a sound (instruction return). The training procedures

are: Call the dog's name enthusiastically and give the command: "Come! Come!" If he walks towards you, praise it. When the dog catches up, give him some snacks and compliments immediately. After several successful exercises, try to say the word alone. The dog understands the sound in previous training when rescuers click the recall button and remotely generate the same sound signal. When the success continues, try the same action with the remote microphone. The proposed device mechanism is vocal sound, the inertia of signal/ action built in early time excites the dog's physiological reasoning.

- Guidance: rescuer asks the survivor person to prompt the dog to get the direction (Figure 13). The recorded path and target's response enabled speed up of target finding.
- Dog's background: The dog's characteristics and the training performance can access by the trainer, through specific description data helps conduct of those tasks with specific dog.



(a)

```
status: {
    'fix': 0,
    'fixquality': 0,
    'lat': 0,
    'lon': 0,
    'speed': 0,
    'alt': 0,
    'temp': 0,
    'hum': 0,
    'sound': 0,
    'alert': 0,
    'updatetime': 1529736163638,
}
```

(b)

Figure 12: The main page of (a) App interfaces; and (b) code of Arduino board; where fix: successful positioning; speed: moving speed; lat/lon: latitude and longitude; Temp/Hum: temperature and humidity; Sound: volume of barking; alert: notifying back.

4.5. User survey

Taichung established a search and rescue dog team in 2016, and had passed the Mission Readiness Test international mission rescue certification. Training require a lot of human resources including build society contact to successfully complete search tasks. Among the team, "Tie Xiong" found survivors in the rubble during the Hualien earthquake in 2018. We took the chest strap to the training site and test it by the commander during the exercise. At the verification stage, we made a prototype with soft materials for user test including magnitude of pressure, the shear force on skin, and the loading. The normal stress around chest does decrease due to force distribute on supporting surface. Per Figure 7 and 8, the most considerable stress occurs on the arch ribs due to pull.

The App interface is properly matched with the mobile phone, and the touch interface for quick setting and operation. The timely response is helpful for urgent handling in emergency situations. The current Arduino modules are large in size and easily interfere with the movement of the dog; in the future, this problem can be overcome through the miniaturization of a customized circuit board.

```
#include "device.h"
#define GPS_TX 10
#define GPS_RX 12

//GPS
SoftwareSerial GPS_Serial(GPS_TX, GPS_RX);
Adafruit_GPS gps(&GPS_Serial);

char ssid[] = "CPH1707"; //"TOTOLINK N100RE"; //CPH1707
char pass[] = "99999999";
IPAddress CloudServer(192, 168, 43, 191);
#define ServerPort  1234
```

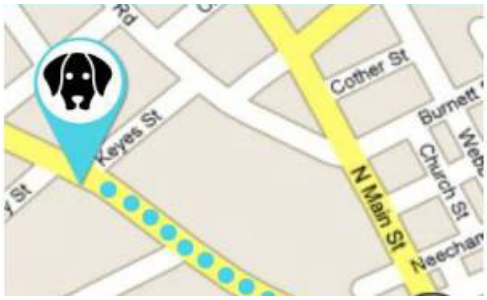


Figure 13: GPS positioning help target located on the trainer mobile App.

5. Conclusions

Based on situational stories and on-site observation data, we have unified the special needs in the rescue situation. The main design proposal includes a strap and a smart collar; the strap that meets the rescue dog's situation. It can be adjusted to suit individual needs. Smart sensing improves task performance; IoT chest strap helps communication and increases the efficiency of search and rescue. The smart collar records the status, and remote voice paging helps to send voice commands. Arduino's rich development modules including GPS and speakers are conducive to the prototype. Researchers found that the dog's wearable sensor improves the dog-human interaction during training.

Many technologies reported; this work looks more at user's needs through multi-disciplinary aspects of view. The conceptual situation requires the realization of the prototype to show the suitability of the concept. From the perspective of design education, we show the main parts through cross-field integration (mechanical/electronic, hardware/software). This proved that designers need to enhance cross-domain sentiment to stand at the high point of design thinking in order to propose appropriate solutions.

References

[1] L. Lit, C.A. Crawford, "Effects of training paradigms on search dog performance," Applied Animal Behaviour Science, **98**(3-4), 277-292, 2006.
[2] K.E. Jones, K. Dashfield, "Search-and-rescue dogs: an overview for veterinarians," Journal of the American Veterinary Medical Association, **225**(6), 854-860, 2004.

[3] C. Zeagler, C. Byrne, G. Valentin, "Search and rescue: dog and handler collaboration through wearable and mobile interfaces" Proceedings of the Third International Conference on Animal-Computer Interaction, **6**, 2016.
[4] M.W. Trevino, "Ergonomic pet harness with enhanced retractable leash" United States Patent Application US 15/398,603. 2018.
[5] S.L. Sokolowski, C. Bettencourt, "Investigation of 3D Functional grip shape to design products for dog walking and hiking," International Conference on Applied Human Factors and Ergonomics, 597-604, 2020.
[6] Y.H. Yun, "X-strap dog harness," United States Patent Application US 29/395,500. 2018.
[7] I.P. Hoskinson, "Ergonomic pet carrier for single-person use," United States Patent Application US 16/104,847, 2019.
[8] I. Greatbatch, J.G Rebecca, and S. Allen, "Quantifying search dog effectiveness in a terrestrial search and rescue environment" Wilderness & Environmental Medicine, **26**(3), 327-334, 2015.
[9] J.M. Alcaidinho, "The internet of living things: Enabling increased information flow in dog-human interactions," Ph.D. diss., Georgia Institute of Technology, 2017.
[10] A.C. Jones, R.A. Josephs, "Interspecies hormonal interactions between man and the domestic dog (Canis familiaris)," Hormones and Behavoiur, **50**(3), 393-400, 2006.
[11] E. Kubinyi, "Dog and owner demographic characteristics and dog personality," Behavioural Processes, **81**(3), 392-401, 2009.
[12] Z. Horvàth, "Affiliative and disciplinary behaviour of human handlers during play with their dog affects concentrations in opposite directions," Hormones and Behaviour, **54**, 107-114, 2008.
[13] J.M. Ley, P.C. Bennett, "Understanding personality by understanding companion dogs," Anthrozoös, **20**(2), 113-124, 2008.
[14] C. Mancini, S. Li, J. Valencia, D. Edwards, "Towards multispecies interaction environments: Extending accessibility to canine users," in ACM International Conference Proceeding Series, Association for Computing Machinery, 2016, doi:10.1145/2995257.2995395.
[15] F.L. Chao, "Evaluation of industrial usefulness of pet companion for elderly," Eco-design, 2013.
[16] S. Mealin S, M. Foster, K. Walker, "Creating an evaluation system for future guide dogs: a case study of designing for both human and canine needs," In Proceedings of the Fourth International Conference on Animal-Computer Interaction, 1-6, 2017.
[17] Y.C. Chen, I. Alexsander, C. Lai, and R.B. Wu, "UWB-assisted high-precision positioning in a UTM prototype," 2020 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT), 42-45, 2020.
[18] L. Bruno, P. Robertson, "Wislam: Improving footslam with wifi," In 2011 International Conference on Indoor Positioning and Indoor Navigation, 1-10, 2011.
[19] M.R. Endsley, Designing for situation awareness: an approach to user-centered design.
[20] B. Ipaki, Z. Merrikhpour, "Designing bike shift lever with user-centric design approach," Journal of Ergonomics, **6**(3), 43-54, 2018.
[21] S. Mazzetto, "A practical, multidisciplinary approach for assessing leadership in project management education," Journal of Applied Research in Higher Education. 11(1), 50–65, 2019.
[22] H. Gu, X. Cai, J. Zhou, F. Zhang, "A coordinate descent method for multidisciplinary design optimization of electric-powered winged UAVs," In2018 International Conference on Unmanned Aircraft Systems (ICUAS), 1189-1198, 2018.
[23] R. Brugarolas, S. Yuschak, D. Adin, L. Sherman, A. Bozkurt, " Simultaneous monitoring of canine heart rate and respiratory patterns during scent detection tasks," IEEE Sensors Journal, **19**(4), 1454-1462, 2019.
[24] M. Poturalski, J.P. Hubaux, "Distance bounding with IEEE 802.15. 4a: Attacks and countermeasures," IEEE Transactions on Wireless Communications, **10**(4), 1334-1344, 2011.
[25] R.B. Brufau, "Towards automated canine training: wearable cyber-physical systems for physiological and behavioral monitoring of dogs," 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 398-402 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Method for Improving the Quality of the Product Obtained by Abrasive Treatment with Impregnated Tools

Viktor Butenko[1], Liana Gusakova[1], Dmitry Durov[2], Boris Safoklov[*,2], Oleg Dolgov[2]

[1]*Department Design of Special Aviation Complexes, Moscow Aviation Institute (National Research University), Moscow, 125993, Russia*

[2]*Department Technological Design and Quality Management, Moscow Aviation Institute (National Research University), Moscow, 125993, Russia*

ARTICLE INFO

Article history:
Received: 28 July, 2020
Accepted: 31 October, 2020
Online: 20 November, 2020

Keywords:
Abrasive Tools
Impregnation
Chromium Diiodide
Surfactants

ABSTRACT

The article describes a way to increase the efficiency of an abrasive tool by its impregnation in an aqueous solution of chromium diiodide. The device scheme for impregnating the instrument is proposed. Research results of durability of tools impregnated with chromium diiodide and structural condition of material of a surface layer of details after processing are resulted. Within of the researches presented in the article the grinding processing a shafts made from various iron-carbon steels with prepared abrasive tools was conducted. As a result of the machining the scattering fields of the batch sizes of the parts within the grinding wheel resistance were determined. This data for processing with standard grinding wheels and impregnated by chromium diiodide were done. Using data from controlled grinding of workpieces, approximate values of the K_{Δ} coefficient were determined depending on the diameter of the workpiece surface, the required size accuracy of part and the method application of the surfactants. It is shown that regardless of the material being processed, the use of a grinding wheel impregnated with chromium diiodide leads to a reduction in the specific value of the accumulated deformation energy E_{spec}. Taking into account these results it is possible to predict up to twofold reduction of the cost of machining with the abrasive tool of grinding parts, which will reduce the cost of their manufacture by 20-25%. The developed process of impregnating grinding wheels with chromium diiodide can be used without large economic costs in many machine-building company also during repair works.

1. Introduction

The problem of increase of efficiency of abrasive processing of surfaces of machine parts remains actual in machine-building manufacture and one of the perspective directions of its decision is impregnation (saturation) of abrasive tools with surfactants.

There are known methods [1] of abrasive instruments [2] impregnation [3,4] with paraffin, ozokerite, phenolformaldehyde resin, sulfur, rosin, cerezin, stearin, potassium ethylxantogenate, sodium silicate solution, aromatic chlorine-containing compounds and other substances [5], compositions of several elements including fillers [6] (graphite, molybdenum disulfide, fluoroplastics, micropowders, etc.). Shortcomings of these compounds are reflected in the limited applicability and the need to select an impregnating compound depending on the material and tool, in some cases, the toxicity to humans and the use of impregnating compounds solely on a water-based, corrosive activity and insufficient retention of chemicals in the pores of the grinding wheel, reducing its hardness, etc, possible increase in quality control points and increase in the prime cost of the final product.

When grinding with impregnated wheels, there are fast processes of destruction and restoration of solid lubricating films in the contact zone between the grinding tool and the workpiece due to high speeds of molecules movement at high temperatures and pressures. The substances located in the pores of the wheel can pass into different aggregate state, and the impregnation efficiency

[*]Corresponding Author: Boris Safoklov, Email: safoklovbb@mai.ru

depends on the products of its decomposition, their penetration ability and the activity of interaction with the processed material [7].

High reactivity at interaction with metals, especially with high adhesion to abrasive material, halogens [8] and among a large number of types of surfactants, as well as methods of their application in the blasting of surfaces of machine parts the most effective is the method of impregnation of grinding wheels in an aqueous solution of chromium diiodide [9, 10], molecular formula CrI_2 [11], we have considered all the various ways to improve the efficiency of an abrasive tool, and developed a new effective method for impregnating grinding wheels.

2. Materials and Methods

To increase the efficiency of the abrasive tool there was developed an effective method of impregnating grinding wheels in the aqueous solution of chromium diiodide [11], the main essence of which is that the abrasive tool on a ceramic bond is placed in a container with aqueous solution of chromium diiodide (20-25 g per liter of water) and within one minute it is informed about the vertical oscillating motion with a certain frequency ν, which provides its complete immersion in the aqueous solution and complete removal from it [12,13]. After impregnation, the impregnated instrument is subjected to convective drying when rotating at a speed of (0.3-0.5) s^{-1} for 1.5-2 hours at a temperature of (40-50) ºC.

Vertical vibrating movements of the impregnated grinding wheel with a given frequency v and amplitude A were carried out using a special installation, shown in Fig. 1 and containing an electric motor 1, gearbox 2, cam 3, pusher 4, rocker arm 5, mobile support 6, rod 7, to which the impregnated abrasive tool 8 is attached, and a container with an aqueous solution of chromium diiodide 9.

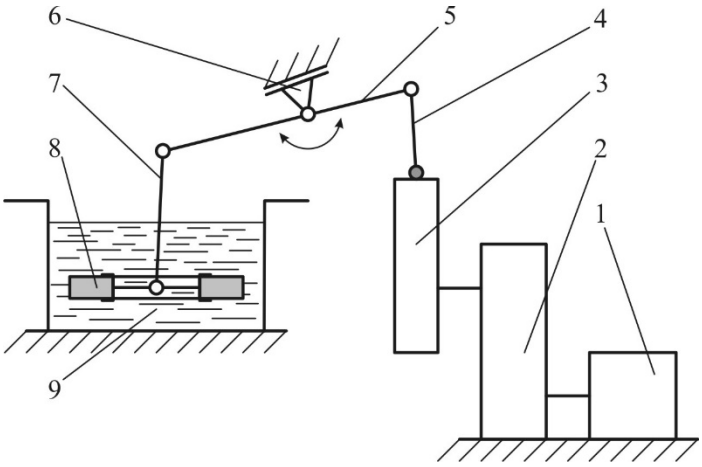


Figure 1: Plant diagram for impregnating grinding wheels

Research of efficiency of application of grinding wheels impregnated with chromium diiodide was carried out by processing of samples from steel 12CrNi3A (Table.1) with diameter 30 mm and length 200 mm after finishing turning (surface roughness before grinding Ra = 8.0-12.5 μm). Grinding of samples was carried out on a circular grinding machine mod. 3U12VF11 PP 400×50×203 14AF60K7V GOST R 5281-2007 with cutting speed V_{gw} = 35 m/s, sample rotation speed V_m = 0.314 m/s, longitudinal feed S_{lon} = 0.02 m/s, cutting feed S_{cut} = 0.01 mm / double stroke, the number of double strokes m = 3. The 5% aqueous solution of emulsion Ukrinol-14 was used as a cooling lubricant [14,15]. The change of firmness of grinding wheels T after their dressing was determined.

Table 1: Steel Type 12CrNi3A, chrome-nickel structural alloy steel products GOST 4543-2016.

C, %	Si, %	Mn, %	P, %	S, %	Cr, %	Ni, %
0.09-0.16	0.17-0.37	0.3-0.6	≤0.025	≤0.025	0.6-0.9	2.75-3.15

With the use of scanning electron microscope "Quanta-200" were carried out spectral studies of the state of the surface layer of samples made of steel 12CrNi3A, subjected to grinding by impregnated chromium diiodide abrasive tool.

In [12,16,17,18] we have shown that the most important characteristic of the operational properties of the surface layer material of processed parts is its energy state, which can be estimated by the specific value of the accumulated energy of deformation E_{spec}. For experimental determination of the value of E_{spec} a special unit was developed, the description of which is given in [17, 19]. One of the most important indicators of the effectiveness of the use of grinding wheels is to ensure the necessary dimensional accuracy of the workpieces during the entire period of tool life. In this regard, a controlled treatment was carried out by grinding shafts of various iron-carbon alloys.

Dispersion fields for batch sizes of parts within the grinding wheel resistance were determined. A large number of shafts with diameters from 20 mm to 40 mm and tolerances of IT6-8 were subjected to controlled machining over a long period of time. Based on the obtained results of controlled machining, the scattering fields of diameters of the parts processed without the use of surfactants ω_1 and with the use of various methods of feeding the impregnative agent – chromium diiodide into the machining zone (impregnation, additive to lubricant-and-cooling technological agent, introduction into the abrasive mass) ω_2 were determined. The scattering fields were determined after processing each batch of 10 pieces. The scattering fields were compared when the first batch ω_1^b, ω_2^b was processed and the last batch ω_1^e, ω_2^e was processed.

3. Results

3.1. The effectiveness of the use of chromium diiodide

The results of research into the effectiveness of the use of chromium diiodide impregnated grinding wheels are presented in the form of histograms in Fig. 2, from the analysis of which it follows that the resistance of chromium diiodide impregnated grinding wheels in the process of dressing decreases: first significantly, then on average by 3-5% after each dressing, and then starts to grow again. This is apparently due to the continuous washing out by the lubricating and cooling fluid of the impregnator chromium diiodide from the intergranular space of the grinding wheel.

3.2. The analysis of spectrograms

The analysis of spectrograms of electronic microscope "Quanta-200" of surfaces of details from a steel 12CrNi3A in an

initial condition and after grinding by impregnated chrome diiodide [17] speaks that in the course of grinding active iodine interacts with juvenile surfaces of details, forming on them the iron iodides having layered structure and possessing low coefficient of a friction [9, 20].

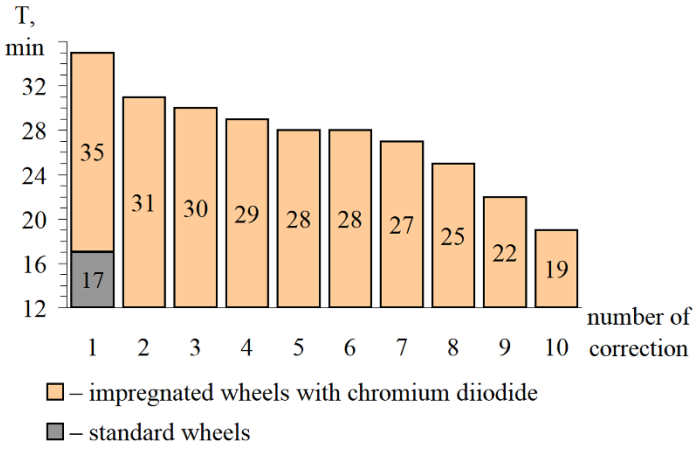


Figure 2: Histograms of changes in resistance of impregnated grinding wheels as they are corrected

3.3. Results of change of scattering fields.

Results of changing scattering fields of sizesits details ω_1 and ω_2 for the period of tool firmness at an admissible range of change of sizes of processed details in the generalized form are presented in fig. 3, which allows to trace the pattern of their change.

It was found that in all cases $\omega^b > \omega^e$, which seems to be associated with a decrease in the number of cutting grains and an increase in smoothing and pressing [21]. This is demonstrated by decreasing in the roughness parameter R_a of the treated surfaces of parts (Fig. 4) and an increasing in the value of technological residual tensile stresses in the material of the surface layer σ_{rem} (Fig. 5) with simultaneous reduction in metal removal.

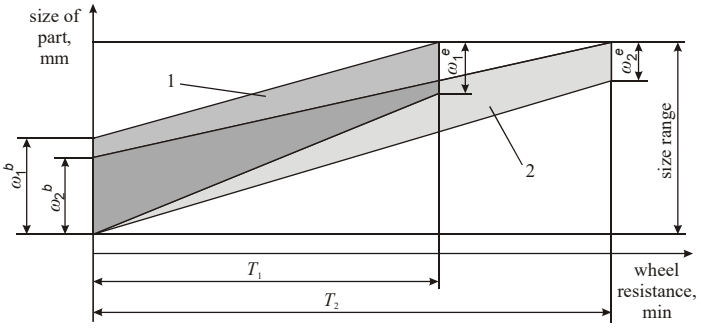


Figure 3: Diagrams of scattering fields for batch sizes of parts within the wheel resistance, processed with standard (1) and impregnated (2) tools.

Analysis of the data shown in Fig. 3, indicates that $\omega_1^b > \omega_2^b$ and $\omega_1^e > \omega_2^e$. If we take $\Delta\omega_1 = \omega_1^b - \omega_2^b$ and $\Delta\omega_2 = \omega_1^e - \omega_2^e$, then $\Delta\omega_1 \geq \Delta\omega_2$.

Using the coefficient $K_\Delta = \Delta\omega_1 / \Delta\omega_2 \geq 1$, then by its value it will be possible to decide on the degree of influence of the adopted type and method of surfactant application on the accuracy of the obtained dimensions of the processed part. Using the data of controlled machining by grinding of workpieces, approximate values of coefficient K_Δ depending on diameter of processed surface, required qualification of size accuracy and application method of impregnator (in this case, chromium diiodide) were determined, which are given in Table 2. These K_Δ values can be used during the selection a surfactant application method.

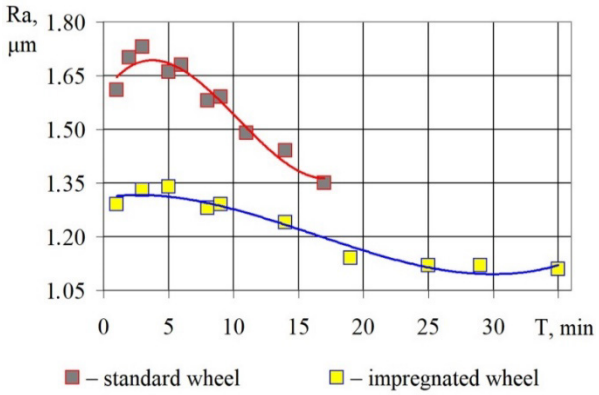


Figure 4: Changing the roughness parameter of the machined surface of the part for the period of the grinding wheel resistance T when using standard and impregnated tools (steel 12CrNi3A)

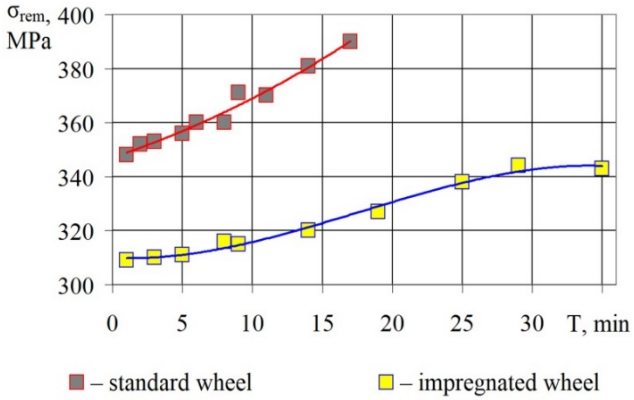


Figure 5: Variation of the value of technological residual stresses in the material of the surface layer of the machines parts for the period of grinding wheel resistance T during the use of standard and impregnated tools (steel 12CrNi3A)

Table 2: Values of coefficient K_Δ by processing parts from iron-carbon alloys with yield point 250-400 MPa

Workpiece diameter, mm	Accuracy qualification	How to use the surfactant	Coefficient K_Δ
20-30	6	Impregnation	1.4
		Additive to cooling lubricant	1.6
		Introduction to abrasive mass	1.5
	7	Impregnation	1.6
		Additive to cooling lubricant	1.7
		An injection into the abrasive mass	1.8
	8	Impregnation	1.9
		Additive to cooling lubricant	1.7
		An injection into the abrasive mass	2.0
31-40	6	Impregnation	1.5
		Additive to cooling lubricant	1.7
		An injection into the	1.6

		abrasive mass	
	7	Impregnation	1.8
		Additive to cooling lubricant	1.7
		An injection into the abrasive mass	1.8
	8	Impregnation	2.0
		Additive to cooling lubricant	1.7
		An injection into the abrasive mass	2.1

To determine the degree of influence of grinding wheel characteristics, its condition and composition of the used cooling lubricant on the specific accumulated energy of deformation E_{spec}, a significant amount of experimental studies on samples from various iron-carbon alloys, subjected to grinding on the regimes recommended in the reference [22], the results of which are reflected in Table 3. From their analysis it follows that the energy state of the material of the surface layer of the workpiece being grinded depends to a great extent on the state of the abrasive tool used and the composition of the cooling lubricant. Regardless of the material to be treated, using of a grinding wheel impregnated with chromium diiodide leads to a reduction in the specific value of the accumulated deformation energy of E_{spec}. At the same time, the injection of 20 grams per liter of chromium diiodide into the composition of the cooling lubricant also leads to a decrease in the value of E_{spec}.

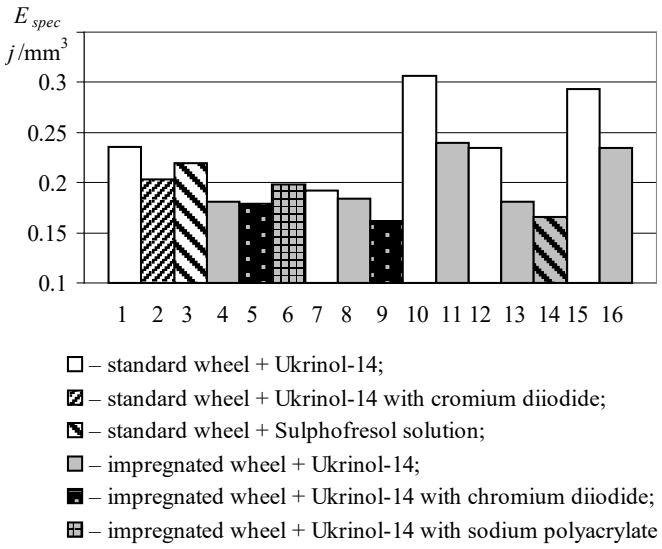


Figure 6: E_{spec} histogram of changes from the state of the grinding wheel and the composition of the used lubricating and cooling process agent

For clarity on the figure 6 is shown a histogram of E_{spec} value changes from the state of the wheel and the composition of the used the cooling lubricant when grinding samples from steel 50 (1-6), steel 12CrNi3A (7-9), alloy 45Cr25Ni20Si2 (10, 11) [23] with a wheel of 14AF60K7V and steel 50 (12-14) and alloy 45Cr25Ni20Si2 (15, 16) when grinding with a wheel of 54AF60 M7V. The histogram shows that the minimum values of E_{spec} can be achieved with the simultaneous use of chromium diiodide impregnated grinding wheel and the cooling lubricant, which consists of 5% aqueous solution of emulsion "Ukrinol-14" with the addition of 20 g/l of chromium diiodide. Apparently in this case, chromium diiodide, which is a part of lubricating and cooling technology, not only reduces the coefficient of friction in the area of contact of the abrasive grain with the processed material, but also contributes to the permanent restoration of chromium diiodide in the impregnated wheel [24-31].

Table 3: Depending of E_{spec} values on the material to be treated, grinding wheel characteristics and the composition of the lubricant and coolant used.

# Ex.	Processed material	Grinding wheel characteristic	Grinding wheel condition	Composition of the lubricant and coolant used	$E_{spec.}$ j / mm^3
1	Steel 50	14AF60K7V	standard	5-% emulsion aqueous solution Ukrinol-14	0.236
2				5-% emulsion aqueous solution Ukrinol-14 + 20 g/l chromium diiodide	0.203
3				5-% sulphofresol solution	0.219
4			impregnated	5-% emulsion aqueous solution Ukrinol-14	0.181
5				5-% emulsion aqueous solution Ukrinol-14 + 20 g/l chromium diiodide	0.179
6				5-% emulsion aqueous solution Ukrinol-14 + 20 g/l sodium polyacrylate	0.198
7	Steel 12CrNi3A		standard	5-% emulsion aqueous solution Ukrinol-14	0.192
8			impregnated	5-% emulsion aqueous solution Ukrinol-14	0.184
9				5-% emulsion aqueous solution Ukrinol-14 + 20 g/l chromium diiodide	0.162
10	Steel		standard	5-% emulsion aqueous solution Ukrinol-14	0.306
11	45Cr25Ni20Si2		impregnated		0.240
12	Steel 50	54AF60M7V	standard		0.235
13			impregnated		0.181
14				5-% sulphofresol solution	0.166
15	Steel		standard	5-% emulsion aqueous solution Ukrinol-14	0.293
16	45Cr25Ni20Si2		impregnated		0.235
17	Steel 50	64AF70P10V	standard		0.230
18			impregnated		0.177
19				5-% sulphofresol solution	0.172
20	Steel		standard	5-% emulsion aqueous solution Ukrinol-14	0.298
21	45Cr25Ni20Si2		impregnated		0.234

4. Conclusions

- The energy state of the material of the surface layer of the grinded part depends largely on the condition of the abrasive tool used and the composition of the lubricant coolant. Regardless of the material to be treated, the use of a grinding wheel impregnated with chromium diiodide leads to a reduction in the specific value of the accumulated deformation energy of E_{spec}. The injection of 20 grams per liter of chromium diiodide into the lubricating and cooling process also leads to a decrease in the value of E_{spec}.
- Thus, the processing of parts with impregnated chromium diiodide wheels allows to increase up to two times the resistance of used abrasive tools, it is quite natural to increase the size accuracy of the processed parts, to improve the

quality of the surface layer and to create on the processed surfaces of parts – rotational bodies, such as rollers, axles, fingers, etc., structures with a low coefficient of friction within 0.08-0.12 depending on the operating conditions, which is very important for the parts of tribosystems.

- Taking into account these results it is possible to predict up to twofold natural reduction of the cost of machining with the abrasive tool of grinding parts, which will reduce the cost of their manufacture to 20-25%. The developed process of impregnating grinding wheels with chromium diiodide can be used in any machine-building production, including repair, without large economic costs.

Nomenclature

E_{spec} – specific latent energy of deformation j/mm^3;
K_Δ – coefficient $K_\Delta = \Delta\omega_1 / \Delta\omega_2$;
R_a – arithmetic mean deviation of the profile (μM);
V_{gw} – cutting speed, m/s;
V_m – rotation speed, m/s;
S_{lon} – longitudinal feed, m/s;
S_{cut} – cutting feed, mm / double stroke;
Ukrinol-14 – Cooling and lubricating fluids (oil);
ω_1 – the scattering fields of diameters of the parts processed without the use of surfactants;
ω_2 – the scattering fields of diameters of the parts processed with the use of the impregnation.

14AF60K7V GOST R 5281-2007 – Grinding wheel Russian standard is adopted from the European, code example*:*

14A*F60K7V* (type of abrasive, ***14A*** – normal electro corundum);
*14A****F60****K7V* (***F*** **–** standard FEP (Federation of European Producers of abrasives) and ISO 8486, **F60** average grain size is 0.26 mm);
*14AF60****K****7V* (Hardness, ***K*** – medium soft);
*14AF60K****7****V* (***7*** – structure number indicates how much of the grinding wheel volume is occupied by abrasive grains, **7** – 48...50%);
*14AF60K7****V*** (type of bundle, ***V*** – a ceramic bundle).

References

[1] Y. Rudometov,. "Application of the abrasive instruments impregnated with impregnator suspensions," (in Russian) STIN, **11,** 34-37, 2012.

[2] W. Kapłonek, K. Nadolny, M. Sutowska, M.Mia, D.Y. Pimenov, M.K. Gupta, "Experimental Studies on MoS2-Treated Grinding Wheel Active Surface Condition after High-Efficiency Internal Cylindrical Grinding Process of INCONEL® Alloy 718," Micromachines **10**(4), 255, 2019, doi.org/10.3390/mi10040255

[3] G.V. Chirkov, "Impact of grinding wheel impregnation on processing quality," Engineering Technology, **2**, 22-23, 2007.

[4] A.V. Nikitin, "Grinding of Hard Materials by Impregnated Wheels as a Way to Increase Their Cutting Properties," Tools and Technologies, **28**, 52-58, 2010.

[5] W. Kapłonek, K. Nadolny, K. Rokosz, J. Marciano, M. Mia, D.Y. Pimenov, O. Kulik, M.K. Gupta, "Internal Cylindrical Grinding Process of INCONEL® Alloy 600 Using Grinding Wheels with Sol–Gel Alumina and a Synthetic Organosilicon Polymer-Based Impregnate," Micromachines **11**(2), 115, 2020, doi.org/10.3390/mi11020115

[6] M.-Y. Tsai S.-X. Jian "Development of a micro-graphite impregnated grinding wheel," International Journal of Machine Tools and Manufacture, **56**, 94-101, 2012, doi.org10.1016/j.ijmachtools.2012.01.007

[7] V.A. Nosenko, A.P. Mitrofanov, A.A. Krutikova. "Increase of Grinding Efficiency with the Use of Halogen-forming Impregnator," Izvestia vysokhraneniya. Mechanical Engineering. **8** (665), 65-72, 2015, doi.org/10.18698/0536-1044-2013-1-50-54

[8] Morrison R., Boyd R. Organic Chemistry, Prentice Hall India, 2002.

[9] V.I. Butenko, Finishing of the parts surfaces: methods, devices, tools (in Russian), Rostov-on-Don: DSTU Publishing Centre, 2016.

[10] V.I. Butenko, L.V. Gusakova. Increase of the machine parts surface grinding efficiency (in Russian) Taganrog: SFU Publishing House, 2012

[11] "Royal Society of Chemistry APIs" [Chromium(II) iodide]. Available Online at: http://www.chemspider.com/Chemical-Structure.13318420.html

[12] V.I. Butenko, L.V. Gusakova, A.D. Kulinskiy "Perspective directions of increase of efficiency of grinding of surfaces of machine parts," Vestnik of Bryansk State Technical University, **2**(50), 112-120, 2016, doi.org/10.12737/20254

[13] H.T. Nguyen, Q.C. Hsu, "Surface Roughness Analysis in the Hard Milling of JIS SKD61 Alloy Steel," Applied Sciences, **6,** 172, 2016, doi.org/10.3390/app6060172

[14] V.I. Butenko, D.S. Durov, R.G. Shapovalov. "Science-intensive technologies for creation of the high resource machine parts," Taganrog SFU Publishing House, 2014.

[15] E. Benedicto, D. Carou, E.M. Rubio, "Technical, Economic and Environmental Review of the Lubrication/Cooling Systems Used in Machining Processes," Procedia Engineering, **184**, 99-116, 2017, doi.org/10.1016 / j. proeng.2017.04.075

[16] V.I. Butenko, Structure and potential of the polycrystalline materials tribocontact (in Russian), Rostov-on-Don: DSTU Publishing Centre, 2018.

[17] V.I. Butenko, Technological compatibility of the functional layers and coatings (in Russian), Rostov-on-Don, 2020.

[18] V.A. Lebedev, Technology of the dynamic methods of the surface plastic deformation (in Russian), Rostov-on-Don: DSTU Publishing Center, 2006.

[19] W. Li, Q Tian, B. Zhang. "Impregnant for impregnated grinding wheels and application thereof in preparing impregnated grinding wheels," Patent no. 101791786 CN, B 24 D 3/34. 2010.

[20] V.I. Butenko, K.V. Davidenko "Method of impregnation of abrasive tools," Patent RF № 2703063. MPK B24 D 3/34/– Bulletin 29, 2019.

[21] A.A. Ryzhkin, Physical bases of the metal processing by cutting: a training manual (in Russian) Rostov-on-Don: Publishing centre. DSTU, 1996.

[22] A.M. Dalskiy, A.G. Suslov, A.G. Kosilov "Directory of the machine-building technologist" Moscow: Mashinostroenie-1, 2003.

[23] F. Ju "Special Steel Co.," Available Online at: https://www.jfs-steel.com/ru/steelDetail/-1Cr25Ni20-Si2/GB-1Cr25Ni20-Si2.html

[24] B.K. Sato, J.C. Lopes, A.E. Diniz, A.R. Rodrigues, H.J. de Mello, L.E.A. Sanchez, P.R. Aguiar, E.C. Bianchi, "Toward sustainable grinding using minimum quantity lubrication technique with diluted oil and simultaneous wheel cleaning," Tribology International, **147**:106276, 2020, doi.org/10.1016/j.triboint.2020.106276.

[25] S. Agarwal, "On the mechanism and mechanics of wheel loading in grinding," Journal of Manufacturing Processes, **41**, 36-47, 2019, doi.org/10.1016/j.jmapro.2019.03.009

[26] H. González, A. Calleja, O. Pereira, N. Ortega, L.N. López de Lacalle, M. Barton, "Super Abrasive Machining of Integral Rotary Components Using Grinding Flank Tools," Metals, **8**, 24, 2018, oi.org/10.3390/met8010024

[27] K. Nadolny, M. Wojtewicz, W. Sienicki, D. Herman, "An analysis of centrifugal MQL supply system potential in the internal cylindrical grinding process," Archives of Civil and Mechanical Engineering 15, 639–649, 2015, doi.org/10.1016/j.acme.2014.08.009

[28] M. Wojtewicz, W. Sienicki, K. Nadolny, "Vitrified grinding wheels impregnation methods with use solid lubricants," Mechanik. **September** 1108-1109, 2016, doi.org/10.17814/mechanik.2016.8-9.271.

[29] B.P. Ravuri, B.K. Goriparthi, R.S. Revuru, V.G. Anne, "Performance evaluation of grinding wheels impregnated with graphene nanoplatelets," The International Journal of Advanced Manufacturing Technology, **85**(9), 2235–2245, 2016, doi.org/10.1007/s00170-015-7459-6.

[30] R.B. Pavan, A.V. Gopal, M. Amrita, B.K. Goriparthi "Experimental investigation of graphene nanoplatelets–based minimum quantity lubrication in grinding Inconel 718", Proceedings of the Institution of Mechanical Engineers, Part B, Journal of Engineering Manufacture. **233**(2), 400-410, 2019, doi.org/10.1177/0954405417728311

[31] W. Kapłonek, K. Nadolny, M. Wojtewicz, W. Sienicki, "Characterisation of abrasive tools active surface after the impregnation process by modified ARS method based on imaging and analysis of the scattered light". International Journal of Machining and Machinability of Materials, **17**(5), 397 – 417, 2015, doi.org/10.1504/IJMMM.2015.073149

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 403-407 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Level of Resilience and Family Functionality in Adolescents of two Educational Institutions of a Vulnerable Area in Lima Province

Rosa Perez-Siguas[1], Hernan Matta-Solis[1], Eduardo Matta-Solis[1], Melissa Yauri-Machaca[2], Anika Remuzgo-Artezano[3], Lourdes Matta-Zamudio[4]

[1]*Research and Intellectual Creativity Direction, Universidad María Auxiliadora, Lima, 15408, Perú*

[2]*Research and Technology Direction, BM Tech, Lima, 15408, Perú*

[3]*Faculty of Health Sciences, Universidad María Auxiliadora, Lima, 15408, Perú*

[4]*Health Sciences, Instituto Peruano de Salud Familiar, Lima, 15408, Perú*

ARTICLE INFO

Article history:
Received: 04 August, 2020
Accepted: 28 October, 2020
Online: 20 November, 2020

Keywords:
Resilience
Family
Adolescent
Family functionality

ABSTRACT

Family functionality is when the family meets the needs of its members, especially adolescents. Resilience is the adolescent's ability to face situations in which there are difficulties, life problems and thus be able to overcome them. The purpose of the study is to determine the Level of Resilience and Family Functionality in Adolescents of two Educational Institutions of a Vulnerable Area in Lima Province. It is a non-experimental study, descriptive of a quantitative approach, it is a cross-sectional correlational investigation, with a population of 204 adolescents, also demographic data and the resilience scale (CD-RISC) and family APGAR. In the results regarding to the level of resilience and family functionality, adolescents present a mild family dysfunction with a medium level of resilience (58.0%), in comparison of the Educational Institutions, in both cases, the majority of students are located in the medium level of resilience, in the case of students from Honorio Manrique Nicho school 48.5% and in the case of students from the Andrés Avelino Cáceres school 54.2%. It is recommended in these institutions to carry out activities that promote the social formation of the adolescent, since this will allow them to communicate with society and the environment.

1. Introduction

Family plays a fundamental role in the adolescent, since it is an influential factor for the relationship with society [1]. One of the stages where one seeks to find changes in society is adolescence, adolescents try to know each other better and also seek difference from themselves with other family members, this process that the adolescent does is a change abruptly, but it is the way in which they will become autonomous and independent people, this process has passed all the members of the family [2].

Family functionality is understood when the family meets the needs of its members, especially adolescents, since it is a stage where there are differentiated changes with other family members and also adapt to situations of change [3]. So, a family is considered functional when it is able to look for solutions to any problem, so that it does not affect the satisfaction that family members have, including the adolescent [4].

Resilience is understood as the adolescent's ability to face situations in which there are difficulties, life problems and thus be able to overcome them [5]. Resilience is very important for adolescents, because it will allow them to adapt and recover from adversities of the problem so that it continues projecting itself in its future [6].

In [7], the author from Portugal presented that it was evidenced that adolescent males are those who have higher levels of resilience compared to adolescent females; in the self-

*Corresponding Author: Rosa Perez-Siguas, Mrs., +51 1 996826968 & rosa.perez@uma.edu.pe

confidence component, males obtained a score of 35.08% and in the relationship with peers, 26.28%.

In [8], carried out in Taiwan, the author mentioned that in children and adolescents, it was found that there is a risk of family dysfunction where approximately 40% of families have accumulated family risk due to family stress that occurs within their home and also 30% of the mothers have low levels of education or health problems. Lastly, around 7% of the participants, children or adolescents, come from large families that have 4 or more children living in a home.

In [9], carried out in South Korea, the author described that it was found that family dysfunction related to domestic violence and parental addiction was related to smartphone addiction, it was found that the observed association between family dysfunction and Smartphone addiction was not significant in adolescents with self-control or high friendship quality, therefore, adolescents exposed to domestic violence or with addicted parents had a higher probability of smartphone addiction, but had high self-control or friendship quality that could protect them.

The objective of the study is to determine the level of family resilience and functionality in adolescents from two educational institutions of a vulnerable area in Lima Province, in which it will allow to observe what is the level of adolescent resilience and family functionality. This study is important since it will give relevant and real data about family resilience and functionality as adolescents enter a stage of sudden changes where they seek solutions, where they can adapt and manage the situation in the most appropriate way.

In the study, the data collection instrument Connor-Davidson Resilience Scale (CD-RISC) and Gabriel Smilkstein's Family APGAR were applied, which have been useful to assess the level of resilience and family functionality of adolescents from the secondary level of educational institutions. The data collection was processed through the survey of adolescents from educational institutions, the data to be entered was performed in a data matrix that will be designed in the statistical program SPSS (Statistical Package for the Social Sciences) in its version 24.0, in which it will allow a better data processing to make statistical tables so that they can be described and interpreted in results and discussions, respectively.

The research work is structured as follows: In section II, the development of the data collection process of each nursing student will be presented, as well as the guidelines to consider so that they are in the research work. In section III, the results will show the level of resilience and family functionality of adolescents according to the specified dimensions of the instruments in the measurement of the variables. In section IV, we present the discussions and conclusions of the research work.

2. Methods

In this section, the type and design of the research will be developed, as well as the population that will be carried out in the research work, the inclusion and exclusion criteria will be explained in detail, and finally the technique and the instrument for collecting data.

2.1. Research type and design

The present research work is non-experimental, descriptive of a quantitative approach, it is a correlational research [10].

2.2. Population

The population is made up of 204 adolescents from the two educational institutions.

Inclusion criteria:

- Adolescents who normally attend the educational institution.
- Adolescents who have informed consent and assent.
- Adolescents who have the authorization signed by their parents and agree to voluntarily join the study.

Exclusion criteria:

- Adolescents who do not regularly attend to the educational institution.
- Adolescents who do not have the authorization of their parents.

2.3. Technique and Instrument

The technique used is the survey, using the questionnaire or data collection instrument (CD-RISC) and Family APGAR which aims to measure the level of family resilience and functionality in adolescents from two educational institutions of a vulnerable area in Lima Province.

The Connor - Davidson Resilience Scale (CD-RISC) was built by Connor and Davidson in the United States in 2003. Its application is for the adolescent and adult population from 11 to 80 years old, the scale evaluates the following dimensions: Persistence-tenacity-self-efficiency (8 items); Control under pressure (7 items); Adaptability and Ability to Recover (5 items); Control and purpose (3 items) and Spirituality (2 items). It consists of 25 structured items, on a summative Likert scale where 0 "absolutely", 1 "rarely", 2 "sometimes", 3 "often" and 4 "almost always". The scale range goes from 0 to 100, the higher score means more resilience [11].

The family APGAR was designed by Gabriel Smilkstein in 1978 at the University of Washington, it evaluates the following dimensions: adaptation, participation, resources, affectivity and capacity, it consists of 5 items structured on a summative Likert scale where 0 "never", 1 "almost never", 2 "sometimes", 3 "almost always" and 4 "always". The range is from 0 to 20, the higher the score mean greater the family functionality [1].

2.4. Place and Application of the Instrument

The questionnaire was carried out to measure the level of family resilience and functionality in adolescents, it was carried out at the educational institutions Andres Avelino Caceres and Honorio Manrique Nicho.

First, the necessary permissions for research are carried out, coordinated with the directors of both educational institutions to carry out the surveys, then, the permission to each classroom teacher explaining about the survey and why the research work is being carried out to they have knowledge about what is going to be done.

Table 1: Resilience level in adolescent students of both Educational Institutions of Lima Province (N=204)

		Resilience Score				
		Count	%	Mean	Minimum	Maximum
Resilience Dimension (grouped)	Low	46	22,5%	52,28	32,00	60,00
	Medium	105	51,5%	70,28	61,00	79,00
	High	53	26,0%	85,75	80,00	95,00
	Total	204	100,0%	70,24	32,00	95,00

Table 2: Family functionality of adolescent students of both Educational Institutions of a Vulnerable Area in Lima Province (N = 204)

		Family Functionality Score				
		Count	%	Mean	Minimum	Maximum
Family Functionality Dimension (grouped)	Family dysfunction	10	4,9%	7,10	2,00	9,00
	Moderate family dysfunction	29	14,2%	11,90	10,00	13,00
	Mild family dysfunction	100	49,0%	15,59	14,00	17,00
	Good family function	65	31,9%	19,15	18,00	20,00
	Total	204	100,0%	15,78	2,00	20,00

The questionnaire was carried out in consecutive shifts during the morning, at the same time as the adolescents from these institutions, the questionnaire was taken to the adolescents with an approximate time of 25 minutes in each section of secondary level (according to the inclusion and exclusion criteria) in the research work, concluding with good satisfaction when collecting the questionnaires as the adolescents supported our research.

It is important to emphasize the presence of the teacher from each surveyed classroom and the medical staff at the time of filling questionnaire of the adolescent, since most of them showed signs of having a health problem, because both educational institutions are in an area where they are exposed to natural or anthropogenic events where the probability of an event that puts adolescents in educational institutions at risk can happen and also the damage can be considerable.

3. Results and Discussions

The following tables from the surveys carried out following the guidelines corresponding to the research work will be shown in table 1.

In Table 1, the level of resilience in adolescents of both Educational Institutions of a Vulnerable Area in Lima Province, it can be seen that of the 204 adolescents surveyed, 105 (51.5%) of adolescents present a medium level of resilience, followed by 53 (26.0%) of the adolescents present a high level of resilience and lastly 46 (22.5%) of the adolescents present a low level of resilience.

In Table 2, the family functionality of the adolescents of both Educational Institutions of a Vulnerable Area in Lima Province, it can be seen that of the 204 adolescents, 100 (49%) of the adolescents present a mild family dysfunction, followed by 65 (31.9 %) adolescents present a good family function, 29 (14.2%) of the adolescents present a moderate family dysfunction and 10 (4.9%) of the adolescents present a family dysfunction.

Table 3 shows that adolescents with good family function have a high level of resilience (49.22%), adolescents with mild family dysfunction have a medium level of resilience (58.0%), adolescents with a moderate family dysfunction has a medium level of resilience (55.2%) and adolescents with family dysfunction have a low level of resilience (60.0%).

To determine the level of resilience and family functionality, the non-parametric Person's chi-square test was applied, which reported an $X^2 = 44.842$ and a p-value = 0.000 per l indicating that there is a significant relationship between these variables.

In Table 4, both cases, most students are located at the medium level of resilience, in the case of students from Honorio Manrique Nicho, 48.5%, and in the case of students from the Andrés Avelino Cáceres school, 54.2%. So, students with a high level of resilience are located at the Honorio Manrique Nicho school with 30.9% compared to students from Andrés Avelino Cáceres who are only 21.5%.

Both studies with the results of the research work can interpret that the levels of resilience and family functionality in adolescents

Table 3: Relationship between Resilience and Family Functionality in adolescent students of both Educational Institutions of a Vulnerable Area in Lima Province (N = 204)

		Resilience Dimension (grouped)							
		Low		Mild		High		Total	
		Count	%	Count	%	Count	%	Count	%
Family Functionality Dimension (grouped)	Family dysfunction	6	60,0%	4	40,0%	0	0,0%	10	100,0%
	Moderate family dysfunction	13	44,8%	16	55,2%	0	0,0%	29	100,0%
	Mild family dysfunction	21	21,0%	58	58,0%	21	21,0%	100	100,0%
	Good family function	6	9,2%	27	41,5%	32	49,2%	65	100,0%
	Total	46	22,5%	105	51,5%	53	26,0%	204	100,0%

Pearson's chi-square tests		
Family Functionality Dimension (grouped)	Chi-square	44,842
	df	6
	Sig.	,000*

Results are based on nonempty rows and columns in each innermost sub table.
*The chi-square statistic is significant at the .05 level.

Table 4: Comparison of Resilience Level of adolescent students between both Educational Institutions of a Vulnerable Area in Lima Province (N = 204)

		Evaluated School					
		Honorio Manrique Nicho		Andrés Avelino Cáceres		Total	
		Count	%	Count	%	Count	%
Resilience Dimension (grouped)	Low	20	20,6%	26	24,3%	46	22,5%
	Medium	47	48,5%	58	54,2%	105	51,5%
	High	30	30,9%	23	21,5%	53	26,0%
	Total	97	100,0%	107	100,0%	204	100,0%

are linked to the promotion of mental and family health of adolescents, taking into account that resilience is very relevant because it allows adolescents to have the ability to adapt and also in making decisions about the negative situations that arise. It also focuses on family functionality because in the adolescent, it will allow it to maintain an affective interaction with the members of its family, and also to has the ability to make changes in order to overcome the difficulties that arise in the family. It is vitally important to have a high level of resilience and, have good family functionality since the development of adolescents must be observed, since there is little knowledge about the topics which allows us to obtain a review of the scientific information related to the topic.

In the present study, the issue of level of family resilience and functionality is raised from the point of view of promoting adolescent health, which seeks to contribute to educational institutions as part of adolescent training, which it seeks to include programs that allow adolescents to find ways to enhance their adaptation capacities and make their own decisions, both in society and family, so that they can improve their lives in the future.

In the level of resilience in adolescents from educational institutions, the medium level predominated, followed by the high and low. These results can be based on how adolescents live with their parents sharing their own experiences, in which it is a factor that favors the adolescent since it allows them to easily excel and overcome any anticipated problem. In the study by Sulimani Y. [12], on the results of resilience, he referred to the adolescent's distinction between formal and informal social ties and, also, the ability to seek help and guidance from adult figures, as well as strong relationships and supporting parents and family members is a debatable factor as a major component of resilience for adolescents.

In relation to family functionality in adolescents from Cajatambo educational institutions, mild family dysfunction predominated, this is due to the lack of empathy, understanding and sensitivity towards the adolescent by the family is scarce, where scenes such as intra-family conflicts and parental misbehavior are reflected in the adolescent adapting to such actions, allowing the adolescent to think that this normally happens in all other families. Suarez P. and collaborators [2], argue that the development of values and actions of adolescents will be reflected in the role played by the family in their environment, factors such as care, affectivity and fluid communication are criteria that support the adolescent development.

In relation to the level of resilience and family functionality in adolescents from Cajatambo's educational institutions, mild family dysfunction predominated with a medium level of resilience, it is important to emphasize the intervention of the family in the development of resilience in adolescents, since it is mainly the parents who will help build the adolescent's resilience by providing the necessary support that will allow them to overcome certain obstacles. The family will influence positively or negatively in the construction of the adolescent's own resilience, but some adolescents mostly present a high resilience. Tian L. et al. [13], mention that parental support was more strongly related to adolescent resilience, and that interventions should be made to promote adolescent resilience with greater emphasis on parental support in the future of their children.

In comparison the level of resilience between the two Educational Institutions in Lima Province, it can be seen that in the Honorio Manrique Nicho Educational Institution and in the Andres Avelino Caceres Educational Institute, the middle level predominates more, this is because after the family as Resilience builder is also the school, at school the adolescents experiences certain conditions that make them promote resilience, because the school also has the capacity to build it in the adolescent. The way in which resilience can be built is based on the trust that the adolescent has in front of the parents, teachers or friends that generate security to solve any obstacle without doubts.

It is concluded that the particular way in which adolescent education is carried out, whether in the family or at school, should be taken into account; it could be investigated about resilience and family function through other studies to determine this study in greater depth.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] L. Higuita-Gutiérrez, J. Cardona-Arias, "Percepción de funcionalidad familiar en adolescentes escolarizados en instituciones educativas públicas de Medellín (Colombia), 2014," Revista CES Psicología, **9**(2), 167–178, 2016, doi:10.21615/cesp.9.2.11.

[2] P. Suárez, M. Vélez, "El papel de la familia en el desarrollo social del niño: una mirada desde la afectividad, la comunicación familiar y estilos de educación parental," Revista Psicoespacios, **12**(20), 173–198, 2018.

[3] D. Paternina, M. Pereira, "Funcionalidad familiar en escolares con comportamientos de riesgo psicosocial en una institución educativa de Sincelejo (Colombia)," Salud Uninorte, **33**(3), 429–437, 2017, doi:10.14482/sun.33.3.10932.

[4] A. Adebayo, E. Ajuonu, B. Betiku, "Family functionality and parental characteristics as determinants of sexual decision-making of in-school youths in a semi-urban area of Southwest Nigeria," International Journal of Adolescent Medicine and Health, **28**(4), 413–418, 2016, doi:10.1515/ijamh-2015-0029.

[5] H. Bravo, J. López, "Resiliencia familiar: una revisión sobre artículos publicados en español," Revista de Investigación En Psicología, **18**(2), 172, 2015, doi:10.15381/rinvp.v18i2.12089.

[6] M. Simón-Saiz, R. Fuentes-Chacón, M. Garrido-Abejar, M. Serrano-Parra, E. Larrañaga-Rubio, S. Yubero-Jiménez, "Influence of resilience on health-related quality of life in adolescents," Enfermeria Clinica, **28**(5), 283–291, 2018, doi:10.1016/j.enfcli.2018.06.003.

[7] C. Albuquerque, J. Almeida, M. Cunha, A. Madureira, A. Andrade, "Protective Resilience Factors in Institutionalised Portuguese Adolescents," Procedia - Social and Behavioral Sciences, **171**(1), 276–283, 2015, doi:10.1016/j.sbspro.2015.01.122.

[8] L. Yi-Ching, S. Dong-Chul, "Cumulative family risks across income levels predict deterioration of children's general health during childhood and adolescence," PLoS ONE, **12**(5), 1–13, 2017, doi:10.1371/journal.pone.0177531.

[9] K. Hye-Jin, M. Jin-Young, M. Kyoung-Bok, L. Tae-Jin, Y. Seunghyun, "Relationship among family environment, self-control, friendship quality, and adolescents' smartphone addiction in South Korea: Findings from nationwide data," PLOS ONE, **13**(2), 1–13, 2018, doi:10.1371/journal.pone.0190896.

[10] C. Fernández, P. Baptista, Metodología de la Investigación, 6ta edición, Mc Graw Hill Education/Interamericana, México, 2014.

[11] G. Guihard, L. Deumier, B. Alliot-Licht, L. Bouton-Kelly, C. Michaut, F. Quilliot, "Psychometric validation of the French version of the Connor-Davidson Resilience Scale," Encephale, **44**(1), 40–45, 2018, doi:10.1016/j.encep.2017.06.002.

[12] Y. Sulimani-Aidan, "Promoting resilience among adolescents in care from their social workers' perspectives," Children and Youth Services Review, **94**(1), 43–48, 2018, doi:10.1016/j.childyouth.2018.09.005.

[13] L. Tian, L. Liu, N. Shan, "Parent–Child Relationships and Resilience Among Chinese Adolescents: The Mediating Role of Self-Esteem," Frontiers in Psychology, **9**(1), 1–11, 2018, doi:10.3389/fpsyg.2018.01030.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 4, 457-462 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Assessing the Operator's Readiness to Perform Tasks of Controlling by the Unmanned Aerial Platforms

Dmytro Kucherov[1,*], Olha Sushchenko[2], Andrii Kozub[3], Volodymyr Nakonechnyi[4]

[1]*Department of Computerized Control Systems, National Aviation University, Kyiv, 03058, Ukraine*

[2]*Department of Aerospace Control Systems, Institute, National Aviation University, Kyiv, 03058, Ukraine*

[3]*National Space Facilities Control and Test Center, State Space Agency, Kyiv, 01010, Ukraine*

[4]*Department of Cyber Security and Information Protection, Taras Shevchenko National University, Kyiv, 01033, Ukraine*

ARTICLE INFO

Article history:
Received: 09 July, 2020
Accepted: 19 July, 2020
Online: 09 August, 2020

Keywords:
Operator of unmanned aerial platform
Assessment criteria
Quality of operator activity
Method of hierarchical analysis

ABSTRACT

Together with the intensity of development in the field of technology of unmanned platforms and their effective use for solving various tasks of peacetime and war, the requirements for the training of the operator managing the platform also increase. This fully applies to personnel providing the flight of manned means. Nevertheless, there are significant differences in the requirement of operator preparedness for an unmanned platform. This is, first of all, control in conditions of remoteness from an unmanned platform, orientation on the display of the control panel, delays in the passage of information, and possibly a complete loss of communication. In such conditions, the requirements for the reaction time of the operator increase, he must also have the ability to anticipate the development of the situation, be able to work with available equipment for a long time. These and several other criteria determine the general criterion of operator productivity, which is introduced in the work. The productivity criterion is a linear convolution function of particular criteria with some weighting factors, the exact values of which are unknown. A detailed analysis of particular criteria made it possible to establish their inconsistency and heterogeneity. The article proposes an approach that allows us to eliminate the inconsistency of local criteria by separately calculating weight coefficients for each part based on the hierarchy analysis method. The basic properties of the proposed approach are also given; modeling confirms the correctness of the solutions. This approach can be useful in the certification of operators of various fields of activity.

1. Introduction

Modern successes in the field of microelectronics and electromechanics have made it possible to create unmanned aerial vehicles that are effectively used to solve various problems [1]. The control of an unmanned vehicle can be performed manually, automatically and semi-automatically, where the role of the human as the machine operator is very important both in preparing and conducting the flight. The operator's remoteness from the control means creates several difficulties in flight control, primarily related to control delay, inadequate presentation of information about the environment, and the dynamics of the control process itself.

A typical flight control system includes an operator, on-board receiving and control equipment, and a ground-based transmitting station, usually a tablet computer with a software interface, and the aircraft itself. The control of a military unmanned vehicle usually involves two human operators, one of them is responsible for controlling the movement of the vehicle (aircraft operator), and the second ensures the fulfillment of the target task (operator of the target task load). Therefore, the management of an unmanned aerial platform is not a simple task but requires training the operator in skills that will help him cope with the control difficulties in conditions of natural and artificial interference, loss of a communication channel, loss of line of sight, delays and distortions of transmitted information, control of many devices, etc.

[*]Corresponding Author: Dmytro Kucherov, Dept. of Computerized Control Systems, Faculty of Cybersecurity, Computer and Software Engineering, National Aviation University, Kyiv, Ukraine, d_kucherov@ukr.net

Completing the target at the terminal point requires significantly more time to prepare the platform for flight in automatic than in manual or semi-automatic modes. These conditions imply the fulfillment of many requirements for the practical skills of operators that can ensure the implementation of the target in real flight conditions.

Our contribution consists of developing the versatile method's assessment of the productivity of operators based on measurement of their indicators such as reaction time, the accuracy of the operation, stress resistance, the fitness of mind and hands, preparedness, and foresight. The proposed method is suitable not only for assessing the productivity of UAV operators but operators of some other industries, where high requirements imposed on the organization of labor.

The rest paper is organized as follows: a literature review is given in the next section, the control system structure and typical operator actions are presented in Section 3, the problem statement is given in Section 4, the solution method is presented in Section 5, numerical examples and discussion of input data are given in Section 6, and conclusions are discussed in Section 7.

2. Review of Previous Researches

This work is an extended version of the report [1] presented at the 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), where a method for evaluating the operator's activity taking into account the study of his psycho-physical reactions is proposed.

An analytical review of the influence of the human factor on the efficiency of using unmanned aerial vehicles (UAVs) in the current and immediate future is presented in [2]. The study emphasizes the need for special training by the UAV human operator, which, in principle, differs from that to which the pilots are involved. The review also notes the need for additional scientific research regarding the skills, knowledge, and abilities of UAV operators and ways to evaluate them.

In [3], the importance of the influence of a human operator on the efficiency and quality of control of unmanned aerial vehicles is noted. The paper notes that about 70% of all UAV-related accidents are caused by erroneous human activities. Despite the high level of automation, a person is not excluded from the operation of the control system, but rather, on the contrary, takes a leading place in the preparation, conduct, and implementation of tasks, eliminating the risks associated with the death of the pilot. The presence of several operators involves the coordination of their actions and interaction to achieve the ultimate goal of using one or a group of devices.

In [4-8], factors are considered that affect the preparation and quality of exploitation of unmanned platforms by operators. So, the effect of spatial capabilities obtained by the operator through the use of UAVs was evaluated in [4]. A sense of space gives confidence to the operator when acting in whichever situation at any time of the day and for any type of purpose.

In [5], a simulator description is proposed for a UAV operator with wide simulation capabilities, namely: flight control, wireless communications, and ground equipment. However, the article does not provide the parameters and methods for evaluating the operator.

In [6], the UAV operator errors were presented in the form of a chi-square curve of the Pearson distribution, obtained by constructing a histogram of the probability density of measurements of errors in operator activity.

The authors of [7] propose evaluating the capabilities of the UAV operator according to the profile generated by special software. As estimates, metrics are proposed that are based on a combination of fuzzy logic methods and cluster analysis. The proposed techniques allow us to evaluate the operator not only of a single device but when managing a group.

Errors of a technical nature were analyzed in [8]. It indicates the technical problems that the operator encounters, namely, the limited field of view, delays in the transfer of information, poor operator understanding of the situation in which the unmanned vehicle is located, and errors on-board equipment.

The methodical complex for accelerated training of UAV operators is presented in [9], the components of which are control situations with corresponding difficulties and the intensity of their implementation for a certain time. It is assumed that during the preparation of the operator by this complex, the generalized indicator of operator readiness assumes some steady-state value.

In the study the authors of [10] the temporal norms of operator training for the formation of the required psycho-physiological characteristics of personnel for unmanned aviation systems. The study relies on training programs for the air force. An assessment of the quality of preparedness of operators and methods for their certification in the study is not considered.

The usefulness of practical developments aimed at training operators who are ready to work in conditions when the equipment is on the verge of failure is proposed in [11].

In [12], a system for modeling dynamics is given, which allows one to analyze factors affecting the training of the professional skills of coal mine operators. The work explores operator training over a long time. Also, priority is determined, an important place among which is the ability to work in difficult conditions.

For some industries, it is important to maintain operator productivity in conditions when the order of actions changes [13, 14]. In this case, there is a need to redistribute the procedure and change the training schedule, and then it is useful a loss training procedure is proposed. The authors also introduced an indicator of operator experience [13]. The skill indicator in the study is represented by a multi-purpose function that takes into account the skill indicator and working time. In research [14] proposes an additional indicator, the standardization point.

The Italian concern Terna [15] has developed three main programs for training and certification of operators in real-time: the initial program, sliding (continuous) training, and the general program for other interested parties. The training is configured for the specific role of each operator in real-time and covers all its operational activities. Besides, an operating training simulator (OTS) has been developed by the company.

The authors of [16] evaluated the mental workload of operators whose main professional activity is air traffic control and students by comparing continuous exposure close to the infrared range and a mobile touch panel of the same range worn on the operator's head. The study revealed the possibility of using a flexible touch panel to assess the mental workload of operators involved in air traffic control based on the repetition of measurements and the ANOVA technique.

The work [17] presents a methodology for the development of psychomotor skills through the use of teaching means. The results of the study showed a positive influence on the development of psychomotor skills in people of different categories.

The article [18] considers the assessment of human hearing by methods by a system of clinical audiometry. A biotechnological system for assessment and stimulating the human auditory analyzer under the influence of various factors was developed and tested, which allowed us to study the effect of sound frequencies on the patient.

The method of training a group of agents is presented in [19]. Training is based on the introduction of a Q-function, the result of which depends on the joint actions of a pair of agents. Unfortunately, the result is characterized by the limited elements involved in the learning process, and simple scenarios are considered.

In [20], [21], the UAV control system is studied, which can become the basis for the construction of a simulator designed for training, preparing, and certification of operators.

Based on the papers review, we can conclude that there is strong usefulness in developing and implement appropriate methods that will make the operator certification process more accurate.

3. Features of the Controlling of Unmanned Platform

The section briefly discusses the generalized structural composition of a typical unmanned platform, the main problems of flight control, and the requirements for operator actions.

3.1. System structure

The structure of an automated UAV control system with a human operator in the control loop has the form shown in Figure 1. The operator's workplace is a remote control that provides several controls that are available on the display screen.

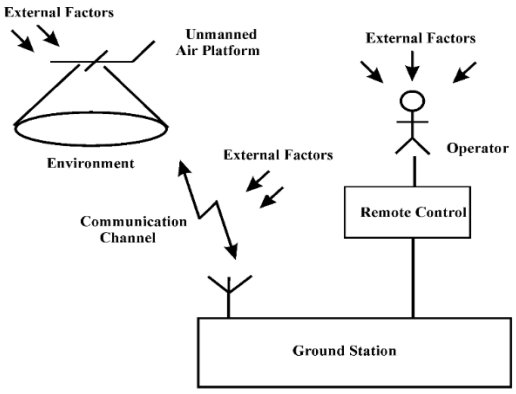


Figure 1: The structure of an automated control system of UAV

3.2. Flight Control Problems

A vulnerable place to control the UAV platform is the flight control channel. The problem here is the limited bandwidth of the radio control channel since the narrow-band path is considered susceptible to interfering signals that occur in the radio channel. In-flight, communication losses are also possible; in places with strong reflections, erroneous or false signals and radio control commands can be received. Besides, the complexity of radio command control lies in the low data rate.

3.3. Operator Actions

The operator controlling the flight of the UAV must know the modes and parameters of the control channel that ensure a stable flight of the platform; be able to make adjustments to flight control; know the conditions for switching manual and semi-automatic modes to autopilot; be prepared to exit autopilot with appropriate authorization.

The operator responsible for carrying out the useful task should be well aware of the capabilities of the equipment used, be able to control existing cameras, be able to extract useful information from the received data, and also closely cooperate with the flight control operator.

4. Problem Statement

Long-term work with the console puts forward some specific requirements for the operator. The situation displayed on the remote control screen is similar to a computer game, and therefore the operator's actions are similar to the reactions of an experienced gamer. The main ones among them are the reaction to the event, its accuracy, perception, and understanding of the displayed situation with a certain delay, fitness of mind and hands, characterized by resistance to long work; stress resistance, determined by the degree of resistance to negative influences and unexpected situations; foresight is the possibility of seeing or feeling probable future events that may occur.

The following performance indicators of the operator are offered:

- the reaction time t is considered the main parameter characterizing the behavior of the operator; it is determined by the time the irritant occurs before the reaction to the event;
- the accuracy of the operation ε, which is the degree to which the data read by the operator corresponds to the values read from the screen;
- exposure τ is the time the operator understands the displayed situation;
- fitness of mind and hands ν, characterized by the frequency of errors that occur in the work of the operator;
- stress resistance k shows the degree of resistance to negative impacts of pressure factors;
- preparedness Σ is a set of knowledge, skills, and conditions of mental and physiological functions that establish its ability to perform specified actions with a certain quality;
- foresight λ is the ability to see or feel the possible future that will take place soon.

The analysis of the selected indicators shows their heterogeneity. There are time indicators estimated by absolute values, this is the reaction time t and the exposure time τ, measured in seconds. The frequency of errors ν, we will measure in relative units, while $\nu \le \nu_{max} \le 1$, the value ν_{max} corresponds to an unacceptable level of errors. In the interest of generality, we write $\nu_{min} \le \nu \le \nu_{max}$, where the desired value is $\nu_{min} \approx 0$. Stress tolerance k, preparedness Σ, and foresight λ should be evaluated in scores. The accuracy of the operation ε when using, for example, a stylus or other manipulator, we will evaluate the distance between the true value and the exact coordinates, in the scale of the display screen this value is measured in millimeters.

Assuming that the real indicators of the operator can be in some intervals allowed by various standards, we will suggest that the measured parameter is in the intervals

$$\underline{\xi} \le \xi_i \le \overline{\xi}, \tag{1}$$

where $\underline{\xi}$, $\overline{\xi}$ are the minimum and maximum values of the measured parameters, and i=1..N, N is the number of indicators.

The heterogeneity of the given parameters is also manifested in the inconsistency of the values of the indicators. For example, the values of reaction time t, accuracy ε, the fitness of mind and hands ν should be small, while exposure τ, stress resistance k, preparedness Σ, and foresight λ for a good operator should correspond to higher values in their ranges.

It is proposed to evaluate the generalized indicator of operator suitability by the relation

$$J(\xi_i \mid i = 1...N) \ge J_{\min}, \tag{2}$$

where J_{min} is a certain threshold value of the indicator by which a decision is made on the operator's admission to the activity. We will call the J as the performance of the operator.

The task is to find a method for determining the value J_{min} of expression (2), in the presence of measured parameters ξ_i satisfying inequality (1), which will allow us to establish inequality (2), which will give an objective assessment of the operator's readiness to carry out the task of controlling an unmanned platform taking into account the influence heterogeneities ξ_i.

5. Technique of Solving

The section discusses the approach to calculating the criterion and the methodology for calculating operator productivity.

5.1. Operator Assessment Criteria

The operator's productivity estimation based on the introduced set of indicators will be calculated in the form of a linear function that written as follow

$$J = \sum_{i=1}^{N} a_i \sigma_i, \tag{3}$$

where a_i is the ith weight coefficient, σ_i is the ith normalized parameter, i.e.

$$\sigma_i = \frac{\overline{\xi} - \xi_i}{\overline{\xi} - \underline{\xi}}, \tag{4}$$

where ξ_i is the measured value of the parameter, i = 1, ..., N. The restriction is imposed on the weighting coefficients

$$\sum_{i=1}^{N} a_i = 1. \tag{5}$$

Now the task of assessing the quality of operator training is reduced to determining the value of J_{min} with unknown weights a_i.

Since the introduced indicators have different effects on the value of J_{min}, and for the most part are contradictory, for example, indicators such as reaction time t, accuracy (error) ε and fitness ν must be small and the others should be as large as possible, it makes sense to separate these factors into two opposite groups, and solve the maximization problem taking into consideration the equal contribution of each component, i.e. define a threshold value as

$$J_{\min} = \max_{a_i}\left(\frac{1}{m_1}\sum_{i=1}^{m_1} a_i \sigma_i^{+} + \frac{1}{m_2}\sum_{i=m_1+1}^{m_2} a_i \sigma_i^{-} \right), \tag{6}$$

where $m_1 + m_2 = N$, in the case the considered parameters N=7.

Expressions (3), (4), (6), and taking into account (1) - (5) as well, provide finding the threshold J_{min}.

5.2. Calculation Approach

The computational procedure involves this sequence of steps

1. Measured are the values of the parameter vector ξ satisfying inequality (1).
2. We calculate the values of the coefficients a_i by the method of hierarchical analysis, including the preparation of the matrix A of pairwise comparisons of size $p \times p$, where p is the number of criteria. We set the diagonal elements of the matrix A equal to 1, calculate the values of the elements a_{ij}, taking into account the importance of the measured parameter so that the parameter takes a value in the range, for example, from 1 to 9. In case, when we deal with the most important parameter it takes the value 9, otherwise the least important will be 1. The element $a_{ji} = 1 / a_{ij}$.
3. Next, the fulfillment of the consistency condition for the matrix of pairwise comparisons is checked by calculating the consistency coefficient CR

$$CC = \frac{p(p_{\max} - p)}{1.98(p-1)(p-2)}. \tag{7}$$

If $CC \le 0.1$, the level of consistency of matrix A is considered acceptable; otherwise, matrix A is checked.

4. Determine the value of J_{min} (6).

5. We measure the parameters of the vector ξ for a particular human operator, apply (2), (1) and (4).

Corollary 1. The quantity J satisfies the condition $0 < J \leq J_{max}$.

Corollary 2. The value $J_{max} = 1$.

6. Simulation

An important point in modeling is the selection of suitable parameters and the evaluation of the results. These issues are addressed in this section.

6.1. Some information about parameter values

The reaction time of the operator is determined by the delay times of the signal t_i in the ith control device and the delay of the signal from the moment the operator recognizes the situation to the corresponding reaction t_{del}, or

$$t = \sum_{i=1}^{n} t_i + t_{del}, \quad (8)$$

where n is the total number of controls. The delay entered by the operator far exceeds the control delay of the computer. This value is in the range of 100 ... 500 ms, which corresponds to the reaction of the gamer. The best indicator is the smaller value.

The exposure time τ of the operator's eye is also determined by two factors: the ability to notice a new luminous object and the time of awareness of this object. A trained eye is capable of detecting flashes of light lasting 5-10 ms, however, this parameter depends on the surrounding light. American scientist B. Libet found that conscious human activity manifests itself after 300-500 ms of the fact that he decided to move his hand [22].

The accuracy ε of the operator's actions is determined by the thickness of the user device and the size of the spot on the screen, which the operator must reliably overlap by the device used. The operator's action will be performed in case of a 50% capture of the screen spot with the stylus tip. For example, if the diameter of the stylus tip is 1.9 - 2 mm, then approximately 1 mm is the accuracy of the operator [23].

The level of operator errors ν depends on the degree of preparedness, workload, and fatigue due to long work. According to experimental data obtained during the processing of visual signals [24], the following results were obtained: for beginners who do not have the experience, the error rate is 0.27, for operators who have completed 4 or more training, this indicator is 0.018. Since there are, as a rule, no untrained operators, the maximum value of the indicator can be taken as 0.15.

Indicators evaluated on a scale of points, which include stress resistance k, preparedness Σ, foresight λ, can also be assessed on a categorical scale in terms: excellent, good, satisfactory, poor. Having chosen this scale as a basis, it is possible to make such a correspondence. The category "excellent" is estimated at 8-9 scores, "good" - 5-7 scores, "satisfactory" - 3-4 scores, "poor" - 1-2 scores.

6.2. Numerical examples

The initial data accepted as the norm are given in Table I. Following the procedure described in the solution section and formulas (3), (4), (6) for $N = 7$, we found weight coefficients that satisfy the consistency condition for the initial matrices of pairwise comparisons (Table 1).

Table 1: Normal Data

para meter	$\underline{t}$	$\underline{\varepsilon}$	$\underline{\tau}$	$\underline{\nu}$	$\overline{k}$	$\overline{\Sigma}$	$\overline{\lambda}$
value	.25 s	.9 mm	.4 s	.07	3	3	2
σ_i	.625	.5	.5	.6	.25	.25	.25

$$a_{\underline{t}} = 0.1017,\ a_{\underline{\varepsilon}} = 0.1835,\ a_{\underline{\tau}} = 0.2602,$$

$$a_{\underline{\nu}} = 0.4545,\ a_{\overline{k}} = 0.5972,\ a_{\overline{\Sigma}} = 0.2847,\ a_{\overline{\lambda}} = 0.1181$$

Using the hierarchy analysis method [25], weights were determined that satisfy the consistency condition for the initial matrices of pairwise comparisons A. Then, following (6), we obtained the value $J_{min} = 0.2788$. It should be noted that condition (5) is only satisfied with the components of expression (6), enclosed in brackets.

Now we will check the weights for other, obviously worse values of the considered parameters, for example, for the data in Table 2.

Table 2: The Second Set of Data

para meter	t	ε	τ	ν	k	Σ	λ
Value	.6 s	1.5 mm	.55 s	.2	2	2	2
σ_i	-.25	-2.5	-.25	-.38	.88	.88	.88

We obtain the value $J = 0.1536$, which does not satisfy condition (2).

Now consider the best indicators than in Table 2 and evaluated these results, Table 3.

Table 3: The Third Set of Data

para meter	t	ε	τ	ν	k	Σ	λ
Value	.2	.85	.35	.05	8	8	8
σ_i	.75	.75	.75	.76	.125	.125	.125

We obtain the value $J = 0.8783$, which fully satisfies condition (11). The results obtained completely coincide with the proposals put forward, and also confirm the study conducted in [26]. Also, we expect that the developed methodology turns out to be enough effective in practical application with any set of measurement indicators.

7. Conclusions

A study of the expansion of research on unmanned platforms confirms the need for special training of operators for maintenance, operation and the performance of tasks as intended.

The paper proposes an approach based on the entering of a generalized indicator of the operator's productivity, which is a linear function of partial heterogeneous quality criteria with unknown weight coefficients. The described method is a modification of the method for analyzing process hierarchies, based on the identification of indicators of the same type, the effectiveness of which is confirmed by numerical examples. The considered approach can be used in the certification of the operator after passing the training process. Our further research we will plan to be directed to the development and implementation of software for the simulator complex with assessing the training of UAV operators following the developed methodology.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors thank both the authorities of the State Space Agency of Ukraine, Taras Shevchenko National University, and National Aviation University the especially leadership of the Faculty of Cybersecurity, Computer and Software Engineering for their support during the preparation of this paper. The authors also thank the anonymous reviewers, whose comments significantly improved the content of the paper.

References

[1] D. Kucherov, O. Sushchenko, A. Kozub, "Operator Training for Unmanned Aerial Vehicles Control" in 2019 IEEE 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 22-24 Oct., Kiev, Ukraine, 1-4, 2019. https://doi.org/10.1109/APUAVD47061.2019.8943918

[2] J. S. McCarley, C.D. Wickens, "Human Factors Implications of UAVs in the National Airspace", 2019. https://www.researchgate.net/publication/228358350_Human_factors_implications_of_UAVs_in_the_national_airspace, last accessed 2019/03/13.

[3] R. M. Taylor, "Human Automation Integration for Supervisory Control of UAVs" in Virtual Media for Military Applications, RTO-MP-HFM-136, 12-1– 12-10, 2006.

[4] J. Y. C. Chen, "Effects of operator spatial ability on UAV-guided ground navigation" in 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2-5 March 2010, Osaka, Japan, Proceedings, 139 – 140, 2010. https://doi.org/10.1109/HRI.2010.5453227

[5] J. Wu, W. Wang, J. Zhang, B. Wang, "Research of a kind of new UAV training simulator based on equipment simulation" in 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, 12-14 Aug., Harbin, China, Proceedings, 4812 – 4815, 2011. https://doi.org/10.1109/EMEIT.2011.6024116

[6] O. V. Kozhokhina, V. M. Gribov, L. V. Blahaia, "Processing statistical data about UAV operator errors", in IEEE 3rd International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 13-15 October, Kyiv, Ukraine, Proceedings, 124 – 127, 2015. https://doi.org/10.1109/APUAVD.2015.7346578

[7] V. Rodriguez-Fernandez, H.D. Menendez, D. Camacho, "Automatic profile generation for UAV operators using a simulation-based training environment", Lect. Notes Artif. Int., **5** (1), 37 – 46, 2016. https://doi.org/10.1007/s13748-015-0072-y

[8] K. W. Williams, "Human Factors Implications of Unmanned Aircraft Accidents: Flight-Control Problems", Federal Aviation Administration, April, 1 – 6, 2006. DOT/FAA/AM-06/8. https://doi.org/10.1016/S1479-3601(05)07008-6.

[9] Y. A. Gunchenko, S. A. Shvorov, N. D. Rudnichenko, V. D. Boyko, "Methodical complex of accelerated training for operators of unmanned aerial vehicles" in 2016 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), 18-20 Oct., Kiev, Ukraine, 130-133, 2016. https://doi.org/10.1109/MSNMC.2016.7783124

[10] S. M. Zlepko, L. G. Koval, D. Kh. Shtofel, V. O. Homolinskyi, M. I. Palamarchuk," Peculiarities of Psycho-Physiological Selection of Operators for Unmanned Aviation Systems" in 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 25-29 Feb., Lviv-Slavske, Ukraine, Ukraine, 628-631, 2020. https://doi.org/10.1109/TCSET49122.2020.235508

[11] J. Dagle, ""Partial panel" operator training: advanced simulator training to enhance situational awareness in off-normal situations" in 2006 IEEE Power Engineering Society General Meeting, 18-22 June, Montreal, Que., Canada, 1-2, 2006. https://doi.org/10.1109/PES.2006.1709229

[12] D. Hong-bin, G. Tao-li, D. Xiao-li, "A System Dynamics simulation analysis of the Influencing Factors of coalmine special operators' competency by training" in 2nd IEEE International Conference on Information Management and Engineering, 16-18 April, Chengdu, China, 367-370, 2010. https://doi.org/ 10.1109/ICIME.2010.5477580

[13] H. Haraguchi, T. Kaihara, N. Fujii, D. Kokuryo, "A study of designing for the operator training under an order change in cell manufacturing system" in 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 20-23 Sept., Tsukuba, Japan, 802-806, 2016. https://doi.org/10.1109/SICE.2016.7749216

[14] H. Haraguchi, "A Study on Operator Allocation Method Considering the Productivity and the Training Effect in Labor-Intensive Manufacturing System" in 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 15-18 Dec., Macao, Macao, Macao, 1236 – 1239, 2019. https://doi.org/10.1109/IEEM44572.2019.8978613

[15] C. Coluzzi, M. Contu, G. Pecoraro, S. Orlandi, F. Allella, A. Pascucci, S. Moroni, "Real time operators training: Terna experience in implementing the Commission Regulation (EU) 2017/1485" AEIT International Annual Conference (AEIT), 18-20 Sept. 2019, Florence, Italy, Italy, p. 1 – 5, 2019. https://doi.org/10.23919/AEIT.2019.8893413

[16] H. Ayaz, P. A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, B. Onaral, "Optical brain monitoring for operator training and mental workload assessment", NeuroImage **59**, 36–47, 2012. https://doi.org/10.1016/j.neuroimage.2011.06.023

[17] K.G. Selivanova, O.V. Ignashchuk, L.G.Koval et al. "Computer-aided system for interactive psychomotor testing," Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, Proc. SPIE, vol. 10445, 2017. https://doi.org/10.1117/12.2280815.

[18] M.V. Bachynskiy, O.Yu. Azarkhov, D.Kh. Shtofel et al. "System and algorithm for evaluation of human auditory analyzer state," Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, Proc. SPIE, vol. 10445, 2017. https://doi.org/10.1117/12.2280738.

[19] Wiem Zemzem, Ines Hosni, "A New Distributed Reinforcement Learning Approach for Multiagent Cooperation Using Team-mate Modeling and Joint Action Generalization", Adv. Sci. Technol. Eng. Syst. J. **5**(2), 1-12, 2020. https://doi.org/10.25046/aj050201

[20] D. P. Kucherov, A. M. Kozub, "Mathematical model of UAV as agent of multi-agent system" in Dependable Systems, Services, and Technologies (DESSERT-2018), May 24-27, Kyiv, Ukraine, Proceedings, 358 – 362, 2018. https://doi.org/ 10.1109/DESSERT.2018.8409156

[21] D. Kucherov, O. Sushchenko, A. Rasstrygin, S. Zhdanov, A. Kozub, "Synthesis of the switching control law for a quadrotor autopilot", Int. J. of Eng. and Tech. (UAE), **7**(4), 3065-3069, 2018. https://doi.org/10.14419/ijet.v7i4.16368

[22] B. Libet, Mind time: The temporal factor in consciousness, Perspectives in Cognitive Neuroscience, Harvard University Press, 2004.

[23] Dektak stylus capabilities. https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/SurfaceAnalysis/StylusProfilometry/ApplicationNotes/AN526-Dektak_Stylus_Capabilities_How_to_Choose_the_Corre.pdf

[24] A methodological approach to determining the influence of the human factor on the performance of information systems. (in Russian) https://habr.com/ru/company/cognitive/blog/209266/

[25] T. L. Saaty, "Relative Measurement and its Generalization in Decision Making Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors the Analytic Hierarchy/Network Process" Rev. R. Acad. Cien. Serie A. Mat., **102**(2), 251 – 318, 2008. http://www.rac.es/ficheros/doc/00576.PDF

[26] A. A. Avetistova, "Psychological features of players in computer games", Psychology. J of the Higher School of Econ., **8** (4), 35 – 58, 2011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 458-463 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Application of Polynomial Regression Analysis in Evaluating the Techno-Economic Performance of DSPV Transformers

Bonginkosi Allen Thango*, Jacobus Andries Jordaan, Agha Francis Nnachi

Department of Electrical Engineering, Tshwane University of Technology, Emalahleni, 1034, South Africa

ARTICLE INFO

Article history:
Received: 22 December, 2020
Accepted: 08 February, 2021
Online: 17 March, 2021

Keywords:
Transformers
Distributed Solar Photovoltaic
Total Ownership Cost

ABSTRACT

To this extent, the delineation of techno-economic evaluations for transformers becomes more intricate through a lens of Distributed Solar Photovoltaic (DSPV) market in South Africa. Essentially, the transformer price and loss evaluation techniques should be tailored for calculating the Total Ownership Cost (TOC) of transformers facilitating decentralized energy systems. In South Africa, the traditional coal power generation and renewables operate concurrently under liberalized energy markets but have distinct operational requirements and therefore have distinct methods for evaluating their generating states, service loss costs and TOC. As a result, their techno-economic evaluations should be different. In this work, new formulae have been developed to contemplate on a comprehensive technique for calculating the transformer prices and losses necessitated to estimate the cost of service losses and TOC for DSPV transformers. These formulae are based on experimental studies undertaken on a fleet of DSPV transformers ranging from 1.25 to 250MVA. In order to substantiate these new formulae, 4 case studies have been presented. The calculated losses and associated cost results against the pragmatic values from the case studies yield an error of estimation of less than 1% and 2% respectively in all cases. Further, these results are used to calculate the cost of losses and TOC using a methodology that has been proposed in previous work exclusively for power producers who are proprietors of DSPV generation systems.

1. Introduction

Transformers make up some of the fundamental power-consuming equipment in the Distributed Solar Photovoltaic (DSPV) systems. The use of transformers has enabled the renewable energy market to grow steadily, with the transformer market expected to increase annually by approximately 16% between 2019 and 2020 [1]. The increased attention by the government to broaden investments in renewable energy resources to satisfy environmental and sustainable objectives is the main driver of this market growth [2]-[5].

The cost-effective measures derived from the choice of enhanced energy efficiency criterion at the time of installation and commissioning of new transformers or replacement of existing units can culminate in a considerable amount of losses decreasing and some benefits for DSPV systems. Contemporary energy-efficient transformers are designed to operate at minimum total losses viz. no-load loss, load loss and auxiliary loss. Transformers manufactures have modernised the manufacturing procedures and new type of optimised core material with thinner laminations [6], [7] to bestow cost-productive and energy-efficient transformers to utility owners such as the Independent Power Producers (IPP's). Substantially, they reduce power consumption and therefore reduce energy generation and the resultant greenhouse gas (GHG) emissions. Therefore, as renewable energy power systems investment keeps on growing, IPP's are to a larger and larger degree fascinated in installing energy-efficient transformers in their power networks.

The initial purchase price of energy-efficient transformers is relatively high; however, they consume less power during their intended service life in comparison to low-efficiency transformers. The choice regarding buying cheap but high loss transformer or a more expensive, with low loss transformer is mainly a lucrative one. The clarification for picking out one transformer over another should be founded on the initial purchase price plus the operational cost experienced over its intended service lifetime [8], [9], and [10]. A habitual procedure for evaluating the service lifetime cost

*Corresponding Author: Bonginkosi Allen Thango, thangotech@gmail.com

of operation, and consequently the best purchase decision for transformers, is established on a service lifetime loss evaluation procedure that gives the ensuing Total Ownership ($TOC - ZAR$) of Transformers. It is pivotal to appreciate that each power utility's operation conditions, to the extent that the transformers service lifetime loss evaluation methods are concerned, could be not similar. Even within one country, different power utilities may have unique operation objectives and financial goals. For instance, service lifetime loss evaluation procedure for DSPV systems engrosses a thorough knowledge of the generation profile based on the availability of solar irradiation and estimation of the incurred losses as a result of this effect and other factors including possible additional losses due to harmonic currents.

On the other hand, for wind plants, a comprehensive understanding of the factors like wind output speed and generation hours per annum are essential [11] – [12]. In the mainstream, the transformers service lifetime loss evaluation method regarded scheming device where its execution predominantly depended on every renewable energy power system's circumspection. Therefore, there might be significant inconsistencies when describing electrical network cost and load factors employed loss evaluation procedures. These inconsistencies are based on the unique operation objectives and financial goals allotted by each renewable energy power system and the extent of analysis involved. The constituents undertaken in the procedures include the network's load profile, rate of interest and energy tariffs, and other operational expenditures.

This work is an extension of previous work presented in [13]. The aim of the current work is to develop a new formulae for establishing the cost of various transformers and their associated losses to advance the adjudication of competing transformer bid offers, evaluation of techno-economic parameters of transformers for new DSPV projects, defer replacement and strategic planning to supersede existing units in service. The advancements include: a) a new formula for calculating the transformer price per MVA and ZAR/MVA of a wide range of transformers from 1.25 – 250MVA, (b) a new formula for calculating the associated no-load and load losses per MVA and $Watts/MVA$. The proposed loss evaluation procedure by the authors and associated annualised cost of energy, interest rate, service lifetime and the generating and non-generating state factors have been adopted from previous work presented in [13] in order to evaluate the cost of service losses and TOC of the studied transformers. The results yield an error of estimation of less than 1% and 2% for the losses and associated cost; and the TOC respectively.

2. Economic Evaluation of Transformer Service Lifetime

The service lifetime loss evaluation of transformers is a procedure employed by transformer manufactures and IPP's to account for the summation of the present cost of unit of electrical energy in kilowatt (kW) of the transformer over its intended service lifetime. The loss of electrical power in transformers are defined as the no-load loss ($P_{NL} - kW$), the load loss ($P_{LL} - kW$), and the auxiliary loss ($P_{Aux} - kW$). Therefore, when determining these transformer losses, their assessment is established on the basis of the plant's energy cost ($E_C - ZA/kW$) and energy demand ($E_D - ZAR/kW$) constituents. The energy cost ($E_C - ZA/kW$) is the present cost of energy unit of electrical energy in kilowatt (kW) that will be consumed by the transformer over its service lifetime within the renewable power plant [14]. Moreover, the energy demand ($E_D - ZAR/kW$) is the present value of electrical energy in kilowatt (kW) of installing an electrical load to service the energy used by the losses [14]. Insofar, the energy and demand factors of the losses are the predominant factors in the procedure of determining the cost of the renewable power plant and energy necessary to supply the service lifetime losses of transformers.

The energy and demand constituents are properly annualised to give the total loss value (ZAR/kW) which include the sum of equivalent discounted value of unit of electrical energy in kilowatt (kW) of the transformer over its intended service lifetime. The loss evaluation procedure ensue the discounted Total Cost of Losses ($TCL - ZAR/kW$) of the transformer over its intended service lifetime. The TCL of transformers can be expressed as follows in Eq. (1) as an arithmetic sum of the cost of, the cost of no-load loss ($C_{NL} - ZAR$), the cost of load loss ($C_{LL} - ZAR$), and the cost of auxiliary loss ($C_{Aux} - ZAR$).

$$TCL = C_{NL} + C_{LL} + C_{Aux} \tag{1}$$

The rate ($F_{NL}(E_D, E_C) - ZAR/kW$) that represents the cost of unit of electrical energy in kilowatt (kW) of the transformer no-load loss over its intended service lifetime is employed as follows in Eq. (2).

$$C_{NL} = F_{NL}(E_D, E_C) \times P_{NL} \tag{2}$$

The rate ($F_{LL}(E_D, E_C) - ZAR/kW$) the represents the cost of nit of electrical energy in kilowatt (kW) of the transformer load loss over its intended service lifetime expressed is employed as follows in Eq. (3).

$$C_{NL} = F_{LL}(E_D, E_C) \times P_{LL} \tag{3}$$

The rate($F_{Aux}(E_D, E_C) - ZAR/kW$) that represents the cost of unit of electrical energy in kilowatt (kW) of the transformer auxiliary loss over its intended service lifetime is employed as follows in Eq. (4).

$$C_{Aux} = F_{Aux}(E_D, E_C) \times P_{Aux} \tag{4}$$

2.1. Transformer TOC Method

The total ownership Cost (ZAR) of transformers in therefore attained using the initial buying price ($BP - ZAR$) of the transformer and its TCL as expressed in Eq. (5) [15] – [16].

$$TOC = BP + TCL \tag{5}$$

The TOC's value is an economic estimation employed to give the IPP's and investors with the overall operational and maintenance costs of their transformer expenditure. Additionally, the TOC is a planning device that can be useable according to the following conditions.

a) Optimization of the transformer design philosophy

It is widely accepted that the transformer losses in the renewable energy power systems is relatively high [17]-[18], and

the manufactures design philosophies must be imbedded with data such as a complete harmonic spectrum to design the most optimised offers. The effect of such implementation will inescapably increase the transformer price. In any event, the TOC model fortifies the reality that reducing transformer losses with expensive and lower losses transformer would comprehensively indicate the reduction of the transformer operational, maintenance and consequently the ownership cost. The consequence of this approach will have significant energy preservation for the IPP's and put off the necessity for renewable power utility energy tariffs increase.

b) Competing transformers during a bid window

The service lifetime loss evaluation procedure and the resulting TOC allows the IPP's to appraise competing transformer manufacturer's offerings in preparing to the highest degree the excellent choice of buying transformers between competing transformers and so affirm buying of the most economical units. Employing the rates $F_{NL}(E_D, E_C)$, $F_{LL}(E_D, E_C)$ and F_{Aux} that represents the cost of a unit of electrical energy in kilowatt (kW) of the DSPV transformer losses over their intended service lifetime, the financial reward of high initial buying price but low-loss units can be easily compared with lower initial buying price and high-loss units.

c) Estimation of the transformer service lifetime

The service lifetime loss evaluation procedure give knowledge to the IPP's of the estimated and most suitable time to replace an existing transformer with more economical and lower losses. This knowledge takes into consideration the financial variability between load-growth consequences under old and new units.

3. Proposed Method for Evaluating the Transformer Price and Losses

The power rating (MVA) of DSPV transformers and replacement of units in service is undertaken instinctively on the basis of the technical or economical determination. Integrating these two aspects in the decision-making process can be even more rewarding for the IPP. In this work, new formulae for evaluating the transformer price and losses for IPPs are derived. These formulae will enable IPPs to optimally use existing units in service and establish a strategic plan to replace these units. Further, these formulae are critical when selecting a suitable power rating (S) and the most economical service losses when purchasing new transformers. In the South African renewable energy market, there are a number of circumstances concerning the acquisition and replacement of transformers, including:

- Acquisition of a new transformer for the construction of a new solar project,
- Deferred replacement of transformer until end of designed service lifetime, and
- Superseding a unit in service with a larger rating.

In any case, the service lifetime of a transformer is reliant on its rating, loading profile and related annual growth. The option for IPPs to defer the replacement of transformers can only slow down the need for new capital investment, however, with the sacrifice of the service lifetime and higher cost of the service losses of the unit.

The derived formulae in this work are based upon empirical studies conducted on a fleet of transformers intended to be of service to various DSPV systems in South Africa that are ranging from 1.25MVA to 250MVA. The evaluation of the relationship between the transformer power rating against transformer prices, no-load losses and load losses is carried out by applying polynomial regression model.

3.1. Polynomial Regression model

Polynomial regression modelling has been chosen in this work to establish the relationship between various transformer ratings against the transformer price and the guaranteed design no-load and load losses.

$$\hat{Y} = b_0 + b_1 x_1 + b_2 {x_2}^2 + \cdots + b_k {x_k}^k$$

here,

$\hat{Y}$ – Transformer anticipated outcome

b_0 – Y intercept

$b_{1,2,\ldots k}$ – Regression model coefficients

$x_{1,2,\ldots k}$ – Regression model independent variables/predictor

A typical graphical representation of the polynomial regression model is shown in Figure 1.

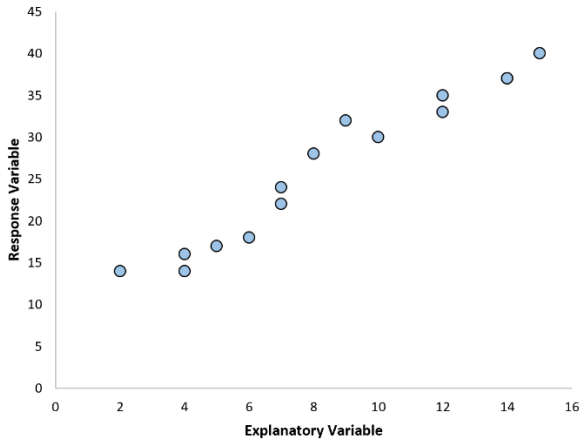


Figure 1: Typical Polynomial regression model

Some general proposition;

- The fitted data model is more credible when built on a fleet of transformer data.
- The transformer data is not extrapolated beyond the boundaries of observed data.
- The independent variables data selected are not too large as they will trigger overflow with higher order polynomials.

3.2. Transformer Price (TP)

The transformer price is dependent on various factors including ratings, materials and technical specification supplied by the customer, materials and manufacturing labour. A polynomial

regression model using a quintic function to evaluate the cost of DSPV transformers is proposed and presented as follows in Eq. (6) below. This model is built upon a fleet of transformer price data that has been analysed using the general proposition presented in the previous subsection.

$$TP = 822069 + 371895S - 10435 \times S^2 + 139.4S^3 - 0.758S^4 + 1{,}39e-3 \times S^5 \quad (6)$$

here,

TP – Transformer price in ZAR

S – Transformer rating in MVA

The rating in the studied transformers considers the highest voltage systems, top oil of 55℃, mean winding temperature of 60℃ and hotspot temperature of 78℃. The relationship between the ratings the TP are demonstrated graphically as shown in Figure 2.

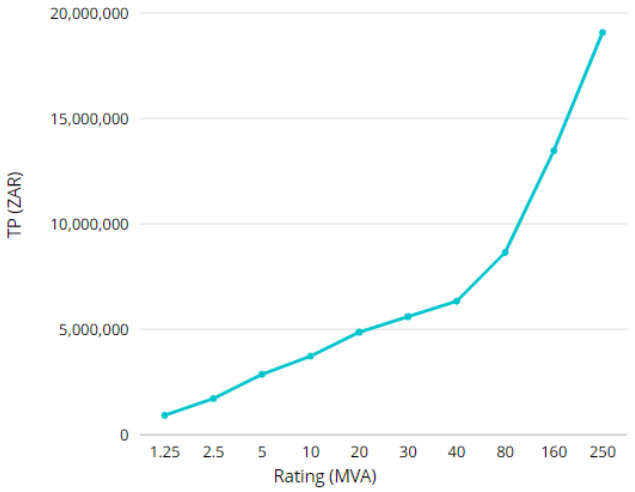


Figure 2: MVA vs ZAR

The percentage contribution of the various components and materials that makes up the transformer price are summarised in Table 1 below. At the present time, the manufacturing of transformers is distinguished by a large range of designs required to meet a variety of customer with distinct technical requirements. Raw material prices will have an effect on these prices, but nonetheless, the percentage contribution will be somewhat as tabulated below.

Table 1: Total transformer manufacturing cost

Item	Percentage
Bushings	3%
Tapchanger	8%
Cooling System	7%
Tank Steel	4%
Windings copper conductor	17%
Cellulose insulation	5%
Liquid insulation	4%
Transformer core	5%
Fittings and accessories	5%
Factory acceptance tests	5%
Manufacturing cost	38%

The increase in the transformer price is observed to be reliant on the rating of the transformer. Another phenomenon is observed with regards to this relationship. In Eq. (7), the ZAR/MVA function is also derived and presented as a quintic function.

$$TP_{\frac{ZAR}{MVA}} = 799349 - 53395S + 1605S^2 - 21.58S^3 + 0.101S^4 - 2e-4S^5 \quad (7)$$

Correspondingly, the graphical representation of this relationship in presented in Figure 3 below.

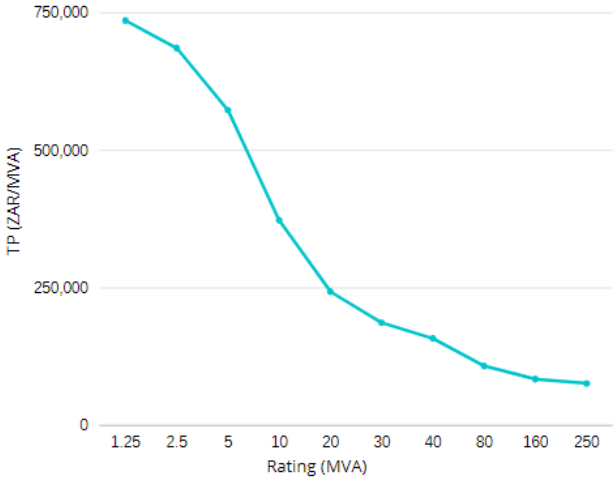


Figure 3: MVA vs ZAR/MVA

It is observed the function ZAR/MVA decreases as the rating of the transformer increase and slows down at about 160MVA and above.

3.3. Transformer Losses

The service no-load and load losses vary with the rating of the transformer. In the technical evaluation of competing bid offers, the service losses are critical parameters for the IPP. The quintic regression models derived for the no-load and load losses are expressed in Eq. (8) and Eq. (9) below.

$$P_{NL} = 2714 + 337.3S + 3.43S^2 - 0.036S^3 + 6{,}09e-6S^4 + 2.91e-7S^5 \quad (8)$$

$$P_{LL} = -2302 + 7837S - 112S^2 + 1.121S^3 - 0.00567S^4 + 1{,}04e-5S^5 \quad (9)$$

The relationship between the service losses and the rating is demonstrated graphically as shown in Figure 4.

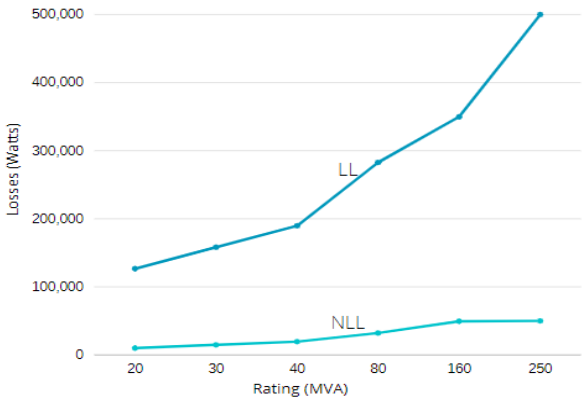


Figure 4: MVA vs $Watts$

Correspondingly, the graphical representation of the relationship between the $Watts/MVA$ and the rating is as demonstrated in Figure 5 below.

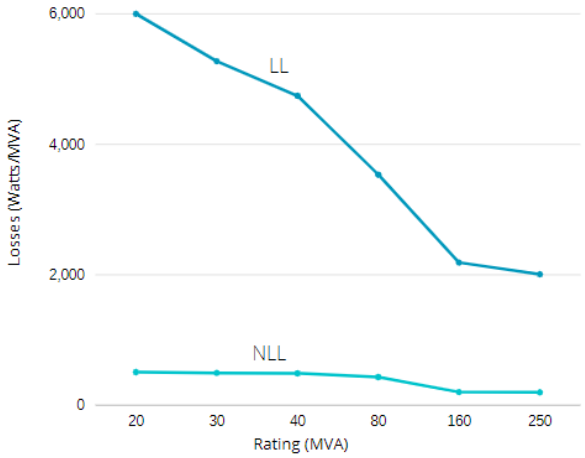


Figure 5: MVA vs $Watts/MVA$

Finally, a set of new transformers intended for a DSPV project are studied in the next section to substantiate the new formulae.

4. Case Scenarios

In order to demonstrate the validity of the proposed method in this work, realistic techno-economic offers of various transformers have been provided subsequent to a request to bid issued by various IPPs within South Africa is presented in Table 2 below. The financial particulars considered in this case scenarios are adopted form our previous work [13].

Table 2: Bid offers

S	TP	P_{NL}	P_{LL}
10	3 729 213	6 437	65 506
20	4 864 152	10 541	117 661
40	6 332 296	19 533	189 746
160	13 468 698	32 230	349 712

In the next sub-section, the assessment of these offers is carried out for the individual sizes.

4.1. Application and Validation of Proposed Method

To facilitate a bona fide comparison of the proposed (prop) no-load and load loss formulae against supplied data in Table 2, Eq. (8) and Eq. (9) are applied to estimate these losses as indicated in Table 3.

Table 3: Benchmarking of no-load and load losses

S	$P_{NL(Act)}$	$P_{NL(Prop)}$	$Error$	$P_{LL(Act)}$	$P_{LL(Prop)}$	$Error$
10	6437	6394	0,67	65506	65916	0,63
20	10541	10546	0,05	117661	117667	0,00
40	19533	19441	0,47	189746	190020	0,14
160	32230	32229	0,00	349712	349714	0,00

The results indicate that the new formulae have an error of estimation of less than 1% in all cases. The proposed loss evaluation procedure by the authors and associated annualised cost of energy, interest rate, service lifetime and the generating and non-generating state factors have been adopted from a previous work [13] in order to evaluate the cost of service losses and TOC of the studied transformers. The TOC of the units based upon the actual values is evaluated and tabulated as shown in Table 4 below. In our previous work [13], we have derived harmonic loss factors of DSPV transformers in which IPPs can apply to evaluate the cost of C_{NL} and C_{LL} as shown below. Additionally, the TOC is evaluated using Eq. (5) as expressed in section sub-section 2.1 above. Hence, Table 4 tabulates the calculated TOC of individual transformer sizes described in Table 2.

Table 4: Assessment of bid offers (Actual)

S	C_{NL}	C_{LL}	TOC
10	222 697	1 389 795	5 341 705
20	364 680	2 496 331	7 725 164
40	675 771	4 024 987	11 033 054
160	1 115 040	7 419 596	22 003 334

On the other hand, the TOC of the units based upon the new formulae is evaluated and tabulated as shown in Table 5 below.

Table 5: Assessment of bid offers (Proposed)

S	C_{NL}	C_{LL}	TOC
10	221 209	1 389 795	5 340 217
20	364 853	2 496 331	7 725 337
40	499 606	4 025 709	10 857 610
160	1 115 006	7 419 638	22 003 342

The error of estimation for Table 4 and Table 5 above is tabulated as shown in Table 6 below.

Table 6: Assessment of bid offers (Comparison)

S	C_{NL}	C_{LL}	TOC
10	0,67%	0,00%	0,03%
20	0,05%	0,00%	0,00%
40	1,37%	0,02%	1,59%
160	0,00%	0,00%	0,00%

The results yield an error of estimation of less than 2% in all of the studied transformer ranges. It is recommendable that IPPs use the new formulae in this work when planning to deploy new solar projects to indicate the cost of the units, losses, cost of operation, and the TOC.

5. Conclusion

It is well documented that the prevailing factors in evaluating the service lifetime cost of DSPV transformers are the power rating and the associated service losses. These components need to be adjusted when major considerations such as rates of exchange for materials – provided that they affect the overall acquisition of materials and manufacturing cost.

More particularly, this work defines new formulae to determine the techno-economic parameters of DPSV transformers within the South African renewable energy market by deriving the contribution of the transformer price and the service losses to the TOC. In previous work, a method to evaluate the TOC was

proposed. This method and associated price indexes are applied in the current work to determine the cost of no-load, load loss and Total Ownership Cost of various transformers. The results yield an error of estimation of 2% against actual values for all the studied costs.

Future work will explore the development of regression models for evaluating the cost of losses and the TOC.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] 6Wresearch, "South Africa Power & Distribution Transformer Market (2020-2026), No. 36005, May 2020".

[2] S. Mamphweli, "South Africa's renewable energy landscape", February 2019.

[3] NDC Partnership," Making renewable energy affordable: The South African Renewables Initiative", ndcpartnership.org/case-study/making-renewable-energy-affordable-south-african-renewables-initiative

[4] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Effects of Current Harmonics on Maximum Loading Capability for Solar Power Plant Transformers," 2020 International SAUPEC/RobMech/PRASA Conference, Cape Town, South Africa,2020,doi:10.1109/SAUPEC/RobMech/PRASA48453.2020.9041101

[5] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Step-Up Transformers for PV Plants: Load Loss Estimation under Harmonic Conditions," 2020 19th International Conference on Harmonics and Quality of Power (ICHQP), Dubai, United Arab Emirates, 2020, doi: 10.1109/ICHQP46026.2020.9177938.

[6] D. Pavlik, D. C. Johnson and R. S. Girgis, "Calculation and reduction of stray and eddy losses in core-form transformers using a highly accurate finite element modelling technique," in IEEE Transactions on Power Delivery, **8**(1), Jan. 1993, doi: 10.1109/61.180342.

[7] X. Yan, X. Yu, M. Shen, D. Xie and B. Bai, "Research on Calculating Eddy-Current Losses in Power Transformer Tank Walls Using Finite-Element Method Combined with Analytical Method," in IEEE Transactions on Magnetics, **52**(3), March 2016, doi: 10.1109/TMAG.2015.2494375.

[8] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Service Life Estimation of Photovoltaic Plant Transformers Under Non-Linear Loads," 2020 IEEE PES/IAS PowerAfrica, Nairobi, Kenya, 2020, doi: 10.1109/PowerAfrica49420.2020.9219912.

[9] E. Müllerová, J. Hrůza, J. Velek, I. Ullman and F. Stříska, "Life cycle management of power transformers: results and discussion of case studies," in IEEE Transactions on Dielectrics and Electrical Insulation, **22**(4), August 2015, doi: 10.1109/TDEI.2015.005025.

[10] Q. Guo-Hua, R. Zheng, S. Lei, Z. Bo, X. Jian-Gang and Z. Xiang-Ling, "A new life cycle cost model of power transformer and its comprehensive sensitivity analysis," 2014 International Conference on Power System Technology, Chengdu, 2014, doi: 10.1109/POWERCON.2014.6993659.

[11] D. Ambach and P. Vetter, "Wind Speed and Power Forecasting - A Review and Incorporating Asymmetric Loss," 2016 Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO), Beer-Sheva, 2016, doi: 10.1109/SMRLO.2016.29.

[12] R. S. Tarade and P. K. Katti, "A comparative analysis for wind speed prediction," 2011 International Conference on Energy, Automation and Signal, Bhubaneswar, Odisha, 2011, doi: 10.1109/ICEAS.2011.6147167.

[13] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Total Ownership Cost Evaluation for Transformers within Solar Power Plants," 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 2020, doi: 10.1109/ENERGYCon48941.2020.9236613.

[14] J. F. Baranowski and P. J. Hopkinson, "An alternative evaluation of distribution transformers to achieve the lowest TOC," in IEEE Transactions on Power Delivery,**7**(2), April 1992, doi: 10.1109/61.127057.

[15] IEEE Std. C57.120-2017, "IEEE Guide for Loss Evaluation of Distribution and Power Transformers and Reactors," in IEEE Std C57.120-2017 (Revision of IEEE Std C57.120-1991), 18 Oct. 2017, doi: 10.1109/IEEESTD.2017.8103991.

[16] H. Wen, H. Zou, C. Ouyang, X. Bao, H. Li and X. Li, "Exploration on energy-saving effect of amorphous transformers extended in Nanning based on theoretical analysis and Total Owning Cost methods (TOC)," 2008 China International Conference on Electricity Distribution, Guangzhou, 2008, doi: 10.1109/CICED.2008.5211678.

[17] B. A. Thango, J. A. Jordaan and A. F. Nnachi, "Contemplation of Harmonic Currents Loading on Large-Scale Photovoltaic Transformers," 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 2020, doi: 10.1109/ENERGYCon48941.2020.9236514.

[18] B. A. Thango, J. A. Jordaan and A. F. nnachi, "A Weighting Factor for Estimating the Winding Eddy Loss in Transformers for High Frequencies," 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunis, Tunisia, 2020, doi: 10.1109/ENERGYCon48941.2020.9236472.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 464-471 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design Approach of an Electric Single-Seat Vehicle with ABS and TCS for Autonomous Driving Based on Q-Learning Algorithm

Jason Valera[1, *], Sebastian Herrera[2]

[1]*Mechanical Engineering Section, Engineering Department, Pontifical Catholic University of Peru, 15088, Peru*

[2]*Electronic Engineering Section, Engineering Department, Pontifical Catholic University of Peru, 15088, Peru*

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 26 February, 2021
Online: 17 March, 2021

Keywords:
Autonomous Vehicle
Steering System
Safety Features
Vehicle Design
Artificial Intelligence
Sustainable Energy

ABSTRACT

Compared to other types of autonomous vehicles, the single-seat is the simplest when designing, since its compact design makes it an option that can simplify different mechanical aspects and enhance those of greater importance such as the steering and the braking system. Likewise, the electronic and electrical design may be a great improvement on the vehicle. It enhances the safety on road by interacting with the mechanical parts of the vehicle and increasing the driver's perspective or reaction in a larger range of scenarios. For an electric vehicle is also important to clarify that, as an internal combustion engine vehicle, it needs to be regulated and have all the necessary equipment to circulate on the streets. Other interesting information is that an electric vehicle can be recharged with electricity and it can come from renewable energy, diminishing its already lower carbon footprint. Therefore, to achieve autonomy over the detection and evasion of objects, the application of intelligent algorithms is dispensable. To achieve the obtained result, a Q-Learning algorithm was applied on the complete 3D model of the vehicle in a simulation environment, which allows finding the best parameters of forward and turning speed. In this way, by reaching a design that meets the requirements and applying the results obtained in the aforementioned algorithm, it allows their interaction in a real environment to be satisfactory.

1. Introduction

This work is an extension based originally on the groundwork presented in the 4th IEEE Sciences and Humanities International Research Conference (SHIRCON-2019) [1]. This paper will focus on the improvements of the mechanical, electronic and electrical design, as well as the implementation of the Q-Learning algorithm while Ref. [1] focuses on the preliminary design of the vehicle, the development of the vision algorithms to be used and the development structure in the simulation environment.

Obstacle detection and avoidance in autonomous vehicles is mandatory for these smart systems to work fully. To the present, there are multiple algorithms of imminent detection and evasion developed by the automotive industry with a really appreciable rate of precision. Thus, in order to get closer to achieving these results, it is essential that the mechanical, electronic and electrical design of the vehicle has specific criteria in the most important considerations that are related to the objective of detection and evasion. That is why, for example, the design of the steering system must provide the closest modeling to reality, in order to obtain the best parameters to include in the simulation environment. On the other hand, for the overall vehicle design to respond according to the results obtained in the simulation, the electronic and electrical design must take into account the appropriate selection of components, as well as their interaction.

Consequently, Figure 1 shows the final design of the vehicle. Despite the changes made, the safety factor has maintained its value of 1.8, this is because when considering the improvements, the distribution of mass which allowed the most critical areas to be subjected to less stress.

For the additional calculus in the present paper, the approximate value of power of the vehicle is 1911 watts. The mechanical efficiency is 96%, effective power of approximately 2000 watts is required. According to the previously obtained value of power, a 2.5 HP PCM motor was selected for the impulse of the vehicle, this motor has a maximum speed of 2800 RPM. Additionally, another motor appointed to control the car's steering

*Corresponding Author: Jason Valera & Email: jason.valera@pucp.pe

system was required. For this case, a 350 watts DC motor was selected.

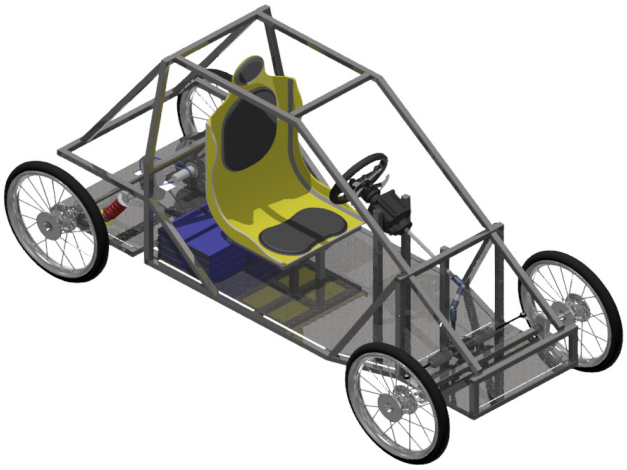


Figure 1: Isometric view of the 3D model of the chassis and components.

With the details described above, they allowed the previous research to obtain a preliminary mechanical, electronic and electrical design based on the considerations of making a low-cost prototype. Likewise, the baselines for the recognition of the environment that will allow the artificial intelligence algorithm to be trained. However, the design in general had not been developed in depth and in detail on the most relevant considerations for the intelligent algorithm, such as the vehicle's turning design and the brake system, as well as its integration into the electronic and electrical design.

This research has the following outline: Section 2 will introduce the improved vehicle's mechanical design. Next, the improved vehicle's electronic and electrical design is presented in Section 3. Then, Section 4 will present the Q-Learning algorithm results. Finally, Section 5 will discuss the conclusion of the findings.

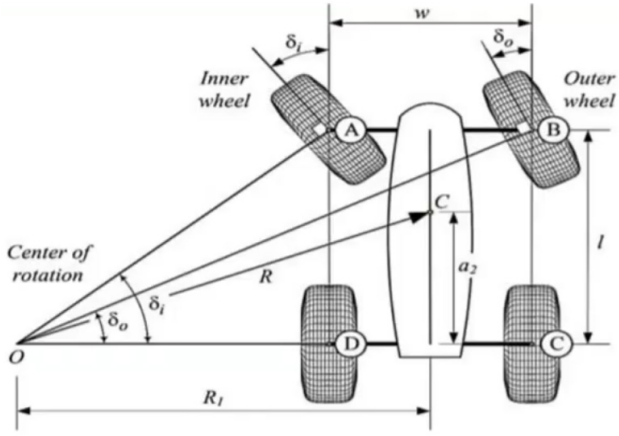


Figure 2: Ackerman Principle [2]

2. Mechanical Design Enhancement

2.1. Steering System

The direction of a vehicle is responsible for orienting some or all of its wheels so that it takes the desired path. It was found that for a better turning behavior, in the case of a four-wheel vehicle with its steering system on the front axle, it is necessary for the inner wheel to turn a greater angle than the outer one, because the inner wheel follows a smaller turning radius than the outer wheel. It was also found that to experience even better turning behavior, the projections of the axes perpendicular to the steering axle of the wheels must intersect at the same point. This behavior is called the Ackerman principle, after Rudolph Ackerman patented it in 1817 for use in horse carts [2]. The Ackerman Principle is shown in the following figure.

To guarantee a close point of rotation of both wheels, which ensures almost zero skid angles at low speeds when the vehicle goes around a curve, the Ackerman configuration shown in by the following expression.

$$cot(\delta i) - cot(\delta o) = w/l \quad (1)$$

Where δo is the steering angle of the outer wheel, δi is the steering angle of the inner wheel, w is the distance between the axes of the wheels, called Track, and l is the distance between the front and rear axles of the vehicle, called Wheelbase, as shown in the figure [2].

In the same way, an expression can be obtained that describes the radius of gyration of the center of mass of the vehicle in steady state. This is calculated by the following expression.

$$R^2 = a2^2 + l^2 cot^2(\delta) \quad (2)$$

Where $a2$ is the distance on the vehicle's axis of travel between the rear axle and the center of mass and δ is the cotangent average of the internal and external angles of the wheels.

Therefore, to determine the closeness of a mechanism to the behavior of the Ackerman configuration, an error function must be determined, in this manner by minimizing its value, get the closest expected result according to the configuration that was found. This function can be expressed as a mean square error (RMS) value. In this case, the error function cannot be explicitly defined and must be evaluated for n values of the mentioned angles. The expression that determines the mean square error value of a set of discrete values of er is defined by the following equation [3].

$$er^2 = \int_{\delta 1}^{\delta 2} (\delta Do - \delta Ao)^2 \, d\delta i \quad (3)$$

Where δDo is the angle of the external wheel of the developed mechanism and δAo is the angle of the external wheel of the Ackerman configuration according to the characteristics of the vehicle for a range of internal wheel angles δi determined.

The drawback in this case is to be able to determine the external wheel angle δo for an internal wheel angle value δi in a given mechanism. Showers and Lee present in their document "Design of the Steering System of an SELU Mini Baja Car" a methodology to determine internal and external wheel angles for a steering mechanism used corresponding to that of a Daewoo Damas model car [4].

Finally, according to the proposed mechanical design, there is a value of w equal to 0.994 meters and a value of l equal to 1.960

meters, as well as a value of $a2$ approximately 0.784 meters, for which when performing the iterations in MATLAB, it was obtained that the radius of gyration R corresponds to a value of 0.803 meters, which is within the established limits.

2.2. Braking System

The principle of a braking system is the reduction of kinetic and potential energy that a vehicle presents when it is in motion. This energy is transformed into heat energy and is dissipated in the brake discs, produced by the friction of the brake pads. Therefore, to select a component that meets the vehicle's requirements, it is necessary to find the maximum torque provided [5]. A figure illustrating this is shown below.

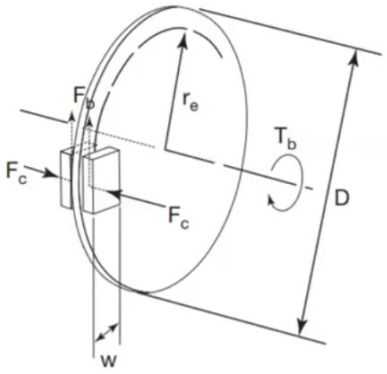


Figure 3: Brake Disc Diagram [5]

where Fc is the clamping force of the brake mechanism and re is the effective radius. The braking torque is calculated by the following expression.

$$T = 2(Fc)(re)(u) \quad (4)$$

According to the selected braking system, there is a value of *Fc* equal to 1000 Newton and a value of *re* equal to 0.09 meters, as well as a value of *u*, coefficient of friction, approximately 0.2, with these values it was obtained that the braking torque T corresponds to a value of 36 Newton-meters. On the other hand, the maximum torque generated by the vehicle's engine is equivalent to a value of around 32 Newton-meters. This value is the result of the chain drive reduction used; the ratio of 15 teeth on the pinion and 72 teeth on the spur gear, delivers 1:4.8, which equates to a maximum speed of approximately 583 RPM. Therefore, this torque value is less than the brake system torque.

3. Electronic and Electrical Design Enhancement

The electronic and electrical enhancement of this electric vehicle (EV) is focused in two main critical points. First, safety features like anti-lock brake system (ABS) and traction control system (TCS) were added. Second, mandatory equipment like head lights, direction indicator lights and others were considered in the design. In addition, electronic controls and better understanding of the connections are reviewed.

3.1. Anti-lock Brake System and Traction Control System

The simplest anti-lock brake and traction control systems are based on the information acquired through speed sensors that are placed in each wheel. Both of them may use this information to control and stabilize the vehicle by assuring that the wheels are spinning as intended.

The ABS prevents the brakes from locking up the wheels. It maintains the steerability and controllability of the vehicle under braking circumstances [6]. The electronic control unit (ECU) receives the information of the speed sensors and performs control techniques in order to regulate the brake pressure applied on the pedal. Figure 4 shows a common ABS control loop of an internal combustion engine (ICE) vehicle.

The TCS prevents the wheels from overspinning. It maintains the stability and traction when accelerating [7]. In this case, the ECU receives the same information as the ABS, but, depending on the design, TCS may have two variants. On one hand, the TCS with brake intervention activates the brakes of the overspinning wheels until they match the speed of the other ones. On the other hand, the TCS with engine intervention reduces the amount of power the overspinning wheels are receiving until they match the speed of the other ones. A comparison is shown in Figure 5.

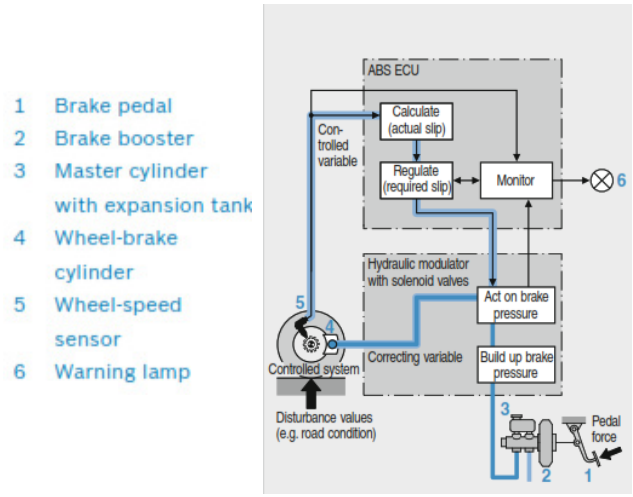


Figure 4: ABS control loop [7]

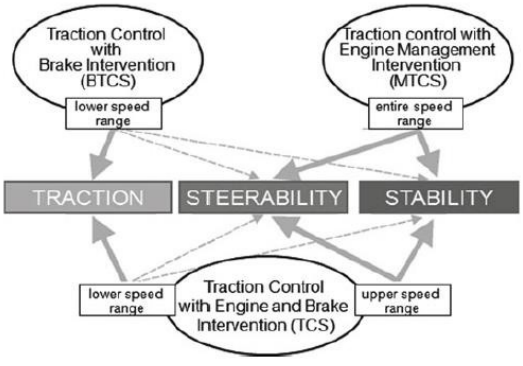


Figure 5: TCS variants comparisons [8]

These systems are tested in different types of environments and scenarios. For example, a vehicle equipped with ABS and TCS must not lose control on certain types of surfaces (e.g. wet surface), must be stable during acceleration or braking and must deal with curves without losing traction. Apart from that, it is important to know that these systems should be calibrated according to the vehicle itself because parameters like the torque of the motor or

the size of the wheels may change its dynamics, thus the control algorithm for these systems.

For this work, ABS and TCS with engine intervention are included in the vehicle's design as safety features. They are meant to be simple, effective and complementary. The objective is to increase the variety of environments and scenarios the vehicle is capable to deal with.

The connection diagram of the EV is presented to have a better understanding of the vehicle's internal distribution (see Figure 6). The main components are the 2000 watts' permanent magnet synchronous motor (PMSM), the motor controller, the 72 V lead-acid battery pack, the gear shift lever, the ignition key, the brake and the throttle. These components are enough to make the EV move forward or backward. In addition, two DC-DC converters are necessary. On one hand, a 72 V to 12 V converter is used to power other components that run on 12 V and later on 5 V for the microcontroller or sensors. On the other hand, a 72 V to 24 V converter is used to power the steering motor.

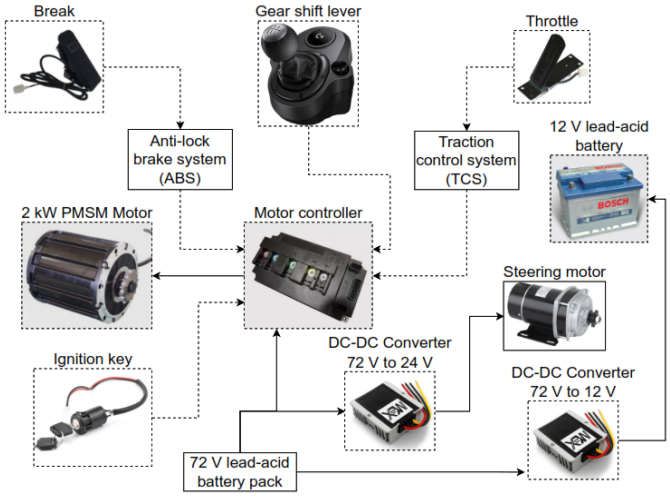


Figure 6: EV connection diagram

The throttle and the brake are different from each other. On one hand, the throttle uses hall sensors to measure its position. It is powered with 5V and delivers a signal from 1 V to 4.8 V into the throttle pin of the motor controller when manipulated. On the other hand, the brake is a simple switch that connects 12 V to the brake pin of the motor controller (electric brake). The ABS and TCS concept for this design relies on activating or deactivating the throttle and the brake when necessary. For this purpose, a window comparator circuit is implemented in Arduino and simulated in Proteus (see Figure 7).

The main components used for the simulation are the Arduino MEGA with ATmega 2560 microcontroller, the IRF9530 MOSFET transistor and the 2N3904 BJT transistor. The input signals are attached to the interrupt pins (18, 19 and 20) of the microcontroller so it doesn't miss any data while sensing. These signals simulate the speed sensors (Hall effect sensors). LED D1 and LED D2 represent the throttle and the brake (ON state means enabled and OFF state means disabled). It is important to notice that the ABS and TCS can be activated or deactivated depending on the user's decision. The user only needs to close the switch to bypass any of those systems.

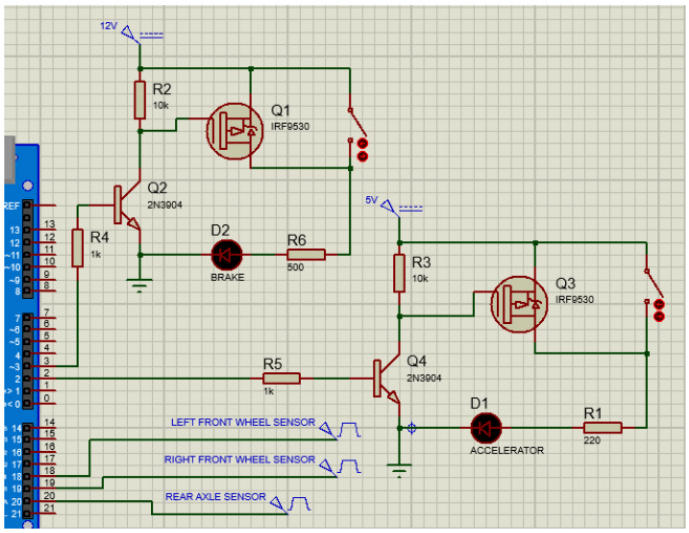


Figure 7: ABS and TCS simulation

The speed sensor used for the design is the A1104 Hall effect sensor (see Figure 8). Two are placed in the front wheels and one is placed on the rear axle. This sensor will commute when detecting the south pole of the magnetic ring (48 north poles and 48 south poles) attached to the rotary axles of each wheel. The time is measured between each pulse and compared with each sensor to determine if the wheels are locked up or overspinning. First, if the wheels lock up while braking, the microcontroller deactivates the brake (ABS) but keeps the throttle activated. Second, if the wheels start overspinning, the throttle is deactivated (TCS) but keeps the brake activated. Third, if the speed of each wheel is almost equal, no control is applied and both remain activated. A difference threshold is considered for the simulation in case the sensors do not match properly. For TCS a threshold of 5 RPM to 10 RPM was estimated and for ABS a threshold of 30 RPM or higher was considered. These values may vary depending on the road conditions and the level of the systems' effectiveness. They must be calibrated based on that information.

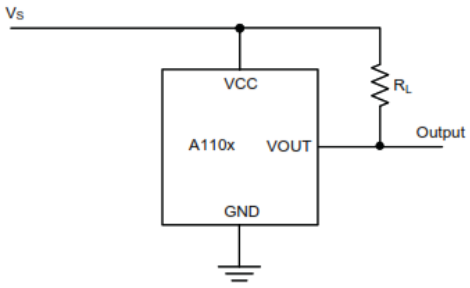


Figure 8: Schematic of the A1104 sensor [9]

The following calculations were made to assure a correct behavior of the device. The A1104 sensor has a slew rate of 2.5 V/us and it works on 5 V. This gives a 2 us rise and fall time for this voltage. It means that the signal could have a period four times the rise time, in other words, a frequency of 125 kHz before it gets distorted. The maximum speed of the motor reaches 5000 RPM (84 Hz) without load and with flux weakening. As mentioned before, there are 48 south poles that will trigger the sensor and gives a maximum frequency of 4032 Hz. This frequency is far

below the 125 kHz and determines that will work properly. The complete connection diagram is showed in Figure 9.

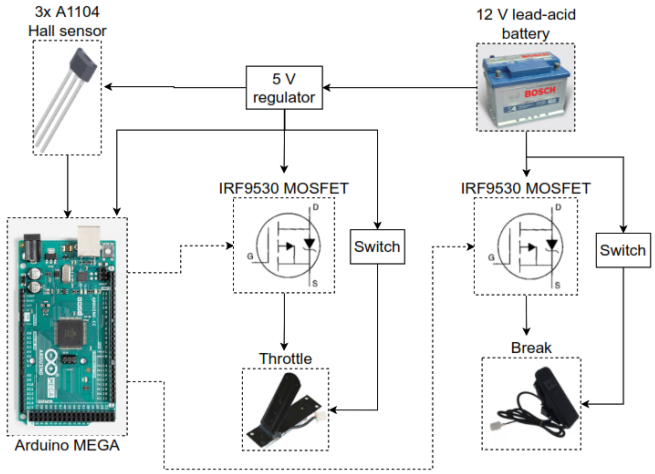


Figure 9: ABS and TCS connection diagram

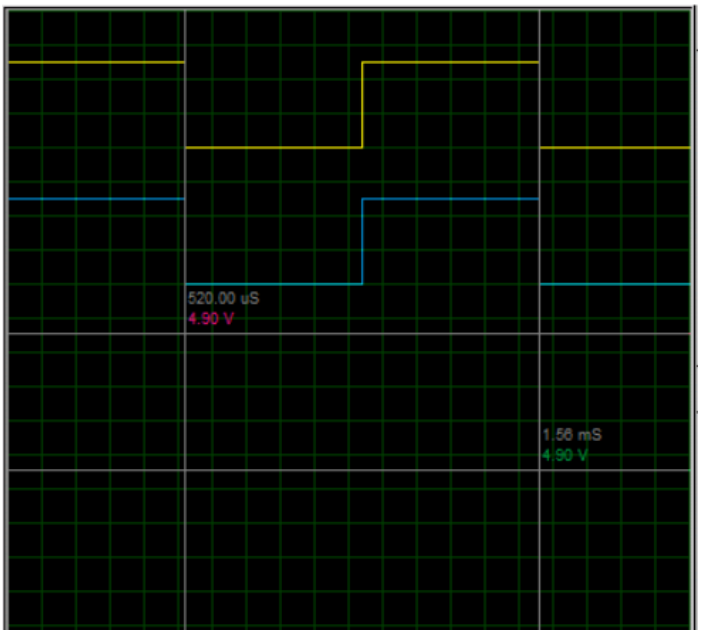


Figure 10: Normal state

The results can be seen in Figure 10, Figure 11 and Figure 12 (Proteus' oscilloscope). The yellow signal represents the left front wheel sensor (only using one as a reference), the blue signal represents the rear axle sensor, the red signal controls the activation of the break and the green signal controls the activation of the throttle. First, Figure 10 shows a normal state in which both sensors match the same frequency. This means that the wheels are spinning at the same speed and no control is applied. The break and the throttle are activated (4.90 V and 4.90 V respectively). Second, Figure 11 shows an overspinning situation in which the rear axle sensor is measuring a higher frequency than the front wheel sensor. This means that the rear wheels are spinning faster than the front wheels. The brake remains activated and the throttle is deactivated (4.90 V and 0 V respectively). Third, Figure 12 shows a lock up situation in which the rear axle sensor is measuring a lower frequency than the front wheel sensor. This means that the rear wheels are spinning slower than the front wheels. The brake is deactivated and the throttle remains activated (0 V and 4.90 V respectively).

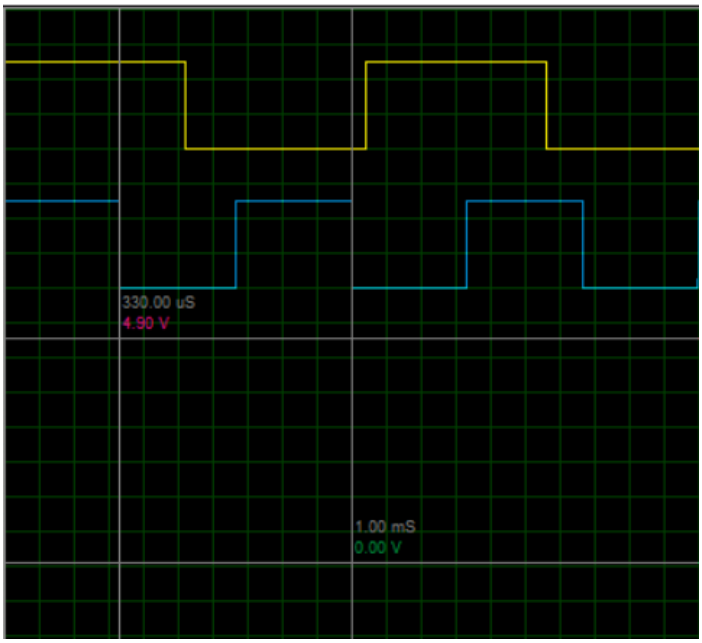


Figure 11: Overspinning situation

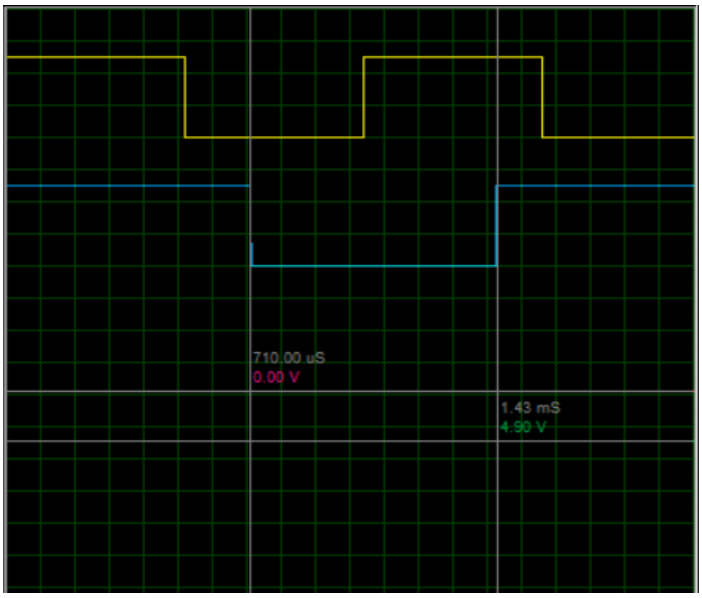


Figure 12: Lock up situation

3.2. Equipment

All vehicles are enforced to follow certain specifications in their design to be admitted to circulate on the streets. Minimum requirements include head, reverse, indicator and brake lights. These lights were included in the design and are graded to be used on almost any vehicle (see Figure 13). White reverse lights are activated with the gear shift lever, red brake lights are activated if the head lights switch is closed or the brake is applied, and amber indicator lights, head lights and the cabin light are activated with a switch.

This vehicle is powered by six 12 V lead-acid batteries connected in series (total of 72 V) and a 12 V lead-acid battery for

the electronics, accessories and other equipment. Each battery has 55 Ah of capacity and the total energy of the vehicle is estimated to be 330 Ah or 4 kWh (including the six pack only). Eventually, the batteries will lose energy when being used and will need to be recharged. In this case, a commercial 72 V charger and a solar charging system is being included in the design (see Figure 14).

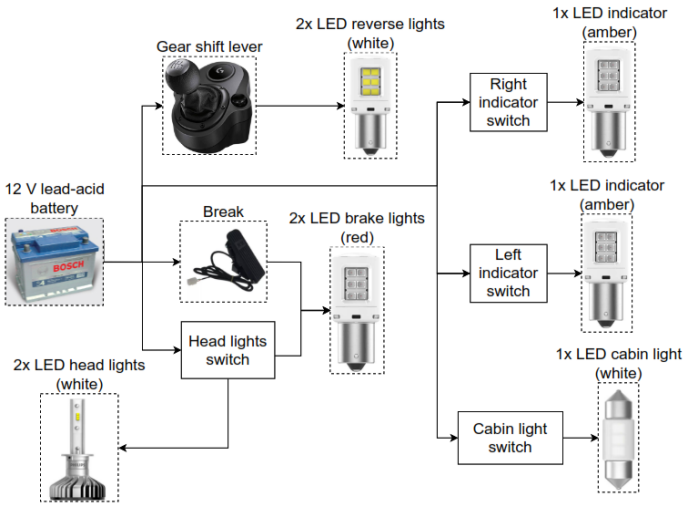


Figure 13: Lights connection diagram

The charger itself is capable of delivering 12 A, so it will take approximately 30 hours to completely charge the vehicle. It is recommended not to discharge these batteries below 40% of their capacity because they could be damaged and decrease their performance.

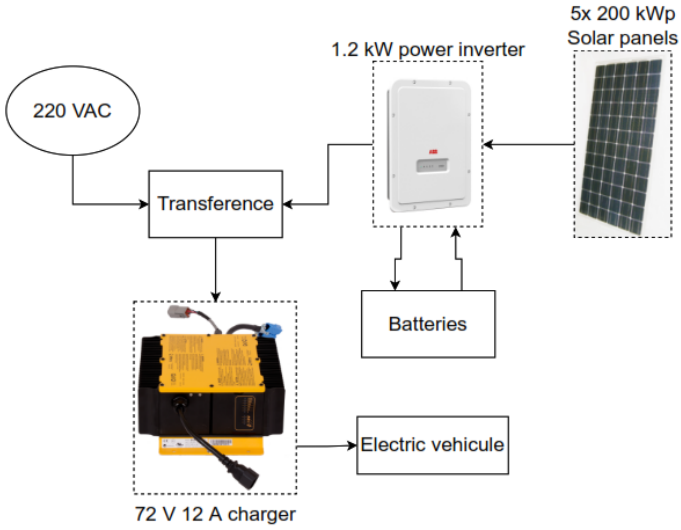


Figure 14: Charging infrastructure

The solar charging system is calculated based on a full charge for the EV assuming that it won't be discharged below 40% (4 kWh*0.6 = 2.4 kWh). All the parameters for this system are obtained with the software Stringsizer from ABB and it will be considered On the grid (Lima, Peru). The peak sun hour (PSH) is determined considering the mean during a year in Lima and is equal to 5.71 kWh/m^2. The performance ratio (PR) is a constant value applied to give a margin in case more energy is required and is equal to 0.8. The kilo Watt-hour (kWh per day) is the amount of energy needed, for this design, it is considered to be 2.4 kWh that was calculated previously. The following expression is used to establish the power needed for the solar panels.

$$kWh = (kWp)(HSP)(K) \quad (5)$$

After making the calculations, the results determine that a 0.52 kWp (kilo Watt peak) solar charging system is required. Considering a 2 m^2 and 200 Wp solar panel, the system requires a minimum of 3 solar panels. To increase the efficiency due to inverters, a group of 5 panels with a total of 1.2 kWp and a 1.2 kW power inverter are chosen. A battery bank with the same vehicle's capacity is recommended but not essential.

4. Applied Intelligence

Reinforcement learning (RL) is defined as a sequential decision-making problem of an agent that has to learn how to perform a task through trial and error interactions with an unknown environment which provides feedback in terms of numerical reward [10,11].

The Markov process is the theoretical basis of reinforcement learning, which can be expressed by the Q-learning algorithm, which is the most classical and commonly used algorithm to solve the MDP problem. It can only be used to solve MDP problems with actions and state spaces both discrete and finite. With an initial Q-table and a predefined policy, in each step, the agent selects an action based on the current state, which according to the reward is fed back and the observation of the next state, the Q table is updated using the algorithm described in the section 4.1, to finally repeat the steps until the Q function converges to an acceptable level [12].

With the Q Learning algorithm, we can allow intelligent agents to operate in environments with discrete action spaces. The discrete action space refers to actions that are well defined, such as movement from left to right, up or down.

For example, in the context of autonomous driving, while the dynamics of the autonomous vehicle is clearly a Markov decision process, the next state depends on the behavior of other elements such as vehicles, pedestrians or cyclists, which are not necessarily an MDP [13].

Therefore, in the case of the design proposed as a starting point for local research in our region, it is convenient to dispense with those elements that are not MDP. Thus, in the context of autonomous driving, given the unpredictable behavior of vehicles or pedestrians, these elements will not be considered for the reinforcement learning algorithm.

4.1. Algorithm

We start by defining the appropriate set of desires for driving based on the desired speed of the vehicle and the angle of rotation that the turn allows. On the other hand, a cost function is needed on the conduction paths that corresponds to its location by means of coordinates in a plane. The cost that is assigned to a trajectory corresponds to the weighted sum of the individual cost assigned to the speed and position. Additionally, weights are assigned to each of the costs described above to obtain a single objective function to control the trajectory. Therefore, the policy consists of a

mapping of the state of desires and a mapping of the requested trajectory. This last mapping is implemented by solving an optimization problem whose cost depends on the set of desires [14].

Algorithm 1: Q-Learning

function: Q-Learning Agent (percept) **returns** an action
Initialization;
inputs: percept, indicating the current state *s'* and reward *r'*
persistent:
Q, a table of action values indexed by state and action, initially zero
N*sa*, a table of frequencies for state-actions pairs, initially zero
s, a, r, the previous state, action and reward, initially null
if Terminal? (s) **then** Q[*s*, none] ← *r'*
  if *s* is not null **then**
    Increment N*sa*[*s*, *a*]
    Q[*s*, *a*] ← Q[*s*, *a*] + *alpha*(N*sa*[*s*, *a*])(*r* + *y* Q[*s'*, *a'*]- Q[*s*, *a*])
    s, *a*, *r* ← *s'*, argmax((Q[*s'*, *a'*]- Q[*s*, *a*]),N*sa*[*s'*, *a'*]), *r'*
return *a*

The Q values correspond to the utility of executing each action *a* in state *s*, that is, the action of advancing or stopping in a state of object detection, as well as the action of evading by turning in a state mentioned above. Also, *alpha* and *y* are the hyper parameters for tuning the algorithm.

4.2. Results

During the development of the algorithm in the GAZEBO simulation environment, it was important that the algorithm be able to calculate the "road parameters", these correspond mainly to the speed of the vehicle, which includes acceleration or deceleration, and the steering angle.

After having tried different rates of learning, no great progress has been made. This is because the optimal local solution could be due to the magnitude gap between the reward value and the punishment, leading to a large number of training epochs. on the way. Finally, the learning procedure converges after approximately 250 episodes. The vehicle can evade fixed obstacles by avoiding sudden changes in acceleration by means of a speed and angle of turn.



Figure 15: The evaluation of the algorithm

The figures below show the implementation of the model in the simulation environment together with objects that can be detected by the LIDAR sensor and the camera. Figure 16 shows how the vehicle with the sensor region and camera field of view, Figure 17 shows its interaction with the environment and Figure 18 and 19 shows the decision and evasion sequence to avoid colliding with an object.

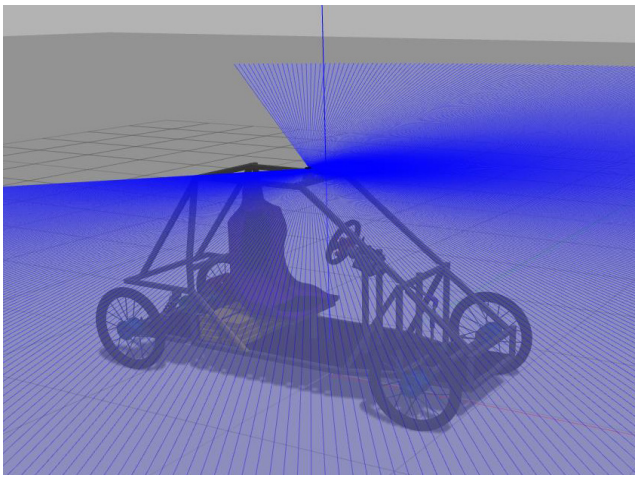


Figure 16: Implementation of the 3D model in the simulated environment

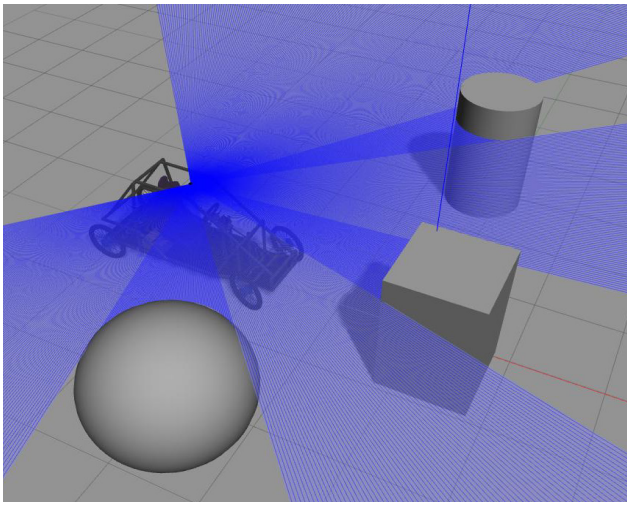


Figure 17: Interaction of vehicle with the environment

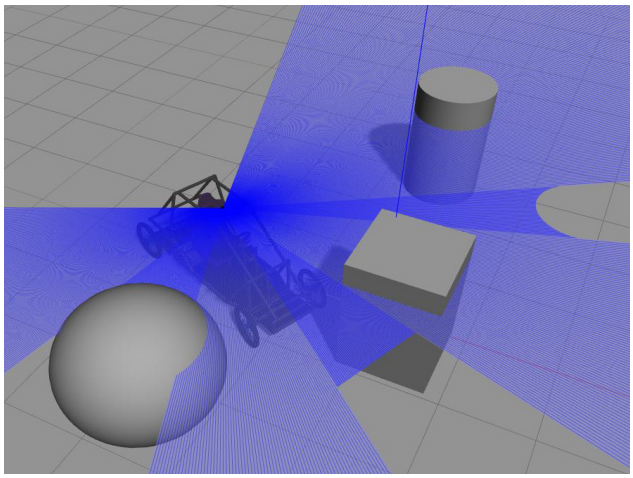


Figure 18: Actions of the vehicle with the environment

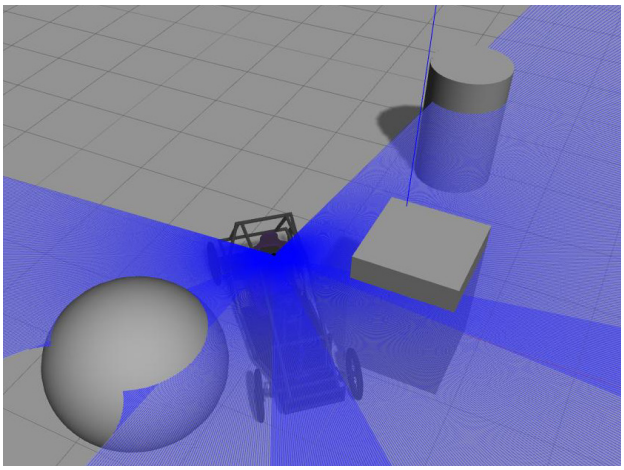


Figure 19: Actions of the vehicle with the environment

5. Conclusions

It was achieved that the mechanical design, taking into account the manufacturing and budget limitations, is based on the methodologies for the correct operation of the system in general such as the steering system. On the other hand, design improvements will be analyzed and developed according to the completion of manufacturing. It is appropriate to mention that, as a consequence of the pandemic that has been affecting the world population, the implementation and manufacture of the vehicle has been stopped to safeguard the health of the members.

The electronic and electrical design enhancement has proved its benefits. First, safety on road has increased due to the addition of the anti-lock brake and traction control systems. The simulations determined the effectiveness of the control algorithm design for this vehicle. This algorithm was aimed to work better on wet surfaces, but it was not simulated for other kinds of terrain, like mud or gravel. It enhanced the driver response and the stability of the vehicle. Second, mandatory equipment has been added to allow the circulation of this vehicle on the streets. This also increased safety while driving among other vehicles. A future insertion to the cities automotive fleet is closer. Finally, an EV has a low carbon footprint in comparisons to the ICE vehicles. To keep this track, a solar charging system demonstrated that it is not necessary to utilize the city power, mostly if this energy is generated by hydrocarbons.

Finally, the classical Q-learning algorithm is unfeasible for more complex control problems, since each additional dimension contributes exponentially to the size of the state space, with it a significant high-dimensional preprocessing. Thus, there are other techniques such as DDPG, which combines the DQN algorithm, Deterministic Policy Gradient (DPG) and Actor-Critic [15]. The basic idea consists of the repetition of experiences to eliminate relevance, since at the moment of storing the experiences of an agent and then randomly extracting batches to train a neural network, a more solid learning for the specific tasks can be guaranteed. For example, the results obtained by Chen, Seff, Kornhauser and Xiao use the algorithm described above as a baseline, and show the penalties it gets and the large amount of processing required to reach an optimal policy (around 80,000 iterations) [16]. However, the proposal of this paper includes simplifications in the mechanical and electrical design that allow reducing the dimensionality of the Q-learning algorithm, which results in a much faster dimensional processing and penalties close to 0 compared to the DDPG algorithm. Therefore, the best driving parameters according to the proposed vehicle design are obtained by the Q-learning algorithm. The subsequent optimization will be the training with controlled situations implemented in the simulation environment, such as street generation and route indication. In this way, once the prototype has finished its manufacture, perform the integration with ROS so that the simulation environment allows the agent to obtain the actions closest to a real situation.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This implementation was funded by IEEE RAS Chapter Initiative Grants and by Faculty of Sciences and Engineering of the Pontifical Catholic University of Peru.

References

[1] J. Valera, L. Huaman, L. Pasapera, E. Prada, L. Soto, L. Agapito, "Design of an autonomous electric single-seat vehicle based on environment recognition algorithms," 2019 IEEE Sciences and Humanities International Research Conference (SHIRCON), 1–4, 2019, doi:10.1109/SHIRCON48091.2019.9024852.

[2] H. Heiser, Vehicle and Engine Technology, Society of Automotive Engineers, 1999.

[3] R.N. Jazar, Vehicle Dynamics: Theory and Application, 2nd ed., Springer-Verlag New York, 2014, doi:10.1007/978-1-4614-8544-5.

[4] A. Showers, H.-H. Lee, "Design of the Steering System of an SELU Mini Baja Car," International Journal of Engineering Research & Technology (IJERT), **2**(10), 2396–2400, 2013.

[5] G. Rill, Road Vehicle Dynamics: Fundamentals and Modeling, 1st ed., CRC Press, 2012.

[6] K. Reif, ed., Brakes, Brake Control and Driver Assistance Systems: Function, Regulation and Components, Springer Vieweg, 2014, doi:10.1007/978-3-658-03978-3.

[7] K. Reif, ed., Fundamentals of Automotive and Engine Technology: Standard Drives, Hybrid Drives, Brakes, Safety Systems, Springer Vieweg, 2014, doi:10.1007/978-3-658-03972-1.

[8] B. Heißing, M. Ersoy, eds., Chassis Handbook: Fundamentals, Driving Dynamics, Components, Mechatronics, Perspectives, Vieweg+Teubner Verlag, 2011, doi:10.1007/978-3-8348-9789-3.

[9] Allergo, Continuous-Time Switch Family A110X, 1–11, 2012.

[10] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Pearson, 2020.

[11] H.J. Vishnukumar, B. Butting, C. Müller, E. Sax, "Machine learning and deep neural network — Artificial intelligence core for lab and real-world test and validation for ADAS and autonomous vehicles: AI for efficient and quality test and validation," 2017 Intelligent Systems Conference (IntelliSys), 714–721, 2017, doi:10.1109/IntelliSys.2017.8324372.

[12] C.J.C.H. Watkins, Learning from Delayed Rewards (PhD. thesis), University of Cambridge, 1989.

[13] L. Wang, N. Zhang, H. Du, "Real-time identification of vehicle motion-modes using neural networks," Mechanical Systems and Signal Processing, **50**–**51**, 632–645, 2015, doi:10.1016/j.ymssp.2014.05.043.

[14] N. van Hoorn, J. Togelius, D. Wierstra, J. Schmidhuber, "Robust player imitation using multiobjective evolution," 2009 IEEE Congress on Evolutionary Computation, 652–659, 2009, doi:10.1109/CEC.2009.4983007.

[15] T.P. Lilicrap, J.J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, "Continuous control with deep reinforcement learning," 4th International Conference on Learning Representations (ICLR), 1–14, 2016.

[16] C. Chen, A. Seff, A. Kornhauser, J. Xiao, "DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving," 2015 IEEE International Conference on Computer Vision (ICCV), 2722–2730, doi:10.1109/ICCV.2015.312.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 5, 469-477 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Interpretation of Machine Learning Models for Medical Diagnosis

Nghia Duong-Trung*,[1], Nga Quynh Thi Tang[2], Xuan Son Ha[1]

[1]*Software Engineering Department, FPT University, Can Tho city, 94000, Vietnam*

[2]*National Taiwan University of Science and Technology, Taipei city, 106335, Taiwan*

ARTICLE INFO

Article history:
Received: 13 August, 2020
Accepted: 18 September, 2020
Online: 24 September, 2020

Keywords:
Interpretable Machine Learning
Medical Diagnosis
Black-box Model Selection
Mobile Application

ABSTRACT

Machine learning has been dramatically advanced over several decades, from theory context to a general business and technology implementation. Especially in healthcare research, it is obvious to perceive the scrutinizing implementation of machine learning to warranty the rewarded benefits in early disease detection and service recommendation. Many practitioners and researchers have eventually recognized no absolute winner approach to all kinds of data. Even when implicit, the learning algorithms rely on learning parameters, hyperparameters tuning to find the best values for these coefficients that optimize a particular evaluation metric. Consequently, machine learning is complicated and should not rely on one single model since the correct diagnosis can be controversial in a particular circumstance. Hence, an effective workflow should effortlessly incorporate a diversity of learning models and select the best candidate for a particular input data. In addressing the mentioned problem, the authors present processes that interpret the most appropriate learning models for each of the different clinical datasets as the foundation of developing and recommending diagnostic procedures. The whole process works as (i) automatic hyperparameters tuning for picking the most appropriate learning approach, and (ii) mobile application is developed to support clinical practices. A high F1-measurement has been achieved up to 1.0. Numerous experiments have been investigated on eight real-world datasets, applying several machine learning models, including a non-parameter approach, parameter model, bagging, and boosting techniques.

1 Introduction

This paper is an extension of work initially presented in IEEE ACOMP 2019 [1] as an invited paper.

The advancement in clinical and healthcare practice today is due to the vast database explosion. Hospitals, health professionals, and treatment centers disclose medical data to the community to call for mutual support and benefit. The data availability is expanding in many forms, both the number of known attributes and the number of new observations [2]. We can see an excellent example from the widely published new coronavirus (COVID-19) data [3]. However, due to privacy concerns, medical data is abundant, but it is also very sparse, which we only have on specific individuals. A country's medical data may not be relevant to develop solutions for the same disease in another country. That creates challenges and difficulties for traditional medical diagnosis because many observations are lost due to a lack of information. The central issue is that health data is characterized by the considerable complexity of detecting new symptoms and sparsely presenting diseases due to insufficient data collection across a population or community. As an added difficulty, medical data exists in many different structures. For example, numbers, categories, text, images, and time-series make medical diagnosis even more difficult. However, from the benefit of big data to health care, leveraging medical data collection offers an excellent opportunity to improve the efficiency of healthcare provider [4, 5]. Current challenges in health practice include information overload, confounding attributes, and noise data in different populations, making manual analysis of experts ultimately difficult [6, 7]. The cost of using medical professionals to examine multiple clinical cases accumulated over time is very high. From a medical diagnosis, one can argue that there is an urgent need to work with a wide variety of data, leverage useful knowledge, and recommend a system that can make the most of information from that data.

Classification is an essential task in health diagnostic systems. We use observed metrics to classify samples into different diseases, or separate degrees of the same illness [8, 9]. Clinicians are trained

*Corresponding author: Nghia Duong-Trung, Software Engineering Department, FPT University, Can Tho city, 94000, Vietnam. Contact No. (+84)0939.657.063. Email: duong-trung@ismll.de

and have the solid knowledge to classify diseases based on sample data accurately. We can agree that correctly distinguish the right disease is one of the fundamental bases for effective treatment of the disease. However, as more data appears in clinical fields, the existing manual diagnostic process that relies on expert expertise cannot be applied quickly and effectively. As a result, the course of treatment may have to be done more slowly or eliminated. Addressing mentioned problems lies in the interactions between healthcare and automation decisions based on machine learning, applying to the classification problem. Thus, one can argue that the classification design is an iterative process between machine learning and available clinical sources to demonstrate the implications of a medical diagnostic procedure. Consequently, the authors turn our attention to proposing processes that automatically select the most suitable machine learning models for each of the different types of data as the foundation of building and recommending diagnostic procedures.

The classic definition of machine learning is "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". It was initially described as a computer program that learns to automatically perform a required task or make a decision from data without explicitly programmed. This definition is comprehensive and can cover almost any set of data-driven approaches. Machine learning has been dramatically advanced over several decades, from theory context to a general business and technology implementation. Especially in healthcare research, it is obvious to perceive the scrutinizing implementation of machine learning to warranty the rewarded benefits in early disease detection, service recommendation, and patient-oriented information offering [10]–[13]. There are two substantially interrelated questions in medical diagnosis: How can computer scientists build machine learning programs that automatically improve through experience, e.g., through data? How can practitioners and clinical experts incorporate with these machine learning programs? The first question can be addressed by marriage between machine learning models and vast medical data. The more data we feed into the learning algorithms, the more accurate the prediction is. Whatever the machine learning models are used, the fundamental expectation is to generalize the knowledge from training data to unseen observations. It is the generalization capacities of learning approaches on how robust they are to their modeling assumptions or the errors in the test set. However, many practitioners and researchers have eventually recognized that there is no absolute winner approach to all kinds of data. The reasons can be broken down into many considerations. The prediction accuracy is diminished when the quantity and quality of clinical data are incomplete.

Different regions expose unique characteristics of a particular disease, which may affect the generalization capacities of learning models [13]. The privacy of medical records also subtracts the high availability of data for research. It is inadvertently possible that racial biases might be built into healthcare systems [14]. The reason might be placed at the characteristics of machine learning models themselves. Even when implicit, the learning approaches generally reply to learn parameters, hyperparameters tuning to find the best values for these coefficients that optimize a particular evaluation metric. Consequently, the use of machine learning is complicated and should not rely on one single model since the correct diagnosis in a particular circumstance can be controversial. Hence, an effective workflow should effortlessly incorporate a diversity of learning models and select the best candidate for a particular input data. It comes to the design idea of black-box models [15] in unintended consequences of machine learning in medicine. The workflow should work with various types of input data and transparently amalgamate diverse models. In that setting, different learning approaches for medical diagnosis are evaluated to select the most accurate one. More importantly, the result's interface should also be attractive to enhance hospital experts and end-users' incorporation.

Today, smartphones are one of the most ubiquitous communication devices and the fastest growing technology industry sectors. An increasing number of mobile applications have been developed to perform a comprehensive spectrum of daily tasks and entertainment. Its impact on medical treatment has already been significant on a global scale. In this paper, the authors deploy an Android mobile application to illustrate how the proposed workflow integrates with hospital experts and patients as end-users. Mobile health is a new concept that describes services supported by mobile communication devices such as smartphones, tablets, smartwatches, patient monitoring devices. However, the discussion of mobile health is out of our research scope. Interesting readers might refer to several mindful papers in the literature [16]–[18].

To be the best of our knowledge, the authors have made several contributions as follows:

- We sharp the connection between the automatic selection of machine learning models for classification in medical diagnosis.
- We extend the pool of machine learning models, including a single approach, bagging algorithm, and advanced boosting technique. We prove that there is no winner model to address medical data sources.
- We extend the background of machine learning algorithms in much more details.
- We extend the experiments section where we carefully describe experimental results, reproducibility, and mobile app development.

The rest of the paper is organized as follows. Section 2 gives a general idea of how our proposed workflow works and how the best learning model is selected. Then in Section 3, the authors introduce required machine learning materials and methodology that need to comprehend the experiments. More specifically, we discuss several well-known machine learning models and evaluation metrics. Our intensive comparison is then discussed in Section 4 in which we go through data collection, experimental results, reproducibility, and mobile app development. Finally, several final thoughts are presented in Section 5.

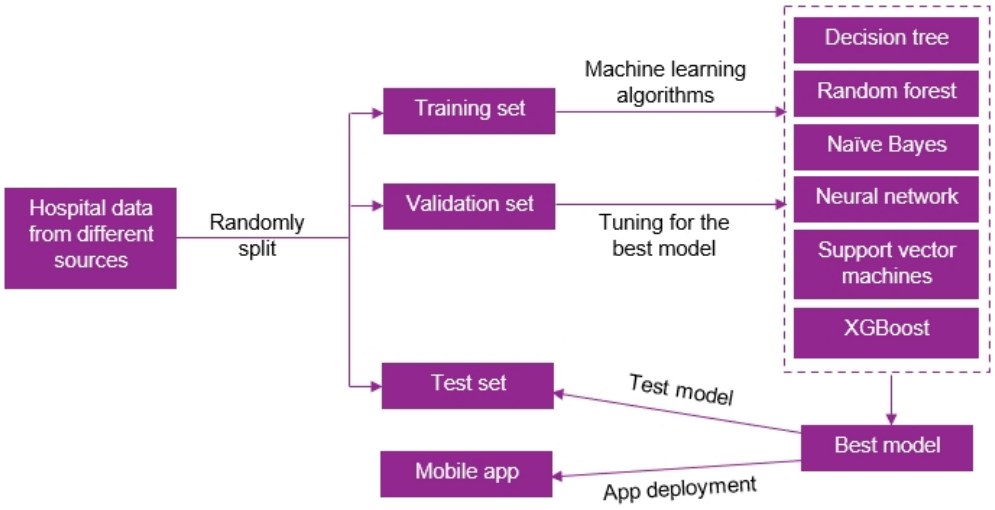


Figure 1: Automatic Selection of Machine Learning Models.

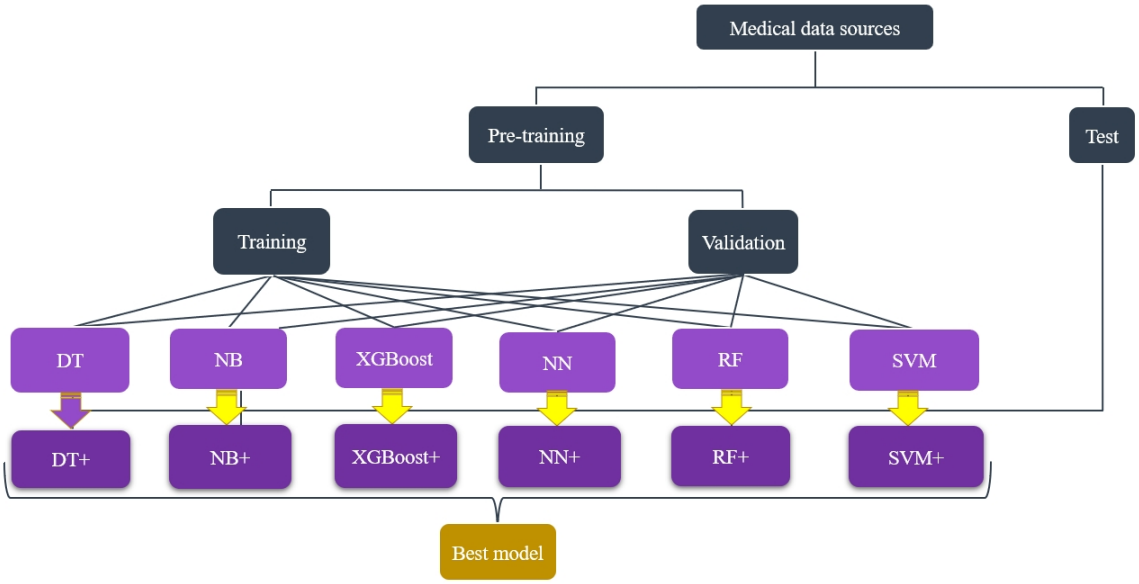


Figure 2: Hyperparameters tuning for picking the most appropriate learning approach.

2 Interpretation of Machine Learning Models

As mentioned in the previous discussion, the authors propose processes that interpret the most appropriate learning models for each of the different clinical datasets as the foundation of developing and recommending diagnostic procedures. The whole process works as (i) hyperparameters tuning for picking the most appropriate learning approach, and (ii) a mobile application is developed to support clinical practices, see Figure 1. Here, clinical experts might not need to understand the technical part because it is done automatically. Experts feed hospital data from different sources to the system. First, data preparation starts with splitting sources into training-validation-test schemes. The training and validation sets are feed into a black box of several predefined learning models. The best model is selected based on an evaluation metric's optimization, performing on the validation set. Finally, the instances that are considered as unseen data belong to the test dataset. Figure (3) illustrates a practical splitting protocol in machine learning. This selection mechanism is described in Figure (2). The authors apply decision trees (DT), naïve Bayes (NB), artificial neural networks (NN), random forest (RF), support vector machines (SVM), and extreme gradient boosting (XGBoost).

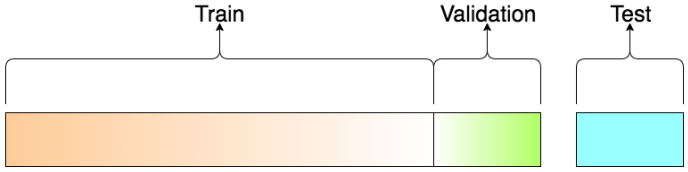


Figure 3: A visualization of the dataset splitting.

3 Materials and Methodology

In this section, the authors briefly present the implemented machine learning approaches. We do not pretend that our discussion will be an overview of models. Interesting readers might refer to several textbooks for further self-learning and comprehension [19]–[22]. Furthermore, other interesting papers that connect the gap between machine learning and medical research can be found at [23]–[26].

3.1 Support Vector Machines

SVM constructs a hyperplane in a high dimensional space. Assuming that the linear model is $v^T x$, then the following constraints need to be satisfied in case of the data points are linearly separable:

$$y_i(v^T x_i + b) \geq 1. \tag{1}$$

SVM seeks to find a good hyperplane where the margin to different classes' data points is maximum. Hence, SVM optimizes:

$$\min \frac{1}{2}\|v\|^2 \tag{2}$$

In case of non-linear separation, the constraints cannot be fulfilled. Then, soft margin is applied. The optimization in Equation 2 becomes

$$\begin{aligned} &\min \frac{1}{2}\|v\|^2 + C\sum_i \sigma_i, \\ &\text{subject to } y_i(v^T x_i + b) \geq 1 - \sigma_i\ , \sigma_i \geq 0. \end{aligned} \tag{3}$$

The Lagrangian is calculated as follows:

$$L = \frac{1}{2}\|v\|^2 - \sum_i \beta_i(y_i(v^T x_i + b) - 1 + \sigma_i). \tag{4}$$

Setting the respective derivatives to 0, then the dual form of the optimization is as follows:

$$\begin{aligned} &\max \sum_i \beta_i - \frac{1}{2}\beta_i\beta_j y_i y_j x_i^T x_j, \\ &\text{subject to } \sum_i \beta_i y_i = 0, \beta_i \geq 0. \end{aligned} \tag{5}$$

where the ranking model $v = \sum_i \alpha_i y_i x_i$ is achieved by solving the quadratic programming. The optimality in Equation 5 is applied by the Karush-Kuhn-Tucker (KTT) conditions. Then we get:

$$\beta_i = 0 \quad \Rightarrow \quad y_i(v^T x_i + b) \geq 1, \tag{6}$$

$$0 < \beta_i < C \quad \Rightarrow \quad y_i(v^T x_i + b) = 1, \tag{7}$$

$$\beta_i = C \quad \Rightarrow \quad y_i(v^T x_i + b) \leq 1. \tag{8}$$

3.2 Naïve Bayes

Due to its easy implementation and high performance, many machine learning practitioners consider Naïve Bayes is a simple but effective machine learning model [27]. We denote a vector $\mathbf{x}$, a set τ and $y = s$ as model's parameters and accompanying label respectively. Then, we can define a generative model $\mathbf{x}$ as follows:

$$P(y = s|\mathbf{x}, \tau) = \frac{P(y = s|\tau)P(\mathbf{x}|y = s, \tau)}{\sum_{s'} P(y = s'|\tau)P(\mathbf{x}|y = s', \tau)}\ , \tag{9}$$

where $P(\mathbf{x}|y = s, \tau)$, $P(y = s|\mathbf{x})$ are class-conditional density and class posterior. $P(y = s)$ is class prior. Equation (9) can be proportionally calculated as follows:

$$P(y = s|\mathbf{x}, \tau) \propto P(y = s|\tau)P(\mathbf{x}|y = s, \tau). \tag{10}$$

Then, the class-conditional density in Equation (9) is estimated as follows:

$$P(\mathbf{x}|y = s, \tau) = \prod_{i=1}^{D} P(x_i|y = s, \tau_{is})\ , \tag{11}$$

which is the Naïve Bayes classifier. Because it is not expected that the features should be independent, Equation (11) can be re-written depending on each feature's type, e.g. binary, real-valued, or categorical attributes. Specifically, in case of binary attributes, the Bernoulli distribution can be utilized as follows:

$$P(\mathbf{x}|y = s, \tau) = \prod_{i=1}^{D} \mathcal{B}(x_i|\mu_{is})\ , \tag{12}$$

where μ_{is} means the probability of attribute i objected to class s. In case of real-valued attributes, the Gaussian distribution can be computed as follows:

$$P(\mathbf{x}|y = s, \tau) = \prod_{i=1}^{D} \mathcal{N}(x_i|\mu_{is}, \sigma_{is}^2)\ , \tag{13}$$

where μ_{is} and σ_{is}^2 mean the probability of attribute i objected to class s and its variance respectively. In case of categorical attributes, multinoulli distribution is used as follows:

$$P(\mathbf{x}|y = s, \tau) = \prod_{i=1}^{D} \mathcal{C}(x_i|\mu_{is}, \sigma_{is})\ , \tag{14}$$

where $x_i \in \{1, \ldots, K\}$ is categorical attributes and σ_{is} is a histogram over K.

In our research, disease classification is basically to classify input vector into different categories. Depending on the representation as a binary or real-valued matrices, Equation (12 or 13) is applied to make prediction.

3.3 Artificial Neural Networks

We apply Multilayer Perceptrons as the version of artificial neural networks applied in the experiments. We form the the model as follows:

$$p(y|x, \theta) = \mathcal{N}(y|\mathbf{w}^T\mathbf{g}(\mathbf{x}), \sigma^2)\ , \tag{15}$$

where $\mathbf{g}(\mathbf{x})$ is called the hidden layer. It is defined as follows:

$$\mathbf{g}(\mathbf{x}) = f(\mathbf{V}x) = \left[f(\mathbf{1}^T\mathbf{x}), \ldots, f(\mathbf{v}_H^T\mathbf{x})\right] , \tag{16}$$

where H is the number of hidden units, f is a logical function. $\mathbf{V}$ is the weight matrix from the inputs to the hidden nodes, while $\mathbf{w}$ is the weight vector from the hidden nodes to the output. In our experiments, we deploy an artificial neural network with 2 hidden layers due to the hardware constraints. A sigmoid function is activated on the output if the classification is binary.

$$p(y|x, \theta) = \mathcal{B}(y|\text{sigmoid}(\mathbf{w}^T\mathbf{g}(x))) . \tag{17}$$

Regarding multi-class classification, sum-to-one constraint is applied:

$$p(y|\mathbf{x}, \theta) = C(y|\mathcal{S}(\mathbf{W}\mathbf{g}(\mathbf{x}))) . \tag{18}$$

3.4 Decision Trees

We denote R as the number of regions, w_r is the weight response in the r region. A decision tree is formed as follows.

$$f(x) = \sum_{r=1}^{R} w_r \phi(\mathbf{x}, \mathbf{v}_r) , \tag{19}$$

where $\mathbf{v}_r$ is the choice of variable to split on.

To find the best partitioning of the input data, the greedy procedure is used in common. There is progress that measures the quality of a split in the classification setting. Given a threshold t, we fit a multinoulli model to the data that satisfies the condition $X_j < t$ by estimating the class-conditional probabilities as follows:

$$\hat{\pi}_c = \frac{1}{|L|} \sum_{i \in L} \mathbb{I}(y_i = c) , \tag{20}$$

where L is the data in the leaf. Then the misclassification rate is calculated as follows:

$$\frac{1}{|L|} \sum_{i \in L} \mathbb{I}(y_i \neq \hat{y}) = 1 - \hat{\pi}_{\hat{y}}. \tag{21}$$

Note that the most probable class is $\hat{y}_c = \text{argmax}_c\, \hat{\pi}_c$. Moreover, the Entropy can be measured as follows:

$$\mathbb{H}(\hat{\pi}) = -\sum_{c=1}^{C} \hat{\pi}_c \log \hat{\pi}_c . \tag{22}$$

We leave the Entropy as the default setting for the decision tree in our experiments.

3.5 Bagging Aggregation with Random Forest

Although DT is one of the most effective and speedy models, it is highly variable due the the splitting. At first, DT is trained on a complete dataset. Then that dataset is split into two portions. DT is applied on the two portions and interestingly, they return different results. The idea of bagging technique helps to reduce the variance in any model [28]. An example of bagging is the generation of many decision trees in parallel. We can train T different trees on different subsets of the data, chosen randomly with replacement, and then compute the ensemble as follows:

$$f(x) = \sum_{t=1}^{T} \frac{1}{T} f_m(x) , \tag{23}$$

where f_t is the t'th tree. This helps reduce the variance of the predictions.

3.6 Boosting Technique with XGBoost

XGBoost, see Figure 4, is a popular and efficient machine learning implementation of the gradient boosted trees [29]. It has been widely applied in some data competitions. The general idea of gradient boosting is to predict a class by combining several weak learners. XGBoost used a regularized objective function (L1 and L2) that combines a convex loss function (emerging from the difference between the ground-truth and prediction) and a penalty term for controlling model complexity. A gradient descent algorithm is used to minimize the loss when adding new learners. The training proceeds iteratively, adding new trees that predict the prior trees' residuals associated with previous trees to perform the final prediction. Regularization is included to reduce overfitting. The authors denote the i-th instance with an associated label as $x_i \in \mathbb{R}^d$, $\hat{y}$ as the prediction given x_i. T is the number of trees. Then we define:

$$\hat{y}_i = \sum_{t=1}^{T} f_t(x_i) , f_t \in \mathcal{F} . \tag{24}$$

Machine learning is basically the procedure of learning parameters $\theta = \{w_j | j = 1, \ldots, d\}$. The objective function is follows:

$$H(\theta) = L(\theta) + \Omega(\theta) , \tag{25}$$

where $L(\theta)$ is the training loss and $\Omega(\theta)$ is the regularization configuration. Optimizing $L(\theta)$ results in high prediction accuracy, while optimizing $\Omega(\theta)$ balances the simplicity of model. For each iteration j, we define the XGBoost objective function by expanding Equation (25).

$$H^{(j)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{j-1} + f_j(x_i)) + \Omega(f_j) . \tag{26}$$

We can see that we cannot optimize Equation (26) by using traditional optimization methods because XGBoost objective is a function of functions. However, we can transform the original objective function to the Euclidean domain by Taylor approximation [30].

$$f(x + \Delta x) \simeq f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2 . \tag{27}$$

The first order gradient statistics is defined as $g_i = \partial_{\hat{y}^{t-1}} l(y_i, \hat{y}^{t-1})$, while the second order gradient statistics of the loss function is defined as $h_i = \partial^2_{\hat{y}^{t-1}} l(y_i, \hat{y}^{t-1})$:. Hence, Equation (26) can be rewritten as follows:

$$H^{(j)} = \sum_{i=1}^{n} [g_i f_j(x_i) + \frac{1}{2} h_i f_j^2(x_i)] + \Omega(f_j) \tag{28}$$

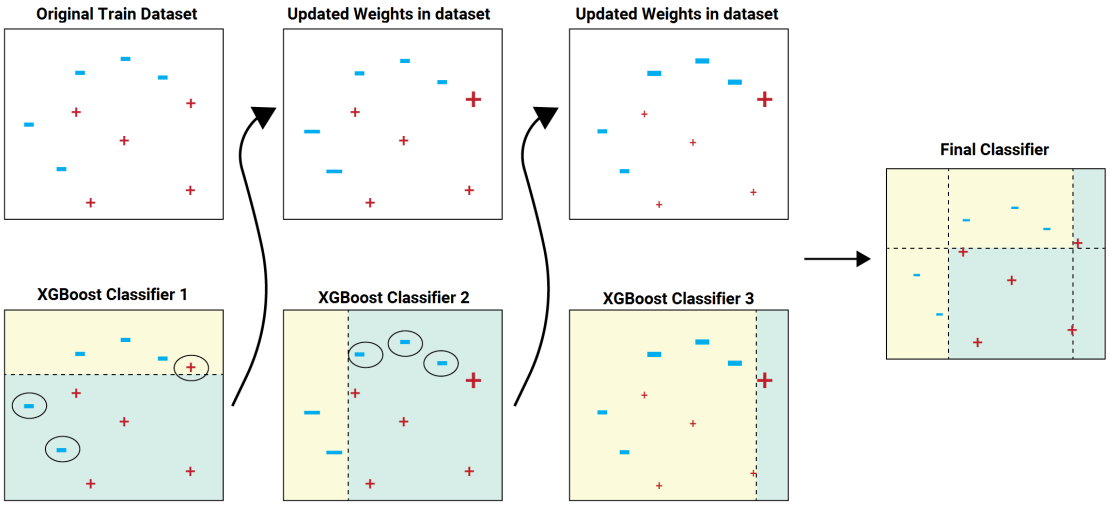


Figure 4: The sequential ensemble methods in XGBoost.

4 Experiments

4.1 Evaluation Metrics: F1-measure

We denote TrueP, FalseP, and FalseN as true positive, false positive, and false negative. With the classification problem where the classes' data sets are very different from each other, there is a logical operation commonly used as Precision-Recall. First of all, consider the problem of binary classification. We also consider one of the two classes to be positive and the other to be negative. With a way of determining a class to be positive, precision (Pre) is defined as the ratio of the number of true positive points to those classified as positive (TrueP + FalseP). The recall (Rec) is defined as the ratio of the number of true positive points to positive (TrueP + FalseN). Mathematically, Precision and Recall are two fractions with equal numerators but different denominators:

$$\text{Pre} = \frac{\text{TrueP}}{\text{TrueP} + \text{FalseP}} \tag{29}$$

$$\text{Rec} = \frac{\text{TrueP}}{\text{TrueP} + \text{FalseN}} \tag{30}$$

We combine Equations 29 and 30 to compute F1-measure as follows:

$$\text{F1-measure} = 2\frac{\text{Pre} \times \text{Rec}}{\text{Pre} + \text{Rec}} \tag{31}$$

4.2 Dataset Collection

In this paper, eight datasets related to medical data are selected from the UCI Machine Learning Repository. More specifically, the experimental datasets are Breast Cancer Data[1], Heart Disease Data[2], Dermatology Data[3], Vertebrae Data[4], Hepatitis Data[5], Cervical cancer (Risk Factors) (RF) Data[6], Autism Adolescent Data[7], and Diabetes Data[8]. The authors prudentially select several categorization datasets that are suitable for the research's scope. Categorical attributes are converted into numerical representation using a one-hot encoding process. Data imbalance is accepted as the native characteristics of classification data sources. The eight experimental datasets are shown in Table (1).

Table 1: Hospital data from different sources

#	Dataset	Prediction Type	# Samples	# Attributes
1	Breast Cancer	Binary	286	9
2	Dermatology	Multi class	365	22
3	Heart Disease	Multi class	303	13
4	Hepatitis	Binary	155	19
5	Vertebrae	Multi class	310	6
6	Cervical Cancer (RF)	Binary	858	32
7	Autism Adolescent	Binary	768	20
8	Diabetes	Binary	104	8

In the experiments, the authors set up the splitting scheme by the following ratios. First, we randomly split without replacement the dataset into 70% and 30% for the pre-training and test parts respectively. Next, the pre-training portion is randomly split without replacement into 80% and 20% for the training and validation

[1] https://archive.ics.uci.edu/ml/datasets/Breast+Cancer
[2] https://archive.ics.uci.edu/ml/datasets/Heart+Disease
[3] https://archive.ics.uci.edu/ml/datasets/Dermatology
[4] https://archive.ics.uci.edu/ml/datasets/Vertebral+Column
[5] https://archive.ics.uci.edu/ml/datasets/Hepatitis
[6] https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+(Risk+Factors)
[7] https://archive.ics.uci.edu/ml/machine-learning-databases/00420/
[8] https://archive.ics.uci.edu/ml/datasets/diabetes

portions respectively. The role of these sets is introduced above, see Figure (2).

4.3 Experimental Results

As mentioned above, the model's performance is judged by its ability to predict unseen data accurately. Hence, the best scores are considered on the test sets. The F1-measure of experimental models performing on datasets are summarized in Table (2). Overall, there is no absolute best approach. However, the random forest model achieves remarkable performance on four different datasets, e.g., Breast Cancer, Cervical Cancer, Autism, and Diabetes. The neural network approach gets the least performance since it does not win any best prediction.

One interesting point to note is that Naïve Bayes algorithm achieves the F1-measure of 1.0 on the Dermatology data. Meanwhile, XGBoost gains the F1-measure of 0.94 on the Diabetes dataset. RF is the best models performing on four different data, namely Breast Cancer, Hepatitis, Cervical Cancer (RF), and Autism Adolescent. Observing all experimental datasets, Vertebrae is the most challenging. Even the best model, e.g., SVM, only achieves the F1-measure of 0.54. Statistics in Table (1) show that Cervical Cancer has the most instances and the most attributes, e.g., 858 and 32, respectively. Even though Cervical Cancer has the most observations, one can argue that only partial representations of properties might not be enough to learn coefficients for 32 attributes. The performance of XGBoost is not stable. While it is the winner performing on Diabetes dataset, the prediction capacity is much more insufficient than other models on other datasets.

Table 2: The report of F1-scores. The best scores are in **bold**.

#	Dataset	Pool of Learning Models					
		DT	RF	XGBoost	NB	NN	SVM
1	Breast Cancer	0.76	**0.79**	0.26	0.75	0.77	0.75
2	Dermatology	0.95	0.96	0.18	**1.00**	0.90	0.91
3	Heart Disease	**0.63**	0.61	0.49	0.58	0.52	0.58
4	Hepatitis	**0.80**	**0.80**	0.21	**0.80**	0.70	0.74
5	Vertebrae	0.50	0.52	0.38	0.53	0.51	**0.54**
6	Cervical Cancer (RF)	**0.94**	**0.94**	0.11	0.93	0.91	0.87
7	Autism Adolescent	0.59	**0.93**	0.15	0.87	0.90	0.87
8	Diabetes	0.61	0.92	**0.94**	0.53	0.63	0.50

4.4 Reproducibility

A grid search strategy on tuning models' hyperparameters is provided for the ease of reproducing the experimental results. The number of estimators and the allowable depth of the trees are investigated regarding tree-based models. We investigate the performance of Gaussian and Bernoulli Naïve Bayes algorithms. Regarding the neural network, the authors examine the effect of the number of units in its hidden layers. We take into account the effect of C and kernel configuration for SVM. Other hyperparameters leave default settings by scikit-learn library [31]. The experimental environment is as follows: windows 10, CPU Intel Core i5-2410M, 2.30GHz, and 8GB of RAM. Table (3) presents our hyperparameters setttings. Hyper-parameters are parameters that are not directly learned within estimators. They are passed as the constructor's arguments of the estimator classes. The best hyperparameters' combination for each machine learning model, and the total search time is presented in Table (4).

Table 3: Hyperparameters space.

Models	Hyperparameters	Settings
DT	Max depth (Max)	$1 \rightarrow 101$
	Min samples leaf (Min)	$1 \rightarrow 101$
RF	Max	$1 \rightarrow 51$
	N estimators (N)	$1 \rightarrow 51$
	Min	$1 \rightarrow 51$
NB	Algorithm (A)	Gaussian (G), Bernoulli (B)
XGBoost	Max	$1 \rightarrow 101$
	gamma (Gm)	$0.1 \rightarrow 0.9$
	# of Parallel trees (Nt)	$1 \rightarrow 51$
	# of jobs (Nj)	$1 \rightarrow 51$
	Sketch Eps (Ep)	$0.1 \rightarrow 0.9$
NN	Hidden layer size (S)	$(1 \rightarrow 101, 1 \rightarrow 101)$
SVM	Kernel (K)	Linear (Li), Poly, RBF
	C	$1 \rightarrow 101$

4.5 Mobile App Development

A mobile application, named Medical Diagnosis, is developed for disease diagnosis in the clinic. The app supports several preliminary features such as disease information and input questions for inspection. Currently, the application runs on the Android mobile operating system and works on a client-server architecture. Machine learning approaches are trained on a regular laptop, which acts as a server. After training the models on a server, the most current classifiers are updated into the client devices. Android Studio 3.2 is used as the IDE to develop the mobile application. The emulation and debugging are done on Genymotion version 5.4.2. The proposed interpretation procedure automatically selects the best model for an investigated input dataset. Then, the models are saved as pickle classifiers[9] [32, 33].

The client is an Android mobile device that is equipped with saved pickle classifiers. A friendly graphical user interface (GUI) provides questions to get input from users on the client-side; either they type in the required information or select from pre-defined options. The main components of the GUI are designed as follows. Textview is used to display messages, comments, and headings. Gridview is used to create a list of two image columns to choose diseases that need an easy diagnosis. Listview creates a list of histological properties for the user to select. The display of symptoms for users to choose on Listview helps users conveniently review the selected symptom. HorizontalScrollView is used to display the properties that the user has selected or entered. Besides, on this screen, the user can copy or delete the selected properties. Toast in Android helps users recognize the entered or selected properties, showing errors connected to the API. The properties are described clearly for the convenience of the user entering and selecting information using Dialog. Each disease's attribute is handled separately with different Dialogs depending on the value of the feature that the user enters or selects. The Dialog can also be used to provide diagnostic

[9] https://docs.python.org/3.6/library/pickle.html

Table 4: The best hyperparameters combination.

Dataset	Learning Models																			
	DT			RF				XGBoost						NB		NN		SVM		
	Ma	Mi	Search time	Ma	N	Mi	Search time	Ma	Gm	Nt	Nj	Ep	Search time	A	Search time	S	Search time	K	C	Search time
Breast Cancer	11	1	35s	1	2	23	3h03m	1	0.1	1	1	0.1	1h06m	G	8s	(8,13)	3h17m	Li	6	1m22s
Dermatology	6	1	50s	4	10	1	3h20m	1	0.1	1	1	0.1	1h35m	B	7s	(49,39)	56m	Li	1	1m47s
Heart Disease	4	5	50s	8	18	2	3h22m	1	0.1	1	1	0.1	2h03m	B	4s	(32,28)	31m	Li	1	2m41s
Hepatitis	1	1	40s	5	11	1	3h07m	1	0.3	20	44	0.2	1h12m	B	2s	(43,72)	45m	Li	1	1m05s
Vertebrae	1	1	20s	2	47	22	3h08m	1	0.3	20	14	0.2	1h33m	G	3s	(7,90)	1h15m	Li	3	3m38s
Cervical Cancer (RF)	5	1	9s	1	1	1	2h05m	1	0.1	1	1	0.1	2h05m	B	8s	(37,33)	50m	Li	2	2m30s
Autism Adolescent	3	28	44s	1	8	1	2h54m	1	0.1	1	1	0.1	0h43m	B	3s	(73,53)	1h19m	Li	1	1m09s
Diabetes	4	48	58s	12	7	3	3h17m	1	0.3	20	5	0.2	1h15m	G	3s	(34,14)	57m	Li	2	2m18s

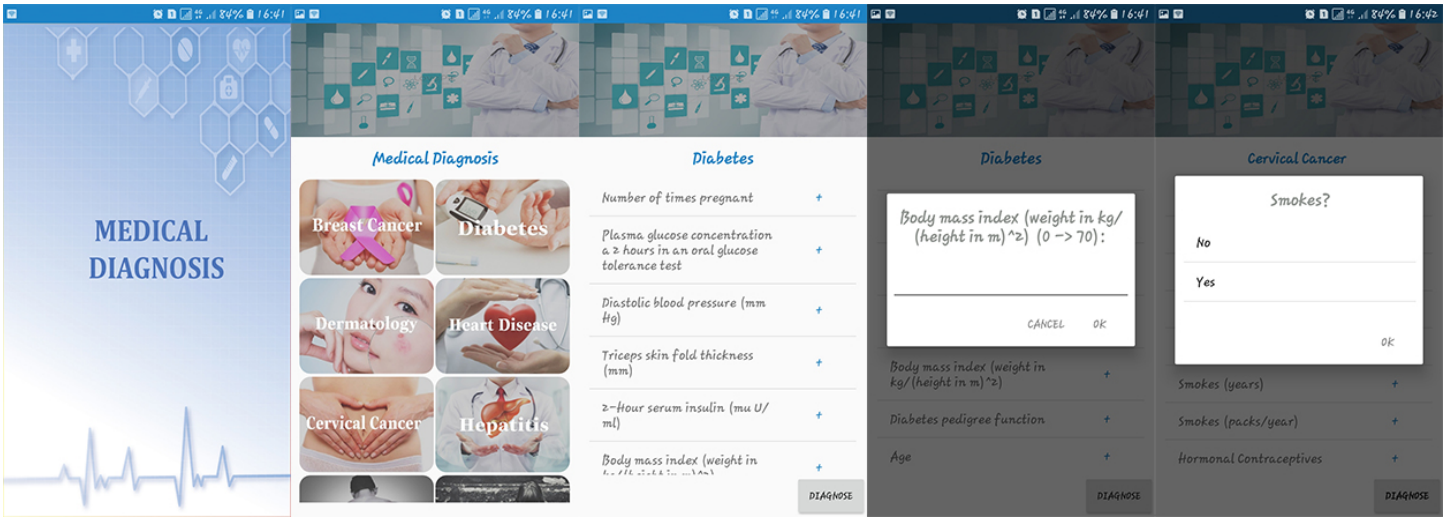


Figure 5: Several screenshots of our mobile application.

returns or report missing symptoms of the user. The application is designed to suit all types of screens of different mobile phones. The application can run on Android OS 5.0.0 and above.

The server holds an up-to-date trained model, configuration, a database, and complete Android packages that are really to install and update to the client. When this paper is conducted, the authors merely develop a mobile application for Android only. Several screenshots of our mobile application are shown in Figure (5).

5 Conclusion

In this research, the authors have described the interpretation of machine learning models for the task of classification in medical diagnosis. We propose processes that interpret the most appropriate learning models for each of the different clinical datasets as the foundation of developing and recommending diagnostic procedures. The whole process works as (i) hyperparameters tuning for picking the most appropriate learning approach, and (ii) a mobile application is developed to support clinical practices. We also explain an urgent need to optimize medical processes and regular experts' workflows to support healthcare services while reducing investment costs and improving efficiencies. The experimental results, interpretation of models, and reproducibility are thoughtfully discussed. A mobile application is also developed. We believe that our work has substantially extended our previous paper has encouraged further research machine learning, healthcare, and medical diagnosis.

References

[1] N. Duong-Trung, X. N. Hoang, T. B. T. Tu, K. N. Minh, V. U. Tran, T.-D. Luu, "Blueprinting the Workflow of Medical Diagnosis through the Lens of Machine Learning Perspective," in 2019 International Conference on Advanced Computing and Applications (ACOMP), 23–26, IEEE, 2019, doi: 10.1109/ACOMP.2019.00011.

[2] W. H. Crown, "Potential application of machine learning in health outcomes research and some statistical cautions," Value in health, **18**(2), 137–140, 2015, doi:10.1016/j.jval.2014.12.005.

[3] C. Cheng, J. Barceló, A. S. Hartnett, R. Kubinec, L. Messerschmidt, "COVID-19 Government Response Event Dataset (CoronaNet v. 1.0)," Nature Human Behaviour, **4**(7), 756–768, 2020, doi:10.1038/s41562-020-0909-7.

[4] Y. Wang, N. Hajli, "Exploring the path to big data analytics success in healthcare," Journal of Business Research, **70**, 287–299, 2017, doi:10.1016/j.jbusres.2016.08.

[5] C. H. Lee, H.-J. Yoon, "Medical big data: promise and challenges," Kidney research and clinical practice, **36**(1), 3, 2017, doi:10.23876/j.krcp.2017.36.1.3.

[6] M. Swan, "The quantified self: Fundamental disruption in big data science and biological discovery," Big data, **1**(2), 85–99, 2013, doi:10.1089/big.2012.0002.

[7] A. Holzinger, I. Jurisica, "Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions," in Interactive knowledge discovery and data mining in biomedical informatics, 1–18, Springer, 2014, doi:10.1007/978-3-662-43968-5_1.

[8] R. Freeman, L. Frisina, "Health care systems and the problem of classification," Journal of Comparative Policy Analysis, **12**(1-2), 163–178, 2010, doi:10.1080/13876980903076278.

[9] K. A. Wager, F. W. Lee, J. P. Glaser, Health care information systems: a practical approach for health care management, John Wiley & Sons, 2017.

[10] T. B. Murdoch, A. S. Detsky, "The inevitable application of big data to health care," Jama, **309**(13), 1351–1352, 2013, doi:10.1001/jama.2013.393.

[11] H. M. Krumholz, "Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system," Health Affairs, **33**(7), 1163–1170, 2014, doi:10.1377/hlthaff.2014.0053.

[12] Z. Obermeyer, E. J. Emanuel, "Big data, machine learning, and clinical medicine," The New England journal of medicine, **375**(13), 1216, 2016, doi:10.1056/NEJMp1606181.

[13] M. Chen, Y. Hao, K. Hwang, L. Wang, L. Wang, "Disease prediction by machine learning over big data from healthcare communities," Ieee Access, **5**, 8869–8879, 2017, doi:10.1109/ACCESS.2017.2694446.

[14] D. S. Char, N. H. Shah, D. Magnus, "Implementing machine learning in health care—addressing ethical challenges," The New England journal of medicine, **378**(11), 981, 2018, doi:10.1056/NEJMp1714229.

[15] F. Cabitza, R. Rasoini, G. F. Gensini, "Unintended consequences of machine learning in medicine," Jama, **318**(6), 517–518, 2017, doi:10.1001/jama.2017.7797.

[16] S. R. Steinhubl, E. D. Muse, E. J. Topol, "The emerging field of mobile health," Science translational medicine, **7**(283), 283rv3–283rv3, 2015, doi:10.1126/scitranslmed.aaa3487.

[17] A. Jutel, D. Lupton, "Digitizing diagnosis: a review of mobile applications in the diagnostic process," Diagnosis, **2**(2), 89–96, 2015, doi:10.1515/dx-2014-0068.

[18] C.-K. Kao, D. M. Liebovitz, "Consumer mobile health apps: current state, barriers, and future directions," PM&R, **9**(5), S106–S115, 2017, doi:10.1016/j.pmrj.2017.02.018.

[19] K. P. Murphy, Machine learning: a probabilistic perspective, MIT press, 2012.

[20] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2016.

[21] M. Mohri, A. Rostamizadeh, A. Talwalkar, Foundations of machine learning, MIT press, 2018.

[22] N. Duong-Trung, Social Media Learning: Novel Text Analytics for Geolocation and Topic Modeling, Cuvillier Verlag, 2017.

[23] Y.-Y. Song, L. Ying, "Decision tree methods: applications for classification and prediction," Shanghai archives of psychiatry, **27**(2), 130, 2015, doi:10.11919/j.issn.1002-0829.215044.

[24] D. Walk, "Using Random Forest Methods to Identify Factors Associated with Diabetic Neuropathy: A Novel Approach," 2017, doi:10.1093/pm/pnw311.

[25] X. Liu, H. Zhu, R. Lu, H. Li, "Efficient privacy-preserving online medical primary diagnosis scheme on naive bayesian classification," Peer-to-Peer Networking and Applications, **11**(2), 334–347, 2018, doi:10.1007/s12083-016-0506-8.

[26] P. Naraei, A. Abhari, A. Sadeghian, "Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data," in 2016 Future Technologies Conference (FTC), 848–852, IEEE, 2016, doi:10.1109/FTC.2016.7821702.

[27] T. Li, J. Li, Z. Liu, P. Li, C. Jia, "Differentially private Naive Bayes learning over multiple data sources," Information Sciences, **444**, 89–104, 2018, doi:10.1016/j.ins.2018.02.056.

[28] A. M. Prasad, L. R. Iverson, A. Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction," Ecosystems, **9**(2), 181–199, 2006, doi:10.1007/s10021-005-0054-1.

[29] T. Chen, C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785–794, ACM, 2016, doi:10.1145/2939672.2939785.

[30] A. Guzman, Derivatives and integrals of multivariable functions, Springer Science & Business Media, 2012.

[31] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, **12**, 2825–2830, 2011, doi:10.1016/j.patcog.2011.04.006.

[32] M. Lutz, Learning python: Powerful object-oriented programming, O'Reilly Media, Inc., 2013.

[33] J. Avila, T. Hauck, Scikit-learn cookbook: over 80 recipes for machine learning in Python with scikit-learn, Packt Publishing Ltd, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 472-479 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Teaching/Learning Strategies in Context of Education 4.0

Irina Golitsyna[*,1], Farid Eminov[2], Bulat Eminov[3]

[1]*Institute of Information Technology and Intelligent Systems, Kazan (Volga) Federal University, Kazan, 420008, Russian Federation*

[2]*Automated Information Processing Systems and Control Department, Kazan National Research Technical University, Kazan,420111, Russian Federation*

[3]*GBU PDD, Kazan, 420059, Russian Federation*

ARTICLE INFO

Article history:
Received: 22 December, 2020
Accepted: 22 February, 2021
Online: 17 March, 2021

Keywords:
Education 4.0
educational technologies
training of IT-specialists
educational content
enterprise infrastructure

ABSTRACT

Coronavirus pandemic and transition to distance learning have significantly accelerated the introduction of Education 3.0 - 4.0 technologies into traditional educational process. This paper discusses questions of training of IT- specialists in context of Education 4.0. Based on our experience, approaches to the organization of the educational process of IT-students are considered. It is discussed, what elements of mobile learning, self-directed learning and informal learning are used by students. Informal learning in traditional educational process of IT- students is considered in such aspects as the source of knowledge, personalization, teaching/learning methods. The paper discusses a stage-by-stage approach to formation of interdisciplinary educational content for IT-students. In conclusion, the strategies of teaching / learning of Education 4.0, useful for forming of competencies for Industry 4.0, are discussed.

1. Introduction

This paper is an extension of work originally presented on the 12th International Conference on Developments in eSystems Engineering (DeSE)[1].

Pandemic changed traditional forms of educational process and accelerated introduction of e-learning technologies of Education 3.0 - 4.0. Education 3.0 is web-based, student-centered learning. Education 4.0 places the learner at the center of the educational ecosystem. Education 4.0 refers to the Industry 4.0 curricula. "The term '4.0' is related to the new technical possibilities for which the concept of E-learning or digital learning already exists"[2].

The infrastructure of Education 3.0-4.0 forms on the base of mobile technologies, social software, cloud services, online educational resources, mass open online courses (MOOC) [1]. Education 4.0 summarizes every problem that arises in 21st century education. This can be educational content, didactics, or methods [2].

Next aspects are important in advanced Training 4.0: learning on-demand, online learning in web-communities, learning regardless of time and place, mobile learning, self-organized and self-directed learning, a combination of different training methods and environments [2].

In this paper we consider strategies of teaching/learning and formation of educational content for IT-students in the Education 4.0 concept, based on our teaching experience.

The hypothesis of the research is: in order to develop the competencies of modern specialists, it is necessary to keep in mind the experience of using modern technologies by students, regardless of teachers; on the other hand, it is necessary to develop approaches to the educational content formation and teaching/ learning methods, taking into account the needs of modern industry.

Based on the survey of IT students, we shown that they independently use such the Education 4.0 technologies as mobile learning, self-directed learning and informal learning.

In conditions of the Industry 4.0 development, specialists must be ready for successful professional activities in a modern enterprise. As one of possible approaches to the educational content formation for IT students we propose the interdisciplinary

[*]Corresponding Author: Irina Golitsyna, 35, Kremlevskaya Street, Kazan, 420008, Russia, (843) 221-34-33 & Irina.Golitsyna@gmail.com

approach based on the infocommunication infrastructure of an enterprise.

2. Literature review

The main goal of Industry 4.0 is to improve the efficiency and flexibility of enterprises so that they can adapt to the demands of the future. Technologies such as robotics, simulation, horizontal and vertical integration systems., the industrial internet of things, cybersecurity, cloud computing, additive manufacturing, augmented reality, big data and analytics can be used to achieve this goal [3]. All these technologies allow enterprises to use the vast amount of information to make more effective decisions. The development of digital technologies increases the importance of specialists who are able to interact in network environment and maintain complex production [4], [5]. In conditions of rapid development of technologies, a professional of Industry 4.0 must constantly learn from other professionals with different education and experience [3]. An important task for modern education is to identify the competencies necessary for future specialists to effectively enter the labor market [3], [6], [7]. Knowledge and competencies required for Industry 4.0 include: information technologies, software engineering, programming, data communications and computer networks, interfaces, network protocols, internet of things, systems understanding, cloud solutions, sensors and electronics, lean production. Among the competencies that new participants must have in order to implement Industry 4.0 are the following: technical skills, process understanding, understanding IT security.

Real-world experience show that the most corresponding skills and competencies are 21st century skills or competencies for lifelong learning. They contribute to the continuous and progressive self-renewing and self-regulation, and can be transferred and directly applicable in various social and professional business environments [8].

The concept of Industry 4.0 legitimizes changes in the production sphere, in the labor market, but, basically, in all spheres of society [9], including education. Education 4.0 combines various methods and approaches by which institutions of higher education align their services and curricula to prepare future graduates for work [10]. An obvious idea of universities 4.0 is: they must develop the 'learning to learn' meta-competence. This meta-competence seems to be the key to a good job and a good life; this can lead to optimization of a person and overcoming his obsolescence [11]. At the same time, Education 4.0 places more emphasis on self-directed learning using technology [12].

Education 4.0 is education for Industry 4.0, on the other hand, its infrastructure is the natural development of the infrastructure of e-learning technologies Education 1.0 - 3.0.

The formation of the Education 4.0 infrastructure leads to development of e-learning technologies as shown in Table 1 [13].

Table 1: Educational Technologies in Education 4.0

Characteristics	Education 4.0	Educational technologies
Technological platform	Mobile devices	Mobile learning.
Software	LMS (Learning Management Systems), MOOCs (Massive Open Online Courses), social software, cloud resources, professional software, open electronic platforms for online interaction	Distance learning. Educational self-organization of students in social networks.
Learning management	Collaboration of teachers and students in electronic information-educational environment (EIEE)	Flexible learning. Blended learning.
Methods of teaching/learning	Productive, in context of EIEE	Self-directed learning. Informal learning.

3. Methodology

The purpose of this study is to identify the factors contributing to the development of key competencies of IT specialists in the context of the development of Education 4.0 technologies. We consider strategies of teaching/learning in context of Education 4.0 in the next aspects:

1. Educational technologies, that develop independently of educational institutions on the basis of the Education 4.0 infrastructure, are reviewed based on literature and our experience in training of specialists of bachelor programs 09.03.02 Information systems and technologies and 09.03.03 Applied Informatics.

Quantitative descriptive research as full-time classroom survey was made to determine what elements of mobile learning, self-education and informal learning are used by students in the traditional educational process. We made a survey of students of bachelor programs 09.03.02 Information systems and technologies and 09.03.03 Applied Informatics of Kazan Federal University in 2018. In the questioning took part 96 students of 3rd and 4th years of study. From this survey we have considered such aspects, as sources of knowledge, forms of learning, self-directed learning. Questions in the questionnaire were identified and based on the literature review and the expert opinions. Our colleagues, the teachers involved in the training of IT specialists acted as experts. The survey results are presented in the form of histograms. Most of these histograms have been discussed at conferences [1], [14], or published in the paper [15]. The histograms show the ratio of positive answers on the proposed questions to the total number of respondents as a percentage.

2. We propose approach to formation of educational content for IT-students on the base of the enterprise infocommunication infrastructure. This approach has developed in our experience in training of students of bachelor programs 09.03.02 Information systems and technologies and 09.03.03 Applied Informatics in Kazan Federal University and in Kazan National Research Technical University for several years. This approach has been discussed at conferences [1], [16]. The educational content and

methodological support of disciplines within the framework of the declared approach has been published in the number of educational and methodological manuals, e.g.[17]-[19].

Feedback has come also from personal conversations with students and discussions at the alumni meetings of Kazan National Research Technical University over the years.

4. Results and Discussions

4.1. Strategies of Teaching/ Learning

We consider below, how the teaching / learning strategies outlined in Table 1, are implemented in the traditional educational process of IT professionals. We used the mentioned above survey of students of bachelor programs 09.03.02 Information systems and technologies and 09.03.03 Applied Informatics.

4.1.1. Educational Technologies of Education 4.0 in Traditional Educational Process

Mobile learning is being developed on the basis of mobile devices and technological communication channels. Immediate switch of educational institutions on distant learning stimulated to use of mobile devices in the pandemic context.

According to [20], the mobility concepts in higher education can be divided into three important directions: mobility of technology, mobility of learners and mobility of learning. Mobile devices are used by students for educational communication, for example, to discuss educational issues and exchange ideas with classmates. In general, most students think that mobile devices can help them to learn.

We can see from histogram in Figure 1 students' answers to the question: "What forms of mobile internet do you use during educational process?".

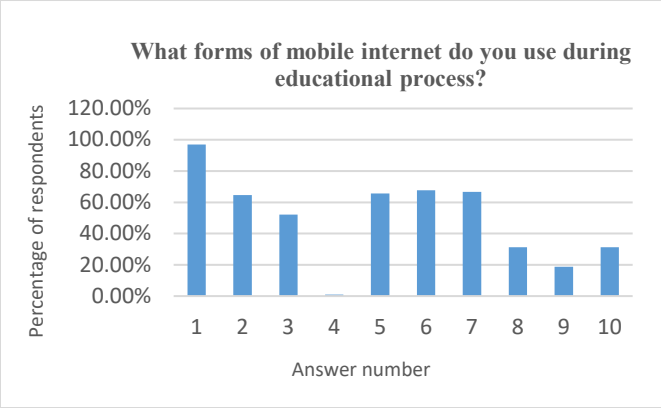


Figure 1: Students' answers to the question: "What forms of mobile Internet do you use during educational process?"

Answer options:
[1] Search engines
[2] Wikipedia
[3] Cloud storage resources
[4] Mobile access to the university library
[5] Electronic textbooks
[6] Social networks
[7] Messengers
[8] Specialized programming forums
[9] Integrated development environments (IDE)
[10] Mathematical problem-solving resources

Survey shows, that students actively use mobile access in the educational process to available web services, such as search engines, Wikipedia, cloud storage resources, electronic textbooks, social networks and messengers. This infrastructure of mobile learning has been developed independently of educational institutions, and teachers have been able to use the ready-made infrastructure in the educational process in the current conditions. Students use the virtual environment for access to educational resources from mobile devices, thereby using the opportunities of mobile learning.

Students form a mobile personal learning environment and become full participants in the formation of EIEE of educational institusions in accordance with the modern education development trends.

Distance learning has become an only possible form of education in the context of coronavirus pandemic. This situation accelerated the introduction of e-learning technologies into the traditional educational process. Along with LMS and MOOC, open electronic platforms for online interaction (such as Microsoft Teams, Google Zoom etc.) have been widely used in current conditions.

Flexible learning refers to the use of online learning in a traditional learning process. IT- students actively use the electronic professional environment for learning and solving assigned tasks, therefore, technology of flexible learning can be effectively used for their education. Students use web-resources in educational activities, form a personal-oriented EIEE. Teachers receive the opportunities to develop new teaching methods, to personalize learning and organize students' independent work [21].

Blended learning is developing based on the use of e-learning elements in a traditional educational process [22]. The concept of blended learning also involves learning in a workplace, which allows teachers to master and use vocational training technologies.

Self-directed learning. As noted above, self-directed learning is one of the key competencies for Industry 4.0.

We can see from histogram in Figure 2 students' answers to the question: "Are you engaged in self-directed learning?".

We can conclude, that students are independent in self-directed learning, they actively use specialized books and textbooks and programming forums, learn on educational online resources, some of them attend educational courses.

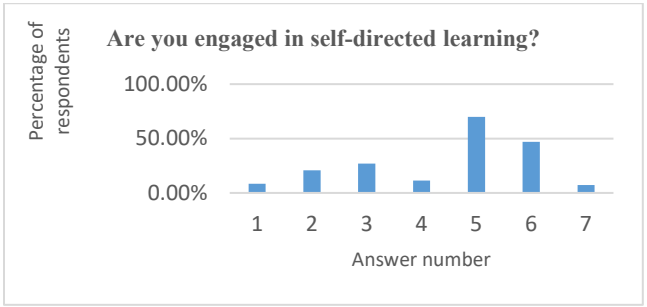


Figure 2: Students' answers to the question: "Are you engaged in self-directed learning?"

Answer options:
[1] No
[2] I attend educational courses
[3] I learn on educational online resources

[4] I use professional forums
[5] I read specialized books and textbooks
[6] I use programming forums
[7] Other

Students widely use web-resources for self-directed learning and for solving academic tasks. Table 2 shows web- resources that IT-students use in various types of educational activities to form professional competencies [1].

Table 2: Learning Activities of Students with Web-resources

Learning activities	Web-resources
Educational communications of students and teachers	E-mail, social networks, cloud services
Information search	Search engines
Studying of new educational content	Electronic textbooks, educational resources, Wiki -resources
Solving mathematical tasks	Mathematical online resources, online calculators, computer algebra systems
Programming	IDE, programming forums, online professional communities

The use of web resources in traditional educational process and in self-directed learning leads to formation of personal learning environment for each student. This environment is a continuation and extension of EIEE of educational institutions, allowing teachers and students to develop teaching / learning strategies in context of requirements of Education 4.0.

4.1.2. Informal Learning

As noted above, competencies of lifelong learning are the most relevant competencies for Industry 4.0. Lifelong learning includes formal, nonformal, and informal learning. The role of informal learning has increased in current conditions. Informal learning is an unstructured learning in a family, community or/in the workplace. Mastering of informal learning experiences prepares students for lifelong learning. Development of this competence is important for professional activity of specialists in conditions of the fourth industrial revolution.

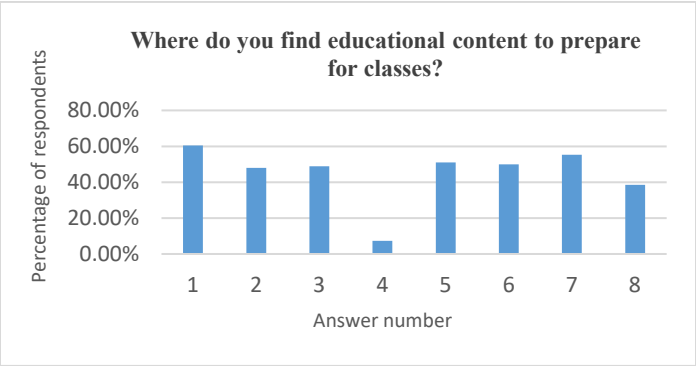


Figure 3: Students' answers to the question: "Where do you find educational content to prepare for classes?"

Answer options:
[1] I use summary of lectures or seminars
[2] I use an educational content recommended by a teacher
[3] I discuss study questions with other students
[4] I use resources of the university library
[5] I read electronic textbooks
[6] I use e-guides
[7] I read articles on Wikipedia
[8] I use specialized forums

Elements of informal learning are changing the traditional educational process of IT-students in such aspects as the source of knowledge, personalization, teaching/learning methods and teacher activities [14].

We can see from histogram in Figure 3 students' answers to the question: "Where do you find educational content to prepare for classes?" [14].

Survey shows, that students use the summaries of lectures or seminars and educational content, recommended by a teacher in educational process as often as electronic educational resources: electronic textbooks, e-guides, specialized forums. They use Wikipedia often, but only 7% of students tend to use resources of the university library.

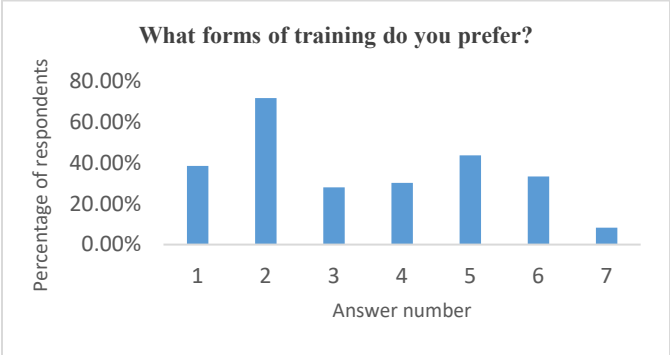


Figure 4: Students' answers to the question: "What forms of training do you prefer?"

Answer options:
[1] Traditional lectures and classroom workshops
[2] Project work in a small group
[3] Independent work outside the classroom
[4] Solving individual tasks
[5] Independent development of software applications
[6] The solution of creative tasks without ready-made solutions
[7] Other

Students give preferences to forms of training, relevant for informal learning in frames of traditional educational process. We can see from histogram in Figure 4 students' answers to the question: "What forms of training do you prefer?"[14].

Histogram shows that students prefer learning interactions (project work) in small groups and implementation of specific learning activities (independent development of software applications) rather, than traditional forms of learning. These forms of training are relevant for informal learning.

Thus, students personalize the sources of educational content, forms and methods of training, place and time of self-directed learning. Studying in formal academic groups, university students actively shape their personal learning environment and try to develop individual learning strategies.

Four factors are changing the activities of teachers: permanent changes in the subject area of education, development of sources of educational information, transformation of the educational

activities forms of students and formation of EIEE. Obviously, teachers should take into account these factors under developing of teaching strategies.

Teachers can use open educational resources for teaching and create new educational resources in collaboration with students. Web - resources and EIEE allow to restructure information in its accumulation and development, provide the ability to share and edit educational documents [1]. Use of educational and professional web-resources allows teachers to expand the content of IT disciplines, organize collaboration with students, promote the creative educational activities of students and ensure the personalization of learning.

4.2. Enterprise Infocommunication Infrastructure in Educational Content

In [23] it is proposed to combine all major engineering disciplines in the Industry 4.0 knowledge areas with their respective functions: "Designing CPS (Cyber Physical Systems), Modularity, Interoperability (Objects, machines and people need to be able to communicate through the Internet of Things and the Internet of People), Virtualization, Decentralization, Real-Time Capability, Service-Orientation".

Kondakov distinguishes following stages of the educational content formation in the era of fourth industrial revolution [24]:

- *Pre-subject stage* includes development of general theoretical ideas about the structure and composition of educational content, taking into account prospective requests from families, society, state, peculiarities of the educational process organization in accordance with curriculum and with extracurricular activities.
- *Subject stage* includes definition of composition of academic disciplines, their content and allocation by educational levels.
- *Development of educational materials*.
- *Organization of teaching/learning strategies.*
- *Stage of assignment by students of new educational content, forms and types of activities, skills and competencies*.

Every stage is considered further based on our experience in teaching of IT-disciplines.

4.2.1. Pre-subject Stage of Formation of Educational Content

At the pre-subject stage of formation of educational content of IT-stidents, it should be borne in mind that modern enterprises have a permanently developing complex infocommunication infrastructure. Specialists in such infrastructure have to work in the network environment, they must be proficient in network technologies and enterprise information systems [16].

Enterprises use information technologies in different areas of their functioning: the business-processes automation, support of collaboration with partners, the electronic document flow management etc. Structure of the enterprise IT-department is built as a system of the processes managing services. This service-oriented structure allows management of IT-department to provide high quality services and efficient business solutions for an enterprise. IT-department provides services to functional departments, and connections between them are based on the "provider - consumer" model. Functional departments set their requirements for the services of specified types and their quality. Infocommunication infrastructure is supported and developed by IT-departments according to required services and set quality. Thus the enterprise IT-department is an important resource for the enterprise management system. Table 3 shows the hierarchy of enterprise management levels and basic management functions corresponding to each level.

Table 3: Hierarchy of the Enterprise Management Levels

Level of enterprise management	Functions
Financial and economic activity	Financial planning, customer relationship, supply chain, human resources, etc.
Production activities	Operational planning, quality management, order deadlines control, productivity control, etc.
Manufacture management	Technological processes, production processes
Industrial controllers	Data collection, data processing
Sensors	Scanning of sensors, operating mechanisms

Each level performs specific management functions and uses its own computing network environment. The lowest level is sensory. At this level, a network of sensors and of various actuating mechanisms is implemented. In the context of the development of Industry 4.0, sensors will be the basis for the development of the Internet of Things [25]. The level of industrial controllers is implemented by the controller network. The industrial and office local networks are used at the manufacture management level. The local or the enterprise office networks are used at the production management levels and at the levels of financial and of economic activities of an enterprise.

The information networks have specific characteristics of functioning at every management level. The network infrastructure management must be considered during training of IT-students.

4.2.2. Subject Stage of Formation of Educational Content

At the subject stage, we offer teaching of several disciplines within the curriculum on an interdisciplinary basis as one of the educational approaches to the formation of educational content of IT- students.

A hierarchical model of enterprise architecture can be a basis for interdisciplinary connections in the professional training of IT-students. Enterprise architecture includes structured description of document flow and business processes of an enterprise, applications and automation methods that support business processes, as well as information, technology and infrastructure.

Table 4 shows levels of enterprise architecture in accordance with model TOGAF (The Open Group Architecture Framework) [17], their components and disciplines for training of IT-students.

Table 4: Disciplines for Training of IT-students in Accordance with TOGAF Model

Enterprise architecture levels	Components of enterprise information systems architecture	Disciplines for training of IT-students
Business architecture	Business applications	Information systems design. Enterprise information systems
Application architecture	Enterprise information systems	Information systems design. Enterprise information systems
Data architecture	Physical data stores, data management tools, logical and physical data models	Database. Database management systems. Enterprise information systems
Technological architecture	Infocommunication infrastructure	Infocommunication systems and networks. Network technologies. Information security

Content of studied disciplines is defined by the enterprise architecture type. Its own network computing environment with own characteristics of functioning corresponds to each hierarchical level of infrastructure. In training of IT-students, attention should be paid to the management of information network infrastructure [16].

4.2.3. Creation of Educational Materials

As noted above, knowledge related to the elements of Industry 4.0 include: networks, information technologies, interfaces and communication protocols, understanding of systems, sensors and electronics. The composition of IT disciplines can be formed in accordance with the enterprise infocommunication infrastructure. To study disciplines in which the information network environment is considered, students have to acquire next basic knowledges:

- distributed computing; distributed information-processing system;
- basic networking concepts;
- the cable and wireless technologies in local and enterprise networks; networking standards; communication interfaces;
- data communications technologies; data access, coding of signals and data;
- network services;
- life cycle management of an enterprise; service-oriented computing of business processes;
- the system development life cycle; functional subsystems of an enterprise.

4.2.4. Organization of teaching/learning strategies

An interdisciplinary approach to teaching allows organizing the following forms of learning activities of IT students:

- Creation of infrastructure of local network on real computers as practical activities. Students are invited to install the cable network, computers and communication hardware, they have to set up and test of elements of infocommunication system and network.
- Each student invited to create a virtual environment on his/her real machine to master the administration of networks. Next step is creation a virtual environment that includes a server and one (or more) networked workstations. After creating the virtual environment, the student should master the skills of configuring different network services.
- In order to master the formation and configuration of the complex services parameters, students should study the setup of terminal access, methods of organizing of the secure data transmission tunnels, the formation of help services and other issues of the enterprise infrastructure management.

4.2.5. Assignment of new educational content, forms and types of activities, skills and competencies by students

The rapid transformation of the professional environment of IT specialists requires constant updating of the content of professional disciplines. Forms and methods of teaching / learning are developing in accordance with the updating of educational content.

In our teaching activities, we constantly update the content of the taught disciplines in accordance with the development of information systems and network technologies in enterprises (e.g. [17]-[19]).

Feedback. Kazan National Research Technical University annually holds meetings of students with representatives of enterprises, and often our graduates are representatives of enterprises. Discussion of our approach to the formation of educational content and to forms of practical classes with students and alumni shows:

- the topics covered in this paper seem extremely interesting to them;
- students are interested in practical applications of technologies that they get acquainted with in classrooms;
- forms of practical classes are especially fascinating for students.

5. Conclusion

According to [6], it is too early to draw conclusions about the long-term impact of the global pandemic on education, and emergency online teaching should not be equated with pedagogically effective online teaching in accordance with the needs of students and potential employers. But it is clear, that coronavirus pandemic and rapid transition to distance learning significantly accelerated an introduction of Education 3.0 - 4.0 technologies into traditional educational process.

Development of e-learning methods takes place on this basis, and these methods are actively applied by the pedagogical community in modern situation. Introduction of e-learning technologies leads to development of EIEE of educational

institutions and to active development of personal-oriented EIEE of students. All these aspects of advanced learning are realized by students through educational self-organization in traditional educational process.

The results of the survey of IT students presented in this paper show that they independently use such technologies of Education 4.0 as mobile learning, self-directed learning and informal learning. The survey results are consistent with our teaching experience. These processes contribute to the 'learning to learn' meta-competence formation and to preparation of students for lifelong learning in context of fourth industrial revolution.

Teachers need to perform the following activities in modern conditions [13]: participation in the continuous development of the EIEE of educational institution in collaboration with students; constant updating of the educational content of disciplines, which should include interdisciplinary educational content; developing teaching/learning strategies in accordance with the state of the professional environment of modern specialists.

For developing educational content in teaching IT students, we propose to use the enterprise infocommunication infrastructure as the system-forming factor. Professional activities in the conditions of permanent development of the infocommunication infrastructure of modern enterprises requires specialists to master enterprise information systems and their network environment. Teaching on an interdisciplinary basis can be considered as one of the perspective approaches to training of such specialists. Our experience shows that this approach allows to update the necessary educational content and teaching/ learning strategies in accordance with the development of the professional environment of IT professionals in modern conditions.

Table 5 shows, what possible teaching/learning strategies in context of Education 4.0 can be used for formation of competencies of Industry 4.0.

Table 5: Teaching/learning Strategies of Education 4.0 for Formation of Competencies of Industry 4.0

Competencies of Industry 4.0	Teaching/learning strategies in context of Education 4.0
Lifelong learning	Informal learning
Self-directed learning	Mobile learning. Personally-oriented EIEE of students
Programming	Informal learning. Mobile learning. Personally-oriented EIEE of students
Technical skills; information technologies; process understanding	Teaching on an interdisciplinary basis
Data communications and networks; interfaces and communication protocols; understanding of systems; understanding IT security	Teaching on an interdisciplinary basis. Enterprise infocommunication infrastructure, as the system-forming factor of educational content

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors are grateful to colleagues for the expert assessment and useful discussion of our work. The authors would like to thank Ekaterina Golitsyna for help in writing and redaction of the paper text.

References

[1] I.N. Golitsyna, F.I. Eminov, B.F. Eminov, "Education 4.0 in teaching/learning strategies," in Proceedings - International Conference on Developments in eSystems Engineering, DeSE, 2019, doi:10.1109/DeSE.2019.00046.

[2] M. Klonn, J. Abke, "'Learning 4.0': A Conceptual Discussion," in Proceedings of 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2018, 2019, doi:10.1109/TALE.2018.8615244.

[3] M. Hernandez-de-Menendez, R. Morales-Menendez, C.A. Escobar, M. McGovern, "Competencies for Industry 4.0," International Journal on Interactive Design and Manufacturing, **14**(4), 1511–1524, 2020, doi:10.1007/s12008-020-00716-2.

[4] S. Poloskov, A. Zheltenkov, I. Braga, I. Kuznetsova, "Adaptation of high-tech knowledge-intensive enterprises to the challenges of industry 4.0," E3S Web of Conferences, **210**, 1–10, 2020, doi:10.1051/e3sconf/202021013026.

[5] M.T.A. Jima'ain, F.N.A. Hassan, K.A. Razak, A. Hehsan, J. Junaidi, "The emerging challenges of industrial revolution 4.0: A students' perspective," International Journal of Advanced Science and Technology, **29**(6), 1215–1225, 2020.

[6] L.M. Kipper, S. Iepsen, A.J. Dal Forno, R. Frozza, L. Furstenau, J. Agnes, D. Cossul, "Scientific mapping to identify competencies required by industry 4.0," Technology in Society, **64**(October 2020), 2021, doi:10.1016/j.techsoc.2020.101454.

[7] W. Maisiri, H. Darwish, L. van Dyk, "An investigation of industry 4.0 skills requirements," South African Journal of Industrial Engineering, **30**(3), 90–105, 2019, doi:10.7166/30-3-2230.

[8] C. Demartini, L. Benussi, "Do Web 4.0 and Industry 4.0 Imply Education X.0?," IT Professional, **19**(3), 4–7, 2017, doi:10.1109/MITP.2017.47.

[9] P. Kowalikova, P. Polak, R. Rakowski, "The Challenges of Defining the Term 'Industry 4.0,'" Society, **57**(6), 631–636, 2020, doi:10.1007/s12115-020-00555-7.

[10] C.A. Bonfield, M. Salter, A. Longmuir, M. Benson, C. Adachi, "Transformation or evolution?: Education 4.0, teaching and learning in the digital age," Higher Education Pedagogies, **5**(1), 223–246, 2020, doi:10.1080/23752696.2020.1816847.

[11] F.J. GARCÍA-GARCÍA, E.E. MOCTEZUMA-RAMÍREZ, T. YURÉN, "Learning to learn in universities 4.0. Human obsolescence and short-term change," Teoria de La Educacion, **33**(1), 221–241, 2021, doi:10.14201/TERI.23548.

[12] H. Sulaiman, N. Suid, A.B.I. Mohd, "Usability Evaluation of Confirm-A Learning Tool Towards Education 4.0," in 2018 IEEE Conference on e-Learning, e-Management and e-Services, IC3e 2018, 2019, doi:10.1109/IC3e.2018.8632637.

[13] I.N. Golitsyna, "[Education 4.0 in the training of modern specialists]," Mezhdunarodnyy Elektronnyy Zhurnal "Obrazovatel'nyye Tekhnologii i Obshchestvo (Educational Technology & Society)", **23**(1), 12–19, 2020. (In Russ., abstract in Eng.).

[14] I.N. Golitsyna, "Informal Learning in the Educational Process of IT Professionals," in 5th International Conference on Lifelong Education and Leadership for all, Conference Proceeding Book, 93–98, 2019.

[15] I.N. Golitsyna, "[Informal learning as a part of modern educational process]," Mezhdunarodnyy Elektronnyy Zhurnal "Obrazovatel'nyye Tekhnologii i Obshchestvo [Educational Technology & Society] ", **21**(4), 344–350, 2018. (In Russ., abstract in Eng.).

[16] F.I. Eminov, I.N. Golitsyna, B.F. Eminov, "Enterprise infocommunication infrastructure in training of IT-professionals," in Journal of Physics: Conference Series, 2018, doi:10.1088/1742-6596/1015/4/042014.

[17] F.I. Eminov, Information technologies of enterprises management, Master Line, Kazan, 2015. (In Russ.).

[18] B.F. Eminov, F.I. Eminov, Enterprice infomation systems. Study guide. Ministry of Science and Higher Education of the Russian Federation; Kazan National Research Technical University named after A.N. Tupolev, Kazan State Technical University, Kazan, 2019. (In Russ.).
[19] B.F. Eminov, F.I. Eminov, Network technologies. Study guide, KNITU-KAI, Kazan, 2019. (In Russ.).
[20] R.A. Karim, A. Ghani, B. Abu, A. Haimi, M. Adnan, A. Dwi, J. Suhandoko, L. Studies, U.T. Mara, P. Branch, T. Campus, T. Road, "The use of mobile technology in promoting education 4.0 for higher education," Advanced Journal of Technical and Vocational Education, **2**(3), 34–39, 2018, doi:10.26666/rmp.ajtve.2018.3.6.
[21] I.N. Golitsyna, "[Flexible learning in a traditional educational process]," Vysshee Obrazovanie v Rossii [Higher Education in Russia], (5), 113–117, 2017. (In Russ., abstract in Eng.).
[22] T. Wanner, E. Palmer, "Personalising learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course," Computers and Education, **88**, 2015, doi:10.1016/j.compedu.2015.07.008.
[23] L. Jeganathan, A.N. Khan, J. Kannan Raju, S. Narayanasamy, "On a frame work of curriculum for engineering education 4.0," in 2018 World Engineering Education Forum - Global Engineering Deans Council, WEEF-GEDC 2018, 2019, doi:10.1109/WEEF-GEDC.2018.8629629.
[24] A. Kondakov, [School of the era of the fourth industrial revolution], 2017. (In Russ.).

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 499-511 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Open Access Research Trends in Higher Education: A Literature Review

Mariutsi Alexandra Osorio-Sanabria[1,*], Astrid Jaime[2], Tamara Alcantara-Concepcion3, Piedad Barreto[4]

[1]*Universidad Pontificia Bolivariana, Information and Communication Technologies Faculty, Medellín, 050031, Colombia*

[2]*Universidad El Bosque, Engineering Faculty, Industrial Engineering School, Bogotá, 110121, Colombia*

[3]*Universidad Nacional Autónoma de Mexico, Department, Institute, Ciudad de México, 04510, Mexico*

[4]*Universidad Cooperativa de Colombia, Law Faculty, Bogotá, 111311, Colombia*

ARTICLE INFO

Article history:
Received: 16 December, 2020
Accepted: 14 February, 2021
Online: 17 March, 2021

Keywords:
Open access
Higher education
Trends

ABSTRACT

This study is a review of the literature on open access (OA), seeking to identify trends in research on the subject. This review was conducted in the SCOPUS database and focused on the following as the main topics: 1. Financial aspects, 2. Repositories, 3. Education, 4. Academic community's perception of OA resources, 5. Tools, 6. Policies, 7. Institutions, 8. Stakeholders, and 9. Impact. Out of these topics, the financial aspect, especially in OA's publication costs, was identified as driving great interest among researchers in the field. On the other hand, the study of the impact of OA is a subject little examined. Although research on OA in the higher education sector analyzes different perspectives and describes advances, challenges, and concerns, it is fair to conclude that OA encourages the creation and dissemination of knowledge and academic communication.

1. Introduction

At present, scientific knowledge is considered a public good. Information and Communication Technologies (ICT) have not only facilitated its access, transfer, exchange, and reuse [1], but they have also generated new scenarios of control for the academic community over its production [2] in contrast to the monopoly of commercial publishers. From this point of view, open access (OA) is a mechanism and, at the same time, an international movement, whose purpose is that "any person, with an Internet connection, can freely access without economic, technical or legal restrictions to scientific, academic and cultural information" [3].

The development of this movement is supported by various declarations and initiatives such as the Budapest Open Access Initiative [4], the Bethesda Declaration on Open Access Publications [5], the Berlin Declaration on Open Access to Knowledge in Sciences and Humanities [6], the Declaration of Santo Domingo [7] and the Declaration of Salvador [8]. Thus, the implementation of OA represents both an opportunity and a great challenge for society, particularly for Higher Education Institutions (HEIs), where it has aroused strong research interest.

It is not surprising then, that some academics have worked on the research trends observed in this field. In [9], the author present "focuses on the open-access-related conversations embedded within a serials context or OA discussions in publications targeting serials or electronic resource librarians" for the years 2008 to 2009. Likewise, in [10] the author has also tried to identify OA trends. Despite not specifically established, by exploring the bibliography of his work, it can be determined that he bases his work on documents from 1940 to 2013, although the older documents pertain to intellectual property and scientific activity. [11] "presents a conceptual literature review of the subject of open access as it is reflected in the literature relevant to digital library research", through the study of works" (published between 2010 and 2015). As it can be observed, some of these works focus on specific issues and are based on literature published before 2016. For this reason, we are interested in more recent research trends in OA in HEIs, given the central role it plays in OA, to achieve an

*Corresponding Author: Mariutsi Alexandra Osorio-Sanabria, Engineering PhD student, mariutsi.osorio@upb.edu.co

https://dx.doi.org/10.25046/aj060257

overview of the current concerns of the academic community in this field.

In the next section, the studies related to the object of study are presented followed by the methodology that guided the work carried out. Subsequently, the results and the discussion derived from the identification of the topics addressed in the OA studies are introduced prior to the conclusions.

2. Related work

OA refers to the dissemination of scientific research results, in digital format, which has undergone a peer-review process and doesn't charge subscription fees [12]. There are multiple definitions of this notion. For example, [13] bases its work on the Budapest OA Initiative, the Bethesda Declaration on Open Access Publishing, and the Berlin Declaration on OA to Knowledge in Science and Humanities, to propose "the BBB definition of OA" and summarizes the three definitions of OA as follows: "OA literature is digital, online, free of charge, and free from most copyright and licensing restrictions." For his part, in [14] is defines OA as "access online, at no cost, to peer-reviewed scientific content with limited copyright and licensing restrictions." On the other hand, Uribe-Tirado indicates that the OA is the right anyone has, without registration, subscription, or payment restrictions, to be able to read, download, copy, distribute, print, search or link the full texts or educational, scientific, or other digital content, and use them in a legitimate way according to the Creative Commons licenses assumed [12].

Furthermore, from the different definitions of OA, several issues around the OA movement have been studied. A few of these works focused on the research trends that can be observed in this field, as mentioned in section 1. In [9] the author identifies, for the years 2008 to 2009, five issues: (i) the NIH Mandates, (ii) "the increasing number of access policies adopted by universities", (iii) the "Arguments promoting OA" (par. 42), (iv) the strategies deployed to increase acceptance of OA among faculty, and (v) the roles different actors should adopt in supporting the OA movement (par. 43). The study [10] proposed four trends: "New OA publication channels" (p. 82), "legal aspects of academic publishing and OA Publishing" (p. 114), new business models around OA, and "Mandates, both institutional and from funding organizations" (p. 186). Finally, in [11] the author focusing on OA "literature relevant to digital library research" (p. 2), identifies five categories: "Open Access, Authors, Scholarly Communication, Libraries and Librarians, and Developing and Transitional Countries." (p. 2). Table 1 summarizes the trends recognized by these works, identifying the authors that distinguish each one.

Table 1: Research trends identified.

Research trends	Source
Policies and Mandates	[9,10]
Arguments promoting OA	[9]
Acceptance of OA	[9]
Actors and their roles (Authors, Librarians)	[9,11]
OA publication channels	[10]
Legal aspects of OAP	[10]
Business models around OA	[10]
Manifestations, Barriers, and Benefits of OA	[11]
Scholarly Communication	[11]
Libraries	[11]
Developing and Transitional Countries	[11]

As it can be established, there are just a few trends identified by more than one author, which is not surprising given the specificity of the focus of each work. For that reason, we are interested in analyzing current trends in research in the OA field. Although we had recently studied this subject [1], the dynamics of the field has led us to further explore if new trends in OA research in HEIs have emerged.

3. Methodology

There is an increase in research to conceptualize relevant issues and aspects of open access. For this reason, it is essential to summarize and provide an overview of these proposals. The existing literature is examined with a Systematic Mapping of Literature [15,16]. This method makes it possible to systematically and objectively identify the scope of available empirical studies to answer specific research questions. The review process has four stages: (1) Definition of research questions, (2) Search strategy, (3) Inclusion and exclusion criteria; and (4) Classification of documents (figure 1). The study stages are described below in the following subsections.

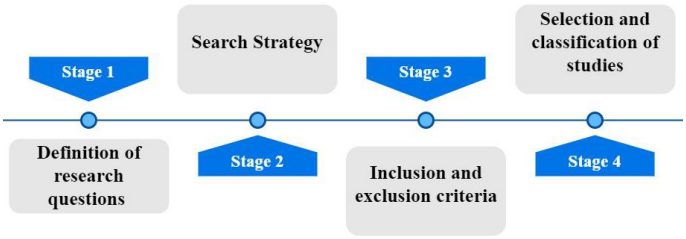


Figure 1: Methodology

3.1. Definition research questions

This study aims to identify the advances, topics, and trends of OA research and which aspects are being analyzed. Three research questions associated with the objective of the study are defined and addressed:

- What is the distribution of the publications of OA studies regarding HEIs?
- What are the topics of OA studies identified in the literature?
- What are the trends in OA considered by the studies identified in the literature?

3.2. Search Strategy

The review intends to find primary studies about OA research using a search strategy in the SCOPUS database. The search was developed through the review of the data needed to answer the research questions.

The terms "open access" and "higher education institution" were defined as keywords. These terms are connected with the conjunction "AND", yielding the search string:

("open access") and ("higher education institution") (1)

The search was limited to the title of the study, and journal articles, excluding articles published in conferences, books, and others. The result of applying the search equation to the selected database was 871 documents. Subsequently, 11 duplicates were identified and removed.

3.3. Inclusion and exclusion criteria

We defined inclusion and exclusion criteria filters and selected relevant studies to answer the research questions.

The inclusion criteria for papers are:

- Research published between 2010 and 2020.
- Research that explicitly mentions the term open access and higher education institution in the title and abstract and present a contribution to the OA issue.
- Research in areas other than the health sciences.

The exclusion criteria of the papers applied are:

- Research focuses which is unrelated to OA.
- Duplicate studies, only the most recent is included.
- Research that is not published as peer-reviewed journal articles.

3.4. Selection and classification of studies

We identified the main articles that provided direct evidence on the research questions. The search is limited to studies published between 2010 and 2020. The initial step was to search SCOPUS, using the search equation (1).

The selection of primary studies implies three eligibility filters to select the relevant results: year of publication, full-text availability, and research of areas other than the health sciences.

Figure 2 summarizes the systematic literature review process in a diagrammatic format according to the PRISMA Statement [17], indicating the number of studies in each phase. The studies were analyzed and selected by four reviewers. Each reviewer opined on whether to approve or reject each study.

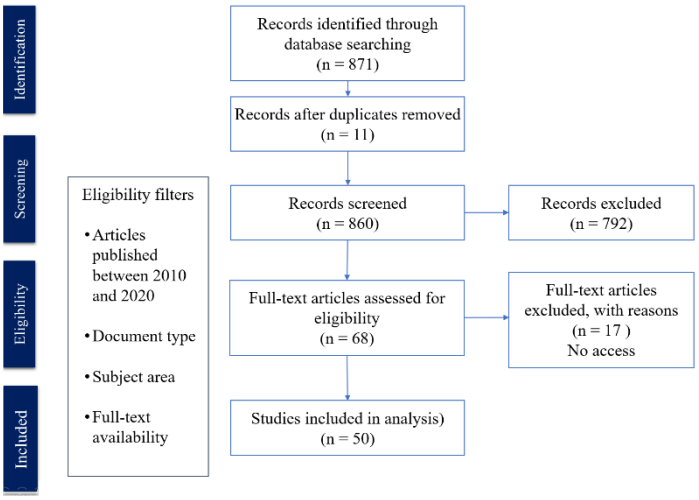


Figure 2: Number of included articles during the study selection process

As a result of the systematic literature review process, 50 primary studies were selected. These studies were organized in different categories as presented in Table 2.

Table 2: Classification according to trend

Trend	Description
Financial aspects	Aspects related to the costs and monetary difficulties linked to OA.
Repositories	Set of systems and services that facilitate the storage, management, retrieval, presentation, and reuse of digital content [18].
Education	Aspects related to the impact of OA on the teaching process of the higher education sector.
Perceptions about the use of OA resources	Surveys and interviews to consult the academic communities' opinion (students, professors, and/or researchers), about their experiences when using OA resources.
Tools	Use of ICTs that allows the transformation from closed access to OA in scientific publications [19].
Policies on OA	Regulations and guidelines about initiatives of OA and licenses of use.
Institutions	Aspects related to the changing roles, policies, and practices of the institutions participating in the OA Ecosystem.
Stakeholders	Actors that participate and influence the OA ecosystem.
Impact	Changes produced on the environment, processes, products or in some population groups, due to a certain action [20].

4. Findings

The research questions stated above (see section 3.1). are responded through the analysis performed. For this reason, this section is organized according to each one of these questions. The first issue addressed is the distribution of the publications, then the topics, and, finally, the trends in OA considered by the studies identified in the literature.

4.1. Distribution of publications on OA regarding HEIs

We analyzed 50 studies on OA in HEIs, published between the years 2010 and 2020. We observed that 2019 was the year with the highest number of publications, with 15 works. These results show a growing interest in the subject of OA for the higher education sector. Figure 3 below shows the distribution of documents by year.

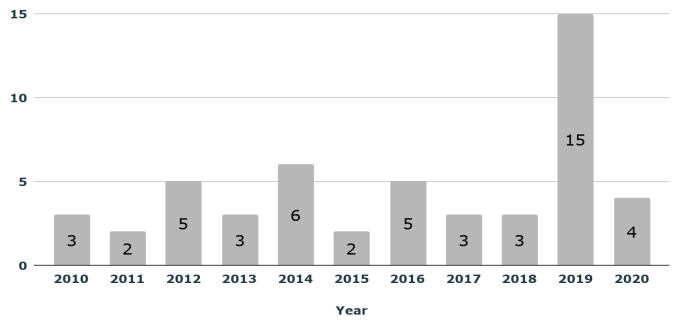


Figure 3: Documents analyzed by year

It is also important to note that the result for the year 2020 does not include the whole year, because the search was done on September 15th, 2020.

On the other hand, we analyzed the distribution of the selected works according to the country in which the study was carried out. We noted that 24% of the studies were conducted in the UK, 10% in Austria, 8% in South Africa and India, and 6% in Brazil and Australia. On the other hand, 38% of the remaining works were conducted in Canada, Egypt, Ecuador, France, Germany, Kenya, Mexico, Nigeria, the United Arab Emirates, Slovenia, United States, Zimbabwe, and countries not defined (Table 3). Furthermore, only one study focuses on Latin America and another one in Europe. In the Latin American context, we observed that OA in HEIs has been analyzed in Argentina, Brasil, Costa Rica, Peru, and Mexico. Figure 4 below shows the documents analyzed by the country mentioned in the research.

Table 3: Classification according to country

Country	Documents	Sources
United Kingdom	12	[21–32]
Austria	5	[19,33–36]
India	4	[37–40]
South Africa	4	[41–44]
Australia	3	[45–47]
Brazil	3	[48–50]
Mexico	2	[51,52]
United States	2	[53,54]
Canada	1	[55]
Ecuador	1	[56]
Egypt	1	[57]
Slovenia	1	[58]
France	1	[59]
Germany	1	[60]
Kenya	1	[61]
Nigeria	1	[62]
United Arab Emirates	1	[63]
Zimbabwe	1	[64]
Latin America	1	[65]
Europa	1	[66]
Country is not defined	3	[2,67,68]
Total	50	

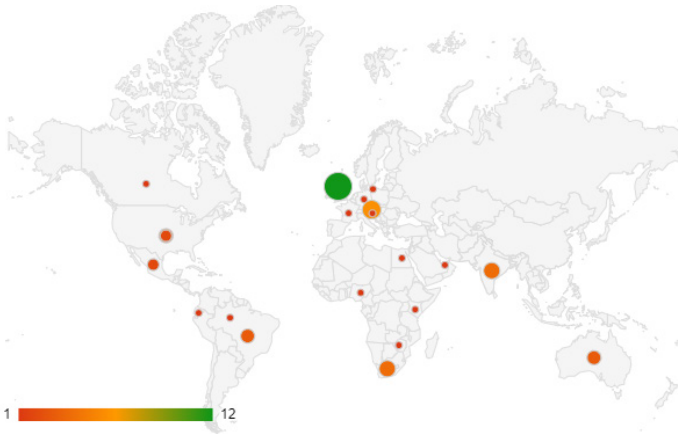


Figure 4: Documents analyzed by country

4.2. Topics emerging from the studies

To identify and classify the selected studies according to the authors' contribution. The trends defined in Table 5 were the ones that emerged from the analysis of the selected works. Table 4 presents the results of the classification. The trend with the most associated studies is focusing on the financial aspects (32%), followed by research related to repositories (22%), and education (12%).

Table 4: Classification according to trend

Country	Documents	Sources
Financial Aspects	16	[2,19,21,22,28,30,33–36,38–43,54,67]
Repositories	12	[29,37,41,44,45,51,52, 56,58,61,62,64]
Education	6	[47,49,55,65,68,69]
Perceptions	5	[32,39,48,57,59]
Tools	5	[38,40,50,58,66]
Policies	4	[25,31,45,53]
Institutions	2	[27,63]
Stakeholders	2	[23,46]
Impact	2	[42,60]
Total	54[a]	

Note: The results show a higher number of documents because four documents support more than one trend.

4.3. Trends

Below, we present the main trends that emerged from the review process as well as some aspects that stand out in each one.

4.3.1. Financial Aspects

Several authors have focused on different aspects related to the financial difficulties linked to OA. They treat aspects such as the costs of OA for universities [22,28,30,42,54,67], strategies used by some publishers for addressing OA [2], financial challenges for implementing policies for incentivizing OA [19,21,22,33–36,54], and transformative agreements [43].

In [28], the authors investigated the administrative costs incurred by 29 UK universities to publish research results in OA in the "Gold" or "Green" route. The analysis considers the author's time, peer review time, administrative staff time, and academic management time. It identifies that the time and cost invested by universities in the Gold Route is 2.5 times greater than in the Green route, which is faster and less costly. While universities recognize the importance of increasing access to their research results through OA, they are also concerned about the costs involved. On the other hand, [30] determined that the costs of Article Processing Charges (APC) and subscription to journals have increased in HEIs from the UK identifying that the increase in the adoption of OA has created pressures in terms of costs for these institutions. Furthermore, they observed that there is a correlation between the price of APC and the quality of the journals.

For his part, in [67] the author analyzes the impact of indexed Gold OA journals to increase available content, citation frequency, and access to content published by recognized publishers, universities, society, or specific groups. It identifies

the growth of Gold OA journals and their inclusion in indexing and abstract databases. However, this research also highlights the fact that costs of access to academic databases offered to universities have not decreased.

In [54]. the author examined the OA Textbooks and Open Educational Resources processes to promote, evaluate, preserve, curate, and facilitate its access as well as the funding required to incorporate them into repositories in the United States. He found that the main role that libraries have is to make this kind of education resources affordable and accessible [54]. Moreover, a study carried out on South African universities, mentioned some challenges associated with the dissemination of knowledge resulting from African research due to the new APC financing model, which implies that institutions must assume this additional cost to traditional research funding, research staff salaries, and journal access licenses available to the educational community. It also highlights risks of low visibility in the absence of resources to assume the costs of the new financing model [42].

On the other hand, in [2] the authors highlight how commercial publishers appropriated academic communication spaces, turning them into an inelastic market that hinders the circulation of knowledge. It finds two sources of financing: either the authors pay APCs; or they find external sponsorships and grants through advertising, donations, and fundraising. These two routes relate to the gold and green access models, respectively.

In [21], the author reported the important steps to the growth of OA in the UK, in 2012. Initially, the Fintech recommendations to the Minister of Universities and Science, followed by the new UK Research Councils policy adopting the Finch Group recommendations. The Fintech report recommended increasing access through three mechanisms – OA journals, extensions to licensing, and repositories. That being said, the Fintech report had some adverse reactions mainly because of the lack of price regulation, and because they estimated some transitions that institutions cannot afford because of the limits of their budgets [21].

Following a survey from 2019 to adopt central funds for the payment of OA in high education institutions, the authors repeated the survey in 2011 finding there were not big advances to encourage the OA publications in 26 UK universities [22]. The possible reasons stated were high APC publication charges, difficulty to associate publishing costs with a related grant, the administrative overhead associated with managing micropayments for APCs, difficulties associated with raising awareness of the fund amongst authors, and perception challenges amongst researchers linking OA with lower quality of the results presented in the investigations which affect current levels of demand [22].

A group of authors focused their attention on the "Austrian Transition to Open Access (AT2OA)", authorized by the Austrian Federal Ministry of Science, Research and Economy [19,33,35,38–41]. According to [29] in this project, running from 2017 to 2020, "the 21 public universities cooperate to promote OA through concerted measures" [29]. This initiative included four subprojects that focused on: "drafting an expert's report about the financial impact of a total adjustment to OA on a national and institutional basis, the extension of existing consortium licenses with an OA component, establishment of an OA publishing fund as well as the promotion of OA monographs and suitable OA infrastructures" [33]. In contrast, in [34] the authors deal with "a number of questions that were discussed in the course of the drafting of a guidance document around funding conditions for publication funds in the project" (p. 1). While in [36] the authors describe "the prerequisites for an equitable and appropriate use of central funds for such transitional business models, and present a case study of a currently running, successful agreement". Finally, in [35] the author summarizes the methods and findings of the final report of the AT2OA-Transition-Study.

Finally, in [43] the author focuses on transformative agreements as an advance towards "a new global publishing regime where scholarly communication is available and free for all readers with internet access". According to the author, this change has been primarily caused by "bottom-up pressures from university systems and researchers as producers of knowledge, but it also represents a reflection of many new adventures towards the imagination of the intricate and complex relationships between science and society" [43]. Special attention is paid to the OA2020 project, which "calls on the publishing houses to flip their journals from being 'pay to read' to 'pay to publish' [43]. Examples of institutions and countries committed to the OA2020 route are the University of California system and China, as well as Sweden, Germany, the Netherlands, and Norway which are putting pressure on Elsevier to "come to the negotiating table to strike up a 'pay-to-publish' model rather than the 'pay-to-read' one" [43]. Another example of an alternative negotiation model is the SCOAP3 [South African OA publishing] project, coordinated by the European Organization for Nuclear Research (CERN). The project includes a centralized fund. "The project negotiated with the publishers of all the key particle physics journals to enter the kind of agreement captured by OA2020. Each country contributes to the central fund an amount, which is directly proportional to its international share of the publications output, which for South Africa is 0.5% of the global output, and therefore a national contribution of 0.5% of the total central fund. This makes particle physics articles, even in the most high-impact, prestigious journals, openly accessible to all. CERN manages the payment for the publication of the articles. It is an example of diamond OA where there is no APCs or 'author facing charges" [43].

4.3.2. Repositories

Works on repositories focus on diagnosing existing repositories [56,70], presenting specific projects or initiatives for putting in place repositories [29,37,41,51,52,56,61], or in studying the challenges faced by researchers and other stakeholders for embracing OA. These challenges could be partly mitigated, if they took advantage of repositories [44,62,64], while others show the challenges for fully taking advantage of repositories [45,58].

In [51], the authors studied the growth of repositories around the world from 2006 to 2012. However, repositories were not well known in that time. That being said, there were universities' efforts to constitute meta-search engines, catalogs, and directories to make available valid and reliable information. In the Latin-American case, the author presents advances of open repositories in Costa Rica, Peru, Argentina, and Mexico. It also includes the

infrastructure development, its geographic distribution, and some policy initiatives. Then again, in [56] the authors executed a diagnosis and classification of 30 institutional repositories in Ecuador. As a result of the study, they identified that most of the repositories correspond to databases managed by private HEIs, with DSPACE being the most widely used software. However, only 11.37% of IRs use a Creative Commons license, and 7.23% do not indicate the type of license they use.

In the same way, in [52] the authors shown an overview of OA repositories of HEIs in Mexico. They analyzed eleven repositories, which have the highest levels of resources in OpenDOAR. The authors conclude that these repositories are an important alternative for the visibility of resources and keep a growing tendency in the quantity and quality of the resources included in OA repositories. Nevertheless, these authors also noted that there are important challenges shared by the repositories analyzed. It is worth mentioning that according to the authors, there is a Mexican Network of Institutional Repositories (REMERI by its initials in Spanish), which has defined General Guidelines for OA. REMERI declares the repository program, aiming at collecting, preserving, and ensuring OA to Scientific, Technological and Innovation Information Resources generated mainly with public resources [52].

At the same time, Singh studies the development of OA repositories in higher education and research institutions in India [37]. The study analyzes the growth, development, geographical distribution, and technical characteristics of the repositories such as software, size, and type of content, OA policy, and protocol for metadata collection. It shows that university libraries have implemented OA IRs to improve service to users and to obtain benefits for both the universities and the academic and research communities in terms of visibility and dissemination of research material to the community.

Researchers from Uganda report an OA project called African Higher Education Research Online (AHERO) to establish a better communication of research among HEIs, making more accessible information resources. This repository is semi-automated and focuses on African education. An online survey to evaluate AHERO was made among 26 countries; and the results showed this repository has been successful especially among Uganda, South Africa, and Kenya users, and could represent the start point to encourage OA among disciplinary communities in Africa [41].

Moreover, McCutcheon and Eadie present the development of IR from the University of Glasgow in the UK as a resource for managing OA [29]. The repository considers the UK's OA policy, the University's OA flow, and the provisions defined by funding funds such as the UK Charities Open Access Fund, the European Union, other funder policies, and internal reporting requirements. Additionally, they considered the national standards for OA Metadata proposed by RCUK (Research Councils UK), REF (Research Excellence Framework), the EU, the CASRAI (Consortia Advancing Standards in Research Administration Information) UK Open Access Working Group, as well as internal requirements.

In 2010, in [61] the author presented that reported to the Kenya Information Preservation Society (KIPS) an initiative to make available bibliographic details and abstracts of electronic theses and dissertations, since 1999. KIPS had a collection of approximately 15,000 citations in the database and 24 organizations contributing records on all disciplines. The author concludes that: "it is asserted that digital theses and dissertations massively increase the impact of institutional research, an essential consideration at a time when evidence of the impact of research is becoming an all-important factor in research evaluation, and therefore in future funding [61].

In [44], the authors focused on OA practices (OAPs) and the challenges faced by African researchers. Among these challenges, there are some for accessing materials, such as the unawareness of "Resarch4Life availability, inadequate staffing by LIS[1] professionals and infrastructure challenges". They also faced challenges to publish their research papers in OA. According to the authors, these, among other issues, "are the reasons why most African academic institutions are unable to embrace OAPs" [44]. Furthermore, indicated "OAPs can provide a roadmap in research paper publication in academic institutions, which then can be discoverable, implementable and achievable, also demonstrate the value of academic libraries to their institutions and hopefully provide continual sources of funding." [44].

The approach proposed by the authors in [64] analyzes the roles of stakeholders in strengthening the OA initiatives among academic libraries in Zimbabwe. According to the study, "faculty participants confirmed that academic libraries in Zimbabwe could effectively use IRs to generate new e-content and explore opportunities to license their content". They also found that "Academic institutions are also benefiting from OA initiatives as shown by the growing strength of a research culture in all institutions. It is imperative to take into consideration several challenges ranging from costs, an unwillingness by authors to generate content, unclear systems to coordinate and promote the sharing of knowledge, lack of support from institutional management, and copyright issues" [26].

In the same sense, the library's role in encouraging open repositories was studied in Nigeria and its possible contributions to reach Vision 20: 2020 objectives of economic growth. An important conclusion of the study established the lack of dialogue among stakeholders to encourage the accessibility and utilization of research outputs and intellectual products at Nigerian academy institutions. Conclusion: there is a clear need for a National Policy to set up OA repositories. It is necessary to establish national coordination and implementation mechanisms for the policy, as the international and national cooperation among high educational institutions adopt best practices and models to develop open repositories and to contribute effectively into Nigeria's economic objectives [62].

In [58], the authors point out that despite the great benefits that the development of a single OA portal brings to society, this is not always possible as indicated by the Slovenia's case in which a complete unification was impossible due to the differences in the

[1] LIS: Library and Information Sciences

control of plagiarism recognition in documents deposited in institutional repositories.

In different circumstances in [45], the authors mentions that one of the greatest challenges for repositories is to insert in the academic culture a favorable environment towards open research data, since collecting, managing, and publishing for future reuse are not phases traditionally carried out. This implies an important change in which, since the commercial publishers are not present because they are not part of their business model, new dialogues between the academic communities are necessary to create stimuli that facilitate their acceptance.

4.3.3. Education

In this category, are the studies that focused on the perception [47,68] and use of OA to support the teaching process in higher education [49,55,69], both classroom-based and distance-learning.

In [47], the authors study OA enabling courses in Australian HEI's[2] institutions "The study examined the following areas: learning barriers faced by students, student engagement and experience in learning, skills developed, and further skills needed while undertaking OA courses, motivation to complete study, career pathway of students, key reasons for selecting particular pathways" [47].

For their part, in [69] the authors analyze the impact of academic literature OA on the teaching process of the higher education sector in the United Kingdom. Among the perceived benefits is the easy and quick access to content to support teaching without the need to pay additional licensing fees. However, knowledge barriers persist around the location, permanence, legality, and licensing of articles. Thus, they recommend encouraging researchers to license their content to facilitate its reuse [69]. Moreover, [49] conducts a case study at the Federal University of Minas Gerais in Brazil to identify whether the OA to information resources provided by university libraries, contributes to distance learning. Although the libraries indeed offer databases, collections, and digital libraries of Theses and Dissertations, only two resources are from OA, which implies making investments to cover the fees for accessing some databases. In this sense, the potential of OA is evident: institutions can find alternative ways to access content and strengthen the democratization of knowledge and education [49].

Similarly, in [68] the authors conducted a study of the experiences and perceptions of OA held by teachers teaching online in different humanities, social science, and science programs at institutions in the United States, Australia, India, and the United Kingdom. The study showed that participants perceive OA in five different ways, specifically as: 1) a teaching resource, 2) a publishing channel, 3) a social movement, 4) open source; and 5) "free to me" (p. 128). Additionally, each form presented six variations of perception: 1) materials, 2) costs, 3) platforms, 4) benefits, 5) drawbacks, and 6) professional context or purpose. Considering that, there are divergent understandings of OA, publishers, university libraries, and higher education administrators are encouraged to propose strategies to improve dissemination, communication, and further study of aspects of OA [68].

In [55], the authors established the emergence of open education in 2001 with the MIT OpenCourseWare project; and make a historical and philosophical account, highlighting the most important courses around the World. After presenting a statistical analysis of 14 current journals in curriculum studies, the study concludes that they reflect globalization and economic systems in the educational offer; however, authors question the possibility to create spaces with a different perspective, encouraging knowledge as a social good [55].

In Latin America, Tzoc concludes that universities in the region should work in three scenarios to promote OA: 1) increase of institutional agreements and research projects between universities; 2) increased participation of Latin American teachers in professional networks or foreign projects, and 3) proliferation of open, distance, or virtual education programs [65].

4.3.4. Perceptions

There is an important group of studies that analyze the perception and adoption of OA by the academic community. This group focuses on attitudes towards OA [32,39,48,59], as well as on the use of resources and services available [48,57].

In [39], the authors study how well the teaching community working in HEIs in Tamil Nadu (India) endorse the OA publishing model. The study shows that most of the teaching staff prefer the OA publishing model to the other commercial and hybrid publication models [39].

In contrast, in [32] the author examined the perceptions, development, and conceptions of OA practices in HEIs in the United Kingdom, through an ethnographic study. The results of the study focused on the barriers perceived by the academic community affecting the adoption of OA. The main perceived barriers are operational, technological, political, or legal, and community knowledge and attitudes towards OA. Addressing these barriers will involve developing different strategies that generate cultural acceptance of possible forms of open knowledge dissemination [32].

In [59] the author analyzes the results of the survey conducted by French consortium Couperin (Unified Consortium of Higher Education and Research Organizations for Access to Numerical Publications). The survey aims to understand researchers' practices and opinions regarding OA and the use of social networks. The results show the importance of promoting information related to editorial policies and intellectual property rights. Also, it identifies the opportunity to link the OA and social networks, to promote open science, and increase the visibility of researchers.

For his part, in [57] the author studied the student's adoption of OA online higher education services in Egypt. The study has divided into two phases: 11 deep interviews of information directors of some universities, and senior managers of the

[2] The search by author, title or year in My Library (2019) OA enabling courses are preparatory [44]

Ministry of Higher Education, and the second phase consisted of student surveys from three universities (public, private, and foreign) [57]. The interviews were structured to test and support the model and the elements that determine the adoption of OA online education. These elements are: a) internal attributes, composed by Facilities and resources, faculty and staff support and quality and risk control b) External, integrated by Socio-economic factors, Market structure, and competition and government support; c) The perception of open innovation attributes has five elements: relative advantage; compatibility; observability; trialability and complexity; d) Faculty innovation adoption and e) Students innovation adoption [57].

Moreover, in [48] the authors presented the results of an OA perception survey of the scientific literature, as well as the use and citation of OA sources. Brazilian researchers responded to the survey. They found that although the researchers know and support the principles of OA, there is low awareness and use of OA repositories. One of the causes identified was the confusion regarding the article's copyright s that could be published in a repository since researchers believe that the papers that should be available are post-prints. Thus, they recommend increasing the awareness of the research community on these issues.

4.3.5. Tools

In this trend, the works related to tools aimed at facilitating the development, use, or access to digital resources are included. A group of works focuses on tools that integrate different resources [38,40,50], while others propose tools to complement the integration efforts [58] and others work on the elements necessary for developing and publishing open academic courses, including the tools that would allow different national standards to be compatible [66].

Swain focuses on Virtual Libraries in Higher Education in modern India and their role in the "formation and dissemination of information and knowledge" [36]. This work focuses on INFLIBNET, which "may be given the status of master repository having linkage with all institutional repositories" [36]. INFLIBNET is the Information and Library Network Centre initiated by the University Grants Commission in 1991[3].

In [50], the authors describe the development of a tool that facilitates data integration between the Institutional Repository (IR) of the Federal University of the Amazon, the scientific initiation management system (Lira portal), and the faculty portal. This work contributed to the automation of the flow of deposit and dissemination of scientific initiation projects in the University's IR as a strategy to promote the OA.

In [38], the authors describe the development of a Research Archive Information Retrieval and Visualization System for Indian universities. This system allows access and consultation in the OA of digital resources in the Bengali language. From the prototype created, the aim is, among other results, to promote OA through the dissemination and international visibility of research results.

In [58], the authors point out that before the integration of institutional repositories in a single search engine, it is important to develop expert plagiarism detection systems based on the design of algorithms that make comparisons between documents. This system was developed in Slovenia, including an exhaustive analysis of documents. This analysis considers linguistic qualities, identification of similarities higher than 60%, and results show. These results show the first five documents that stand out for exceeding the originality threshold. Furthermore, making visible criteria such as visits, downloads, previous readers' ratings, among others.

In [66] the authors reviewed, in Greece, the ASK-CDM-ECTS. This tool facilitates the authoring and publishing of open academic courses using the CDM-ECTS metadata description. Their analysis involves a review of Norway, USA, India, UK, China, and Japan initiatives and the required structure composed of 18 elements to constitute an open course. Another interesting element is the consideration of standards to describe an open course. In this case, a subset in XML format, compatible between two standards was proposed: the CDM standard developed by the University of Oslo for the Norwegian for higher education, and the ECTS standard adopted by high educational institutions in Europe [66].

4.3.6. Policies

The development of the OA has given way to the definition of policies in HEIs. The analyses performed and trends in policy development for OA promotion [45], knowledge communication [31], open courses [25], and commercial exploitation of research results [53].

In [45], the authors show that the OA generated a globalization scenario around education policy thanks to two strategic allies, such as governments and libraries. The first, defined policies with an emphasis on the promotion of OA, compliance frameworks, and infrastructure proposals like the creation of portals of OA; and the second, libraries, with the creation of institutional repositories, and the design of strategies to strengthen the culture of self-archiving in the academic community.

In [31], the authors analyze the use and implementation of the UK Scholarly Communication Licensing policy (UK-SCL) as a strategy that encourages OA and allows control of the content produced in HEIs. The results identify perceived benefits in helping authors gain control of their publications, disseminate publications rapidly and widely, and reduce costs at institutions. On the other hand, the perceived challenges identified are resistance from editors, administrative demands, and confusion among researchers [31]. Also, the factors influencing policy adoption are institutional collaborations, external support, internal communication and engagement, the legal framework, business processes, and the role of the library within the institution [31].

In [53], the author studied the OA advances and situation and proposed an agenda to encourage the legislation to avoid the commercial exploitation of research results and to allow the growth of OA. The proposed agenda encourages higher education researchers and leaders to advance OA by lobbying Congress to pass the FRPAA legislation. The proposed agenda encourages higher education researchers and leaders to advance OA by

[3] https://inflibnet.ac.in/about/index.php Accessed on September 17th, 2020.

lobbying Congress to pass the FRPAA legislation. Furthermore, it works with funding agencies to obtain policies to support OA publication rates; establish institutional funds to support LO publication charges not covered by grants; approve OA mandates on campus and support local IRs in depositing research, and publishers in the formalization of agreements that allow wide and unrestricted distribution of research.

Finally, OA scholarly publishing as mega-journals and academic networking services as in the case of Massive Online Open Courses (MOOCs) currently represent an innovative and disruptive educational form. A study in the UK and USA shows the emergence of policy initiatives to regulate these new forms of academic scope. It is important to notice the economic impact encouraged by this kind of information and educational industry, transforming the traditional academic markets [25].

4.3.7. Institutions

Some authors have focused their attention on the changing roles [27], policies, and practices of the institutions participating in the OA Ecosystem [63].

A case study carried out in the United Kingdom reference how, since 2005, this country has an OA policy [27]. Also, the financing of APC has evolved, since, at first, it was the UK Research Council that was directly in charge of providing the resources individually requested. However, since 2013, these costs are managed and allocated by the research organizations to which the researchers are ascribed. The resources are assigned in a single item, to these organizations. In turn, they oversee transparently managing and allocating them, according to criteria of coverage of disciplines and researchers, at different stages of their career, in compliance with the OA policy.

In [63] the authors study the role of HEIs in the United Arab Emirates (UAE) OA uptake. Their results suggest that although "there seems to be a generally positive perception of OA", there is also a moderate adoption of OA policies and practices. They also point out the need to create a culture that increases priority for OA adoption, as well as "awareness, policies, best practices, and infrastructure" [63]. They also suggest the need for "nation-wide strategies aligned with international initiatives such as OA2020 and Plan S need to be adopted" if the UAE is to increase its presence in the OA movement [63].

4.3.8. Stakeholders

This trend presents analysis on different stakeholder groups, such as New University Presses [23] and Australian universities [46], presenting the challenges and opportunities these stakeholder groups face.

In [23], the authors studied the University of Huddersfield Press (the Press), one of the 21 New University Presses (NUPs) in the UK, the vast majority of which publish OA journals and monographs. Taylor establishes six key areas on which "the Press has focused on since 2016 to enable a professional, efficient and effective publishing process that suits our journal and monograph portfolio: Commissioning; Review; Production; Discoverability; Marketing; and Analytics. Furthermore, the author emphasizes the importance for OA publications to be transparent, given the rise of predatory OA journals, and to be recognized as academic publications of high standards and professional quality. According to [23], the Press aims at optimizing the publishing process to achieve a "streamlined workflow, increased discoverability, an improved level of service for authors and editors, an improved reader experience, and ultimately, a steady and measurable increase in downloads and citations" [59, p. 10]. Finally, The Press has proposed a Press Model for university presses.

In [46] the author carried out in a study carried out in Australian universities, points out how the OA scenario is being unduly exploited by agents who develop practices that could be classified as unfair, which generates new actions on the part of the universities who must accompany researchers by providing advice and new tools to identify those journals that meet the desired quality and reputation standards to send their research results.

4.3.9. Impact

In this trend, we find the works that study how OA has altered research [42] and knowledge management activities [60].

In [42], the authors highlight that the dynamics of the research culture have changed from a linear structure- in which the purpose of an investigation is the publication of its results-, to give way to the creation of virtuous circles, around sharing and reusing research data which can be advantageous for developing countries.

A study published in 2010, analyzed the impact of OA on the Socialization, Externalization, Combination, and Internalization (SECI) model from Nonaka applied for higher education institutions. The study concludes that OA is a positive factor for knowledge Management because it makes scientific literature openly accessible for anyone interested in the creation, retrieval, and transfer of scientific knowledge. The OA contributions are resumed as the acceleration of the spiral of knowledge creation; Improvement of transfer into wide dissemination mechanisms; OA also changes the method of storing and retrieving scientific online information; now in repositories; and finally, a potential decrease in the costs of scholarly communication [60].

5. Discussion

This study looks at the current interests of researchers in the OA field in the HEIs sector. As mentioned before, although we had recently studied this subject [1], updating it, showed new trends in OA research in education. We had previously found that researchers were focusing their work mainly on the repositories, journals, tools, and perceptions of the use of OA resources. However, our new study allowed making explicit new emerging trends. Notably, we found that financial issues related to the implementation of OA are now one of the main subjects in which current efforts are focused. Also, the study of policies intended to foster the appropriation of OA at different levels has been the object of attention. In addition, the technological tools that allow OA to work continue to evolve. Furthermore, the changing role of institutions during the implementation of some policies has also been studied, as well as the changes OA has brought to Education. Finally, the impact OA has had on scientific research has also been studied.

Likewise, some previous works have identified research trends in OA. Table 1 presents the main trends identified by [9–11], where some have similarities with the trends identified in this work. Notably, the Finances trend, presents some similarity with what has previously been identified as Business models around OA, although this trend includes many more aspects such as the costs of OA for universities [22, 28, 30, 42, 54, 67], and transformative agreements [43].

Another trend that emerged from this work was the one related to Repositories, which in [10] in included in the OA publication channels, together with OA Journals and OA Books, and in [11] appeared in a Sub-category called "Manifestations of OA", together with OA journals and "Awareness of Open Access Manifestations". However, in this work, the studies around Repositories appear to be the focus of great interest among researchers.

The Education trend encompasses studies on the perception [47, 68] and use of OA to support teaching processes in higher education [49, 55, 69]. Although previous works mention the participation of Higher Education in OA dynamics [9], they focus on mandates rather than on the teaching processes.

The Perceptions trend was also found as having several works focusing on attitudes towards OA [32, 39, 48, 59], as well as on the use of resources and services available [48,57] by the academic community. Previous studies had identified the Acceptance of OA as well as the Arguments promoting OA [9].

The Tools trend includes works related to the tools aimed at facilitating the development, use, or access to digital resources are included, by integrating different resources [38, 40, 50], complementing the integration efforts [58], presenting the elements necessary for developing and publishing open academic courses, and allowing different national standards to be compatible [66]. Previous works don't seem to address the tools as such but in the relation to the actors that use some OA resources [11] or see them as a part of the technical barriers [11]. On the other hand, we identify in the results associated advances the tools used in repositories of Latin American libraries in developed countries (Argentina, Brazil, Costa Rica, Ecuador, Peru, and Mexico). These advances are associated with the roles, barriers, and benefits presented for developing and transitional countries category [11].

Next, the Policies trend is the one that coincides with previous works that have identified Policies and Mandates [9, 10] as an important research trend.

In terms of institutions and stakeholders, we observed that previous studies have identified actors and their roles [9,11], where Authors and Librarians are included. In addition, in Scholarly Communication, identified by [11] the Commercial publishers are included. Also, the Libraries Sub-category [11] includes the role that Libraries may have as Publishers when they "offer Open Access publishing in Open Access Journals and act as mediators in the depositing of content in Institutional Repositories through Authors Self-archivation" (p. 83). This last aspect shows the changing roles we have identified as being at the essence of the Institutions and Stakeholders trends.

The last trend we identified is the one related to the impact OA has had on research [42] and knowledge management activities [60]. There is a similarity with a previous work [11] that includes Manifestations, Barriers and Benefits of OA as a trend, and mentions "Availability, Usability, and Visibility" of disseminated content (p. 45), which could be seen as an advantage for researchers as it positively impacts the knowledge management cycle. The two impact studies we identified state that OA has accelerated the creation, storing, retrieving, and dissemination processes creating a virtuous circle in benefit of scholarly communication [38]. However, not much work was identified in this trend, which focuses on changes in processes linked to OA scholarly communication, and not to changes in impact factors, where there are many works that have been done [71, 72].

6. Conclusions

The OA is a movement that has grown in recent years. This study focused on the analysis of the research developed on OA in HEIs to identify main research trends. The analysis initial phase was based on the SCOPUS database consultation. The result of this consultation, allowed the identification of nine trends: repositories, financial aspects, education, and perception of the academic community on the use of OA resources, supporting tools, policies on OA, stakeholders, institutions, and impact of publications on OA.

Through this study it is also determined that in addition to technology, the human and financial resources are paramount in the implementation of OA platforms. In this area, it is essential to recognize the central role that HEIs play in the implementation of the OA. Not only because of the institutional efforts concerning platforms, repositories, journals, and the articulation of their internal actors (mainly libraries, academics, and students) but, because of their negotiating power with publishers and their role as public policy actors in the matter. In any case, this work identifies the concern of HEIs about investment in the Gold route and the APC financing model for the dissemination of knowledge.

As we have shown, some works show the academic community and the public sector's acceptance of the orientations that emerged from civil society with the pillar initiatives of the OA, called BBB (Budapest. 2002; Bethesda. 2003; Berlin. 2003). These were nurtured in subsequent years with guidelines that have come from different latitudes, promoting scenarios for their enhanced development, and have therefore, contributed to an important globalization effect in the design of national OA policies and the homologation of standards by educational institutions. In turn, this provides benefits in terms of clarity around the expected contributions of the State and University actors. For example, the State is in charge of approving regulations and financing national OA repositories; while the University is in charge of the construction of institutional repositories thanks to the work of the libraries. In both cases, making strategic contributions to strengthening the OA culture in academic communities.

In the trend of education, we observe that OA is perceived as a resource to support the teaching process. Furthermore, the contents in OA are used both in classroom-based and distance-learning.

It is then not surprising that diverse studies have been made about perceptions. These studies have mainly focus on the opinions of teachers and students, who have expressed their concerns throughout surveys in UK, Egypt, India, Brazil, and France. Universities have led these studies to determine the level of knowledge, attitudes, and perceptions of the academic community towards the services provided by the OA resources, being the institutional repositories, the resources that are mainly analyzed. Although the results of these surveys showed a generalized acceptance of the OA publishing model and its advantages, they also show that there are different ways of perceiving OA, and that some of these perceptions could constitute barriers to its development [9–12] given the operational, technological, political, or legal challenges involved. There is evidence that highlights the academic community's general lack of information and understanding surrounding the appropriate use of publishing licenses and OA repositories, even if it has also been determined that there exists a part of the community that does know and make use of the repositories. In [46], the authors it was noted that there are challenges to improve the visibility of the content that these resources manage, and that it is necessary to improve practices in the use of schemas of indexing descriptors (metadata). Moreover, although progress has indeed been made in the development and use of platforms and software for the management of digital collections, more efforts are necessary on the development and use of interoperability standards that facilitate the integration of OA with existing resources. This will improve visibility and access to content.

In terms of policy development on the OA, efforts led by civil society and the academic community has gotten back some lost domain over their research results. However, it remains a challenge to maintain and fund projects where freedom is not only for society but also for the academic community, releasing it from the responsibility of the APC. Public policies on APC should facilitate the allocation of public resources to finance them. Similarly, it is important to create fiscal policies and incentives to support the private sector.

It is also important to point out that a reorganization of two great forces is taking place: on the one hand, are those publishing companies that have generated an industry from the provision of access services to academic information and therefore, seek to maintain their profitability. On the other hand, HEIs are concerned with a process of advocating and providing information on OA. It is also evident that many of the new forms of OA reported in the reviewed literature, are proposed by the commercial publishers as a response to this global movement, to stay alive in the market in which they are intermediaries. This opens new scenarios on how to connect their research results with society, which shows the rethinking of relationships and discussions about how that knowledge is disseminated and used.

Finally, it has been established the OA has encouraged all steps in the knowledge process, changing the structure of publication and improving the dissemination mechanisms, which are all accelerators of the knowledge spiral and is potentiating scholarly communications.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The Ministry of Science Technology and Innovation Colombia (MinCiencias), the Universidad Pontificia Bolivariana [587C-05/2023]; the Universidad El Bosque, Universidad Autónoma de México, and the Universidad Cooperativa de Colombia [INV2564] supported this research.

The authors also want to acknowledge the valuable comments made by the reviewers; whose input greatly contributed to the strengthening of this work.

References

[1] M.A. Osorio-Sanabria, P.L. Barreto, T. Alcantara-Concepcion, A. Jaime, "Open access research trends in the education sector: A literature review," in Iberian Conference on Information Systems and Technologies, CISTI, IEEE Computer Society, 2020, doi:10.23919/CISTI49556.2020.9140935.

[2] M.R. Ghane, M.R. Niazmand, A. Sabet Sarvestani, "The citation advantage for open access science journals with and without article processing charges," Journal of Information Science, 2019, doi:10.1177/0165551519837183.

[3] Redalyc, "Declaración de México a favor del ecosistema latinoamericano de acceso abierto no comercial," Andes, **29**(1), 2018.

[4] Leslie Chan, Sato Sho, Itsumura Hiroshi, Budapest Open Access Initiative (BOAI), Interlending & Document Supply, **30**(2), 89–96, 2002, doi:10.1108/ilds.2002.12230bab.012.

[5] Bethesda Statement BS, "Bethesda Statement on Open Access Publishing," 2003.

[6] S.M. Planck, Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, Revista Negotium, **0**(10), 89–91, 2008.

[7] OEI, Declaración de Santo Domingo: La ciencia para el siglo XXI: una nueva visión y un marco de acción. Sala de lectura CTS+I, 1999.

[8] ICML, Declaración de Salvador sobre "Acceso Abierto": la perspectiva del mundo en desarrollo, Salvador, Bahía, Brasil, 2005.

[9] M. Collins, "Open Access Literature Review 2008–9," Library Resources & Technical Services, **55**(3), 138–147, 2011.

[10] G.F. Frosio, "Open Access Publishing: A Literature Review," SSRN Electronic Journal, 2015, doi:10.2139/ssrn.2697412.

[11] K.-I. Andersson, Developing a theory of open access: a grounded theory based literature review, Master's Thesis in Library and Information Science, Swedish School of Library and Information Science, 2016.

[12] OECD, Making Open Science a Reality OECD, 25, Paris, 2015, doi:10.1787/5jrs2f963zs1-en.

[13] P. Suber, Open Access, MIT Press, 2012.

[14] FOSTER, Open Access Definition | FOSTER, Feb. 2019.

[15] K. Petersen, S. Vakkalanka, L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," in Information and Software Technology, Elsevier: 1–18, 2015, doi:10.1016/j.infsof.2015.03.007.

[16] B. Kitchenham, S. Charters, D. Budgen, P. Brereton, M. Turner, S. Linkman, M. Jørgensen, E. Mendes, G. Visaggio, Guidelines for performing Systematic Literature Reviews in Software Engineering, Durham, 2007.

[17] PRISMA, PRISMA Statement, 2009.

[18] M.M. Mashroofa, W. Seneviratne B A Senior, A. Librarian, Open access initiatives and institutional repositories: Sri Lankan scenario, 2016.

[19] B. Kromp, M. Seissl, T. Zarka, "Austrian Transition to Open Access (AT2OA) – ein Überblick," Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare, **72**(1), 28–34, 2019, doi:10.31263/voebm.v72i1.2274.

[20] M. Blanca, E. Liberta Bonilla, Impacto, impacto social y evaluación del impacto, 2007.

[21] M. Hall, "Green or gold? Open access after Finch," Insights: The UKSG Journal, **25**(3), 235–240, 2012, doi:10.1629/2048-7754.25.3.235.

[22] S. Pinfield, C. Middleton, "Open access central funds in UK universities," Learned Publishing, **25**(2), 107–114, 2012, doi:10.1087/20120205.

[23] M. Taylor, "Mapping the Publishing Challenges for an Open Access University Press," Publications, **7**(4), 63, 2019, doi:10.3390/publications7040063.

[24] E. Gadd, D. Troll Covey, "What does 'green' open access mean? Tracking twelve years of changes to journal publisher self-archiving policies," Journal of Librarianship and Information Science, **51**(1), 106–122, 2019,

doi:10.1177/0961000616657406.
[25] R. Wellen, "Open access, Megajournals, And moocs: On the political economy of academic unbundling," SAGE Open, **3**(4), 2013, doi:10.1177/2158244013507271.
[26] M. Jubb, "The 'Finch Report' and the transition to Open Access: Long term monitoring of progress in the United Kingdom," Information Services and Use, **34**(3–4), 189–193, 2014, doi:10.3233/ISU-140732.
[27] N. Pontika, D. Rozenberga, "Developing strategies to ensure compliance with funders' open access policies," Insights: The UKSG Journal, **28**(1), 32–36, 2015, doi:10.1629/uksg.168.
[28] R. Johnson, S. Pinfield, M. Fosci, "Business process costs of implementing 'gold' and 'green' open access in institutional and national contexts," Journal of the Association for Information Science and Technology, **67**(9), 2283–2295, 2016, doi:10.1002/asi.23545.
[29] V. McCutcheon, M. Eadie, "Managing open access with EPrints software: a case study," Insights, **29**(1), 1–8, 2016, doi:10.1629/uksg.277.
[30] S. Pinfield, J. Salter, P.A. Bath, "A 'Gold-Centric' Implementation of Open Access: Hybrid Journals, the 'Total Cost of Publication,' and Policy Development in the UK and Beyond," Journal of the Association for Information Science and Technology, **68**(9), 2248–2263, 2017, doi:10.1002/asi.
[31] J. Baldwin, S. Pinfield, "The UK Scholarly Communication Licence: Attempting to Cut through the Gordian Knot of the Complexities of Funder Mandates, Publisher Embargoes and Researcher Caution in Achieving Open Access," Publications, **6**(3), 1–28, 2018.
[32] G.J. Johnson, "Cultural, ideological and practical barriers to open access adoption within the UK Academy: an ethnographically framed examination," Insights the UKSG Journal, **31**(0), 2018, doi:10.1629/uksg.400.
[33] B. Bauer, "Austrian Transition to Open Access" 2017–2020," VOEB-Mitteilungen, 1–6, 2019.
[34] V.C. Capellaro, C. Kaier, "Funding conditions for open access publication funds: Open access - 'All or nothing'?," VOEB-Mitteilungen, **72**(1), 74–88, 2019, doi:10.31263/voebm.v72i1.2278.
[35] V.G. Fessler, "Expansion of open access at Austrian universities: Financial needs for the years 2019-2021. Summary of the final report of the HRSM AT2OA Transition-study," VOEB-Mitteilungen, **72**(1), 35–49, 2019, doi:10.31263/voebm.v72i1.2275.
[36] V.B. Kromp, F. Koren-Wilhelmer, "Funding open access transition models: Guideline for publishing agreements for the open access transformation," VOEB-Mitteilungen, **72**(1), 66–73, 2019, doi:10.31263/voebm.v72i1.2277.
[37] P. Singh, "Open access repositories in India: Characteristics and future potential," IFLA Journal, **42**(1), 16–24, 2016, doi:10.1177/0340035215610131.
[38] B.K. Roy, S.C. Biswas, P. Mukhopadhyay, "BURA: An open access multilingual information retrieval and representation system for Indian higher education and research institutions," Library Philosophy and Practice, **2017**(1), 2017.
[39] G. Ramadoss, D. S, "Open Access Publishing Model: Preferences, Opportunities, and Challenges – An Opinion Survey Among Teaching Staff in Higher Education Institutions in Tamilnadu," Library Philosophy and Practice (e-Journal), 2613, 2019.
[40] S.C. Swain, "Formation and Open Access of Institutional Electronic Research Corners for Promotion of Communication Facilities and Quality Research," International Journal of Emerging Technologies in Learning (IJET), **15**(06), 192–199, 2020, doi:10.3991/ijet.v15i06.12003.
[41] B. Sekabembe, J. Ssempebwa, Bridging the Knowledge Gap for African Researchers through Open Access Publishing: The Case of African Higher Education Research Online (AHERO), 95–103, 2011, doi:10.1007/978-3-642-19715-4_10.
[42] L. Czerniewicz, S. Goodier, "Open access in South Africa: A case study and reflections," South African Journal of Science, **110**(9–10), 2014, doi:10.1590/sajs.2014/20140111.
[43] A.C. Bawa, "South Africa's journey towards open access publishing," Biochemist, **42**(3), 30–33, 2020, doi:10.1042/BIO20200029.
[44] R.T. Enakrire, J.M. Ngoaketsi, "Open access practices: roadmap to research paper publications in academic institutions," Library Hi Tech News, **37**(5), 13–15, 2020, doi:10.1108/LHTN-01-2020-0003.
[45] V. Picasso, L. Phelan, "La evolución del acceso abierto a la investigación y a los datos en la educación superior en Australia," RUSC Universities and Knowledge Society Journal, **11**(3), 128–141, 2014, doi:10.7238/rusc.v11i3.2076.
[46] K. McNaught, "The changing publication practices in academia: Inherent uses and issues in open access and online publishing and the rise of fraudulent publications," Journal of Electronic Publishing, **18**(3), 1–1, 2015, doi:10.3998/3336451.0018.308.
[47] M. Shah, M. Cheng, "Exploring factors impacting student engagement in open access courses," Open Learning, **34**(2), 187–202, 2019, doi:10.1080/02680513.2018.1508337.
[48] A.C.M. Furnival, N.S. Silva-Jerez, "Percepções de pesquisadores brasileiros sobre o acesso aberto à literatura científica," Informacao e Sociedade, **27**(2), 153–166, 2017, doi:10.22478/ufpb.1809-4783.2017v27n2.32667.
[49] E.M. de Oliveira Costa, J. Santa Anna, "Acesso aberto e educação a distância : novas configurações para a democratização do conhecimento," **3**, 536–546, 2019.
[50] D.S. Moreira, L.A.S. da Silva, A.E. Yanai, "Interoperabilidade de sistemas para viabilização do acesso aberto aos relatórios de iniciação científica da UFAM," Ciência Da Informação, **48**(3 (Supl.)), 490–491, 2019.
[51] S.I. Adame, L. Lloréns Baez, M. Schorr Wiener, "Retrospectiva de los repositorios de acceso abierto y tendencias en la socialización del conocimiento educativa: REDIE.," Revista Electrónica de Investigación Educativa, **15**(2), 148–162, 2013.
[52] G.C.T. Sepúlveda, M.M. Reyes, A.S. Martín, "Open access repositories in higher education institutions in Mexico, An initial review using the SCOT methodology," Informacion, Cultura y Sociedad, **40**(junio), 117–130, 2019, doi:10.34096/ics.i40.5317.
[53] P. Renfro, "Open access within reach: An agenda for action," Journal of Library Administration, **51**(5–6), 464–475, 2011, doi:10.1080/01930826.2011.589351.
[54] K. Okamoto, "Making Higher Education More Affordable, One Course Reading at a Time: Academic Libraries as Key Advocates for Open Access Textbooks and Educational Resources," Public Services Quarterly, **9**(4), 267–283, 2013, doi:10.1080/15228959.2013.842397.
[55] J.A. Corrigan, N. Ng-A-Fook, "Mobilizing curriculum studies in a (Virtual) world: Open access, edupunks, and the public good," Canadian Journal of Education, **35**(2), 58–76, 2012.
[56] E.M. Bodero, M. De Giusti, C.D. Radicelle, E.P. Villacrés, "Análisis de los repositorios digitales institucionales de Acceso Abierto en el Ecuador Analyses of institutional Open Access digital repositories in Ecuador Contenido," Revista Espacios, **40**, 1–9, 2019.
[57] M. Mourad, "Students' adoption of an open access online education service An exploratory study in an emerging higher education (HE) market," Online Information Review, **34**(4), 604–617, 2010, doi:10.1108/14684521011073007.
[58] M. Ojsteršek, J. Brezovnik, M. Kotar, M. Ferme, G. Hrovat, A. Bregant, M. Borovič, "Establishing of a Slovenian open access infrastructure: A technical point of view," Program, **48**(4), 394–412, 2014, doi:10.1108/PROG-02-2014-0005.
[59] C. Okret-Manville, "Academic Social Networks and Open Access: French Researchers at the crossroads," LIBER QUARTERLY, **25**(3), 118–135, 2016, doi:10.18352/lq.10131.
[60] S. Bernius, "The impact of open access on the management of scientific knowledge," Online Information Review, **34**(4), 583–603, 2010, doi:10.1108/14684521011072990.
[61] F.C. Ratanya, "Electronic theses and dissertations (ETD) as unique open access materials: Case of the Kenya Information Preservation Society (KIPS)," Library Hi Tech News, **27**(4), 15–20, 2010, doi:10.1108/07419051011083190.
[62] K.N. Igwe, "Open access repositories in academic and research institutions for the realization of Nigeria's Vision 20: 2020," International Journal of Information Science and Management, **12**(1), 33–46, 2014.
[63] M. Boufarss, M. Laakso, "Open Sesame? Open access priorities, incentives, and policies among higher education institutions in the United Arab Emirates," Scientometrics, **124**(2), 1553–1577, 2020, doi:10.1007/s11192-020-03529-y.
[64] C.T. Chisita, B. Chiparausha, "Open Access initiatives in Zimbabwe: Case of academic libraries," The Journal of Academic Librarianship, **45**(5), 1–8, 2019, doi:10.1016/j.acalib.2019.102047.
[65] E. Tzoc, "El Acceso Abierto en América Latina: Situación actual y expectativas," Rev. Interam. Bibliot. Medellín (Colombia), **35**(1), 83–95, 2012.
[66] D.G. Sampson, P. Zervas, "Supporting open access to European academic courses: The ASK-CDM-ECTS tool," Campus-Wide Information Systems, **30**(1), 44–62, 2012, doi:10.1108/10650741311288814.
[67] G. Machovec, "The Impact of Gold Open Access Journals in Aggregators and Abstracting/Indexing Services," Journal of Library Administration, **56**(7), 875–883, 2016, doi:10.1080/01930826.2016.1216227.
[68] A.M. Salaz, N. Johnston, C. Pickles, "Faculty Members Who Teach Online: A Phenomenographic Typology of Open Access Experiences," Journal of Academic Librarianship, **44**(1), 125–132, 2018,

doi:10.1016/j.acalib.2017.09.006.
[69] E. Gadd, C. Morrison, J. Secker, "The impact of open access on teaching-how far have we come?," Publications, **7**(3), 2019, doi:10.3390/publications7030056.
[70] V. Martinovich, "Argentine open access scientific journals with international circulation: An analysis from Pierre Bourdieu's field theory," Informacion, Cultura y Sociedad, **40**(junio), 93–116, 2019, doi:10.34096/ics.i40.5540.
[71] S. Gunasekaran, S. Arunachalam, "The impact factors of open access and subscription journals across fields," Current Science, **107**(3), 380–388, 2014, doi:10.18520/cs/v107/i3/380-388.
[72] I.K. Razumova, A. Kuznetsov, "Impact of open access models on citation metrics," Journal of Information Science Theory and Practice, **7**(2), 23–31, 2019, doi:10.1633/JISTaP.2019.7.2.2.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 500-505 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Simulated Annealing for Traveling Salesman Problem with Hotel Selection for a Distribution Company Based in Mexico

Raúl Jiménez-Gutiérrez*, Diana Sánchez-Partida, José-Luis Martínez-Flores, Eduardo-Arturo Garzón-Garnica

Department of Logistics and Supply Chain Management, UPAEP University, 17 Sur 901, Barrio de Santiago, CP 72410 Puebla, Puebla, México

ARTICLE INFO

Article history:
Received: 16 July, 2020
Accepted: 29 August, 2020
Online: 24 September, 2020

Keywords:
TSP
TSPHS
TSP-HOBS
Simulated Annealing
Metaheuristic

ABSTRACT

A distribution company in Mexico covers the travel expenses for 21 sales representatives. Currently, the routes they follow are not established clearly, which can lead to high costs in this subject. A reduction of such cost is sought after, by optimizing the routes for each one of them. The following research finds an improvement on the routes for the sales representative of a distribution company in Mexico. It was done by using the Traveling Salesman Problem with Hotel Selection or Base selections via a Simulated Annealing algorithm. The results show an improvement in a reasonable timeframe by using the Simulated Annealing. It also shows that the maximum process time was of 156.63 minutes, and the least amount of improvement was 24.44% over the current route selection. Applying this model will be beneficial for the company as the company is trying to reduce costs related to the sales representatives such as; fuel cost, hotel cost, and travel expenses.

1. Introduction

For the agro-industry, to satisfy the specific demands of customers using different strategies in product and service has proven successful in an increasingly competitive global market.

Logistics has been an essential part of the success of the companies. Nowadays, there is more focus on how companies can optimize their resources and do more with less, a perfect balance between efficiency and customer service. According to [1] logistics are vital because it creates value, for customers and suppliers of the firm, and value for the firm's stakeholders. Value in logistics is expressed in terms of time and place.

Logistics have several improvement methods that could be implemented, not only in the logistics areas of the companies but also in commercial areas. Mexico is not an exception where TSP models can be helpful as the lack of road infrastructure and the long distances, besides the facts mentioned there some companies trying to make savings by reducing the sales force and extending the region coverage of the sales representatives.

As an example, we can find the Traveling Salesman Problem (TSP), where rapid growth of the companies raises the need to implement commercial strategies to visit their clients with the highest possible efficiency and quality but with the least cost. It is at this point is where mathematical models show up to create a clear vision and achieve such goals. Moreover, in the case of the TSP, it can help the companies to reduce the fixed costs for the sales representative such as:

1. Fuel costs.
2. Pay toll costs.
3. Maintenance costs.
4. Accommodation costs.
5. Meal expenses.
6. GPS costs.
7. Other travel expenses.

To develop study cases related to having route efficiency is essential for the companies established in Mexico. According to the National Association of supermarkets and department stores, during 2019, Mexico has grown 7.7% [2]. Furthermore, part of this growth is due to the effort of the sales force.

Therefore, for the current research, a variant for the TSP will be implemented, the Traveling Salesman Problem with Hotel Selection (TSPHS) with Simulated Annealing. The main objective for a TSPHS is to have an efficient route, including the stops where the sales representative will rest after their work shifts as they need, in many cases, to visit several clients.

*Corresponding Author: Raúl Jiménez-Gutiérrez, 15 pte 317, +52 2212673274, Raul.jimenez02@upaep.edu.mx

2. Problem Description

A distribution company based in Mexico is responsible for delivering products to a nationwide authorized dealer's network. Goods are sent using established couriers, and the distribution center is based in the central region of the country. All sales are made through authorized dealers to assure the quality of service. The company tries to keep a high level of service for the dealers and has several sales representatives throughout the country. The role of each sales representative is to establish a sales plan with each dealer, take orders, assure such orders have been fulfilled, and present the dealers with technical information and training, both for existing products as well as new products. The company decided to divide the country's sales territory into eleven commercial zones to have an efficient approach, and to know and cover the needs of the dealer's network. Fig. 1 shows the current zones within the country, the distribution company provided the information as part of its current structure of commercial zones.



Figure 1: Sales Representative Zones

Each commercial zone has a sales representative in charge. The sales representative, the primary task is to visit each dealer of the zone he is in charge of and know the needs of each dealer. He must provide feedback about the sales plan that the dealer has for the year, as well as other tasks that include a demonstration of new products, special deliveries (this term is used when a sales representative brings support to the final user and explains how to assemble, use and give maintenance) and training.

It is worth to mention that the company covers the travel expenses for the route or the visits that they need to do. It means that the company covers the gasoline, toll payment, hotels, meals, and other possible expenses that the sales representative incurs during the visits in its entirety. The sales representative needs to make all of the visits as traveling during his working hours, which are set to nine working hours per day. The information provided by the company did not include the time each sales representative spends with the dealer/client. However, the variation in travelling time between the calculated average and the real one was found to be enough to cover the time of the visit. It was decided to leave the visit time out of the calculations, to ease the solving of the model.

Despite the facts above, there is no specific route for the sales representative. Therefore, the order in which to drive to each dealer store to make their visits is entirely up to the sales representative. In several cases, they pass through the same point twice. Being inefficient on the route design could increase the expenses for the company. As an example, the sales representative in charge of the south region is in charge of 19 dealers. Back in 2019, he drove 57081 km, equivalent to MXN 98,011.00 in gasoline expenses only.

The main objective of this research is to find an improvement in the efficiency of the routes that each sales representative travel among their specific commercial zones, using a well-known combinatorial optimization problem, the Traveling Salesman Problem with hotel selection (TSPHS) solved by a simulated annealing method, focused on finding the following objectives:

- Improvement of the routes.
- Find the hotel that allows the sales representative to rest or, in some cases, go back to their homes but still being efficient on the route.

As part of the improvement strategy to have better routes, the distribution company conducted and applied a survey to the sales representatives to know their preferences about staying in a hotel or at their home. All sales representatives agreed that they prefer to get back to their homes, but they know that, in most cases is not possible to return to their homes due to the long distances that they need to drive.

Table 1: Hotel and Base Data

ZONE	HOTEL	HOME
M11	7	1
M12	5	1
M13	3	1
M14	7	1
M21	5	1
M22	5	1
M23	3	1
M24	13	1
M31	4	1
M32	2	1
M33	5	1

The sales representatives commented as well that, in some cases, they suffer from stressful situations, as there are several times when they do not have the chance to get back to their homes, and when they try to book a hotel, there are no available rooms. According to [3], sales representatives are susceptible to stress by the nature of their job. Therefore, this research includes the hotel and the home of the sales representative that for this purpose is called *base*.

The results of the aforementioned internal survey were compiled into a table with the number of hotels commonly used in each zone. The table is the following:

It is worth to mention that, the agent has a list of authorized hotels that the company selected previously. All of those hotels are in the same price range; therefore, the sales representative can choose either one, without it representing a considerable difference in cost.

Hence the limitations for the research will be to stablish the routes for each sales representative.

3. Literature Review: Travelling Salesman Problem with Hotel Selection and Simulated Annealing.

The problem of this research has been studied in several research cases, and it is well described in the literature. The TSP is one of the most well-known combinatorial optimization problems, where the main idea is to find an optimal solution to avoid passing the same point twice. Over the years, several methods based on deterministic or probabilistic heuristics have been proposed to solve the known problem. The method includes branch-and-bound, particle swarm optimization, neural network, and simulated annealing [4].

In simple words, the TSP consists of determining a minimum distance circuit passing through each vertex once and only once. Such a circuit is known as a tour or Hamiltonian circuit [5].

A formal definition of TSP, according to [6]:

Let $K_n = (V_n, E_n)$ be the complete undirected graph with n = [Vn[nodes and $m = |E_n| = \binom{n}{2}$ edges. An edge e with endpoints i and j is also denoted by ij, or by (i,j). We denote by $\mathbb{R}^{E_n}$ the space of real vectors whose components are indexed by the elements of E_n. The component of any vector $z \in \mathbb{R}^{E_n}$ indexed by the edge $e = ij$ is denoted by Z_e, z_i, or $z(i,j)$.

Now, considering that the salesperson, in some cases, needs to travel long distances, the need for a hotel stay arises. Such stay is an everyday activity, to end his daily route and start his route the next day to visit another dealer. Selecting a hotel to end/rest/start the route converts the research into a traveling salesman problem with hotel selection.

According to this research, there are several cases where the Traveling Salesman Problem with Hotel Selection (TSPHS) has been studied; Vansteenwegen [7] introduced the concept of TSPHS, the difference between TSP and TSPH, the last one is where a mobilized entity (salesman) cannot visit all locations (customers) at once without taking any break (break at a hotel location). It is because the salesman has limited working hours; the target at the end is to reduce the total travel time.

The TSPHS can be defined [7] on a complete graph $G = (V, A)$ where V is the set of vertices and $A = \{(k,l) \mid k,l \in V, k \neq l\}$ is the set of arcs. There are s available hotels (which are indexed through $k = 1 \ldots s$) and n customers (which are indexed through $k = (s+1) \ldots (S+n)$) In the set V. Each customer k is assigned service or visiting time T_k where T_k is $k = 0, \forall = 1, \ldots,$. The time kl_c needed to travel between locations k and l is known for all $(k,l) \in A$.

Several examples of TSPHS can be found in the literature as well as new heuristics solution methods. As it is mentioned in [8] that the TSPHS is a difficult optimization problem that arises in many practical situations, in this case, the authors gather the data into four sets:

Table 2: Customer and Hotels Data

Set	Number of Customers	Number of hotels
1	48-288	6
2	10	2
	15	
	30	
	40	
3	52 - 1002	3
		5
		10
4	52 - 1002	10

As can be seen, it is similar to the distribution company based in Mexico, but instead of having four sets, it has eleven commercial zones with a specific number of customers and hotels.

The new metaheuristic that they implemented is a two-phase multi-start procedure that combines a simple GRASP (Greedy randomized adaptive search procedures) and a variable neighborhood descent to improve the solution. They find that for smaller instances, the method was able to find either optimal solutions or best solutions, but also that the difficulty increases if the number of hotels grows.

Another example of solving the TSPHS with a new metaheuristic solution can be found with a Migrating Bird Algorithm (MBO) [9], and the solution approach is based on the flight formation of migrating birds, using this method it is shown that it is efficient for solving the problem in a reasonable computational time. Furthermore, in this paper, it is mentioned that simulated annealing was run, and it got results in a reasonable timeframe as well. For the TSPHS with Migration Bird Algorithm, the result shows that the gap between the MBO and the exact solution is acceptable.

Other methods available include Particle Swarm Optimization [10], [11], Artificial Bee Colony [12], Ant Colony [13], Crow Search Optimization [14], Cat Swarm Optimization [15], Grey Wolf Optimization [16], Whale Optimization [17]. The advantages/disadvantages of these methods are all well documented in literature. Some of these methods have successfully been used to solve TSP [12], [13], [16]. And even some have been applied in combination with Simulated Annealing [11]. In previous (unpublished) work by the authors, the Simulated Annealing algorithm was programmed and tested. The advantages and disadvantages of applying each of the aforementioned methods were evaluated. But the largest consideration was still the cost of implementing, or programming, such algorithms. In all the cases, the cost, expressed mainly in programming time, of implementing the newer algorithms was higher than:

1. The available time the authors could allocate for the research.
2. The deadline the company had to respect.

Both these factors resulted in a cost that was higher than the benefits of implementing a newer algorithm, even if it would eventually yield better results. Still, an improvement could be obtained by implementing a simpler, cheaper (in this case) to apply, method.

A simple example of solving the TSPHS [18] that was formulated as a mixed-integer linear programming (MILP).

$$\min \sum_{d=1}^{m} \sum_{k=0}^{s+n} \sum_{l=0}^{s+n} x_{k,l,d}\, C_{k,l} \tag{1}$$

$$s.t. \sum_{l=0}^{s} x_{0,l,1} = \sum_{k=0}^{s+n} x_{k,0,m} = 1 \tag{2}$$

$$\sum_{h=0}^{s} \sum_{k=0}^{s+n} x_{k,h,d} = \sum_{h=0}^{s} \sum_{l=0}^{s+n} x_{h,l,l} = 1, \qquad d = 1, \dots, m \tag{3}$$

$$\sum_{k=0}^{s+n} x_{k,h,d} = \sum_{l01}^{s+n} x_{h,l,d+1,}\ d = 1, \dots, m-1\ ; h = 0, \dots, s \tag{4}$$

$$\sum_{d=1}^{m} \sum_{k=0}^{s+n} x_{k,i,d} = 1, \qquad i = (s+1), \dots, s+n \tag{5}$$

$$\sum_{k=0}^{s+n} x_{k,i,d} = \sum_{l=0}^{s+n} x_{i,l,d}\ \ d = 1, \dots, m; i = s+1, \dots, s+n \tag{6}$$

$$\sum_{k=0}^{s+n} \sum_{l=0}^{s+n} x_{k,l,d}\, (c_{k,l} + T_l) \le C, \qquad d = 1, \dots, m \tag{7}$$

$$u_i - u_j + 1 \le n-1) \left(1 - \sum_{d=1}^{m} x_{i,j,d}\right)\ i = s+1, \dots, s+n; j = s+1, \dots, s+n \tag{8}$$

$$x_{ijd} \in 0,1, \qquad d = 1, \dots, m;\ i = 0, \dots, s+n;$$

$$j = 0, \dots, s+n \tag{9}$$

$$u_i \in 1, \dots, n, \qquad i = s+1, \dots, s+n \tag{10}$$

The objective function (1) minimizes the total travel time. As mentioned above, the number of trips, m, is not a decision variable in this formulation, but a given parameter of the problem. An alternative formulation could be to minimize the number of trips first and then minimize the travel time. Constraint (2) guarantees that the tour starts and ends in the first starting hotel 0. Constraints (3) ensure that each trip starts and ends in one of the available hotels $(0, \dots, s)$. Constraints (4) guarantee that, if a trip ends in a given hotel, the next trip starts in the same hotel. Constraints (5) ensure that every customer is visited once, and constraints (6) verify the connectivity inside each trip. Constraints (7) limit the time budget of each trip, and constraints (8) are required to prevent sub tours.

For this instance, [18] the goal of the TSPHS is to determine a tour of minimal length that visits all n customers. The tour is composed of connected trips with limited length, and every trip should start and end in one of the available hotels. The hotels that are selected determine to a great extent, the length of the total tour.

Combining the TSPHS with another strategy can yield an advantage, in recent research, it can be found that the simulated annealing as a complementary for effective results. In [19] can be found that Kirkpatrick introduced the implementation of simulated annealing. The process begins by considering a solution space $\boldsymbol{S}$ of an exclusive tour through the set of given cities or points $\mathbf{1}, \mathbf{2}, \dots, \boldsymbol{n}$, with a set of updated solutions $\boldsymbol{x_1'}$ created by randomly switching the orders of two cities. The energy function, or fitness function, which represents the length of the route $\boldsymbol{x_i}$, is denoted by $\boldsymbol{f(x_1)}$. The relative change Δ in cost $\boldsymbol{f}$ between $\boldsymbol{x_i}$ and $\boldsymbol{x_i'}$ is expressed as $\Delta \boldsymbol{f} = \frac{f(x_i') - f(x_i)}{f(x_i)}$.

Besides, it is mentioned in [20] that using the potential usefulness of a fixed temperature algorithm is considered suitable for small problem instances. On the same document [20], where they implement a Simulated Annealing (SA) to solve a TSP, the optimal fixed temperature is chosen experimentally by running simulated annealing many times at each of several fixed temperatures and determining which temperature is best, according to an appropriate optimality criterion. As a result, they see that fixed-temperature simulated annealing tends to outperform the use of a fast cooling schedule, but for TSP instances of size less than 150 cities [20].

4. Solution Approach

The only information the authors had in hand was the list of coordinates of the clients, hotels, and base of each zone. However, a detailed record of the agents' travels was not available. As the travel time was to be limited, because of the maximum work time, and to be able to know at which moment a hotel had to be selected, some kind of transformation was needed. The geographical data had to be converted to travel time. The process is explained next.

A sample set of location pairs, covering as many zones as possible, was selected. For each pair of locations, the traveling time was sought, using Google Maps®. Also, the Euclidean distance for that pair was computed. For all pairs, the relationship between the Euclidean distance, the Google Maps® distance, and the travel time was set up in a table. The traveling time varied between 3 minutes and 917 minutes. The Google Earth distance varied between 650 meters and 1,153,000 meters. The Euclidean distance varied between 413 meters and 954,496 meters. The proportion varied between 0.0006 and 0.007 and averaged 0.0013. Therefore, multiplying the Euclidean distance by the proportion's average should yield an approximation to the travel time between both locations.

A simulated annealing method was implemented instead of getting an exact solution due to the fact - that previous research shows that simple TSP models with 50 locations take too much time to get a solution or, even worse, do not get a solution using currently available computing equipment. Additional to this, using metaheuristics yields a result reasonably close to those of the exact methods [21]. As the solution requires the addition of

the hotel selection, which adds even more complexity to the model solving, a quicker solving method was needed. Such is the reason for the use of heuristics instead of trying to obtain an exact solution.

As was already mentioned, a Simulated Annealing algorithm was used. It was programmed in GNU Octave®. The objective of such an algorithm was to find an improvement in the total time of travel for each sales agent. It was, therefore, improving the operating costs of the company. Some considerations had to be taken to achieve such a goal.

If the agent has to choose a hotel and the travel time from his current location to the nearest hotel is larger than the time to his base, including a 30-minute tolerance, the agent will return to his base.

The sum of all the distances traveled, including the times the traveler returns to his base, were considered to evaluate the objective function. According to [22], SA simulates a collection of atoms in equilibrium at a set temperature. SA starts with an initial state. A random change is made to this state, and the change in energy, ΔE, is calculated. If the new state has lower energy than the previous state, i.e.

$$\Delta E \leq 0$$

the new state is carried into the next iteration. However, if the new state has high energy, it is accepted with probability $P(\Delta E) = \exp(-\Delta E / K_B T)$ where T is the temperature and K_B it is Boltzmann's constant. This acceptance of an uphill move (a move with $\Delta E \geq 0$) is analogous to a higher energy molecule knocking lose a molecule trapped in a state of excess energy Determining the cooling (annealing) schedule in step 3(f) of SA is crucial to the success of the algorithm. The chance of accepting an energy-increasing move is inversely proportional to ΔE and proportional to T, which decreases with time.[22]

The following steps were taken into account in order to generate a new solution,

- The current best solution is obtained, and the hotels and returns to the base are removed. The remaining list includes only client locations.
- Two points are obtained randomly, and an exchange, or flip, is performed.
- The base is added at the beginning and end of the route.
- Each leg is traveled, calculating the travel time through the previously obtained time-distance relationship.
- When the accumulated time to arrive at the next client exceeds the maximum journal duration, a hotel or base is added. It is done by obtaining the middle point between both clients and finding the hotel or base closest to such point. If the travel time to the base is lesser than to a hotel, or even if it is more extensive, up to a tolerance of 30 minutes, the base is selected. Then the resting point is added, and the journal time reset.

5. Problem Solution

The first solution was considered as the order in which each client appears in the database. The SA model was run for each zone, finding an improvement in each case. A graph showing the behavior of the model in one of such runs is shown in Figure 2. Improvements were found for all the zones, the execution time and improvement of all the zones can be observed in Table 3.

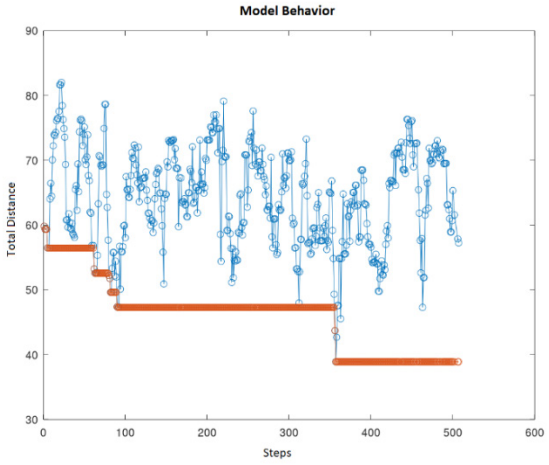


Figure 2: Model Behavior

Table 3: Results of the model's execution

Zone	Init OF	Final OF	Improvement	Runtime (minutes)
M11	54.3782	38.4468	29.29%	114.77
M12	76.8986	56.0846	27.06%	113.06
M13	34.2869	19.4949	43.14%	63.25
M14	102.632	73.683	28.20%	231.82
M21	39.8978	18.2448	54.27%	69.53
M22	41.5431	22.5259	45.77%	89.24
M23	18.4146	12.1156	34.20%	77.32
M24	130.139	84.894	34.76%	266.01
M31	54.583	41.2396	24.44%	156.63
M32	28.2553	14.4192	48.96%	68.57
M33	75.500	48.4674	35.80%	177.82

An example of the resulting routes is shown in Figure 3, where the return of the agent to his base can be appreciated. To ease the reading of this document and to avoid disclosing information that could lead to the identifying of the company or disclosing its clients, the other tables and routes will not be shown.

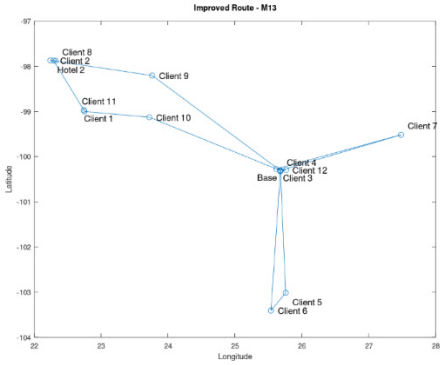


Figure 3: Sample Route

The model was run on a laptop PC. Intel® Core i5, 8Gb RAM. As shown in table 3, the execution times varied from 63.25

minutes to 266.01 minutes, or from roughly one hour to a close to four and a half hours. The full execution of the model for all zones did not take more than two full days. An attempt to solve the model precisely by the use of a linear solver was also made, but the execution took more than 100 hours without results, so the SA results were the only ones available.

6. Conclusions

An improvement in the travel time and hotel stays of the sales representatives for the distribution problem of an international company with a presence in Mexico was sought. The scope of the problem included all the Mexican territory, divided into eleven commercial zones. A Simulated Annealing method was selected to try and find better results than the current ones in a reasonable time frame.

The only information available was the location of each client, hotel, and base, but the problem needed a way to relate those to travel times. A relationship was obtained by calculating the travel time within a small number of locations. In this way, the relationship between Euclidean Distance and Travel Time was obtained.

A Travelling Salesman Problem with Hotel or Base Selection was implemented, using the aforementioned Simulated Annealing method, for each of the commercial zones. The model was run, and new routes were obtained for each zone. The maximum process time was of 156.63 minutes, and the least amount of improvement shows a 24.44% improvement over the current route selection.

6.1. Scope for future research

Using more precise mathematic models to find the relationship between the distance and the travel time can improve the efficiency of the calculation of travel time when only the coordinates are known. It could mean that more straightforward calculations would yield more precise results when travel time is needed.

References

[1] R.H. Ballou, "Business logistics: importance and some research opportunities," Gestão & Produção, **4**, 117–129, 1997, doi:10.1590/s0104-530x1997000200001.

[2] ANTAD, Annual report, Nov. 2019, Accessed on May 2, 2020. [Online]Available on: https://www.antad.net/informeanual/2019/estadisticas.html

[3] J.K. Sager, "A structural model depicting salespeople's job stress," Journal of the Academy of Marketing Science, 1994, doi:10.1177/0092070394221007.

[4] X. Geng, Z. Chen, W. Yang, D. Shi, K. Zhao, "Solving the traveling salesman problem based on an adaptive simulated annealing algorithm with greedy search," Applied Soft Computing Journal, 2011, doi:10.1016/j.asoc.2011.01.039.

[5] G. Laporte, "The traveling salesman problem: An overview of exact and approximate algorithms," European Journal of Operational Research, 1992, doi:10.1016/0377-2217(92)90138-Y.

[6] M. Jünger, G. Reinelt, G. Rinaldi, Chapter 4 The traveling salesman problem, Handbooks in Operations Research and Management Science, 1995, doi:10.1016/S0927-0507(05)80121-5.

[7] C.A. Gencel, B. Keçeci, "Traveling salesman problem with hotel selection: Comparative study of the alternative mathematical formulations," in Procedia Manufacturing, 2019, doi:10.1016/j.promfg.2020.01.270.

[8] C. Marco, S. Kenneth, Vansteenwegen, Pieter, Goos Peter, "A Simple Grasp+VND For The Travelling Salesperson Problem with Hotel Selection," University of Antwerp Faculty of Applied Economics BE Tech. Rep. out, 1-16,2012. https://repository.uantwerpen.be/docman/irua/f171bd/1c42703e.pdf

[9] K. Amirhossein, B. Mahdi, "A Migration Birds Algorithm for the travelling salesperson problem with hotel selection," Shahed University, 2015, **8**, 283-286, 2015 http://confnews.um.ac.ir/images/41/conferences/or8/237_2.pdf

[10] N. Supattananon and R. Akararungruangkul, "The Modified Particle Swarm Optimization for a Special Case of the Assignment Problem: A Case Study in Chicken Transportation," Math. Probl. Eng., **2020**, 1–15, 2020, doi:10.1155/2020/5716985.

[11] T. Bui, T. Nguyen, H.M. Huynh, B. Vo, J. Chun-Wei Lin, T.-P. Hong, "Multiswarm Multiobjective Particle Swarm Optimization with Simulated Annealing for Extracting Multiple Tests," Scientific Programming, 2020, doi:10.1155/2020/7081653.

[12] I. Khan, M.K. Maiti, K. Basuli, "Multi-objective traveling salesman problem: an ABC approach," Applied Intelligence, 2020, doi:10.1007/s10489-020-01713-4.

[13] B.P. Silalahi, N. Fathiah, P.T. Supriyo, "Use of Ant Colony Optimization Algorithm for Determining Traveling Salesman Problem Routes," Jurnal Matematika "MANTIK," 2019, doi:10.15642/mantik.2019.5.2.100-111.

[14] M. Jain, A. Rani, V. Singh, "An improved Crow Search Algorithm for high-dimensional problems," Journal of Intelligent and Fuzzy Systems, 2017, doi:10.3233/JIFS-17275.

[15] A.M. Ahmed, T.A. Rashid, S.A.M. Saeed, Cat Swarm Optimization Algorithm: A Survey and Performance Evaluation, Computational Intelligence and Neuroscience, 2020, doi:10.1155/2020/4854895.

[16] M.L. Taha, B. Al-Khateeb, Y.F. Hassan, O.M.M. Ismail, O.A. Rawash, "Solving competitive traveling salesman problem using gray wolf optimization algorithm," Periodicals of Engineering and Natural Sciences, 2020, doi:10.21533/pen.v8i3.1462.g617.

[17] M.M. Mafarja, S. Mirjalili, "Hybrid Whale Optimization Algorithm with simulated annealing for feature selection," Neurocomputing, 2017, doi:10.1016/j.neucom.2017.04.053.

[18] P. Vansteenwegen, W. Souffriau, K. Sörensen, "The travelling salesperson problem with hotel selection," Journal of the Operational Research Society, 2012, doi:10.1057/jors.2011.18.

[19] A.E.S. Ezugwu, A.O. Adewumi, M.E. Frîncu, "Simulated annealing based symbiotic organisms search optimization algorithm for traveling salesman problem," Expert Systems with Applications, 2017, doi:10.1016/j.eswa.2017.01.053.

[20] M. Fielding, "Simulated annealing with an optimal fixed temperature," SIAM Journal on Optimization, 2000, doi:10.1137/S1052623499363955.

[21] E.A. Garzón-Garnica, D.P. Cruz-Benítez, O.D. Badillo-Valenzuela, D. Sánchez-Partida, J.L. Martínez-Flores, "Automated data acquisition for a large scale Capacitated Vehicle Routing Problem," in IFAC-PapersOnLine, 2015, doi:10.1016/j.ifacol.2015.06.281.

[22] J.W. Pepper, B.L. Golden, E.A. Wasil, "Solving the traveling salesman problem with annealing-based heuristics: A computational study," IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans., 2002, doi:10.1109/3468.995530.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 507-513 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Microcontroller Based Data Acquisition and System Identification of a DC Servo Motor Using ARX, ARMAX, OE, and BJ Models

Mokhlis Salah-eddine*, Said Sadki, Bahloul Bensassi

GITIL Laboratory, Department of physics, Faculy of sciences Ain Chock Hassan II University, Casablanca, 20070, Morocco

ARTICLE INFO

Article history:
Received: 27 August, 2020
Accepted: 05 November, 2020
Online: 20 November, 2020

Keywords:
System identification
polynomial models
PEM
Data acquisition card

ABSTRACT

This paper is carrying out a practical identification of a DC servomotor from input and output data. The system's step response was obtained from an acquisition card that we have developed using a microcontroller. The data is collected and downloaded to the PC Thought RS232 Protocol to process the identification process using polynomial models (ARX, ARMAX, OE, and BJ). Afterward, a comparison between the resulting models has been made using the validation criterions R^2 and FPE. The results show that the ARMAX, OE, and BJ models have reproduced a better fit compared to the ARX model.

1. Introduction

Servomotors have been developed at the origin of different uses. For example, radio-controlled aircraft, biped robots, and robotic arms. And nowadays, they are widely used in the robotics field, due to their cost, compactness, ease of use, and their high torque-to-weight ratio, which made them a good choice to achieve a compact and less expensive mechatronic system [1]. However, DCSM does not provide a good position control when it is in contact with a load and sometimes it provides a chattering motion. Thus, a mathematical model needs to be built to analyze and develop its performances. To do so, we tried to identify the mathematical equation of the system based on polynomial models.

A mathematical model is a multitude of equations that describe the relationship between the variables of a process or a phenomenon. It can take many forms depending on the studied system, and we can distinguish between two major classes.

1.1. Theoretical models

Theoretical models, Also known as whit-box, are models, which mostly used for simulations and theoretical studies. They are based on mathematical and physical laws, which made them often complex and hard to define [2]; however, it can give a complete description of the system.

1.2. Experimental models

Experimental models, also known as grey-box or black-box models are models used to define a mathematical model from the experiment's data. As mentioned in [3], in system identification, we could distinguish between two types.

- non-parametrical model is usually presented with numerical values, a graph, or a table, and it has an infinite number of parameters and an undefined structure.
- parametrical models, which are defined models structure who have a finite number of parameters.

According to [4], identifying a system or a process is about estimating a mathematical model that reproduces a static and dynamic response to a test input as close as possible to the real system. To do so, and to get a valid model, we need to follow a certain procedure [5], described in fig.1. It starts with the experiment preparation by choosing the sampling frequency and the input signal as a first step. Then choosing the model structure type, order, and delay as a second step. After that, selects the criterion that minimizes the error between the estimated model and the experimental data as the third step. Finally, the validation step that consists of applying several tests to prove the usability of the model. The resulting models' main objectives can be to understand the behavior of the studied system [6], or building the system control loop [7].

In this paper, Sections 2 and 3 cover a brief introduction of the linear polynomial models and parameter estimation process. Section 4 and 5 highlight the necessary steps for the implementation of the application. In sections 6 and 7, the results obtained from the experiment are presented and discussed.

*Corresponding Author: Mokhlis Salah-eddine, Email: mokhlis.id@gmail.com

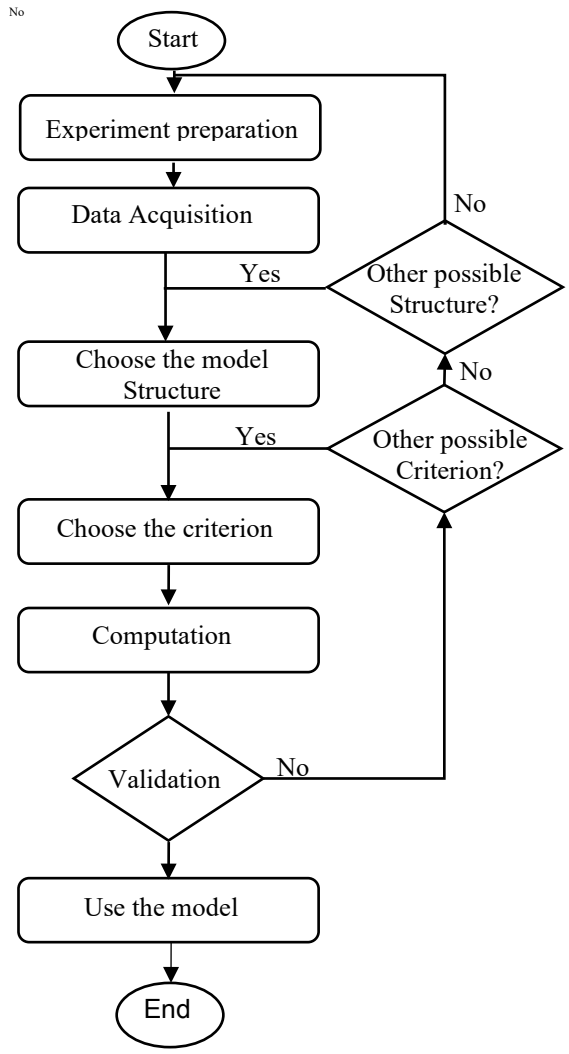


Figure 1: Experimental approach for defining a parametric model of a system

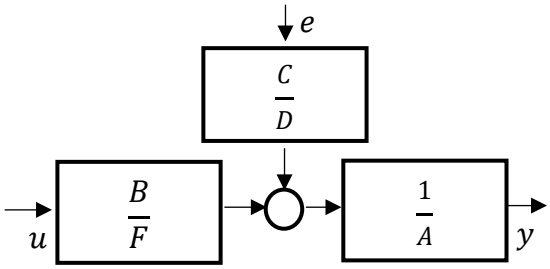


Figure 2: Generalized transfer function Structure

2. Parametrical Models

2.1. Generalities

A discrete linear system can be presented as a general-linear polynomial model, which is considered as a transfer function whose parameters need to be estimated. This estimation is done by recursive and none recursive algorithms like maximum likelihood, instrumental variable or less square. This presentation, also gives flexibility for modelling both stochastic and system dynamics, and it is widely used in many real-world applications from different disciplines, e.g., [8–10].

Polynomial models are described in [11]. They can be defined by the generalized transfer function expressed by (1) and the signal flow represented in figure 2 [5]:

$$A(q)y(k) = \frac{B(z)}{F(z)}z^{-nk}u(k) + \frac{C(z)}{D(z)}e(k) \quad (1)$$

where:

$e(k)$: *is the system disturbance*

nk: *is the system delay*

$$A(z) = 1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a} \quad (2)$$

$$B(z) = b_1 z^{-1} + \cdots + a_{n_b} z^{-n_b} \quad (3)$$

$$C(z) = 1 + c_1 z^{-1} + \cdots + a_{n_c} z^{-n_c} \quad (4)$$

$$D(z) = 1 + d_1 z^{-1} + \cdots + d_{n_d} z^{-n_d} \quad (5)$$

$$F(z) = 1 + f_1 z^{-1} + \cdots + f_{n_f} z^{-n_f} \quad (6)$$

and:

n_a : The order of $A(z)$

n_b : The order of $B(z)$

n_c : The order of $C(z)$

n_d : The order of $D(z)$

n_f : The order of $F(z)$

By setting one or more n_a, n_c, n_d and n_f to 0 can create simpler models such as:

2.1.1. ARX model

ARX structure presented in equation (7) [5] is the result of setting n_c, n_d *and* n_f to 0 the signal flow is presented in fig.3. In this structure the noise model $\frac{1}{A(z)}$ is coupled with the systems dynamics, it is used to obtain simple model with good signal to noise ratio.

$$Y_{arx}(k) = \frac{B(z)}{A(z)}z^{-nk}u(k) + \frac{1}{A(z)}e(k) \quad (7)$$

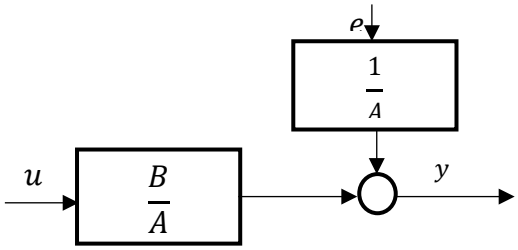


Figure 3: The ARX Structure

2.1.2. ARMAX model

The ARMAX structure presented in the equation (8) [5] is the result of setting n_d and n_f to 0 in the generalized transfer function (1), the signal flow is presented in fig.4. this structure provides more flexibility to model the noise by adding the $C(z)$ parameters

which is considered as a moving average of white noise. It is used when the noise enters in the input.

$$Y_{armax}(k) = \frac{B(z)}{A(z)} z^{-nk} u(k) + \frac{C(z)}{A(z)} e(k) \quad (8)$$

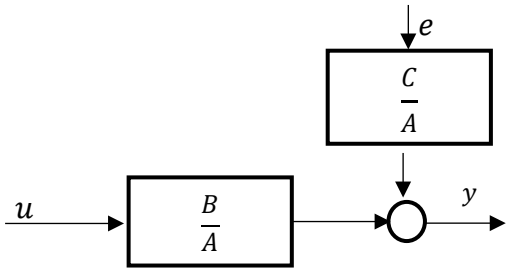


Figure 4: The ARX Structure

2.1.3. Output Error Model

The OE presented in the equation (9) [5] is the result of setting n_a, n_c and n_d to 0 in the generalized transfer function (1), the signal flow is presented in fig.5. In this structure, the noise model is considered as 1. It is used to parameterize the dynamic of the system only.

$$Y_{OE}(k) = \frac{B(z)}{F(z)} z^{-nk} u(k) + e(k) \quad (9)$$

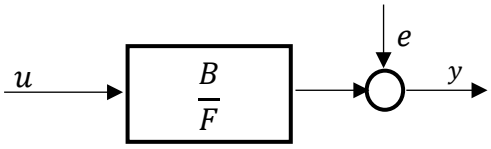


Figure 5: The OE Structure

2.1.4. Box- Jenkins Model

The BJ structure presented in the equation (10) [5] is the result of setting n_d to 0 the signal flow is presented fig.6. This structure provides more flexibility because the noise model is independent of the dynamics model. It is used when the noise enters as a measurement disturbance.

$$Y_{BJ}(k) = \frac{B(z)}{F(z)} z^{-nk} u(k) + \frac{C(z)}{D(z)} e(k) \quad (10)$$

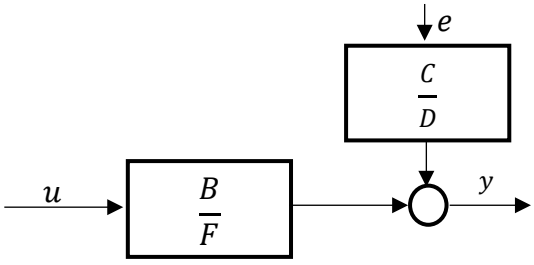


Figure 6: The BJ Structure

3. Parameters Estimation

Parameter estimation is a mathematical process that uses experimental data to estimate the system parameters, in this paper; this process is done by the prediction error method described in [12] and fig 7. this method aims to minimize the error expressed in Eq (13) between the estimated and the real system [13].

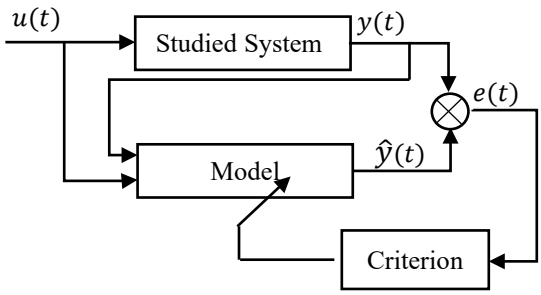


Figure 7: Prediction Error Method

According to [14], to use the Prediction Error Method, we need to choose the model structure, the one-step-ahead predictor $\hat{y}(k|k-1,\theta)$ described with the equation (12) and the criterion function $J(\theta, Z^N)$ given by Eq (11):

$$J(\theta, Z^N) = f\left(R_N(\theta, Z^N)\right) \quad (11)$$

where:

$$R_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} e(k,\theta) e^T(k,\theta)$$

$$\hat{y}(k|k-1,\theta) = H^{-1}(q)G(q)u(k) + (1 - H^{-1}(q))y(k) \quad (12)$$

And the prediction error is given by:

$$e(k|k-1) = y(k) - \hat{y}(k|k-1,\theta)$$
$$= H^{-1}(q)(y(k) - G(q)u(k)) \quad (13)$$

4. Model Validation

Model validation is the final step, which goal is to compare and validate the usability of the estimated model by testing several criterions like those in this paper:

1.3. Akaike Final Prediction Error

This criterion described in the equation (14) [5] was presented by Akaike (1967) as a final prediction error (FPE), based on Akaike's theory, the best model has the smallest FPE.

$$\mathrm{FPE} = \det\left(\frac{1}{\mathrm{N}}\sum_{1}^{\mathrm{N}} \mathrm{e}(\mathrm{t},\theta_{\mathrm{N}})(\mathrm{e}(\mathrm{t},\theta_{\mathrm{N}}))^{\mathrm{T}}\left(\frac{1+\mathrm{d}/\mathrm{N}}{1-\mathrm{d}/\mathrm{N}}\right)\right) \quad (14)$$

where:

N: The number of values in the estimation data set

$e(t)$: The vector of prediction error

θ_N: The estimated parameters

d: The number of the estimated parameters

4.1 Best Fit

The RMSE criterion measures the similarity between the data and the estimated model and can be calculated using equation (15) [15].

$$\text{Best Fit} = 100\left(1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right) \quad (15)$$

where:

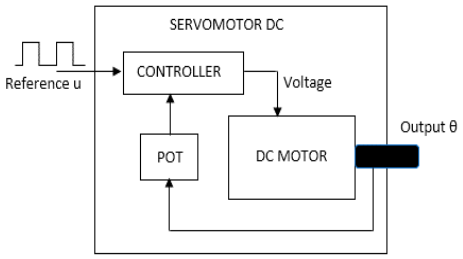
y_i : The measured output

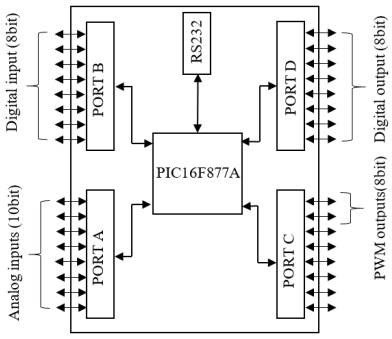
$\hat{y}$: The estimated output

$\bar{y}$: The mean of the output

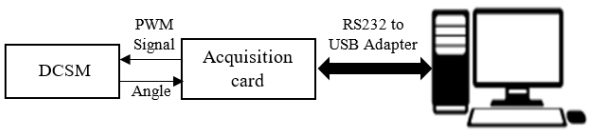
5. Application

A DCSM is composed of a potentiometer and integrated circuit that ensures a dc motor's position control based on a reference signal received from a microcontroller or an external organ (fig.8). However, despite its different applications, there is a major problem in achieving good performance with those servo motors due to the integrated circuit [1]. Therefore, the goal is to estimate all the system's internal and external components' mathematical model, to design a new controller or exploit it in more complex systems. In this application, the reference signal is in the form of a pulse with modulation signal who does not exceed 20ms between two rising edges, and the DCSM supply voltage is 5V.





To extract reliable input and output data, we build the acquisition card with pic16f877a. As shown in fig 9, this microcontroller has eight analogic inputs of 10-bit resolution, one digital input, one digital output of 8-bit resolution, and two PWM output of 8-bit resolution. Moreover, it is linked to the computer using RS232 communication with a speed of 115200bit/s to send and receive data necessary for the experiment. However, today's computer does not support this type of communication. To solve this problem, an RS232 to USB adapter is used, as shown in fig 10.



Afterward, to know the sampling time Ts, which is related to the uncertainty of the internal and external components and the instructions execution time of the microcontroller. The analogic input of the acquisition card is subjected to a periodic signal of known frequency. The samples collected from the acquisition card are represented on a graph. Afterward, the frequency of the test signal is adjusted until a signal similar to that of the input is obtained, which $Ts = 0.0034$ deduced in this application. This will allow having a speed of 300 sps/s.

Next, the analysis and parameter estimation process is applied to the collected data from the acquisition card.

6. Results

Model order determination in parameter estimation is essential because a lower order model cannot adequately describe the system's dynamics, and a higher-order model may have uncertainties. Thus, to extract the best model and to see the effect of changing the order of the parameters, we applied the experimental approach of identifying a parametric model described in fig 1, and we fixed nk to 1 then:

- Over 100 ARX models are calculated where $n_a$and n_b are in the range of 1 to 10.
- Over 125 ARMAX models are calculated where $n_a, n_b \; and \; n_c$ are in the range of 1 to 5.
- Over 25 OE models are calculated where $n_f and \; n_b$are in the range of 1 to 5.
- Over 625 BJ models are calculated where $n_a, n_a, n_a \; and \; n_a$ are in the range of 1 to 5.

Afterward, the Best-fit criterion is used to select the best five structures in each model, and the PFE criterion is used to choose the best structure in each model. Those results are in tables 1 to 4.

7. Discussion

As presented in tables 1 to 4, all the extracted models in each structure have closely a similar FPE. Then the choice of the model is based on the best-fit only. Besides, no significant change occurs in the best-fit criterion by increasing the models' order in each model type. Thus, the selected model will have the lowest order. This will allow a better processing time while doing online parameter estimation in real-time applications.

- The best-estimated ARX model is the ARX331 shown in Figure 11 and Table 1. It has an order of (nb=3 and na = 3), and a fit of 72,44
- The best-estimated ARMAX model is the ARMAX2411 shown in Figure 12 and Table 2. It has an order of (na = 2, nb = 4, nc=1), and a fit of 94,75%.
- The best-estimated OE model is the OE131 shown in Figure 13 and Table 3. It has an order of (nf = 3, nb = 1), and a fit of 95.009.
- The best-estimated BJ model is the BJ12331, shown in Figure 14 and Table 4. It has an order of (nb = 1, nc = 2, nd = 3, nf = 3), and a 95.02% fit.

Based on those results and figure 16. The ARX models provided a fit of 72%, which is a poor fit compared to ARMAX, BJ and OE Models. However, if we want to use a model in real-time applications, the best choice will be OE131 because of the resulting good fit with the measured output and fewer parameters.

Table 1: The estimated ARX models

Model Type	Model Parameters	'Best Fit'	'FPE'
'ARX431'	A(z) = 1 - 1.065 z^-1 + 0.005417 z^-2 + 0.01607 z^-3 + 0.04765 z^-4 B(z) = 3.841e-05 z^-1 + 0.002391 z^-2 + 0.001677 z^-3	72,6272093	0,3764651
'ARX441'	A(z) = 1 - 1.065 z^-1 + 0.005351 z^-2 + 0.01617 z^-3 + 0.04766 z^-4 B(z) = 3.741e-05 z^-1 + 0.002391 z^-2 + 0.002752 z^-3 - 0.001086 z^-4	72,5563214	0,37670021
'ARX421'	A(z) = 1 - 1.065 z^-1 + 0.005566 z^-2 + 0.01616 z^-3 + 0.04759 z^-4 B(z) = 3.772e-05 z^-1 + 0.004052 z^-2	72,5177552	0,37624736
'ARX331'	A(z) = 1 - 1.069 z^-1 + 0.005651 z^-2 + 0.06704 z^-3 B(z) = 0.000315 z^-1 + 0.002366 z^-2 + 0.001609 z^-3	72,4480688	0,37718277
'ARX341'	A(z) = 1 - 1.068 z^-1 + 0.005065 z^-2 + 0.06748 z^-3 B(z) = 0.0003129 z^-1 + 0.002366 z^-2 + 0.002673 z^-3 - 0.001074 z^-4	72,386256	0,37731582

Table 2: The estimated ARMAX models

'Model Type'	Model Parameters	'Best Fit'	'FPE'
'ARMAX3421'	A(z) = 1 - 2.598 z^-1 + 2.202 z^-2 - 0.6041 z^-3 B(z) = 0.001024 z^-1 + 0.002355 z^-2 - 0.008141 z^-3 + 0.004814 z^-4 C(z) = 1 - 1.576 z^-1 + 0.5755 z^-2	94,8596855	0,36096562
'ARMAX2421'	A(z) = 1 - 1.985 z^-1 + 0.9851 z^-2 B(z) = 0.001351 z^-1 + 0.0005912 z^-2 - 0.006229 z^-3 + 0.004419 z^-4 C(z) = 1 - 0.9752 z^-1 - 0.02485 z^-2	94,7854814	0,36124302
'ARMAX4231'	A(z) = 1 - 2.362 z^-1 + 1.557 z^-2 - 0.02015 z^-3 - 0.1743 z^-4 B(z) = -0.0004067 z^-1 + 0.0004638 z^-2 C(z) = 1 - 1.312 z^-1 + 0.1561 z^-2 + 0.156 z^-3	94,7787217	0,36258016
'ARMAX2411'	A(z) = 1 - 1.986 z^-1 + 0.9859 z^-2 B(z) = 0.001361 z^-1 + 0.0006902 z^-2 - 0.006188 z^-3 + 0.004261 z^-4 C(z) = 1 - z^-1	94,7591717	0,3611956
'ARMAX3411'	A(z) = 1 - 2.022 z^-1 + 1.059 z^-2 - 0.03639 z^-3 B(z) = 0.001229 z^-1 + 0.0009276 z^-2 - 0.00656 z^-3 + 0.004529 z^-4 C(z) = 1 - z^-1	94,7457863	0,36131993

Table 3: The estimated OE models

'Model Type'	Model Parameters	'Best Fit'	'FPE'
'OE431'	B(z) = 0.01849 z^-1 - 0.05449 z^-2 + 0.05356 z^-3 - 0.01755 z^-4 F(z) = 1 - 2.947 z^-1 + 2.894 z^-2 - 0.9475 z^-3	95,0341979	6,13354962
'OE131'	B(z) = 7.423e-06 z^-1 F(z) = 1 - 2.933 z^-1 + 2.867 z^-2 - 0.9339 z^-3	95,0098583	6,28005407
'OE341'	B(z) = 0.01923 z^-1 - 0.03923 z^-2 + 0.02007 z^-3 F(z) = 1 - 2.14 z^-1 + 0.8912 z^-2 + 0.6443 z^-3 - 0.3956 z^-4	94,9929819	6,21730334
'OE321'	B(z) = 0.04349 z^-1 - 0.08893 z^-2 + 0.04559 z^-3 F(z) = 1 - 1.984 z^-1 + 0.9845 z^-2	94,9848311	6,21616865
'OE221'	B(z) = -0.001383 z^-1 + 0.001513 z^-2 F(z) = 1 - 1.985 z^-1 + 0.9851 z^-2	94,8945691	6,55918215

Table 4: The estimated BJ models

'Model Type'	Model Parameters	'Best Fit'	'FPE'
'BJ23241'	B(z) = 6.176e-05 z^-1 - 5.46e-05 z^-2 C(z) = 1 - 0.6253 z^-1 + 0.01001 z^-2 + 0.02315 z^-3 D(z) = 1 - 1.629 z^-1 + 0.6427 z^-2 F(z) = 1 - 2.44 z^-1 + 1.393 z^-2 + 0.5362 z^-3 - 0.4889 z^-4	95,0238364	0,360708
'BJ12331'	B(z) = 7.101e-06 z^-1 C(z) = 1 - 1.609 z^-1 + 0.9796 z^-2 D(z) = 1 - 2.46 z^-1 + 2.254 z^-2 - 0.7828 z^-3 F(z) = 1 - 2.936 z^-1 + 2.872 z^-2 - 0.9367 z^-3	95,0160798	0,33251152
'BJ14131'	B(z) = 7.24e-06 z^-1 C(z) = 1 + 0.04146 z^-1 + 0.03341 z^-2 + 0.01694 z^-3 + 0.03648 z^-4 D(z) = 1 - 0.9637 z^-1 F(z) = 1 - 2.934 z^-1 + 2.87 z^-2 - 0.9354 z^-3	95,0135709	0,36037501
'BJ13131'	B(z) = 7.263e-06 z^-1 C(z) = 1 + 0.04056 z^-1 + 0.03335 z^-2 + 0.01701 z^-3 D(z) = 1 - 0.9659 z^-1 F(z) = 1 - 2.934 z^-1 + 2.869 z^-2 - 0.9351 z^-3	95,0131615	0,36038438
'BJ13231'	B(z) = 7.26e-06 z^-1 C(z) = 1 - 0.6241 z^-1 + 0.009762 z^-2 + 0.02454 z^-3 D(z) = 1 - 1.627 z^-1 + 0.6415 z^-2 F(z) = 1 - 2.934 z^-1 + 2.87 z^-2 - 0.9353 z^-3	95,0129791	0,35994338

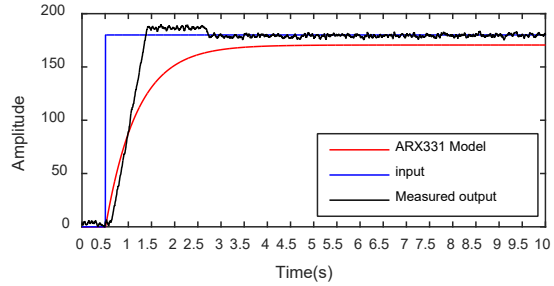


Figure 11: Comparison between the estimated model (ARX331) and measured output

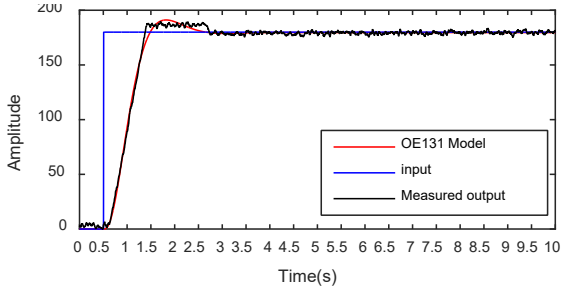


Figure 12: Comparison between the estimated model (OE131) and measured output

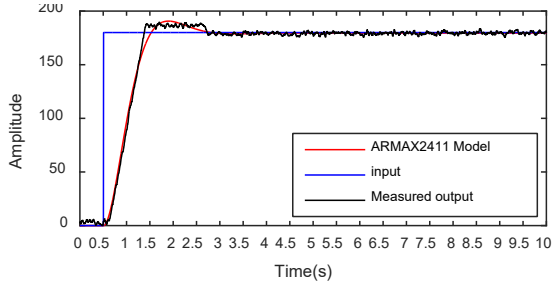


Figure 13: Comparison between the estimated model (ARMAX2411) and measured output

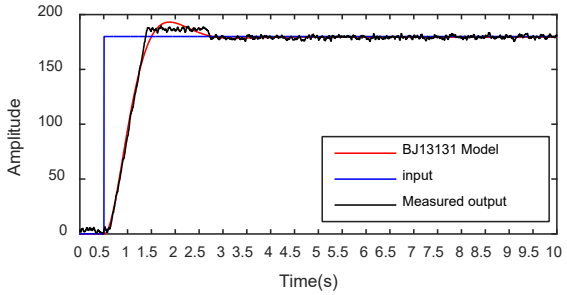


Figure 14: Comparison between the estimated model (BJ1313) and measured output

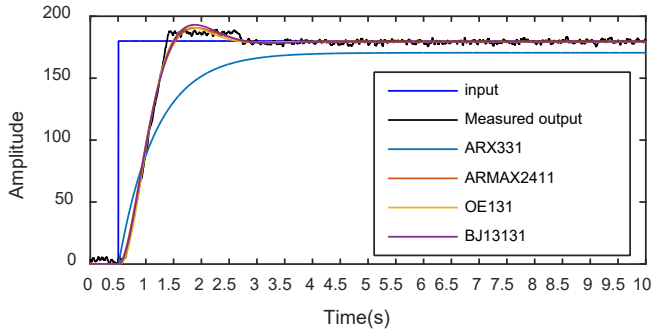


Figure 15: Comparison between the estimated Polynomial models (ARX331, ARMAX2411, OE131, and BJ1313) and measured output

8. Conclusion

In this article, a practical identification of DCSM has been elaborated using polynomial models. To meet application requirements and to collect input-output data, a microcontroller-based acquisition card has been developed. Finally, the comparison between the data and the estimated models using validation criterions shows that the ARMAX, BJ, and OE model has provided a good fit with experimental data and can be used in simulations or a real application.

In future work, the data acquisition system will be based on PLC to be adequate for industrial applications and to provide more sampling frequency and processing time. Moreover, other identification techniques will be applied to more complex non-linear systems to build proper algorithms for the control loop or fault detection.

References

[1] T. Wada, M. Ishikawa, R. Kitayoshi, I. Maruta, T. Sugie, "Practical modeling and system identification of R/C servo motors," in 2009 {IEEE} International Conference on Control Applications, IEEE, 2009, doi:10.1109/cca.2009.5280987.

[2] L. Ljung, "Perspectives on system identification," Annual Reviews in Control, **34**(1), 1–12, 2010, doi:10.1016/j.arcontrol.2009.12.001.

[3] R. Isermann, M. Münchhof, Identification of dynamic systems: an introduction with applications, Springer Science & Business Media, 2010.

[4] M. Villain, Systèmes asservis linéaires, Ellipses, 1996.

[5] L. Lennart, "System identification: theory for the user," PTR Prentice Hall, Upper Saddle River, NJ, 1–14, 1999.

[6] M. Jirgl, L. Obsilova, J. Boril, R. Jalovecky, "Parameter identification for pilot behaviour model using the {MATLAB} system identification toolbox," in 2017 International Conference on Military Technologies ({ICMT}), IEEE, 2017, doi:10.1109/miltechs.2017.7988824.

[7] S. Rai, S.-Y. Yang, T.-C. Tsao, "Wind turbine system identification and individual pitch control," in 2017 American Control Conference ({ACC}), IEEE, 2017, doi:10.23919/acc.2017.7963144.

[8] S. Rachad, B. Nsiri, B. Bensassi, "System Identification of Inventory System Using {ARX} and {ARMAX} Models," International Journal of Control and Automation, **8**(12), 283–294, 2015, doi:10.14257/ijca.2015.8.12.26.

[9] W. Araydah, T.A. Tutunji, I. Al-Naimi, "System identification for a liquid flow process," in 2017 {IEEE} Jordan Conference on Applied Electrical Engineering and Computing Technologies ({AEECT}), IEEE, 2017, doi:10.1109/aeect.2017.8257756.

[10] S.L. Chavan, D.B. Talange, "System identification black box approach for modeling performance of {PEM} fuel cell," Journal of Energy Storage, 18, 327–332, 2018, doi:10.1016/j.est.2018.05.014.

[11] L. Ljung, "Black-box models from input-output measurements," in {IMTC} 2001. Proceedings of the 18th {IEEE} Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No. 01CH 37188), IEEE, 2001, doi:10.1109/imtc.2001.928802.

[12] L. Ljung, "Prediction error estimation methods," Circuits, Systems, and Signal Processing, **21**(1), 11–21, 2002, doi:10.1007/bf01211648.

[13] M. Brunot, A. Janot, F. Carrillo, M. Gautier, "A Separable Prediction Error Method for Robot Identification," {IFAC}-{PapersOnLine}, **49**(21), 487–492, 2016, doi:10.1016/j.ifacol.2016.10.650.

[14] Y. Zhao, B. Huang, H. Su, J. Chu, "Prediction error method for identification of {LPV} models," Journal of Process Control, **22**(1), 180–193, 2012, doi:10.1016/j.jprocont.2011.09.004.

[15] L. Ljung, System identification toolbox. Getting started guide. The Mathworks Inc, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 512-520 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Practical Simulation of Grounded/Floating Lossy Inductors Based on Commercially Available Integrated Circuit LT1228

Tattaya Pukkalanun[1], Pitchayanin Moonmuang[1], Sumalee Unhavanich[2], Worapong Tangsrirat[1,*]

[1]*School of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok 10520, Thailand*

[2]*Faculty of Digital Technology, Chitralada Technology Institute, Bangkok 10300, Thailand*

ARTICLE INFO

Article history:
Received: 19 December, 2020
Accepted: 28 February, 2021
Online: 17 March, 2021

Keywords:
Commercially available IC
LT1228
Inductance Simulator
Lossy Inductor

ABSTRACT

The article suggests four circuit topologies for the practical simulation of grounded and floating lossy inductors. All the suggested circuits use commercially available integrated circuit LT1228 chips as active elements, and only two passive elements, namely one resistor and one capacitor. The first two of the proposed circuits employ only a single LT1228 active element and can realize grounded lossy inductors without the need for element-matching conditions. The last two of the proposed circuits can realize synthetic floating lossy inductors with only two LT1228s. The values of simulated equivalent elements can be tuned electronically by simply adjusting the external DC bias current of the LT1228. Non-ideal transfer error effects of the LT1228 on the synthetic inductor performance are inspected. Sensitivity performance concerning transfer errors and active and passive elements is also demonstrated. PSPICE simulation results and experimental measurements of the commercially available integrated circuit, LT1228, are incorporated to corroborate all our theoretical analyses.

1. Introduction

An inductance simulator circuit is an advantageous element in active circuit design and synthesis, especially for analog signal processing solutions and applications such as active filter, analog phase shifter, chaotic oscillator, and parasitic cancellation circuitry. In the past, there are various publications on the implementation of actively simulated lossy inductance simulators based on several active elements [1-34]. The first ones are intended for grounded lossy inductance simulation [1-22], while the others are for floating ones [23, 34]. However, from a careful inspection of the reported circuit topologies in these references, they have one or more of the following disadvantage features:

1. The use of a relatively large number of active and passive electronic components for their constructions [2, 4-6, 20-21, 23-27, 30-32]. Although the circuits of [1, 7-19, 28-29, 33-34] employ only a single active component, they still require at least three passive elements. Owing to the employment of an excessive number of circuit components, these designs are not suitable for an integrated circuit (IC) implementation point of view.
2. The lack of electronic tuning facility [1–2, 4, 6-18, 20-21, 24-31, 34].
3. The need for strict component-matching conditions or cancellations for the grounded inductance function simulation [1, 6, 9].
4. The use of different types of active electronic devices [20, 23-24, 32], which is inconvenient for further integrated circuit applications.
5. The use of commercially unavailable ICs [1–3, 5-12, 15-17, 19-24, 26, 28-30, 32-34]. Also, the active building blocks used in these publications are complicated active functional blocks. Accordingly, the verification of the proposed realization is mostly performed by only computer simulation results.

Recently, several attempts have been proved that LT1228 IC package has now become a versatile and standard commercial IC in the realization of various types of analog signal processing applications [35-37]. This article reports the four circuit topologies for simulating all grounded/floating lossy inductors employing

*Corresponding Author: Worapong Tangsrirat, Email: worapong.ta@kmitl.ac.th
This paper is an extended version from the proceedings of 2020 8th International Electrical Engineering Congress (iEECON) [38]

commercially available IC LT1228 chips as essential active components [38-39]. The first two circuit topologies ideally provide grounded lossy series and parallel inductors, while the other two topologies can emulate floating lossy series and parallel inductors. All the realized inductors employ canonical active and passive elements and are made adjustable electronically through the externally supplied current of the LT1228. The tracking error effects of the LT1228 on the circuit performance are investigated, and the active and passive sensitivity analyses are also evaluated. The validity of the synthetic lossy inductors is confirmed by the PSPICE simulation results using the macro-model of the LT1228 by Linear Technology as well as the prototype circuit test results using IC LT1228. Also, the circuit performance is demonstrated on an illustrative example of a voltage-mode second-order bandpass filter. The performance of the proposed practical grounded/floating lossy inductance simulators is compared with the previously presented similar works and also summarized in Table 1.

2. Characteristic of IC LT1228

Figure 1 shows the pin configuration and electrical symbol of the LT1228. It is a commercially available IC package, which is internally a combination of an operational transconductance amplifier (OTA) and a current feedback operational amplifier (CFA). The relationships between the appropriate terminal of LT1228 can be expressed through the following matrix expression given below [39]:

$$\begin{bmatrix} i_p \\ i_n \\ i_z \\ v_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ g_m & -g_m & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_p \\ v_n \\ v_z \\ i_o \end{bmatrix} . \quad (1)$$

In matrix (1), g_m is the transconductance gain of the LT1228, which is directly proportional to the external DC supplied current I_B, as defined by the following relation [39]:

$$g_m = 10I_B. \quad (2)$$

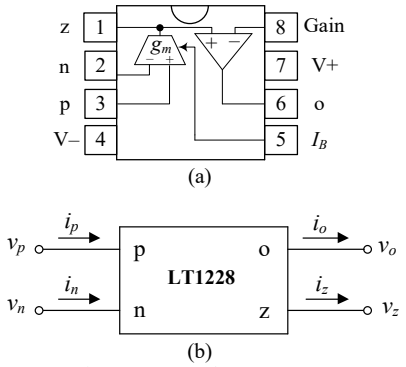


Figure 1: IC package LT1228
(a) internal active element and pin configuration (b) electrical symbol

3. Actively Realizable Lossy Inductance Simulator Circuits

Fig.2 shows the generic circuit configuration for the simulation of lossy grounded and floating inductors. They consist of IC package LT1228s as active elements. Routine analyses of the proposed circuits in Figure 2 yield the inductance functions and the finite quality factor (Q) given in Table 2. Inspection of the table, the four different inductance functions realized by the generic circuit topologies can be described in detail below.

(i) Figure 2(a) can emulate a grounded series RL-type lossy inductor with $R_{eq1} = 1/g_m$ and $L_{eq1} = R_1C_1/g_m$.

(ii) Figure 2(b) can emulate a grounded parallel RL-type lossy inductor with $R_{eq2} = R_2$ and $L_{eq2} = R_2C_2/g_m$.

(iii) The circuit in Figure 2(c) can emulate a floating series RL-type lossy inductor with $R_{eq3} = 1/g_m$ and $L_{eq3} = R_3C_3/g_m$.

(iv) The last circuit in Figure 2(d) can emulate a floating parallel RL-type lossy inductor with $R_{eq4} = 1/g_{m2}$ and $L_{eq4} = C_4/g_{m1}g_{m2}$.

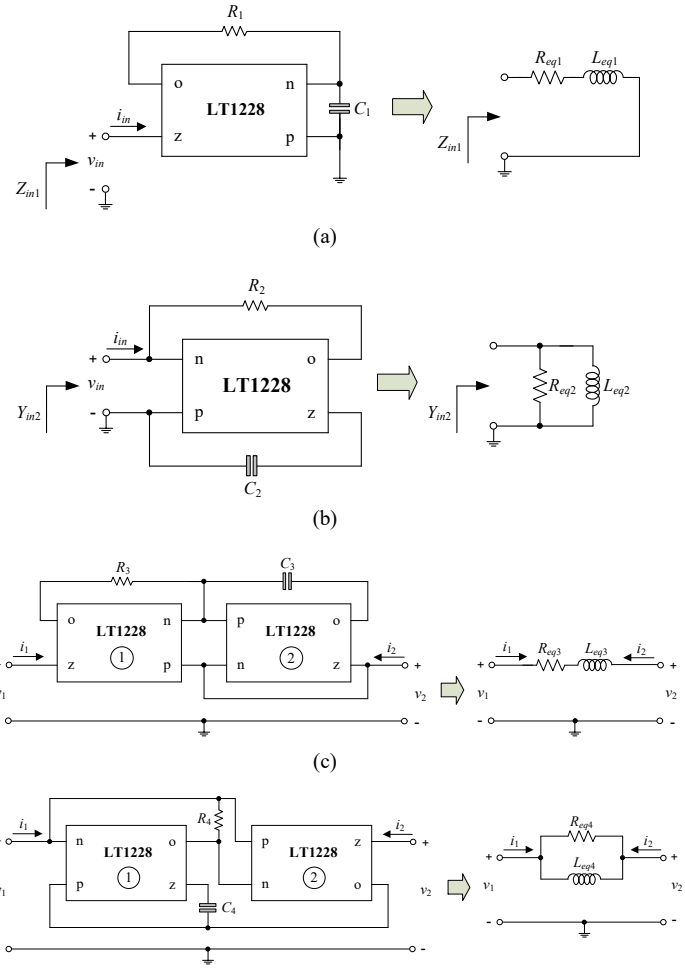


Figure 2: Lossy inductance simulators and equivalent circuits
(a) grounded series RL-type (b) grounded parallel RL-type
(c) floating series RL-type (d) floating parallel RL-type

Note that the proposed grounded lossy inductance sections of Figures 2(a) and 2(b) are no need for critical component-matching conditions. It is known that one can adjust the g_m-parameter by changing the bias current of the LT1228, and then the simulated equivalent elements R_{eq} and L_{eq} of the synthetic lossy inductors are electronically tunable.

Table 1: Performance comparison between the proposed lossy inductance simulators in Figure 2 and the previously published similar works

Reference	Configuration	Inductor Type	Active device number	Passive device number	Commercially available IC	Matching condition	Electronic tuning	Technology	Supply voltage	Power dissipation
[1]	grounded	series, parallel	CCII = 1	R = 4, C = 1	no	yes	no	N/A	N/A	N/A
[2]		series, parallel	CCII = 3	R ≥ 2, C ≥ 1	no	no	no	AD844	±12V	N/A
[3]		parallel (Fig.3)	DO-VDBA = 2	C = 1	no	no	yes	N/A	N/A	N/A
[4]		series, parallel	CCII = 3	R ≥ 2, C ≥ 1	yes	no	no	AD844	±12V	N/A
[5]		series (Fig.3f)	EXCCTA = 2	R = 2, C = 1	no	no	yes	TSMC 0.18-µm CMOS	±0.9V	N/A
		parallel (Fig.3d)								
[6]		series (Fig.2)	CFOA = 2	R = 3, C = 1	no	yes	no	IBM 0.13-µm CMOS	±0.75V (simulation), ±6V (experiment)	3.53mW
[7]		series, parallel	CCIII = 1	R = 2, C = 1	no	no	no	B101 and B102 BJT	±10V	N/A
[8]		series, parallel	CCIII = 1	R = 2, C = 1	no	no	no	TSMC 0.6-µm CMOS	±2.5V	N/A
[9]		parallel (Fig.2c)	DXCCII = 1	R = 2, C = 1	no	yes	no	TSMC 0.35-µm CMOS	±1.5V	N/A
[10]		series, parallel	DXCCII = 1	R ≥ 2, C = 1	no	no	no	TSMC 0.35-µm CMOS	±2.5V, +1.44V	N/A
[11]		series, parallel	DVCC = 1	R = 2, C = 1	no	no	no	MIETEC 0.5-µm CMOS	±2.5V	N/A
[12]		series, parallel	OTRA = 1	R ≥ 2, C ≥ 1	no	no	no	MIETEC 1.2-µm CMOS	±5V	N/A
[13]		series	CFOA = 1	R = 2, C = 1	yes	no	no	AD844	±15V	N/A
[14]		series	CFOA = 1	R = 2, C = 1	yes	no	no	AD844	±10V	N/A
[15]		series, parallel	CFOA = 1	R = 2, C = 1	no	no	no	IBM 0.13-µm CMOS	±0.75V	0.89mW
[16]		parallel	CDBA = 1	R = 2, C = 1	no	no	no	AD844	±12V	N/A
[17]		series (Fig.2e)	DXCCII = 1	R = 2, C = 1	no	no	no	TSMC 0.35-µm CMOS	±1.5V	N/A
[18]		series (Fig.2d)	CFOA = 1	R = 2, C = 2	yes	no	no	AD844	±12V	N/A
[19]		series, parallel	FTFNTA = 1	R = 1, C = 2	no	no	yes	N/A	N/A	N/A
[20]		parallel (Fig.3)	VF = 1, CF = 1	R = 3, C = 1	no	no	no	0.25-µm CMOS, AD844	±1.25V, +0.4V (simulation), ±12V (experiment)	6.87mW
[21]		parallel (Fig.2a)	OTRA = 2	R = 4, C = 1	no	no	no	AD844	±10V	N/A
[22]		series (Fig.2c)	VDCC = 1	R = 1, C = 1	no	no	yes	TSMC 0.18-µm CMOS	±0.9V	0.87mW
[23]	floating	series (Fig.3)	CCCII+ = 2, CCCII- = 1	C = 1	no	no	yes	NR100&PR100 BJT, AD844	±2.5V (simulation), ±5V (experiment)	3.17mW
		parallel (Fig.2)	CCCII- = 2	C = 1						
[24]		series (Fig.2a)	DVCCS = 1, OA = 3	R = 2, C = 1	no	yes	no	N/A	N/A	N/A
		series (Fig.2b)	DVCCS = 1, OA = 2							
		parallel (Fig.2c)	DVCCS = 1, OA = 2							
[25]		parallel	CFOA = 2	R = 3, C = 2	yes	yes	no	AD844	±15V (experiment)	N/A
[26]		series, parallel	CCII = 2	R = 2, C = 1	no	no	no	N/A	N/A	N/A
[27]		series, parallel	CFOA = 2	R = 2, C = 1	yes	no	no	AD844	N/A	N/A

[28]		series, parallel	DDCC = 1	R = 2, C = 1	no	yes	no	TSMC 0.18-µm CMOS	±0.9V, +0.34V	N/A
[29]		parallel	DO-DDCC = 1	R = 2, C = 1	no	no	no	TSMC 0.35-µm CMOS	±1.5V, -0.9V	N/A
[30]		series (Fig.2)	DDCC = 2	R = 2, C = 1	no	yes	no	IBM 0.13-µm CMOS	±0.75V	6.9mW
[31]		series, parallel	CFOA = 2	R = 2, C = 1	yes	no	no	AD844	N/A	N/A
[32]		series (Fig.1)	DVB = 1, ECCII+ = 2, ECCII- = 1	R = 3, C = 1	no	no	yes	EL2082, AD830	±5V	N/A
[33]		parallel (Fig.4)	DDCC = 1	R = 2, C = 1	no	yes	no	IBM 0.13-µm CMOS, AD844	±0.75V	2.08mW
[34]		series (Fig.2), parallel (Fig.3)	DDCC = 1	R = 2, C = 1	no	no	no	IBM 0.13-µm CMOS, AD844	±0.75V, +0.23V (simulation), ±9V (experiment)	2.06mW
Proposed circuits	grounded	series (Fig.2a)	LT1228 = 1	R = 1, C = 1	yes	no	yes	LT1228 from Linear Technology Company	±5V	56.7mW
		parallel (Fig.2b)	LT1228 = 1			no	yes			41.2mW
	floating	series (Fig.2c)	LT1228 = 2			yes	yes			0.114W
		parallel (Fig.2d)	LT1228 = 2			yes	yes			0.117W

N/A : Not available
CCII : second-generation current conveyor, DO-VDBA : dual-output voltage-differencing buffered amplifier,
EXCCTA : extra X current conveyor transconductance amplifier, CFOA : current feedback operational amplifier, CCIII : third-generation current conveyor,
DXCCII : dual X second-generation current conveyor, DVCC : differential voltage current conveyor, OTRA : operational transresistance amplifier,
CDBA : current differencing buffered amplifier, FTFNTA : four terminal floating nullor transconductance amplifier,
VF : voltage follower, CF : current follower, VDCC : voltage differencing current conveyor, CCCII± : positive/negative current-controlled conveyor,
DVCCS : differential voltage controlled current source, OA : operational amplifier, DDCC : differential difference current conveyor,
DVB: differential voltage buffer, DO-DDCC : dual-output differential difference current conveyor,
ECCII± : positive/negative electronically controllable current conveyor

Table 2: Summary of the lossy inductance simulation using the proposed circuit configurations of Figure 2.

Topology	Matching Condition	Realized inductance	Q
Figure 2(a)	no	$Z_{in1} = R_{eq1} + sL_{eq1} = \frac{1}{g_m} + s\left(\frac{R_1C_1}{g_m}\right)$	ωR_1C_1
Figure 2(b)	no	$Y_{in2} = \frac{1}{R_{eq2}} + \frac{1}{sL_{eq2}} = \frac{1}{R_2} + \frac{g_m}{sR_2C_2}$	$\omega C_2/g_m$
Figure 2(c)	$g_m = g_{m1} = g_{m2}$	$Z_{in3} = R_{eq3} + sL_{eq3} = \frac{1}{g_m} + s\left(\frac{R_3C_3}{g_m}\right)$	ωR_3C_3
Figure 2(d)	$R_4 = 1/g_{m2}$	$Y_{in4} = \frac{1}{R_{eq4}} + \frac{1}{sL_{eq4}} = g_{m2} + \frac{g_{m1}g_{m2}}{sC_4}$	$\omega C_4/g_{m1}$

4. Non-Ideal Consideration and Sensitivity Performance

If the non-ideal transfer errors are considered, the terminal relationships of the LT1228 can be described as follows:

$$\begin{bmatrix} i_p \\ i_n \\ i_z \\ v_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \alpha g_m & -\alpha g_m & 0 & 0 \\ 0 & 0 & \beta & 0 \end{bmatrix} \cdot \begin{bmatrix} v_p \\ v_n \\ v_z \\ i_o \end{bmatrix} . \quad (3)$$

here, $\alpha = (1 - \varepsilon_{gm})$ and $\beta = (1 - \varepsilon_v)$, where ε_{gm} and ε_v are respectively the transconductance inaccuracy and the voltage transfer error in which $|\varepsilon_{gm}| << 1$ and $|\varepsilon_v| << 1$. Re-analysis all the topologies in Figure 2 by taking the non-ideal parameters of the LT1228 into account, the various non-ideal characteristic parameters of the synthetic inductors can be evaluated and summarized in Table 3.

Table 3: Non-ideal parameters of the designed inductors in Figure 2.

Topology	Simulated equivalent elements		Q
Figure 2(a)	$R_{eq1} = 1/\alpha\beta g_m$	$L_{eq1} = R_1C_1/\alpha\beta g_m$	ωR_1C_1
Figure 2(b)	$R_{eq2} = R_2$	$L_{eq2} = R_2C_2/\alpha\beta g_m$	$\omega C_2/\alpha\beta g_m$
Figure 2(c)	$R_{eq3} = 1/\alpha_1\beta_1 g_m$	$L_{eq3} = R_3C_3/\alpha_1\beta_1 g_m$	ωR_3C_3
Figure 2(d)	$R_{eq4} = 1/g_{m2}$	$L_{eq4} = C_4/\alpha_1\beta_1 g_{m1}g_{m2}$	$\omega C_4/\alpha_1\beta_1 g_{m1}$

Normalized sensitivities of the simulated equivalent values R_{eq} and L_{eq} with respect to active and passive elements are obtained as:

For Figure 2(a);

$$S^{R_{eq1}}_{R_1,C_1} = 0, \; S^{L_{eq1}}_{R_1,C_1} = 1, \; S^{R_{eq1}}_{\alpha,\beta,g_m} = S^{L_{eq1}}_{\alpha,\beta,g_m} = -1 \quad (4)$$

For Figure 2(b);

$$S^{R_{eq2}}_{R_2} = 1, \; S^{R_{eq2}}_{C_2} = 0, \; S^{L_{eq2}}_{R_2,C_2} = 1, \; S^{R_{eq2}}_{\alpha,\beta,g_m} = 0, \; S^{L_{eq2}}_{\alpha,\beta,g_m} = -1 \quad (5)$$

For Figure 2(c);

$$S^{R_{eq3}}_{R_3,C_3} = 0, \; S^{L_{eq3}}_{R_3,C_3} = 1, \; S^{R_{eq3}}_{\alpha_1,\beta_1,g_m} = S^{L_{eq3}}_{\alpha_1,\beta_1,g_m} = -1 \quad (6)$$

For Figure 2(d);

$$S^{L_{eq4}}_{C_4} = 1, \; S^{L_{eq4}}_{\alpha_1,\beta_1,g_{m1},g_{m2}} = -1, \; S^{R_{eq4}}_{\alpha_1,\beta_1,g_{m1}} = 0, \; S^{R_{eq4}}_{g_{m2}} = -1 \quad (7)$$

From the above, it is clearly seen that the magnitudes of these sensitivity coefficients are less than or equal to unity. Thus, it can be realized that all the proposed inductors exhibit good sensitivity performance.

5. Simulation Results

The proposed lossy inductance simulators in Figure 2 have been simulated by the PSPICE program using LT1228 standard SPICE macro-model obtained from Linear Technology. The LT1228 was biased with symmetrical supply voltages of +V = -V = 5 V. As an example, the proposed floating lossy series inductor of Figure 2(c) has been designed to simulate a floating series RL impedance with R_{eq3} = 1 kΩ and L_{eq3} = 1 mH. For this purpose, the component values are set as: $I_B = I_{B1} = I_{B2}$ = 100 μA for $g_m = g_{m1} = g_{m2}$ = 1 mA/V, R_3 = 1 kΩ and C_3 = 1 nF. Figure 5 shows the simulated transient responses of the input voltage and current (v_{in} and i_1), where a sinusoidal input voltage of 50 mV (peak) at the frequency of 500 kHz is applied to the simulator. The resulting waveforms show that the current i_1 lags the voltage v_{in} by 70°, whereas the theoretical calculation is equal to $\tan^{-1}(\omega L_{eq3}/R_{eq3}) = \tan^{-1}(2\pi\times500\times10^3\times1\times10^{-3})/(1\times10^3) = 72.34^\circ$. With the same component values, the simulated frequency responses corroborating the ideal responses are also plotted in Figure 6. As can be observed from the results, it behaves serial RL impedance well for the frequencies around 1 kHz to 1 MHz.

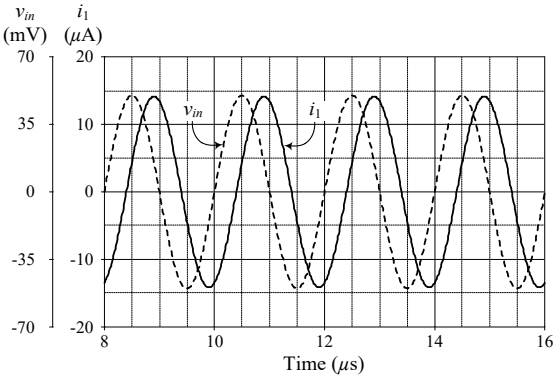


Figure 5: Simulated transient responses of the floating lossy series inductor in Figure 2(c)

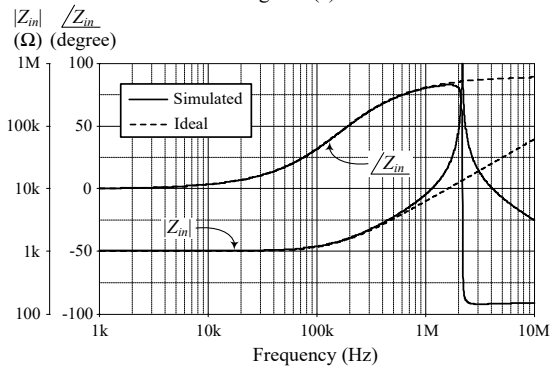


Figure 6: Ideal and simulated frequency responses of the floating lossy series inductor simulator in Figure 2(c).

To show the electronic tunability of the inductor in Figure 2(c), the detailed analysis of frequency characteristic is provided in Figure 7. In the figure, the magnitude responses of the input impedance Z_{in3} are compared with the ideal series RL impedance by changing the bias current for I_B = 50 μA, 200 μA and 400 μA (g_m = 0.5 mA/V, 2 mA/V and 4 mA/V).

On the other hand, the simulated transient responses of the proposed floating lossy parallel inductor of Figure 2(d) are given in Figure 8. In this case, the active and passive components are taken as follows: I_B = 100 μA (g_m = 1 mA/V), R_4 = 1 kΩ and C_4 = 1 nF, which yields R_{eq4} = 1 kΩ and L_{eq4} = 1 mH. In Figure 8, a sinusoidal input signal was applied to the input of the inductor at a frequency of 200 kHz and signal amplitude of 50 mV peak. The phase difference between v_{in} and i_1 is approximately 36°, which is close to the ideal value ($\tan^{-1}(R_{eq4}/\omega L_{eq4}) = \tan^{-1}(1\times10^3)/(2\pi\times200\times10^3\times1\times10^{-3}) = 38.51^\circ$) with a percentage deviation of 6.52%. The tuning capability of the proposed floating lossy parallel inductor has also been tested for I_{B1} = 50 μA, 200 μA, and 400 μA, while keeping I_{B2} constant at 100 μA. The resulting frequency characteristics showing in Figure 10 prove that the proposed inductor circuit can be conveniently tuned by electronic means.

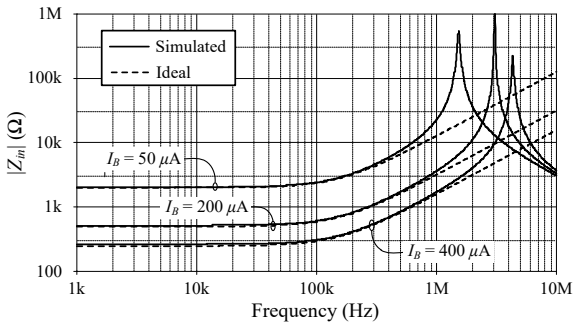


Figure 7: Simulated frequency responses for Z_{in3} in Figure 2(c) with I_B tuning

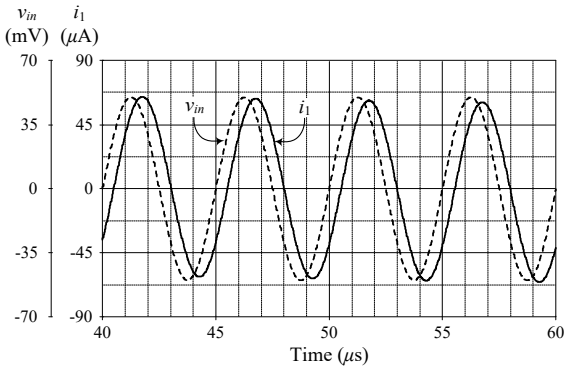


Figure 8: Simulated transient responses of the floating lossy parallel inductor in Figure 2(d)

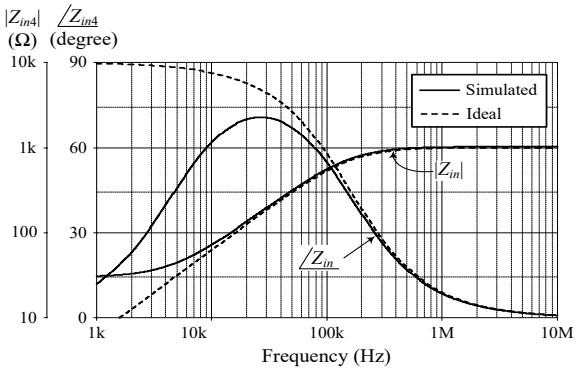


Figure 9: Ideal and simulated frequency responses of the floating lossy parallel inductor simulator in Figure 2(d)

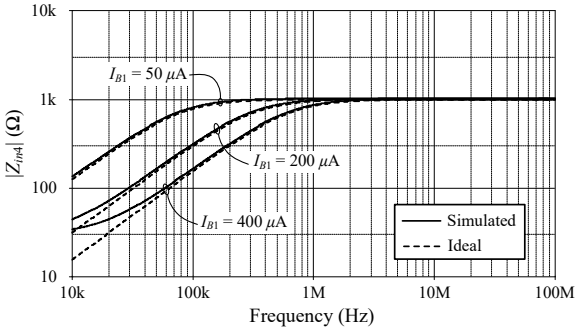


Figure 10: Simulated frequency responses for Z_{in4} in Figure 2(d) with I_{B1} tuning.

6. Experimental Results

To further represent the validity of the presented theory, the proposed circuits of Figure 2 were implemented in prototype hardware using IC LT1228 from Linear Technology. The LT1228 is biased with ±5 V supplies. For the experimental testing of Figure 2(a) and 2(b), the passive components in the schematic diagram are the following values: $R_1 = R_2 = 1$ kΩ and $C_1 = C_2 = 1$ nF. The measured magnitude and phase responses of the proposed grounded series lossy inductor in Figure 2(a) for I_B = 50 μA, 100 μA, 200 μA, and 400 μA are shown in Figures 11 and 12, respectively. The calculated and measurement results for the Z_{in1} frequency responses at a frequency of 100 kHz are summarized in Table 4. According to this table, parasitic gains and signal transfer errors at the corresponding terminals of the LT1228 in conjunction with the tolerances in passive elements used will deviate the impedance frequency responses of the simulated lossy inductor. It should be taken into account the fact that the current and voltage conveying realized by an LT1228 is actually valid under a small-signal operation. However, appropriate adjusting g_m with I_B gives a significant improvement in the magnitude and phase responses.

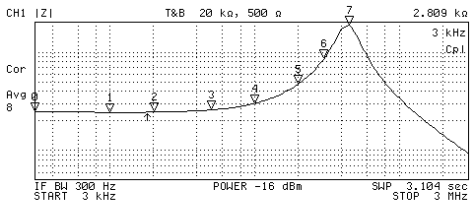
In the same manner, the measured frequency responses for the input impedance Z_{in2} of the proposed grounded parallel lossy inductor in Figure 2(b) are presented in Figures 13 and 14. Regarding the experimental testing for the actively simulated lossy inductor, the results of the Z_{in2}-frequency response at f = 100 kHz is provided in Table 5. Obviously, the practical results are consistent with the theoretical ones. However, the deviation in the magnitude and phase responses is mainly due to the non-ideal parasitic effects and the resistor and capacitor tolerances.

Table 4: Summary of the measured values for the Z_{in1} in Figure 2(a)

I_B (μA)	g_m (mA/V)	$\lvert Z_{in1}\rvert$ (kΩ)		$\angle Z_{in1}$ (degree)	
		Measured	Calculated	Measured	Calculated
50	0.5	2.98	2.36	26.51	32.14
100	1	1.39	1.18	32.40	32.14
200	2	0.84	0.59	44.11	32.14
400	4	0.38	0.29	50.11	32.14

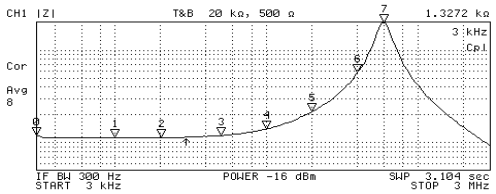
Table 5: Summary of the measured values for the Z_{in2} in Figure 2(b)

I_B (μA)	g_m (mA/V)	$\lvert Z_{in2}\rvert$ (Ω)		$\angle Z_{in2}$ (degree)	
		Measured	Calculated	Measured	Calculated
50	0.5	536.14	782.48	53.94	38.51
100	1	440.83	532.02	63.54	57.86
200	2	232.05	299.72	81.17	72.56
400	4	126.62	155.12	80.73	81.07



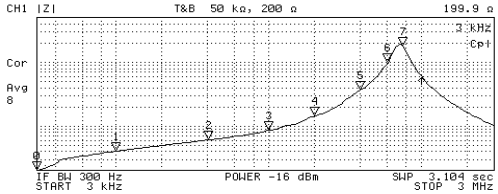
N	SWP PARAM	VAL
0	3 kHz	2.4809 kΩ
1	10 kHz	2.3973 kΩ
2	20 kHz	2.4347 kΩ
3	50 kHz	2.5615 kΩ
4	100 kHz	2.9767 kΩ
5	200 kHz	4.6988 kΩ
6	300 kHz	8.4936 kΩ
7	450 kHz	18.552 kΩ

(a)



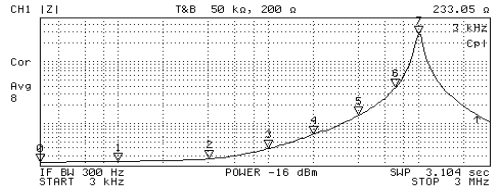
N	SWP PARAM	VAL
0	3 kHz	1.2172 kΩ
1	10 kHz	1.1105 kΩ
2	20 kHz	1.1235 kΩ
3	50 kHz	1.1827 kΩ
4	100 kHz	1.3909 kΩ
5	200 kHz	2.1356 kΩ
6	400 kHz	5.6811 kΩ
7	600 kHz	19.611 kΩ

(b)



N	SWP PARAM	VAL
0	3 kHz	199.9 Ω
1	10 kHz	399.13 Ω
2	40 kHz	597.57 Ω
3	100 kHz	843.77 Ω
4	200 kHz	1.4215 kΩ
5	400 kHz	3.6385 kΩ
6	600 kHz	9.8176 kΩ
7	750 kHz	17.988 kΩ

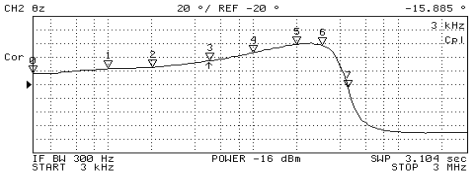
(c)



N	SWP PARAM	VAL
0	3 kHz	233.05 Ω
1	10 kHz	241.17 Ω
2	40 kHz	261.62 Ω
3	100 kHz	381.66 Ω
4	200 kHz	655.13 Ω
5	400 kHz	1.3259 kΩ
6	700 kHz	3.7082 kΩ
7	1 MHz	27.251 kΩ

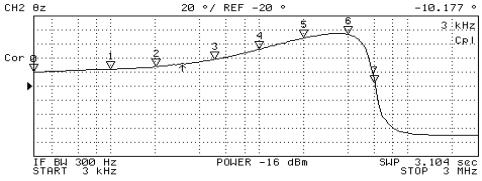
(d)

Figure 11: Measured magnitude responses of the proposed grounded series lossy inductor in Figure 2(a)
(a) I_B = 50 μA (b) I_B = 100 μA (c) I_B = 200 μA (d) I_B = 400 μA



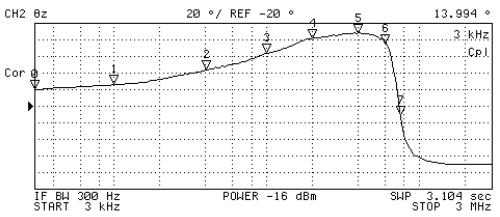
N	SWP PARAM	VAL
0	3 kHz	-3.2885 °
1	10 kHz	3.3293 °
2	20 kHz	5.3911 °
3	50 kHz	14.624 °
4	100 kHz	26.51 °
5	200 kHz	39.152 °
6	300 kHz	36.918 °
7	450 kHz	-24.965 °

(a)



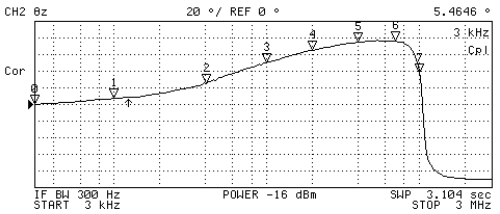
N	SWP PARAM	VAL
0	3 kHz	0.1774 °
1	10 kHz	3.9826 °
2	20 kHz	8.0636 °
3	50 kHz	18.031 °
4	100 kHz	32.397 °
5	200 kHz	48.494 °
6	400 kHz	53.774 °
7	600 kHz	-16.424 °

(b)



N	SWP PARAM	VAL
0	3 kHz	1.7994 °
1	10 kHz	5.9821 °
2	40 kHz	23.974 °
3	100 kHz	44.115 °
4	200 kHz	61.743 °
5	400 kHz	68.106 °
6	600 kHz	55.572 °
7	750 kHz	−31.374 °

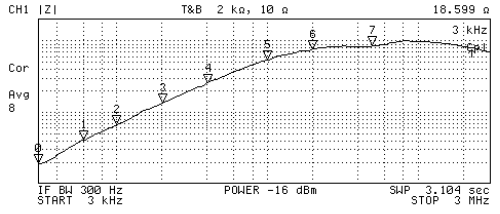
(c)



N	SWP PARAM	VAL
0	3 kHz	2.4646 °
1	10 kHz	7.3182 °
2	40 kHz	25.301 °
3	100 kHz	50.111 °
4	200 kHz	65.223 °
5	400 kHz	74.951 °
6	700 kHz	75.881 °
7	1 MHz	37.342 °

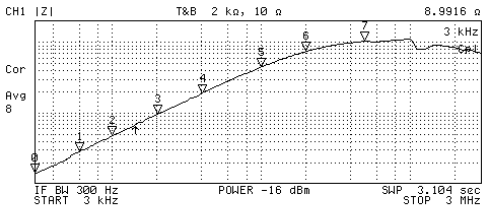
(d)

Figure 12: Measured phase responses of
the proposed grounded series lossy inductor in Figure 2(a)
(a) $I_B = 50\ \mu$A (b) $I_B = 100\ \mu$A (c) $I_B = 200\ \mu$A (d) $I_B = 400\ \mu$A



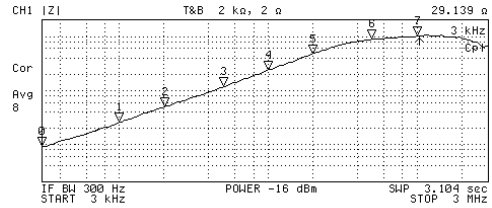
N	SWP PARAM	VAL
0	3 kHz	18.599 Ω
1	6 kHz	40.525 Ω
2	10 kHz	65.485 Ω
3	20 kHz	132.23 Ω
4	40 kHz	251.76 Ω
5	100 kHz	536.14 Ω
6	200 kHz	751.61 Ω
7	500 kHz	838.94 Ω

(a)



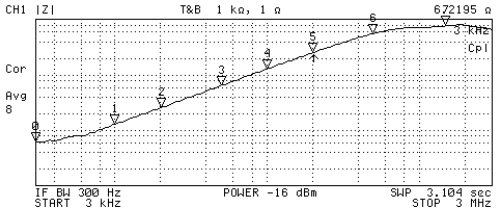
N	SWP PARAM	VAL
0	3 kHz	16.916 Ω
1	6 kHz	28.047 Ω
2	10 kHz	46.916 Ω
3	20 kHz	94.538 Ω
4	40 kHz	187.12 Ω
5	100 kHz	440.83 Ω
6	200 kHz	732.02 Ω
7	500 kHz	1.0219 kΩ

(b)



N	SWP PARAM	VAL
0	3 kHz	9.139 Ω
1	10 kHz	25.64 Ω
2	20 kHz	48.543 Ω
3	50 kHz	114.87 Ω
4	100 kHz	232.05 Ω
5	200 kHz	465.16 Ω
6	500 kHz	864.36 Ω
7	1 MHz	987.67 Ω

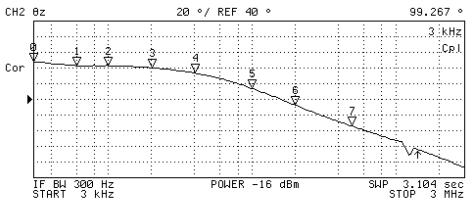
(c)



N	SWP PARAM	VAL
0	3 kHz	6.2195 Ω
1	10 kHz	12.852 Ω
2	20 kHz	25.261 Ω
3	50 kHz	63.414 Ω
4	100 kHz	126.62 Ω
5	200 kHz	248.32 Ω
6	500 kHz	551.03 Ω
7	1.5 MHz	758.66 Ω

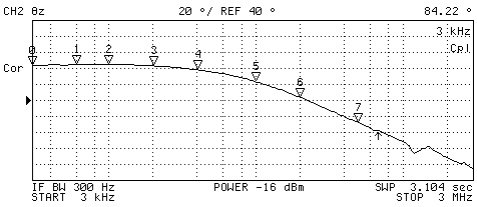
(d)

Figure 13: Measured magnitude responses of the proposed grounded series
lossy inductor in Figure 2(b)
(a) $I_B = 50\ \mu$A (b) $I_B = 100\ \mu$A (c) $I_B = 200\ \mu$A (d) $I_B = 400\ \mu$A



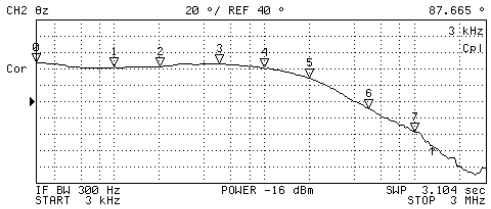
N	SWP PARAM	VAL
0	3 kHz	94.267 °
1	6 kHz	82.793 °
2	10 kHz	83.121 °
3	20 kHz	80.294 °
4	40 kHz	73.487 °
5	100 kHz	53.942 °
6	200 kHz	32.572 °
7	500 kHz	5.8288 °

(a)



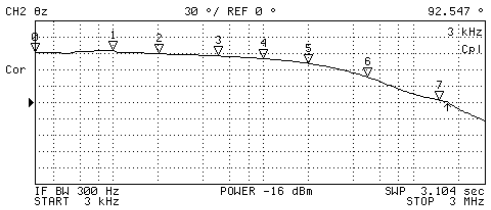
N	SWP PARAM	VAL
0	3 kHz	84.22 °
1	6 kHz	85.306 °
2	10 kHz	85.211 °
3	20 kHz	83.327 °
4	40 kHz	78.37 °
5	100 kHz	63.542 °
6	200 kHz	43.865 °
7	500 kHz	12.439 °

(b)



N	SWP PARAM	VAL
0	3 kHz	87.665 °
1	10 kHz	81.894 °
2	20 kHz	82.902 °
3	50 kHz	86.225 °
4	100 kHz	81.173 °
5	200 kHz	68.506 °
6	500 kHz	31.294 °
7	1 MHz	2.7541 °

(c)



N	SWP PARAM	VAL
0	3 kHz	92.547 °
1	10 kHz	95.763 °
2	20 kHz	89.72 °
3	50 kHz	85.618 °
4	100 kHz	80.735 °
5	200 kHz	71.897 °
6	500 kHz	46.118 °
7	1.5 MHz	3.8374 °

(d)

Figure 14: Measured phase responses of the proposed grounded series lossy inductor in Figure 2(b) (a) I_B = 50 μA (b) I_B = 100 μA (c) I_B = 200 μA (d) I_B = 400 μA

7. Example of Application

To illustrative an application example, the proposed inductor topologies in Figures 2(b) and 2(c) are used for the simulation of serial RL and parallel RL elements of the bandpass filter in Figure 15. The voltage transfer function of Figure 15 is found as:

$$\frac{V_{out}(s)}{V_{in}(s)} = \frac{\left(\frac{R_{eq2}}{L_{eq3}}\right)s}{D(s)} \quad , \tag{8}$$

where the denominator $D(s)$ is equal to

$$D(s) = s^2 + \left(\frac{R_{eq2}}{L_{eq2}} + \frac{R_{eq2}}{L_{eq3}} + \frac{R_{eq3}}{L_{eq2}}\right)s + \left(\frac{R_{eq2}R_{eq3}}{L_{eq2}L_{eq3}}\right) \quad . \tag{9}$$

From Equations (8)-(9), the natural angular frequency (ω_o) and the quality factor (Q) of the proposed filter in Figure 15 can be characterized as given respectively by the following equations:

$$\omega_o = 2\pi f_o = \sqrt{\frac{R_{eq2}R_{eq3}}{L_{eq2}L_{eq3}}} \quad , \tag{10}$$

and

$$Q = \frac{\sqrt{R_{eq2}R_{eq3}L_{eq2}L_{eq3}}}{R_{eq2}(L_{eq2}+L_{eq3}) + R_{eq3}L_{eq2}} \quad . \tag{11}$$

In the simulation, the element values for bandpass filter in Figure 15 with f_o = 159.15 kHz and Q = 0.33 are specified as follows: g_m = 1 mA/V (I_B = 100 μA), $R_2 = R_3$ = 1 kΩ, $C_2 = C_3$ = 1 nF. Figure 16 shows the plots of the theoretical and simulated frequency responses of the filter, in which $R_{eq2} = R_{eq3}$ = 1 kΩ and $L_{eq2} = L_{eq3}$ = 1 mH. From Figure 16, the simulated f_o is 156.15 kHz, where the relative error is found as 1.88%.

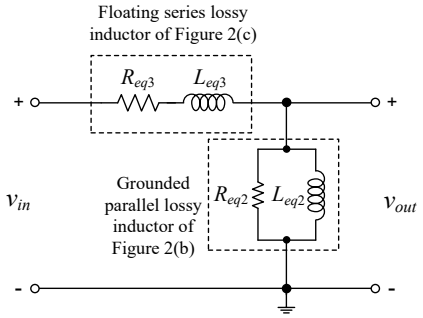


Figure 15: Voltage-mode bandpass filter realization using the proposed actively simulated inductors of Figure 2(b) and 2(c).

8. Conclusions

In this study, the actively simulated grounded/floating lossy inductance simulators using commercially available IC named LT1228 are reported. All of the reported topologies employ a capacitor and a resistor as passive elements and can be tuned electronically through the bias current of the LT1228. The simulation and experimental testing results are achieved to validate their practical performances.

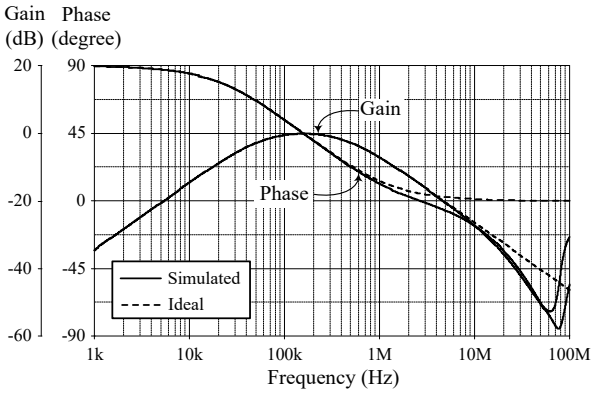


Figure 16: Ideal and simulated frequency responses of the filter in Figure 15

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work is supported by Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang under the contract number 2563-02-01-033.

References

[1] A. N. Paul, D. Patranabis, "Active simulation of grounded inductors using a single current conveyor" IEEE Trans. Circuits Syst., **28**(2), 164–165, 1981. doi: 10.1109/TCS.1981.1084947

[2] M. O. Cicekoglu, "Active simulation of grounded inductors with CCII+s and grounded passive elements" International Journal Electronics, **85**(4), 455–462, 1998. doi: 10.1080/002072198134003

[3] M. Srivastava, K. Bhardwaj, A. Roy, A. Singh, "An active-C realization for simulating electronically controllable lossy grounded inductance" in 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 137–140, 2019. doi: 10.1109/SPIN.2019.8711758

[4] O. Cicekoglu, A. Toker, H. Kuntman, "Universal immittance function simulators using current conveyors" Computers and Electrical Engineering, **27**, 227–238, 2001. doi: 10.1016/S0045-7906(00)00018-5
[5] M. Faseehuddin, J. Sampe, S. Shireen, S. H. M. Ali, "Lossy and lossless inductance simulators and universal filters employing a new versatile active block" Informacije MIDEM, **48**(2), 97–113, 2018.
[6] M. Dogan, E. Yuce, "CFOA based a new grounded inductor simulator and its applications" Microelectronics Journal, **90**, 297–300, 2019. doi: 10.1016/j.mejo.2019.07.002
[7] H. Kuntman, M. Gulsoy, O. Cicekoglu, "Actively simulated grounded lossy inductors using third-generation current conveyors" Microelectronics Journal, **31**, 245–250, 2000. doi: 10.1016/S0026-2692(99)00108-1
[8] H. Y. Wang, C. T. Lee, "Systematic synthesis of R-L and C-D immittances using single CCIII" International Journal Electronics, **87**(3), 292–301, 2000. doi: 10.1080/002072100132192
[9] F. Kacar, A. Yesil, "Novel grounded parallel inductance simulators realization using a minimum number of active and passive components" Microelectronics Journal, **41**, 632–638, 2010. doi: 10.1016/j.mejo.2010.06.011
[10] B. Metin, "Supplementary inductance simulator topologies employing single DXCCII" Radioengineering, **20**(3), 614–618, 2011.
[11] M. Incekaraoglu, U. Cam, "Realization of series and parallel R-L and C-D impedances using single differential voltage current conveyor" Analog Integr. Circ. Signal Process., **43**, 101–104, 2005. doi: 10.1007/s10470-005-6577-8
[12] U. Cam, F. Kacar, O. Cicekoglu, H. Kuntman, A. Kuntman, "Novel grounded parallel immittance simulator topologies employing single OTRA" Int. J. Electron. Commun. (AEU), **57**(4), 287–290, 2003. doi: 10.1078/1434-8411-54100173
[13] E. Yuce, "Novel lossless and lossy grounded inductor simulators consisting of a canonical number of components" Analog Integr. Circ. Signal Process., **59**, 77–82, 2009. doi: 10.1007/s10470-008-9235-0
[14] F. Kacar, H. Kuntman, "CFOA-based lossless and lossy inductance simulators" Radioengineering, **20**(3), 627–631, 2011.
[15] H. Alpaslan, E. Yuce, "Inverting CFOA based lossless and lossy grounded inductor simulators" Circuits Syst. Signal Process., **34**, 3081–3100, 2015. doi: 10.1007/s00034-015-0004-x
[16] J. K. Pathak, A. K. Singh, R. Senani, "New canonic lossy inductor using a single CDBA and its application" International Journal Electronics, **103**(1), 1–13, 2016. doi: 10.1080/00207217.2015.1020884
[17] A. Yesil, F. Kacar, "New DXCCII-based grounded series inductance simulator topologies" Istanbul University-Journal of Electrical and Electronics Engineering (IU-JEEE), **14**(2), 1785–1789, 2014.
[18] E. Basak, F. Kacar, "Lossy/lossless grounded inductance simulators using current feedback operational amplifier (CFOA)" Electrica, **18**(1), 95–99, 2018. doi: 10.5152/iujeee.2018.1815
[19] R. Singh, D. Prasad, "Grounded lossy inductance simulator using single FTFNTA" in IEEE 17th India Council International Conference (INDICON), New Delhi, India, 1–6, 2020. doi: 10.1109/INDICON49873.2020.9342507
[20] H. Alpaslan, E. Yuce, "Current-mode biquadratic universal filter design with two terminal unity gain cells" Radioengineering, **21**(1), 304–311, 2012.
[21] U. Cam, F. Kacar, O. Cicekoglu, H. Kuntman, A. Kuntman, "Novel two OTRA-based grounded immitance simulator topologies" Analog Integr. Circ. Signal Process., **39**, 169–175, 2004. doi: 10.1023/B:ALOG.0000024064.73784.58
[22] F. Kacar, A. Yesil, S. Minaei, H. Kuntman, "Positive/negative lossy/lossless grounded inductance simulators employing single VDCC and only two passive elements", International Journal Electronics Communications (AEU), **68**(1), 73–78, 2014. doi: 10.1016/j.aeue.2013.08.020
[23] S. Minaei and E. Yuce, "Realization of tunable active floating inductance simulators" International Journal Electronics, **95**(1), 27–37, 2008. doi: 10.1080/00207210701809333
[24] R. Senani, "New single-capacitor simulation of floating inductors" Electrocomponent Science and Technology, **10**, 7–12, 1982. doi: 10.1155/APEC.10.7
[25] R. Senani and D. R. Bhaskar, "New lossy/loss-less synthetic floating inductance configuration realized with only two CFOAs" Analog Integr. Circ. Sig. Process., **73**, 981–987, 2012. doi: 10.1007/s10470-012-9897-5
[26] R. Senani, "New tunable synthetic floating inductors" International Journal Electronics, **16**(10), 382–383, 1980. doi: 10.1049/el:19800270
[27] D. R. Bhaskar and R. Senani, "Synthetic floating inductors realized with only two current feedback op-amps" American Journal Electrical Electron. Engineering, **3**(4), 88–92, 2015.
[28] E. Yuce, "New low component count floating inductor simulators consisting of a single DDCC" Analog Integration Circuits Process., **58**, 61–66, 2009. doi: 10.1007/s10470-008-9218-1
[29] M. A. Ibrahim, S. Minaei, E. Yuce, N. Herencsar and J. Koton, "Lossy/lossless floating/grounded inductance simulation using one DDCC" Radioengineering, **21**(1), 3–10, 2012.
[30] E. Yuce, S. Tokat, H. Alpaslan, "Grounded capacitor-based new floating inductor simulators and a stability test" Turkey Journal Electrical Engineering & Computer Science, **23**, 2138–2149, 2015. http://doi.org/10.3906/elk-1301-183
[31] M. T. Abuelma'atti, S. K. Dhar, Z. J. Khalifa, "New two-CFOA-based floating immittance simulators" Analog Integr. Circ. Sig. Process., **91**, 479–489, 2017. doi: 10.1007/s10470-017-0956-9
[32] R. Sotner, N. Herencsar, J. Jerabek, A. Kartci, J. Koton, T. Dostal, "Pseudo-differential filter design using novel adjustable floating inductance simulator with electronically controllable current conveyors" Elektron. Elektrotech., **23**(2), 31–35, 2017. doi: 10.5755/j01.eie.23.2.17996
[33] A. Abaci, E. Yuce, "Single DDCC based new immittance function simulators employing only grounded passive elements and their applications" Microelectron. J., **83**, 94–103, 2019. doi: 10.1016/j.mejo.2018.11.014
[34] A. Abaci, E. Yuce, "Single DDCC− based simulated floating inductors and their applications" IET Circuits, Devices & Systems, **14**(6), 796–804, 2020. doi: 10.1049/iet-cds.2019.0558
[35] N. Roongmuanpha, W. Tangsrirat, "SITO current-mode multifunction biquad using readily available IC LT1228s", in The 6th International Conference on Engineering, Applied Sciences and Technology (ICEAST-2020), July 1-4, Thailand, 108-111, 2020. doi: 10.1109/iceast50382.2020.9165538
[36] R. Sotner, N. Herencsar, J. Jerabek, L. Langhammer, J. Polak, "On practical construction of electronically controllable compact current amplifier based on commercially available elements and its application", International Journal Electronics Communication (AEU), **81**(11), 56-66, 2017. doi: 10.1016/j.aeue.2017.07.002
[37] N. Roongmuanpha, T. Pukkalanun, W. Tangsrirat, "Practical realization of electronically adjustable universal filter using commercially available IC-based VDBA" Engineering Review, **41**(3), 2020. doi: 10.30765/er.1547
[38] P. Moonmuang, N. Roongmuanpha, T. Pukkalanun, W. Tangsrirat, "On the realization of simulated lossy inductors using voltage differencing buffered amplifiers" in 2020 8th International Electrical Engineering Congress (iEECON), Chiang Mai, Thailand, 2020. doi: 10.1109/iEECON48109.2020.229469
[39] Linear Technology, "100MHz current feedback amplifier with DC gain control", LT1228 datasheet, 1994.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 521-525 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Recycling and Reuse of Wastewater Generated in Car-Washing Facilities

Elgaali Elgaali*, Majid Akram

Civil Engineering Program, Higher Colleges of Technology, Dubai Men's College, Dubai, 337150, UAE

ARTICLE INFO

Article history:
Received: 23 December, 2020
Accepted: 08 March, 2021
Online: 17 March, 2021

Keywords:
Sustainability
Wastewater
Recycling

ABSTRACT

Fresh water is already scarce in the world, especially in the Middle East (ME). Desalination industry is the main supplier of the potable water to the municipalities in the ME region. It is well known the high cost of a liter of water produced by the desalination process. Unfortunately, car-washing service consumes substantial amount of this desalinated water. This paper describes a filtration system designed and tested for treatment and reuse of the wastewater generated in car-washing stations. The filtration system assembled from two filters: (1) sand and gravel mix, and (2) activated carbon. The paper is an extension of work originally presented in ASET conference in Dubai. The quality of the effluent (treated wastewater) was investigated and determined in Dubai central laboratories. Wastewater samples were grabbed from different car service stations. Representative samples were prepared and the concentrations of the following parameters were measured in each sample of the effluent: (1) Biological oxygen demand (BOD), (2) Chemical oxygen demand (COD), (3) Total dissolved solids (TDS), (4) Total suspended solids (TSS), and (5) Oil and grease (OG). The results show that the filter system removes the BOD and COD at an efficiency as high as 97.5%, the TSS at 90%, and the TDS and OG at 85.5%. In general, the quality of the effluent was found to fall within the standards set by Dubai regulatory authorities. Further research is recommended to enhance the filtration system performance and make it commercially applicable.

1. Introduction

Water is also termed as white gold in the age of its scarcity and increasing population in the world. The dynamics of geopolitics is changing based on the sharing of this valuable resource among different countries, especially in the Middle East. Desalination industry is the main supplier of the fresh water to the municipalities in the United Arab Emirates (UAE). It is well known the high cost of a liter of water produced by the desalination process ($ 2.0 / gallon) [1], [2]. Besides, desalination process adversely affects the marine environments. For example, brine as a by-product of the desalination process is usually discharged back into the sea degrading the sea water quality and posing hazards to the marine fauna and flora [3]. This makes the water produced from desalination a precious commodity. Unfortunately, considerable amount of the desalinated water produced, is used for non-potable uses (e.g. car-washing) putting more stress on already stressed fresh water resources. Therefore, considerable amount of research has been directed towards exploring methods and techniques to recycle and reuse the wastewater to augment the high fresh water consumption especially for non-potable uses [4].

Globally, wastewater recycling/reuse has been practiced for a while for different purposes. Wastewater, is mainly recycled to minimize the disposed amount in order to alleviate the impacts on environment (degradation and pollution). Recently, it has been used as a lower cost alternative for fresh water in non-potable practices. Ground water recharging one of the first applications has been practiced. Other applications include irrigation of green landscape, air cooling (chillers), and recently toilet flushing [5]. In this region (Middle East), wastewater recycling and reuse is widely used to sustain the valuable and expensive water produced from the desalination of sea water [6]. For example, Dubai municipality recycles and reuses the wastewater for irrigation of the green areas in the landscape of the city [7]. Table 1 shows some standards set for the recycled water to be reused for irrigation.

* Corresponding Author: Elgaali, P.O.Box: 15825, Dubai, UAE, 971 552493085, eelgaali@hct.ac.ae

Table 1: Standards of water recycled for irrigation

Symbol	Unit	Maximum Limit for Unrestricted Irrigation	Maximum Limit for Restricted Irrigation
BOD	mg/L	200	200
COD	mg/L	200	200
TDS	mg/L	2000	2000
TSS	mg/L	15	30
pH	mg/L	6 - 8	6 - 8

Source: Dubai municipality

The practice of car-washing is mounting in this region (ME, UAE). The harsh dusty weather necessitates the extension of this practice especially in the UAE. Unfortunately, car-washing service consumes substantial amount of the desalinated water produced in the emirate of Dubai [2]. The car-washing activity is estimated to consume around 30 gallons to wash small size car, and around 60 gallons to wash big size car. A published survey, extracted from 59 car-washing stations, shows that the water consumption reaches 378770 m^3 per year [2]. This amount represents around 0.2% of the total annual production.

Currently, the whole amount of water wasted in car-washing facilities, is usually discharged into the city sewer systems after brief in situ pretreatment in holding tanks to facilitate settling of large solid particles. In general, the car-washing facilities pay relatively high cost to dispose their wastewater in the municipal sewer system burdening the car-washing facilities with more operational cost. The estimated costs of water usage and disposal for each car-washing facility are $37000 and $4000 respectively [2]. Therefore, considerable amount of research has been directed towards exploring methods and techniques to recycle and reuse the wastewater from car-washing to augment the high consumption of fresh water and reduce the operational cost [8] – [12].

Wide range of technologies, from simple to very complex, are widely used now to treat and recycle the wastewater. For example, the membrane filters (reverse osmosis (RO), micro, ultra, and Nano-filtration) are widely used to remove the dissolved solids (DS) and particulate matter (PM); the membrane bioreactor (MBR) systems are increasingly used now as an efficient technique to reclaim wastewater; the oxidation technique which is advanced to combine ozone, ultraviolet (UV) light, and hydrogen peroxide to create the highly reactive hydroxyl radical (.OH) [4]. The use and level of treatment required, dictates the cost of water recycling [4], [13]-[17].

Almost, all the current treatment methods incur high initial and operational costs. Therefore, this research is a genuine effort to develop and test a system to treat, recycle, and reuse the wastewater generated in car washing facilities in situ. The proposed system incur low initial and operational costs. The proposed system can be well described as a simple filter designed and assembled from natural materials from local sources. The system is assembled from two filter: (1) sand and gravel mix, and (2) activated carbon. For the purpose of this study, the quality of the effluent (treated wastewater) was investigated and determined in Dubai central laboratories. Wastewater samples were grabbed from different car service stations. Representative samples were prepared and the concentrations of the following parameters were measured in each sample of the effluent: (1) Biological oxygen demand (BOD), (2) Chemical oxygen demand (COD), (3) Total dissolved solids (TDS), (4) Total suspended solids (TSS), and (5) Oil and grease (OG), (6) pH. The results are presented and discussed in section 3 of this paper.

2. Materials and Methods

The test program, in this research, is developed to filter and examine the quality of raw and treated wastewater samples.

2.1. Water Samples

The amount of the wastewater tested in this study was collected from 8 stations of car service around Dubai, UAE. Composite sampling was adopted. Composite sampling is widely used to monitor wastewater discharges [15]. In order to collect representative samples, several samples were grabbed throughout the car's washing process. Figure 1 shows the samples collection process. The samples then were thoroughly mixed together into a composite sample (Figure 2). The characteristics of the composite sample (raw wastewater) are presented in Table 2.



Figure 1: Samples collection process

The samples were screened first to remove the large particles or grit that might damage the filter and then passed through the two stage filter.

Table 2: Raw Wastewater Characteristics

Parameter	mg/L
Biological oxygen demand (BOD)	235
Chemical oxygen demand (COD)	976
Oil and grease	88
PH value at 25°C	7.2
Total dissolved solids (TDS)	1100

Total suspended solids (TSS)	112
Cadmium	0.08
Chromium	0.5
Copper	0.38
Lead	0.53
Nickel	0.2
Zinc	1.98



Figure 2: Sample of raw wastewater

2.2. Filter Fabrication

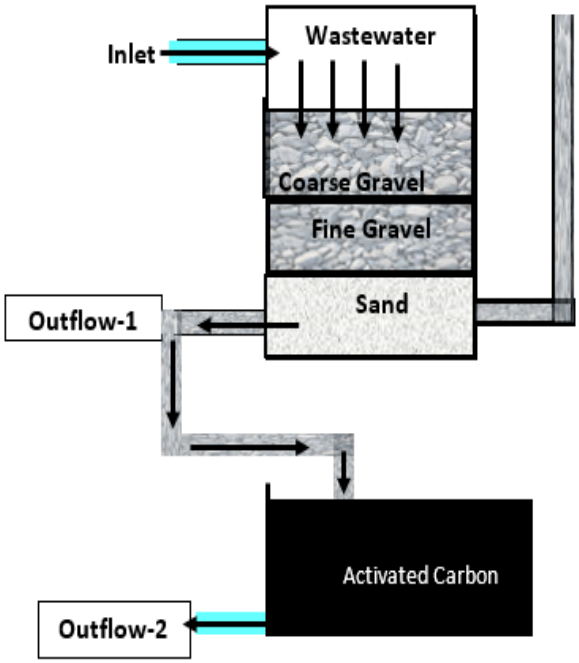
In order to increase the effectiveness of treatment, two filters were assembled and used. The two filter units were fabricated and assembled at the civil engineering workshop in the Dubai men's campus-higher colleges of technology. The two-stage filter was fabricated using materials available from the local market. These materials include: (1) Coarse Gravel, (2) Fine Gravel, (3) Sand, and (4) Activated Carbon. The schematic and the assembly of the two-stage filter are shown in Figure 3-a and 3-b respectively. For better depiction, the unit of the sand and gravel is magnified in the right part of Figure 3-b.

The schematic diagram shown in Figure 3-a, depicts the filtration process. From the entrance (inlet) the water flows through the first stage filter (dual media of sand and gravel) and then through the second stage filter (activated carbon) towards the exit point (outlet). Throughout the filtration process, samples of treated water were at outflow-1 and outflow-2 (Figure 3-a). Since Dubai regulatory authorities stipulate strict standards for the use of recycled water for different beneficial activities, the quality of the treated water was verified in the laboratories of Dubai municipality.

For better illustration of the effectiveness of the filter and the filtration process, the characteristics of the treated water samples were compared to the characteristics of the raw wastewater (baseline). The effectiveness of the filter is further indicated by the removal efficiency which is estimated by a formula taking the following form:

$$R\% = (C_{in} - C_{out})/ C_{in} \quad (1)$$

where R% is the removal efficiency, C_{in} is the initial concentration of the parameter in the influent (raw water), C_{out} is the final concentration of the parameter in the effluent (treated water).



(a)



(b)

Figure 3: Schematic and assembly of the two-stage filter-workshop-higher colleges of technology-Dubai men's campus

3. Results and Discussion

The results of the lab analysis of the raw and treated water samples are presented in Tables 3. The concentrations of the quality parameters in the treated water are depicted and compared to the levels of the parameters in the raw wastewater (baseline). Then the removal efficiency of the designed filter is estimated.

BOD removal: as shown in the table the BOD, as indicator of the strength of the organic pollution in the water sample, is reduced from 235 to 33.8 mg/L with removal efficiency reaches 86%. The huge reduction in the BOD is mainly attributed to the type of sand used in stage1 filter.

COD removal: the COD, as indicator of the strength of the organic pollution in the water sample, is also experienced huge reduction from 976 to 193 mg/L with removal efficiency reaches 81%.

TDS removal: it is well known that the high level of the total dissolved solids (TDS) can result in many problems such as scale

deposit and precipitates on fixtures; corrosion of fixtures, and reduction or malfunctioning of the filter and associated equipment [18]. Successfully, at this stage the level of reduction of the total dissolved solids (TDS) is fairly good (from 1100 mg/L to 249.6 mg/L) with removal efficiency reaches78%.

TSS removal: accumulation of TSS in water systems usually causes blockage of the systems and damages the nozzles widely used in car washing facilities. Moreover, by time TSS create a good medium for bacterial growth which pose potential health risks. Unfortunately, the reduction in the total suspended solids (TSS) is of less magnitude of 57 mg/L (from 112 to 55 mg/L) with removal efficiency reaches 51%.

OG removal: To avoid any blockage to the filter/reuse systems oil and grease (OG) was a parameter of great interest in this research. Effectively, the level of the OG was reduced from 88 to 12 mg/L with removal efficiency reaches 88%.

In general, at this stage the performance of the filter, as measured by the removal efficiency, was fairly good. The removal efficiency was as high as 85% for BOD and as low as 50% for TSS, with average overall average removal efficiency of 73.51%.

The following results can be extracted from Table 3 after stage2:

BOD, the removal efficiency of the BOD improved by magnitude of 8% (from 85.6% to 93.32%);
COD, the removal efficiency of the COD improved by magnitude of 18% (from 80.23% to 98.5%);
TDS, the removal efficiency of the TDS improved by magnitude of 8% (from 77.31% to 85.31%);
TSS, the removal efficiency of the TSS improved by magnitude of 39% (from 50.90% to 90.0%);
OG, the removal efficiency of the OG stayed the same at 86%.

In general, the removal efficiency was as high as 97% for BOD and as low as 85% for TDS, with average overall average removal efficiency of 93%. It is very clear from the results that adding the activated carbon component to the filter assembly enhanced the filter performance a lot. However, within the limits of the parameters (BOD, COD, TDS, TSS, and OG) in the wastewater, it appears that the current levels of TDS had little effect on the removal of parameters examined in this study [18]. However, relative to other studies, the filter system developed, proved to be very effective in producing treated water of high quality and could be reused for car washing safely [8] – [12].

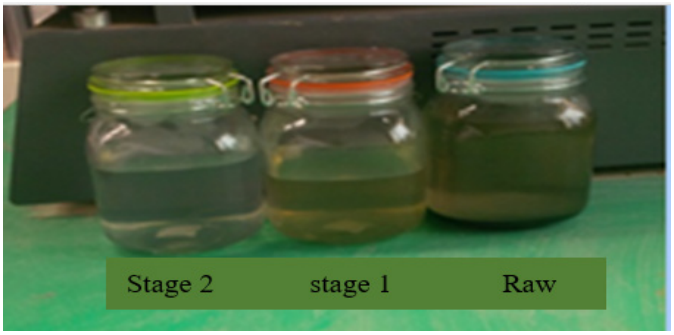
Table 3: Characteristics of raw and treated wastewater

Parameter	Pollution Level			Removal efficiency	
	Original	Stage1	Stage2	Stage1	Stage2
	mg/L	mg/L	mg/L	%	%
BOD	235	33.85	< 7	85.6	97.32
COD	976	193	15	80.23	98.5
OG	88	< 12	< 12	86.36	86.36
TDS	1100	249.6	161.6	77.31	85.31
TSS	112	55	< 11	50.9	90

Figure 4-a, illustrates the high performance of stage 1 of the filter. The color of the sample shown, is greatly diluted from dark black to grey one. This indicates the high removal of the total suspended solid (TSS). Figure 4-b shows from right to left samples of raw wastewater, water after stage 1, and after stage 2.



(a)



(b)

Figure 4: Raw and treated water samples-environmental laboratory-higher colleges of technology-Dubai men's campus

Figure 5 clearly depicts the removal efficiency of the filter system in magnitude (mg/L) and percentage.

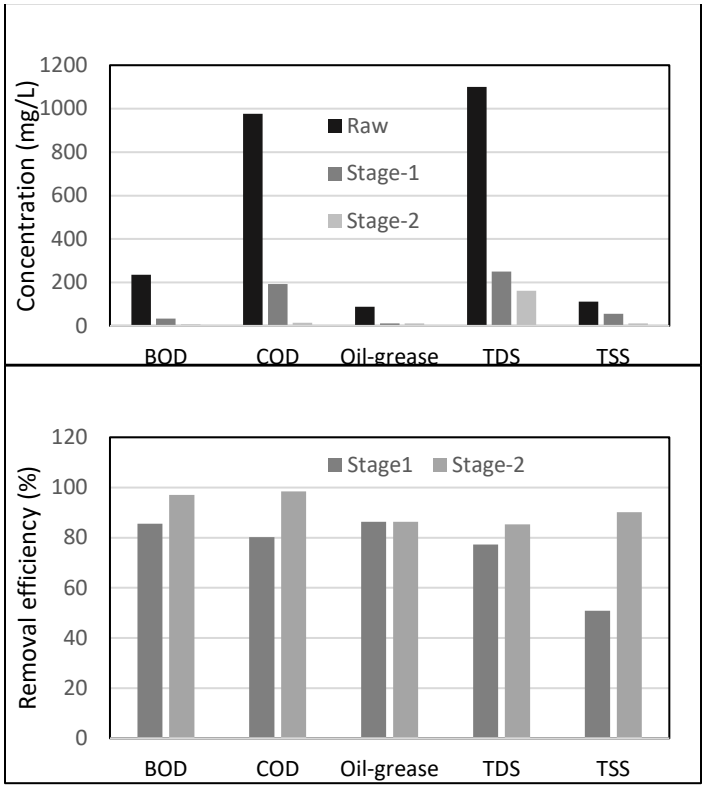


Figure 5: Removal efficiency in magnitude and percentage

4. Summary and Conclusions

This paper describes a filtration system designed and tested for treatment and reuse of the wastewater generated in the car's service stations. The filtration system assembled from two filters: (1) sand and gravel mix, and (2) activated carbon. The system is made of natural materials from local sources. The system is anticipated to incur low initial and operational costs. The quality of the effluent (treated wastewater) was investigated and determined in Dubai central laboratories. Wastewater samples were grabbed from different car service stations. Representative samples were prepared and the concentrations of the following parameters were measured in each sample of the effluent: (1) Biological oxygen demand (BOD), (2) Chemical oxygen demand (COD), (3) Total dissolved solids (TDS), (4) Total suspended solids (TSS), and (5) Oil and grease (OG). The results are presented, discussed, and documented through several figures and table and the following conclusions were drawn:

According to the results, the removal efficiency of the BOD, COD, TDS, TSS, and OG is of magnitudes of 97.32, 98.50, 85.s1, 90.0, and 86.36% respectively. The results suggest that, under the range of the current levels of the parameters investigated the performance of the system was good.

Very clear, adding the activated carbon component to the filter assembly led to a great enhancement in the removal efficiency of the filtration process. However, this cannot be said about the TDS removal. The removal efficiency of the TDS, unlike the other parameters, did not change from stage1 to stage2. This result indicates that activated carbon is not that efficient in filtering out the TDS. Alternative material (e.g. silt) must be sought and tested to improve the removal efficiency of the TDS. However, it can be concluded that the results show that the filter system achieved high performance in removing the parameters tested. In general, the quality of the effluent was found to fall within the standards set by Dubai regulatory authorities (Table 1).

Further research is recommended to enhance the filtration system performance and make it commercially applicable. Significant savings can be realized from treating and reusing of the wastewater generated in the car service stations. For example, sustaining (saving) precious resource (desalinated water) for the future and reducing the carbon foot-print on the planet earth.

Conflict of Interest

No conflict of interest in publishing this paper.

Acknowledgment

The authors would like acknowledge the contribution of Dubai Men's College (DMC) – Higher Colleges of technology, and the following students: Mr. Fahad Yusuf and Mr. Abdulla Salem.

References

[1] M. Akram and E. Elgaali, "Recycling of wastewater from car-washing facilities for landscaping applications," Advances in Science and Engineering Technology International Conferences (ASET), Dubai, United Arab Emirates, 1- 4, 2019, doi: 10.1109/ICASET.2019.8714351.

[2] Dubai Electricity and Water Authority (DEWA), Sustainability Report, 2018, http://www.dewa.gov.ae.

[3] N. Kress, Y. Gertner, E. Shoham-Frider, "Seawater quality at the brine discharge site from two mega size seawater reverse osmosis desalination plants in Israel (Eastern Mediterranean)," Water Research, **171**, 115402, 2020, doi.org/10.1016/j.watres.2019.115402.

[4] E. Mohamed, E. Elgaali, "Sustainable concrete made of construction and demolition wastes using recycled wastewater in the UAE" Journal of Advanced Concrete Technology-Japan, **10**, 110-125, 2012, doi: 10.3151/ jact. 10.110.

[5] Hunter Water Corporation, Annual Report, 2019, https://www.hunterwater .com. au/our-water/sustainability/recycling-and-reuse.

[6] Appropedia organization, sustainability report, water recycling, 2019, http://www.appropedia.org/Water_recycling

[7] Dubai municipality-waste department, Annual Report, 2019, https://www.dm.gov.ae/municipality-business/environment-and-coasts/waste-department.

[8] M. Sarmadi, F. Maryam, H. Saleh, "Efficient technologies for carwash wastewater treatment: a systematic review," Environmental Science and Pollution Research, **27**(28), 2020, doi.10.1007/s11356-020-09741-w.

[9] A. Oghenekome, O. Ovie, "Management of wastewater from car wash sites using locally sourced natural coagulants and simple filtration Process," International Journal of Scientific and Research Publications, **10**(12), 621-625, 2020, DOI: 10.29322/IJSRP.10.12.2020.p10872.

[10] M. Emamjomeha, H. Jamalia, Z. Naghdalib, M. Mousazadeha, "Carwash wastewater treatment by the application of an environmentally friendly hybrid system: an experimental design approach," Desalination and Water Treatment, **160**, 171–177, 2019, DOI:10.5004/dwt.2019.24382.

[11] R. Zaneti, R. Etchepare, J. Rubio, "Car wash wastewater treatment and water reuse - A case study," Water Science & Technology, **67**(1), 82-89, 2013, DOI: 10.2166/wst.2012.492

[12] J. Rubio, R. Zaneti, "Treatment of washrack wastewater with water recycling by advanced flocculation–column flotation," Desalination and Water Treatment, **8**, 146–153, 2009, DOI: 10.5004/dwt.2009.679.

[13] M. J. Hammer Sr., M. J. Hammer Jr., Water and Wastewater Technology, Prentice Hall, 2003, ISBN: 0131911406.

[14] S. Judd, The MBR Book: Principles of Membrane Bioreactors in Water and Wastewater Treatment, Elsevier, UK, 2011.

[15] R. O. Gilbert, Statistical Methods for Environmental Pollution Monitoring, John Wiley, 1987, ISBN: 0471288780.

[16] L. Chen, F. Ling, G. Bakker, W. T. Liue, G. Medemaa, W. Meer, G. Liuab, "Assessing the transition effects in a drinking water distribution system caused by changing supply water quality: an indirect approach by characterizing suspended solids," Water Research, **168**, 115159, 2020, doi: 10.1016/j.watres.2019.115159

[17] I. Husain, M. Alkhatib, M. Jammi, M. Mirghani, Z. Zainudin, A. Hoda, "Problems, control, and treatment of fat, oil, and grease (FOG): a review," J. Oleo Sci., **63**(8), 747-752, 2014, doi: 10.5650/jos.ess13182.

[18] X. Sarah, M. Jason, "Study on the effect of total dissolved solids (TDS) on the performance of an SBR for COD and nutrients removal' Journal of Environmental Science and Health, Toxic/Hazardous Substances and Environmental Engineering. **53**(2), 146 -153, 2017, https://doi.org/10.1080/ 10934529.2017.1383130.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 526-540 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Model for the Application of Automatic Speech Recognition for Generating Lesson Summaries

Phillip Blunt*, Bertram Haskins

Department of Engineering, The Built Environment and Technology, Nelson Mandela University, Port Elizabeth, 6031, South Africa

ARTICLE INFO

Article history:
Received: 23 October, 2020
Accepted: 20 February, 2021
Online: 17 March, 2021

Keywords:
Feature extraction
Machine learning
Continuous speech recognition
Keyword spotting
Transcription
Education
Noise management
Lesson summary

ABSTRACT

Automatic Speech Recognition (ASR) technology has the potential to improve the learning experience of students in the classroom. This article addresses some of the key theoretical areas identified in the pursuit of implementing a speech recognition system, capable of lesson summary generation in the educational setting. The article discusses: some of the applications of ASR technology in education; prominent feature extraction and speech enhancement techniques typically applied to digital speech; and established neural network-based machine learning models capable of keyword spotting or continuous speech recognition. Following the theoretical investigation, a model is proposed for the implementation of an automatic speech recognition system in a noisy educational environment to facilitate automated, speech-driven lesson summary generation. A prototype system was developed and improved based on this model, ultimately proving itself capable of generating a lesson summary intended to bolster students' secondary contact with lesson content. This topic-oriented lesson summary provides students with a lesson transcript, but also helps them to monitor educator-defined keyword terms, their prevalence and order as communicated in the lesson, and their associations with educator-defined sections of course content. The prototype was developed using the Python programming language with a modular approach so that its implemented Continuous Speech Recognition system and noise management technique could be chosen at run-time. The prototype contrasts the performance of CMUSphinx and Google Speech Recognition for ASR, both accessed via a cloud-based programming library, and compared the change in accuracy when applying noise injection, noise cancellation or noise reduction to the educator's speech. Proof of concept was established using the Google Speech Recognition System, which prevailed over CMUSphinx and enabled the prototype to achieve 100,00% accuracy in keyword identification and association on noise-free speech, contrasted with a 96,93% accuracy in keyword identification and association on noise-polluted speech using a noise-cancellation technique.

1 Introduction

Student lesson summaries are a valuable resource for allowing students to focus on the key points of a lesson, boosting secondary contact with lesson content. They allow students to realise which aspects of a lesson may be more important and streamline the study process for courses which employ both formative and summative assessments. Formulating a lesson summary from notes and lesson content alone can be challenging to students especially when covering larger segments of course content over a short period of time. Students are also not all equally capable of creating their own lesson summaries and often rely on available course material and notes made during their lessons as study material. These challenges are also further compounded by a lack of context whenever students attempt to summarise a lesson without integrating with the theoretical focus of the classroom teaching session. Toward providing a solution to these challenges of lesson summary generation; this work is an extension of the paper presented at The 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC) wherein a model was proposed for automated lesson summary generation in a noisy educational environment using keywords in the educator's speech as the prompting mechanism for summary of key points [1]. This extension reintroduces the proposed model, along with the theoretical background it is based in, reinforced by a proof of concept prototype system for lesson summary generation which is used to demonstrate the model's application.

*Corresponding Author: Phillip Blunt, Email: s213308762@mandela.ac.za

Automatic speech recognition (ASR) has been applied in assistive technologies for tasks such as closed captioning, voice search, command detection as well as keyword identification. It therefore stands to reason that these applications of the technology may be applied to the voice of the educator in the educational setting. This work posits that it is possible to generate a lesson summary, by transcribing the voice of the educator during a theory lecture and identifying keywords based on the concepts of the course material being taught. In support of this theory, the objective of this work is to establish an abstract model for the application of ASR technology in the educational setting for generating lesson summaries. Section 2 provides the theoretical base for the model, focusing in three broad areas of theory. First, a number of applications of ASR technologies in education are discussed, particularly as they pertain to disabled learners and distance-based education. Second, recent trends in statistical models used to perform ASR and the underlying machine learning techniques used to design and train these models are addressed. In addition, voice enhancement and noise management strategies are addressed to account for noise pollution in the educational environment known to negatively affect the accuracy of ASR technologies and resulting transcriptions. The model itself is presented in Section 3 and the proof-of-concept prototype is discussed in Section 4. The prototype is evaluated in Section 5, with a discussion of test results in Section 6 and the study's conclusion in Section 7.

2 Background

This section provides an overview of work related to the concepts required to establish and propose the model for lesson summary generation in this work. To that end, Section 2.1 discusses the utility of ASR as applied to an educational context. The topics of feature extraction and noise cancellation are discussed in Section 2.2.2, as these techniques may be used to improve the reliability of speech features in the educator's speech during lesson transcription. Having clearly detectable features is critical in training a machine learning model to recognise, decode and transcribe speech. Neural networks are a widely used technique to train such a machine learning model and Section 2.3 presents a discussion on these techniques, revealing the trends in the improvement of the technology in recent years. The cumulative knowledge addressed in these theory sections provide the theoretical basis for the model and its application.

2.1 Automatic Speech Recognition in Education

ASR is widely known through commercial applications such as Amazon Alexa, Apple Siri, Microsoft Cortana and the Google Assistant [2]. These products enhance the productivity of their users by providing a means of interaction with a variety of applications. The use of ASR is, however, not limited to applications of convenience; it has also found purchase in the domain of education. ASR is one of the key interfaces that humans use to engage with machines in information technology, research in the field has important significance and the interface itself has wide value in application [3, p.84].

A broad overview is provided by 4 of both the underlying literature and experiments conducted with regard to ASR in the field of education. The technology not only holds the promise of helping students surmount the challenges associated with reading, writing and spelling [4, p. 66], but also provides a facility for teaching staff to improve their pedagogical approach [4, p. 69]. A further application of this technology in the educational space is that it holds the promise of improving the interactions of deaf and second language speakers in the classroom [4, p. 66]. Furthermore, many of the studies included in the work by [4] have found utility as an emancipatory tool for those with physical and/or learning disabilities, as it allows them to write tests and complete projects by means of narration; tasks which would otherwise require the services of a human transcriber. With regard to deaf or hard of hearing students, such an ASR intervention could be invaluable if employed in a classroom environment to provide live captioning or to provide a lesson transcript after the fact. This may also alleviate their dependence on a sign language interpreter, transcriber or hearing aid. It has also been shown that it is possible to integrate an ASR system into the mathematics teaching process at a primary school level [5]. In the work conducted by 5, a voice-activated e-learning prototype was used and it was demonstrated that the use of such a tool could facilitate the learning experience at a primary school level, but also demonstrated feasibility up to a tertiary level; being especially helpful to students who have disabilities, learn on-line or are studying in a second language.

Beyond its utility in helping students with disabilities and in overcoming language barriers, ASR is also being applied to aid in other tasks related to the daily activities of students; including lesson reflection, group discussions and oral presentations [4, p. 66]. ASR can aid these activities by allowing students to have access to a lesson transcript, which facilitates a streamlined approach towards note-taking. It has been shown [6] that ASR technology has definite application in synchronous (real-time) cyber classrooms; an approach that has been pulled into the limelight with the advent of the COVID-19 pandemic [7]. Although such a live approach has been shown to be impeded by issues such as latency and bandwidth, resulting in students missing portions of the lecture [6, p. 367-368], the application of ASR could quite easily alleviate this with the lecturer providing a transcription of the lesson, generated by ASR from a full recording of their speech, after the lesson has been delivered.

ASR systems have also been applied to Elicited Oral Response (EOR) testing. EOR is employed to assess the speaking ability of an examinee by having them listen to a phrase and then restate the phrase to the examiner [8, p. 602]. The application of ASR allows such a test to be conducted automatically, with little to no need for a human examiner. The feasibility of such an ASR-based EOR testing process has been demonstrated [8]. It was found that it is a suitable tool for assessing the validity of content, and is able to do so reliably and in a practical fashion. It represents a means by which to support low-stakes decision making, especially when applied to second language learning. All of these technologies, however, have their downsides. The transcripts generated with an ASR-based tool tend to contain many punctuation errors or to eschew punctuation altogether. It also does not account for any recognition errors which may occur as a result of redundancies in repetitive speech [4, p. 68]. These shortcomings require manual human intervention or a third-party grammar analysis application to correct. Furthermore, a common criticism of using ASR-based technology in a noisy en-

vironment such as a classroom, is that the accuracy of the ASR process may be impacted. There are, however, means by which these noise artefacts could be overcome, minimised or filtered out, allowing the required high rate of accuracy to be maintained.

2.2 *Feature Extraction and Speech Enhancement Techniques for Robust Speech Recognition*

This section focuses on prominent feature extraction and speech enhancement techniques such as noise cancellation and noise reduction; with the noise cancelling techniques serving to augment or compliment feature extraction techniques that are not as noise robust.

2.2.1 *Feature Extraction Techniques used in ASR systems*

Although ASR consists of many steps, one of the most important is undoubtedly that of feature extraction; as this step is used to highlight which components of an input speech signal will serve to support the *recognition* aspect of automated speech recognition. For speech-based audio data to be employed as meaningful data, whether for training purposes or in an active ASR system, it needs to be transformed into a less abstract representation which brings to fore the distinguishable components of speech in the input audio signal. The work done in 9, p. 3 distinguishes between two types of features; temporal features (e.g. short time energy and auto-correlation), which exist within the time domain and spectral features (e.g. fundamental frequency and spectral flux) which exist within the frequency domain of a speech signal. A spectrogram is a time-frequency representation of speech data. It is arguably the fundamental feature extraction method, performed by applying the Fast Fourier Transform (FFT) to the speech signal [10, p. 4525] to transform speech energy through time into frequency estimations through time. As there is a need to highlight the frequency intensity related to speech signals, many feature extraction techniques, related to speech processing, are spectral in nature. The words in most languages are built from smaller components of speech, known as phonemes and the various phonemes consist of distinct formants (fundamental frequencies), evident in their pronunciation. Formants are defined [11, p. 5176] as being created by the resonance of the vocal tract and recognisable as the spectral peaks on the frequency-time spectrum of speech. Therefore, techniques used for spectral feature extraction are highly applicable to ASR systems, allowing them to distinguish between the various formants at their distinct spectral peaks.

Another approach is to apply Linear Predictive Coding (LPC) to create observation vectors, based on the frame-based analysis of speech signals. These vectors may provide an estimation of the poles of the vocal transfer function [12, p .495]. During the process of performing LPC feature extraction, a signal is run through a pre-emphasis process to reduce the occurrence of audio pop at the beginning and end of each frame and to reduce signal discontinuity between frames. The first of these issues is addressed by the application of frame blocking and windowing. This is then followed by an auto-correlation analysis, applied window-wise, during which the LPC coefficients are derived as the observational vectors [12][p. 495]. To counter the lack of robustness of LPC with regard to noise, which may cause interference in the calculation of the coefficients, another spectral-based feature extraction technique, Relative Spectral Filtering (RASTA), may be applied to extract features from the spectogram. The intent of RASTA is to enhance the speech characteristics of the signal by means of the reduction of unwanted and additive noise [12, p. 495]. As part of this technique, a spectral analysis is performed, after which static non-linearities are compressed, a filter is applied based on linear band trajectory in the cepstral or log spectral domain, and then finally, the static non-linearities are decompressed, resulting in the set of RASTA features [12, p. 495]. As stated by [12, p. 496] RASTA-based feature extraction finds purchase where speech recognition needs to be performed in a noisy environment.

Another very widely used feature extraction technique is the Mel-Frequency Cepstral Coefficient (MFCC). This technique focuses on audio frequencies in the range 300Hz to 3400Hz, the critical frequency range of the human vocal tract interpretable by the human ear [12, p. 495; 11, p. 5176]. MFCC is associated with a very efficient method of calculation which follows a similar approach to LPC, in that it involves pre-emphasis of the speech signal, frame blocking and windowing [12, p. 495]. After the windowing process, the FFT is applied and the absolute values it returns are placed in a Mel-filter bank; the log of the filter bank values is calculated and the final MFCC feature vectors are created by applying the discrete cosine transform to each Mel-filter bank. Because it relies on auto-correlaton analysis, MFCC shares the trait with LPC that it is not noise robust [13, p. 358]; although there are many MFCC variants, each with their own improvements and compromises [14]. Other feature extraction techniques include Perceptual Linear Predictive Coefficients (PLP), which are often used in conjunction with RASTA for improved performance; Wavelet-based features; and Linear Predictive Cepstral Coefficients (LPCC), an addition to LPC [13–16]. The work performed in [3, p.83] contrasts LPCC and MFCC, demonstrating that LPCC generally results in lower accuracy but has a faster computation rate while MFCC is slower to compute, but often results in improved recognition accuracy.

2.2.2 *Speech Enhancement Techniques for improved ASR*

The effect of noise has always been a major consideration when implementing ASR system, as a noisy input signal may interfere with the feature extraction process. This may yield unreliable speech features, which in turn leads to a low level of accuracy for the underlying ASR model. There are various major sources of noise, namely background noise from a noisy environment, echoes resulting from recording in confined spaces, feedback resulting from two-way communication when a loudspeaker is too close to the recording device, the background hum caused by an amplifier, quantisation noise resulting from analog to digital conversion, noise resulting from the loss of signal quality when compression is applied, and finally, the distorted signal resulting from reverb, which can also be considered noise pollution [11, p. 5176; 17, p. 318]. Within these various noise sources, noise may be classified according to the following types, which can be mitigated via noise cancellation techniques, namely narrow-band noise, coloured noise, white noise, transient noise pulses and impulse noise [11, p. 5177]. The ultimate goal of a noise cancellation technique is to suppress or de-emphasise

the noise components within an input speech signal [11, p. 5178]. Speech enhancement is the process of enriching the spectral characteristics of a speech signal to make it easier to recognise by machine. This process incorporates noise cancellation techniques, but may also transform the speech signal so that the features within it are more distinguished. 9, p. 2 discuss the principles behind speech enhancement, stating that the performance of such a system is measured based on the quality (detected by human ear) of the enhanced speech and the residual noise level that remains in the speech after enhancement. The main objective of a speech enhancement strategy is to remove any additive noise that exists within a speech signal as a result of the recording being performed in a noisy environment [9, p. 8]. The terms speech enhancement and noise cancellation have become largely synonymous, but 9, p. 9 make the point that feature selection is optimised by "selecting the most uncorrelated features" which often determines the effectiveness of the overall speech enhancement strategy.

Noise cancellation techniques fall under two broad categories, namely linear filtering and adaptive filtering [18; 11; 17]. These noise cancellation techniques may occur in the frequency domain, the time domain or in both [11, p. 5179]. Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) are two types of linear filters [11, p.5178]. The purpose of linear filtering is to remove all frequencies which exist outside of the desired frequency domain by moving linearly along the time domain. In the frequency domain these filtering techniques fall under four categories, namely low-pass, band-pass, band-stop and high-pass filters [11, p.5178]. A combination of these filters may be used to remove all frequencies existing outside of the interpretable range of the human vocal tract. This may be done by applying a high-pass filter to remove frequencies below 300Hz and a low-pass filter to remove frequencies above 3400Hz. This approach, however, will not remove noise within the range of the human vocal tract and may also be problematic in noisy environments because the characteristics of the noise may vary (in intensity) over time and as a result it may not be possible to predict the position in the audio stream at which the noise will occur due to its non-stationary nature [18, p. 336; 11, p. 5178]. Linear filtering has utility in cases where the noise levels are more predictable, such as with amplifier or quantisation noise.

Another form of filtering, the adaptive filter, is based on the mathematical principle of cancellation. This process combines two signals to remove the noise from the original. In such an Adaptive Noise Canceller (ANC), the original signal contains the desired speech to serve as input to an ASR process, but also the noise, which may negatively impact the accuracy of recognition. The second signal serves as a representation of the noise and is adaptively filtered from the original speech signal. This second signal is then subtracted from the original signal [11, p. 5178]. This kind of approach accounts for the dynamic nature of a signal containing speech and other audio. This presents an approach where the parameters and band-pass type are adjusted automatically, depending on the signal, rather than relying on pre-set parameters and a specific band-pass type [11, p. 5180]. This approach is conducted by performing audio framing, after which a unitary transform of the time domain for every frame to the given transform domain is performed. This allows filtering to be applied to individual frames; after which the frames are returned to the time domain by applying the inverse unitary transform. As a final step, the frames are converted back into a congruent audio file, representing the noise to be subtracted from the original audio [11, p. 5179]. An alternative to focusing on the time domain is to approach the problem from the frequency domain. Examples of adaptive filters are the Weiner and Kalman filters and the Recursive Least Squares (RLS) algorithm [17, p. 318]. A Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) or Karhunen-Loève Transform may be used to perform frequency domain transforms [18, p. 336]. Of these approaches, the computational efficiency of DFT makes it the most popular [11, p. 5179]. By applying these techniques to an appropriately selected transform domain, better separation may be achieved between the speech and noise signals. This could result in improved filter estimation which may yield superior speech enhancement performance.

This section has discussed a number of feature extraction techniques and the importance of their appropriate selection, especially in noisy environments. As demonstrated in [19], a feature extraction technique can be chosen to work in conjunction with its underlying ASR model along with a speech enhancement strategy for improved ASR performance.

2.3 *Machine Learning Techniques for Automatic Speech Recognition*

Hidden Markov Models (HMM) were widely used in the early development of ASR systems. The HMM applies probability theory to track the likelihood of the phonetic state transitions within words based on spectral templates of phonetic units which are decoded from the speech signal to predict voiced utterances. This process was eventually augmented through the addition of a Gaussian Mixture to each state of an HMM to model the short-time phonetic units. This combined process is referred to as a Gaussian Mixture Model Hidden Markov Model (GMMHMM) [20]. The advent of faster computer processors has made it feasible to train Deep Neural Network (DNN) models for speech recognition. DNN-based models deal very well with the high dimensionality of speech data, using fewer parameters to optimise [21] and have been shown to outperform GMMHMMs [22] leading to the widespread adoption of DNN architectures for the purpose of performing machine learning in ASR. There have been many major improvements since the first implementations of DNNs for ASR. Two of these improved architectures, namely Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are discussed in the following subsections.

2.3.1 *Convolutional Neural Networks*

A CNN depends on two additional logical layers for performing speech recognition, namely convolutional layers, which act as localised filters, and max pooling layers, which normalise spectral variation[20]. The series of filters in a convolutional layer are applied in an over-lapping fashion across acoustic frames that overlap in time over the entire input space and are often referred to as feature detectors [23, p 1534; 20]. By modelling the associations between the frequency and time domains, and using the local filtering and shared weights of the convolutional layers, a CNN maintains the correlations between these domains and provides a superior result

to the input mapping of DNNs [21]. Varying speaking styles also present a challenge to traditional DNNs as they are not inherently designed to model the translation invariance resulting from formant frequency shifts [21]. Convolutional layers combat this through the use of local filtering and the shared weights. The resulting translation invariance improves the robustness of the model on diverse speech signals by allowing speech features to be detected regardless of their location within local input spaces.

The convolutional layers of a CNN work in conjunction with the max-pooling layers. Max-pooling layers are used to reduce the dimensionality of the resulting convolutions by ensuring maximum filter activation at varying points. This is done to reduce the dimensionality of the convolutions [20]. The performance of the speech recognition task, using CNNs, is also improved by performing pooling in frequency and in time, yielding a robustness towards speaking rate [21]. Convolutional and max-pooling layers may also be applied in alternating pairs to further reduce dimensionality. This improves performance in fully connected hidden layers with fewer trainable parameters. This added robustness towards variations in speech styles, provided by the pairing of convolution and max-pooling layers, allow a CNN to learn the acoustic features for various classes, such as speaker, phoneme and gender [23, p. 1534]. This is another feature which adds to the superiority of CNNs over DNNs for analysing speech signals. Studies have also shown that it is possible to add multiple channels of features, like those from a cochleogram and a spectogram, as inputs to a CNN. This approach allows the CNN to learn from multiple sets of features simultaneously, providing improved performance over learning from a single channel [10].

2.3.2 *Recurrent Neural Networks*

RNNs are very useful for language modelling, but are also capable of performing predictions with regard to the likelihood of a feature by making an association based on previously identified spectral speech features. This makes them highly capable of predicting future words or phonetic units based on previously observed words or phonetic states. RNNs make use of Long Short Term Memory (LSTM) cells to keep track of any associations identified in the previous layer with the current layer. RNNs have been employed [24; 25] for creating robust speech recognition models and these approaches have been improved upon [26] by implementing a light gated architecture. It has also been shown [27] that it is feasible to apply convolution and max pooling as inputs of RNN layers to perform local filtering and pooling. The condensed features are then passed to RNN layers, which make use of LSTM cells or Gated Recurrent Units (GRU) to maintain the contextual associations between features through time. HMMs have also been combined with convolution and LSTM to tie phonetic state transitions for speech recognition [28].

3 Model Overview

This section reintroduces the proposed model, which has been updated since its inception, initially published in [1]. This model is theorised through background literature review and exploration; then reinforced by a proof of concept prototype implementation and the experimentation performed to evaluate it. Figure 1 illustrates the model using a process flow diagram which demonstrates the sequence of processes applied and data flow between them that are applied to generate a topic-oriented lesson summary.

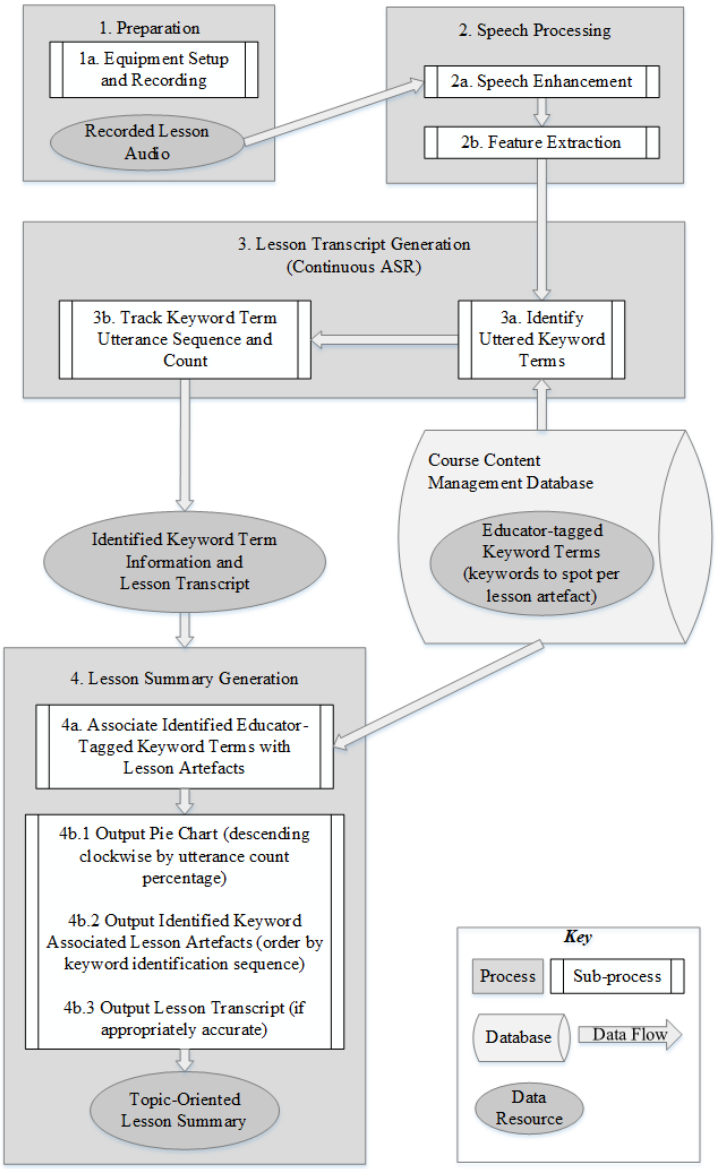


Figure 1: Proposed model for the application of an ASR system used to facilitate topic-oriented lesson summary generation in a noisy educational environment

The subsections that follow discuss the four main processes involved in implementing the proposed model to generate a topic-oriented lesson summary, driven by known keyword terms uttered by the educator during a recorded lesson.

3.1 *Preparation*

To capture the speech of the educator throughout the lesson, a microphone, connected to a recording device is required. There are a number of microphone options, widely available from commercial retailers. This section discusses two variations of microphone setup for use in conjunction with the proposed model. These are

addressed with regard to ease and suitability of use and constraint on budget, as many of these hardware components are expensive and are not all appropriate for a lesson recording setup. There are a number of microphone products to choose from and their quality and capability of audio capture is what sets them apart. Seemingly, the most professional, and consequently the most expensive, are lapel microphones, often used in stage performances and business presentations. Lapel microphones provide a wireless, high fidelity (16000Hz) audio capture solution for the educator, which promotes freedom of movement. These microphones are typically fixed onto the speaker's collar and plugged into a transmitter usually attached at the waist of the speaker, which transmits the stream of their speech signal to a soundboard of sorts (which could be be integrated into a computer or some other audio mixer) for playback and storage. Lapel microphones are intended to capture the primary speaker's voice, with low gain, so as not to capture background noise or over-amplify the speech.

In the medium price range, handheld microphones, with wired and wireless variants, are a more affordable solution for recording the educator's speech during the lesson. Wireless options can still be expensive, but promote freedom of movement compared to their mounted or cabled counterparts. These microphones also transmit the stream of captured speech directly to an audio mixer or computer for storage and playback. Wireless recording solutions provide the most ease of use for the educator or a secondary speaker. The cheapest option for voice capture is the stereo headset, often used for voice communication while gaming or with online voice communication applications like Zoom and Skype. While stereo headsets can be bought cheaply, wireless variants are also more expensive and equally bulky, since the microphone cannot usually be detached from the headset itself and still function. Headsets are also only designed to capture the speech of the primary speaker and cabled variants are more suitable for use while seated, once again limiting the movement of the speaker unless a wireless headset can be used.

Each of these microphone options are also available with hardware-integrated noise cancellation. This additional feature can be costly, but can also remove the need for additional processing for noise management as the lesson is recorded. The choice of which recording devices to use should be made based on budget, fitness of purpose and requirement as well as convenience of use; ideally with a vision of high audio fidelity for a real-world classroom implementation, with the lowest permissible functional costs for testing purposes.

To facilitate voice processing and speech enhancement, discussed in the next section, a combination of microphones can be used. In this case, the primary microphone is used to record the speech of the educator or a student, should they have a question (to be transcribed), in which case they would need to speak into the educator's microphone, with a secondary microphone used to record environmental noise. A mounted wide-band microphone is the cheaper solution to consider for recording environmental noise during the lesson and if appropriately positioned, away from the educator, toward the back of the classroom, will capture noise pollution in the classroom without explicitly recording too much of the educator's speech. These microphones need to record in parallel so that their recordings can be easily aligned and used in combination with a chosen speech enhancement technique whereby the primary recording of the educator's speech is enhanced using the secondary recording of environmental classroom noise to 'cancel' noise captured in the educator's speech signal by the primary microphone.

3.2 Speech Processing

Once all the speech of the educator has been recorded for analysis after the lesson using the equipment setup and an appropriate recording method mentioned in the previous section, the digital speech must be processed to manage environmental noise. The primary requirement of this recording is that it must be of the waveform file format (file extension .wav); the raw, uncompressed audio format that ASR systems are typically developed and trained to be able to recognise speech from using an appropriate feature extraction technique. The quality of the captured audio must be considered in terms of its cost effectiveness. Improved audio quality results in larger waveform audio files as the range of captured frequencies as well as the bit depth and sampling rate increases. These factors contribute to the file size of the lesson audio, and should be considered carefully, especially if a cloud speech recognition service will be used to to transcribe lesson audio, as the file will need to be uploaded for recognition which will take more time if large audio files are uploaded. If a cloud based ASR solution is used to transcribe the recorded lesson, the audio may need to be segmented in overlapping windows of appropriate duration before the ASR system will transcribe the audio due to file size limitations put in place by the ASR service. Resulting segments of lesson transcript will also need to be aligned according to this overlap. In addition, should the audio need to be stored on a web server to be made available to students for review purposes, a large file size will also contribute to storage costs, data usage and buffer time.

Classrooms are notoriously noisy environments owing to the number of students, inevitable chatter, chairs shuffling, corridor activity and a myriad of other possible noise generating events. For this reason, speech enhancement, wherever possible should not be overlooked when implementing speech recognition technology in the educational environment. In the proposed model, speech enhancement is an optional process within the Speech Processing block, and will require the application of a linear or adaptive filter on the recorded lesson audio should this option be taken. This sub-process is optional because if the chosen ASR model uses a noise robust feature extraction technique, applies its own speech enhancement technique or if a noise cancelling microphone is used, then there may not be a need for additional processing for noise management. Over application of speech enhancement can also degrade the educator's speech to a point that it is no longer recognisable by ASR. Speech enhancement should always be considered for classroom applications of ASR technology, but it can be also be accommodated inherently by the feature extraction technique of the ASR system if this technique is noise robust (e.g. RASTA). As stated in the background section, LPC and MFCC are not noise robust feature extraction techniques and will from the application of speech enhancement.

Applying a Speech Enhancement technique will typically require a distinct waveform audio recording of the noise-polluted

speech of the educator during the lesson, recorded by the primary microphone as well as a distinct parallel recording of the noise pollution itself occurring within the classroom, recorded by a secondary microphone. In the ideal recording setup, the noise audio can then be removed from the educator's speech signal almost entirely, resulting in the audible speech of the educator to be analysed by the chosen ASR system, improving its performance and ultimately the accuracy of generated lesson summaries. The chosen speech enhancement technique can be performed during the recording of the lesson, applied to the real-time speech audio stream, using the real-time noise pollution audio stream; or alternatively, the lesson can be recorded in its entirety along with environmental noise and then the speech enhancement technique can be applied to the resulting recording after the lesson once the speech audio and noise audio have been aligned for cancellation. Most importantly, if deemed a necessary intervention, a speech enhancement technique should always be applied prior to ASR.

3.3 Lesson Transcript Generation

The background section covered the capability of ASR systems to transcribe the speech of the educator. Historically, this has been the fundamental reason for implementing speech recognition technology in the educational setting. Lesson transcripts provide many benefits, notably the potential for content reflection and note taking as mentioned in [6, p. 369], as well as improved teaching methods and support of students with disabilities, as described in [4, pp. 65-66], there are clear motivations for lesson transcript generation. In the age of information and with the int eruptions in teaching caused by the COVID-19 pandemic, online learning and video conferencing are becoming more prevalent approaches to education. In these learning environments lesson transcripts are an additional resource for students. Cloud-based ASR systems provide an easily accessible ASR service and can allow researchers to access advanced ASR Models like Google Cloud Speech Recognition and CMU Sphinx. An ASR model can be incredibly challenging to develop from scratch even for an educational institution, due to the mathematical complexities involved in training a machine learning model and the tremendous amount of training data required to optimise it, especially for data as diverse as speech. Depending on the available resources and with recent trends making continuous progress on ASR performance, the task of lesson transcript generation itself might be better suited to a well established, cloud-based ASR model.

3.4 Educator Tagged Keywords and Course Content

The proposed model capitalises on the transcript-generation process by analysing the lesson audio to detect, sequence, count and associate known keyword terms with course content items, prompted by the ASR system's detection of a single utterance of each associated keyword. The count of the number of times each keyword term is uttered throughout the lesson audio can also be maintained to show topic prevalence, allowing students to gauge the importance of various keywords used throughout the lesson and prioritise the amount of time they should spend studying associated topics. The identified keyword terms along with their sequence and utterance counts should only be extracted/calculated based on the new segment of lesson transcript (appended to the overall transcript), after the overlap has been accounted for to avoid duplicate keyword terms being counted.

For the proposed model to be successfully applied, each section within the course content must be tagged with one or more keywords to facilitate the associations between each of the identified keywords in the lesson transcript, with their relevant sections in the course content. This can be achieved by creating meta-tags in a database-bound content management system, storing the relevant keywords for each section of the course. Since the course content is generated from these systems by querying the course content database, the keywords identified through lesson transcription can be queried against the meta-tags within the course content database. Many keywords can be tagged to account for different teaching styles and linguistic preferences. By adding these meta-tags wherever relevant, the identification of a keyword can trickle down through the entirety of the course content and highlight all the sections of the course where the keyword has been tagged. The course content director or the educator should delineate the relevant keywords in each of their associated sections within the course content. The educator must then ensure that they use some of the specific keywords when delivering the lesson so that the model can highlight these associations. The keywords that the educator plans to use are then provided as an input to the transcript generation process to specify which keywords to look for as the speech of the educator is transcribed. This entails sufficient lesson planning and familiarity with subject matter.

3.5 Topic-oriented Lesson Summary Generation

Through the generation of the lesson transcript, the full transcript resulting from segmented ASR performed on the lesson audio, as well as the recorded keyword details were collected. The keyword details include the sequence of its utterance and its utterance count. This information can be summarised and presented to students to help them reflect on what was covered during the lesson, as emphasised in [4, pp. 66-67]. As stated in [4, p. 67] the lesson transcript acts as the primary resource for clarification of what was directly communicated during in the lesson. The additional keyword details captured are supplementary, but allow for the generation of the lesson summary to be structured in accordance with keyword sequence. The identified keyword terms and their utterance counts serve to emphasise the specific topics discussed and their prevalence within the lesson transcript, further bolstering secondary contact with the subject matter.

To finalise the lesson summary and present it to students, one final process must be performed to associate the identified keywords uttered with their relevant sections in the course content. This is achieved by iterating through the list of identified keywords and for each keyword, testing whether it has been tied using meta-tags to a section of the course content, by the educator. If the keyword is assigned then the association is made via the database bound content management system and the section heading along with any additional information stored in the database (such as the page number) is added to the lesson summary. Once all of the keyword information, along with associated course content artefacts have been extracted, the information can be arranged and presented to

students for review. The topic-oriented lesson summary could be structured as follows: first, the details for each keyword term can be stated or plotted on a pie chart (or bar graph) in accordance with their sequence and utterance count to show topic prevalence; second, the sections associated with each keyword can be listed, aligned with the sequence of their related keyword terms; and third, the lesson transcript itself can be added to the lesson summary. It is worth noting that the release of the lesson summary to students could be delayed to allow for the educator to edit the transcript (should it be added to the summary); removing redundancies and adding punctuation where needed to correct mistakes made by the ASR model during transcription.

4 Proof of Concept Prototype Overview

This section discusses the implementation of the prototype based on the proposed model. This prototype was developed in the Python programming language, over versions 3.6 and 3.7, using the Anaconda platform with the Spyder Python editor. Python is a flexible, object-oriented programming language, providing exceptional access to various programming libraries written in Python, C and C++ which are appropriate for handling and manipulating digital speech data. Python was the programming language of choice for the prototype, since it provided access to numerous programming libraries available within the topic area, made available free of charge by its longstanding data science community. The subsections that follow describe the functions of each of the major modules which facilitate the prototype's goal of generating a topical lesson summary. Figure 2 shows the class diagram of the proof of concept prototype developed alongside the model. For simplicity, method input parameters and return types have been omitted.

4.1 Large Vocabulary Continuous Speech Recognition (LVCSR)

The LVCSR Class is responsible for all ASR performed by the prototype. This is achieved through cloud speech recognition services made available by [29], an API which provides access to a number of cloud-based ASR services, some at a cost, and others free for application development, research and testing. The two ASR service providers utilised by the prototype were CMU Sphinx [30] and Google Cloud Speech Recognition [31]. Both of these providers have made their APIs available free of charge for use in the development of speech-driven systems. Regardless of provider, the access methods of this API take as input a single waveform audio file, which the prototype uploads to the given recogniser via the World Wide Web. The chosen recogniser then processes this data on the cloud server and returns the result of transcription of the uploaded audio. Google Cloud Speech Recognition [31] is based in a DNN strategy for ASR and reports to perform speech enhancement technique, while [30] is based in a HMMGMM strategy for ASR and does not perform noise cancellation. Other paid cloud speech recognition solutions made available by the API in [29] include IBM Speech to Text and Microsoft Bing Speech Recognition, among others. What makes this particular ASR service API so useful is that each recognition service is exposed in the same way, making the implementation cross-compatible, regardless of the recogniser chosen at run-time.

The implemented prototype was adapted to generate a lesson transcript by uploading multiple shorter segments of lesson audio to accommodate cloud-based ASR using Google Cloud Speech Recognition and CMU Sphinx. To achieve this, lesson transcripts were generated by sequencing the overlapping segments of lesson audio from start to finish; then performing ASR on overlapping segments of recorded lesson audio with the transcript accumulating as new audio segments were analysed. The overlap was applied to account for sudden cuts in the audio mid sentence owing to the segmentation and also to ensure continuity between the results of transcribed segments of lesson audio, resulting in a more accurate lesson transcript once the transcript was aligned and textual overlap had been accounted for between segments. Through this process of segmentation analysis, the sequence and a counter of each keyword term uttered by the educator were maintained.

4.2 The Wave Handler Module

The Wave Handler module facilitates all the necessary access, storage and segmentation of Waveform Audio Specification files associated with lesson audio analysis. This is achieved using two methods. The first returns the duration of a WAV file specified by file path. The duration of the specified file is calculated by dividing the number of frames in the file by the frame rate of the file, retrieved from its file header. The result of the division is returned as a floating point number representing the duration, measured in seconds, of the specified file. This method primarily serves to determine the boundaries of a WAV file when accessing or segmenting WAV data, but is also used by the control loop to iterate the analysis window over the entirety of the recorded lesson audio.

The second method is used to segment WAV audio and also takes as input a specified WAV file path as well as a start time and an end time. The start time and end time parameters are used to window the speech audio for segmentation and, ultimately, for analysis by the chosen ASR system. This method also accepts three optional parameters: a possible WAV file path to a noise sample WAV file (defaulted to empty), a noise reduction Boolean (defaulted to false) and a noise cancellation Boolean (defaulted to false) indicating whether the segmented audio should have pseudo-random noise injection, noise reduction and/or noise cancellation applied respectively, using the specified noise sample. This allows for any WAV file to be read and segmented according to appropriate time intervals, as desired. When a noise sample is specified for injection and possibly subsequent noise reduction and/or cancellation, a random interval from within the roughly ten hour noise audio file is chosen to be used for injection each time the speech audio is segmented for analysis.

4.3 Database Design for Lesson Summary Generation

In order to establish associations between known keyword terms and sections of course content, a structure was required to retain these associations. The structure of choice to meet this requirement was a MySQL database. This database has three tables: first, a keywords table, which stores a primary key identity and lower-case

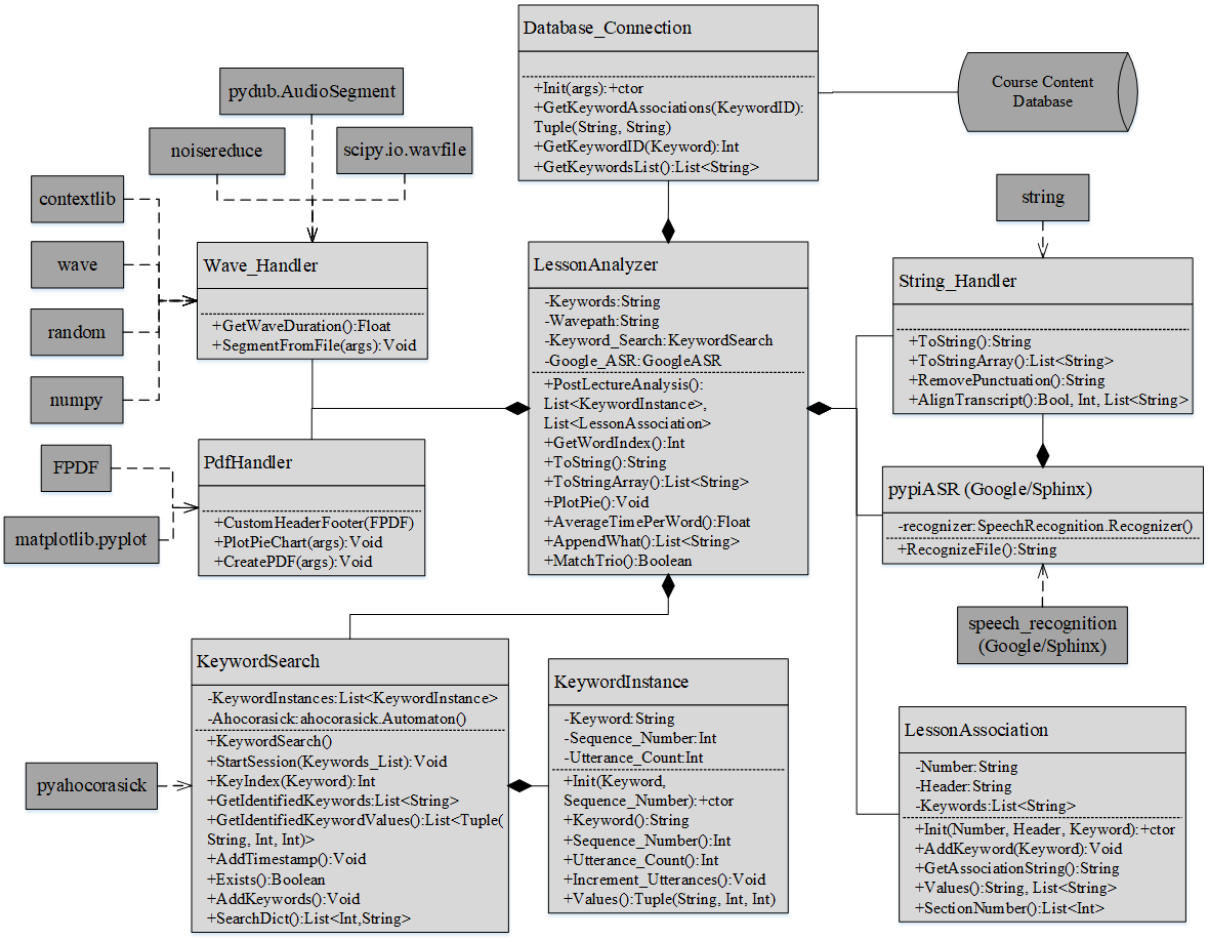


Figure 2: Class diagram of the proof of concept prototype system based on the proposed model

text for each of the known keyword terms the prototype is concerned with; second, a content table, which stores a primary key identity for each course content item, the header text of the course content item, a potential section number as well as a description; and third, a keyword-content relational table that ties each keyword term to one or more course content items. These ties are stored as distinct combinations of keyword term and content item primary key identity pairs which can be queried to yield the course content items associated with any known keyword term that the ASR system has identified.

In the keywords table, spelling is essential and distinctions should be made between standardised (American and British English) spelling. While the associations exist between the identity values of keywords and course content items, the association can only be triggered if the chosen speech recogniser identifies the text of the keyword term while generating the transcript. If the keyword term is misspelled compared to its recognised counterpart, it will not be identified and the association cannot be made. This is also why all string comparisons are performed in lower case since uppercase characters have different values to their lowercase counterparts. The prototype is concerned with seventeen keyword terms in total (four of which are not associated with any course content items), lists fourteen content items (four of which are not associated with a keyword term), and maintains fifteen keyword-content associations, with some associations existing between one keyword and many course content items.

4.4 *The String Handler Module*

The String Handler module deals with all text data transferred by prototype operation between processes. This module has five functions, and all string comparisons implemented by these functions are performed in lower case. The first two methods are used to split a sentence (on space characters) into an array of words, and the second recombines this array back into a sentence in the word order of the array indices. These functions are used to analyse and align word sequences during prototype operation. The third function is defined to remove punctuation and accepts any continuous string as input; then returns the equivalent string with commas, apostrophes, full stops, etc removed. This method is also used to facilitate string comparisons and to ensure that grammatical character differences do not prevent word sequence matches.

The last two methods of the String Handler module are used to align transcript segments. The fourth method performs a sequence match using two string arrays, created by the first method, as input. This method iterates through the words in the first array and determines whether the sequence of words in the first array matches the sequence of words in the second input array. If a match is found,

this function returns a Boolean true; and if not, it returns Boolean false. The fifth method in the String Handler module is used to align transcript segments as the audio is windowed, analysed and transcribed by the speech recogniser. This method applies the previous sequence match method to identify the index where two given string arrays align, and then returns the aligned text from the identified index onward. As input, this method takes two string arrays; the current transcript (or an empty array if this does not yet exist) and the most recently transcribed text (converted to an array) as well as an integer to define how many words in sequence are required to match, and a second integer to define how deep through the array the method should search. The last two parameters are dependant on the duration of the overlap of the speech analysis window and the maximum number of words in the longest keyword term. This alignment method also persists with its search once a match is identified to ensure that the last possible matching index is used to return the overlapping text, and not simply the first identified match. This is to prevent transcribed utterances that are repeated from being matched prematurely, which would result in repeating segments of the transcript. This alignment is performed with each iteration of the prototype over the lesson audio, prior to any keyword term identification and possible association.

4.5 The Aho-Corasick Keyword Search

To identify keyword terms uttered by the educator, an Aho-Corasick keyword search was implemented. This search algorithm, made available by [32] was chosen because it is fast and reliable, and would not result in significant delays when searching for keyword terms between overlapping windows. Technically, any string-search algorithm would suffice here, but the Aho-Corasick search was deemed most appropriate to meet the requirements of the prototype. This search algorithm is implemented as an automaton whereby each keyword term that the prototype is concerned with is added to the watch list of the automaton at run-time. Any known keyword term can be identified in any potential search string. The Aho-Corasick automaton searches for keyword terms in parallel, meaning that the prototype does not need to iterate though each known keyword term and search for it; instead, the automaton has its bound keyword terms set at run-time and will search for any of these terms simultaneously. This behaviour facilitated the identification and accumulation of keyword term utterances stored by a customised Keyword Instance class.

4.6 The Lesson Analyser Program

The Lesson Analyser class combines the logic of each of the modules discussed in the previous subsections to form the execution algorithm of the prototype for lesson summary generation. The prototype is configured at run-time and requires the path to the recorded lesson audio; a directory into which to output the lesson summary; a duration for the analysis window (defaulted to ninety seconds); a duration for the overlap between analysis windows (defaulted to fifteen seconds); a string to specify the speech recogniser to use; an overlap comparison length used by the transcript alignment method (defaulted to twenty five words); and an overlap comparison match count (defaulted to four words). In addition, the algorithm accepts a possible noise audio path to use for noise injection, a noise reduction Boolean and a noise cancellation Boolean to specify which noise management technique(s) to apply.

The prototype's algorithm iterates over the lesson audio in ninety second windows of data with fifteen seconds of overlap in the data between them. If a path to a noise file has been specified, the prototype samples a random ninety second window of data from the noise audio and injects it over the current segment of lesson speech audio. The windowed audio is then written to a temporary directory, where it is uploaded via the World Wide Web to the specified speech recogniser. Once the upload completes, the lesson audio is transcribed and the text result is returned to the prototype. The returned transcript text is then analysed for overlap with the existing transcript text and the new text is appended to the transcript. The newly appended section of the transcript is then analysed for known keyword terms and, upon identification, a database lookup executes to identify the lesson content item(s) that the educator has associated with the keyword term. In addition, the utterance counts for each known keyword are maintained or updated. Identified keyword terms and their associations are then stored in memory and the loop iterates over the next 75 seconds (analysis window minus the overlap window) of lesson audio. This process continues until all the lesson audio has been processed, transcribed and analysed for keyword terms and associations. Once the process completes, the information acquired during analysis is handed off to be ordered and written to a PDF document to be presented to students. This PDF lesson summary document has three sections: first, a pie chart demonstrating the keyword terms uttered during the lesson, the sequence of their utterance and the number of times each was uttered; second, the list of associated course content items for students to review; and third, the lesson transcript. The lesson summary generated by the prototype is then provided to students, either via email or in print to support them in their secondary contact with the lesson material and to point them toward the course content that the lesson was based on.

5 Evaluation of the Prototype

To test the robustness of the prototype for ASR, keyword term identification and subsequent association to lesson content, a series of evaluation metrics were considered and a series of test case scenarios were established to provide quantitative evidence of the prototype's performance. The subsection that follows provides contextual definitions of the chosen recorded (base-line) performance metrics, and the chosen comparative (derived) performance metrics used for statistical evaluation of the prototype. After this, the next subsection defines each of the five established evaluation test scenarios. These scenarios are established to contrast the potential for acceptable lesson summary generation of Google versus CMU Sphinx cloud-based ASR systems against the combined effects of noise injection, noise reduction and noise cancellation.

5.1 Chosen Prototype Performance Metrics

Since the prototype system is driven by the speech of the educator, the lesson summary generation process is dependent on the ASR

system for keyword term identification. Consequently, the accumulation of keyword-term utterances and the association of keyword-terms to lesson content is also dependant on the implemented ASR system. If the prototype fails to identify a keyword-term which is tied (via database relationship) to a lesson association, the association will not be made. In this sense, the proposed model is dependent on the speech recogniser. However, at a functional level, this dependence is binary in terms of the prototype's ability to associate any correctly uttered keyword term with its database-bound relationship to course content, defined by the educator. Established binary predictive machine learning models use base-line metrics known as True Positives, True Negatives, False Positives and False Negatives to measure binary predictive performance.

The four base-line performance metrics are defined in [33] and [34] where the outcomes of the predictive model result in a confusion matrix that describes its positive and negative behaviours. While the prototype system and reflective proposed model presented herein isolate the associations between keyword terms and course content as a database relationship, the ASR component of the prototype system either identifies a known keyword term - resulting in association, or it does not. An example is provided [35] where a binary classifier for diabetes detection is laid out. Contrasted with this example, one can think of each keyword identification as an independent classification resulting in the detection of known lesson content association. The descriptions of base-line performance metrics in the context of the prototype system presented in this work are described in Table 1.

Table 1: List of recorded performance metrics considered to evaluate the prototype

Recorded Metric	Known Keyword Term Recognition	Educator Defined Association
True Positives	Prediction is +ve	Keyword Term has +ve Association(s)
False Negatives	Prediction is -ve	Keyword Term has +ve Association(s)
False Positives	Prediction is +ve	Keyword Term has -ve Association(s)
True Negatives	Prediction is +ve	Keyword Term has -ve Association(s)

Table 1 defines the measurements recorded by the prototype during lesson analysis. A true positive occurs when the prototype identifies a known keyword term uttered by the educator and associates this keyword term to the correct course content item(s). Conversely, false negative occurs when the prototype identifies a known keyword term, but cannot associate the keyword because there is no required database constraint. A false positive occurs when the prototype identifies an unknown, spoken keyword term and associates it with unintended course content item(s). On a functional level, the prototype should not measure any false positives given that the relational database constrains the associations between keyword terms and lesson content. A true negative occurs when the prototype identifies a known, spoken keyword term, but there is no association between the keyword term and course content item(s). Due to the database constraints, true negatives can be eliminated by ensuring that only keyword terms with existing associations can be identified by the prototype.

The number of utterances of the keyword term gives the ASR component multiple attempts at keyword-term identification; however, if the particular term (e.g. the word "Euclidian") is not in the ASR system's lexicon, then the likelihood of detection falls to zero and the association cannot be made until further training of the ASR component. This constraint or dependency of the prototype on its ASR component is a trade-off between the number of times a keyword term is uttered and the reliability of the speech recogniser in identifying the keyword term.

Using the baseline performance metrics described allows for the measurement of derived performance metrics that can be calculated to provide a better indication of the prototype system's performance. Table 2 defines the formulas according to [33] and [34], which are used to calculate each of these comparative performance metrics using the values captured by the baseline performance metrics.

Table 2: List of comparative performance metrics considered to evaluate the prototype

Performance Metric	Measurement Formula
Accuracy	$\frac{TP+TN}{TP+FP+FN+TN}$ (1)
Precision	$\frac{TP}{TP+FP}$ (2)
Recall (Sensitivity)	$\frac{TP}{TP+FN}$ (3)
Specificity	$\frac{TN}{TN+FP}$ (4)
F1-Score	$2\times\frac{Recall\times Precision}{Recall+Precision}$ (5)

As previously addressed, the prototype's design constraints and dependencies, in theory, prevent it from capturing false positives and can prevent it from capturing true negatives with optimal configuration of the prototype's database component. Although these are both measured for, when their values equal zero, they become trivial to some of the comparative metrics that use them in associated formulae. In addition to this factor, the third-party nature of using a cloud-based ASR component abstracts the prototype (as the user) from the base-line metrics measured by the ASR model itself when decoding (transcribing) the educator's speech. Thus, we cannot measure the ASR model's performance directly using these comparative metrics.

Table 3: List of considered comparative performance metrics to use in evaluating the prototype system

Performance Metric	Metric Formula
Average Accuracy (AA)	$\frac{\sum_{n=1}^{n} Accuracy_n}{N}$ (6) where n is the test case number and N is the total number of tests conducted for the test case
Average True Accuracy (ATA)	$\frac{\sum_{n=1}^{n} TAccuracy_n}{N}$ (7) where FP = 0, TN = 0, n is the test case number and N is the total number of tests conducted for the test case
Global Average Accuracy	$\frac{\sum_{n=1}^{n} AA_n}{N}$ (8) where n is the test case number and N is the total number of tests conducted for the test case
Global Average True Accuracy	$\frac{\sum_{n=1}^{n} ATA_n}{N}$ (9) where FP = 0, TN = 0, n is the test case number and N is the total number of tests conducted for the test case

The database constraints on the prototype system coupled with the lack of base-line recognition metrics of the underlying ASR model, mean that some of the comparative metrics from Table 2 do not truly reflect the intention of the metric in the context of the prototype system. Because the count of false positives measured by the prototype will always be zero, precision and specificity were deemed inappropriate metrics for prototype evaluation from the outset. This consequently eliminated F1-Score as a potential measure for contrasting the performance of cloud-based Google and Sphinx recognition services. If we cater for FP = 0 and TN = 0, our measure for sensitivity results in the same value for Accuracy. By this deduction, accuracy was deemed the most appropriate metric to measure prototype performance. Table 3 provides formulae for the types of accuracy, measured for the evaluation of the prototype.

The accuracy metrics in Table 3 will provide a measure of accuracy for the prototype over the series of tests conducted across test cases. To cater for the potential of true negative values being measured as zero, the Average Accuracy (AA), which includes the number of true negatives measured, is contrasted with the Average True Accuracy (ATA) where the number of true negatives measured is assumed to be zero. These accuracy metrics are reported for each test case and the Global Average Accuracy and Global Average True Accuracy is reported across all the test cases to provide a final measure of prototype performance using the given speech recogniser.

5.2 Specification of Test Samples

To test the ASR prototype system, two audio readings of the article 'Speech Recognition for Learning' [36] were recorded using a stereo headset. The participants, one male and one female, were encouraged to enunciate their speech as well as possible and to read at a comfortable pace in their natural voices. In addition to these two audio recordings, a total of twelve hours of captured classroom background noise was acquired to use as sample audio for noise injection. These three audio test samples were all recorded at 16000Hz with a 256kbps bitrate as signed 16-bit PCM encoded single channel (mono) waveform audio (.wav).

5.3 Prototype Test Cases

A series of test cases were established to test the prototype system's performance using the test samples and performance metrics discussed. These test cases simulate alternative classroom equipment setups and environmental noise constraints on the prototype. Table 4 indicates whether the input data had noise injection (NI), noise reduction (NR) and/or noise cancellation (NC) applied for each of the test case scenarios, as well as the recogniser used to identify keyword terms.

Test Case T is established to provide an indicator of the functional performance of the prototype system assuming that the underlying ASR system is completely accurate. Rather than performing speech recognition on the speech audio, this test case instead has the algorithm operate on the raw text of the document which was read aloud. This allows for the algorithm to be tested and modified until it was proven to be working ideally on a functional level, with its dependency on ASR accuracy removed. Test case A, on the other hand, is used to measure the prototype's performance working with the ASR system to provide a measure of the prototype system's accuracy with its ASR dependency, but without any environmental noise. Test case B introduces this environmental noise by isolating a random sample within the specified classroom noise sample audio and overlaying this randomly chosen sample on the speech audio prior to ASR analysis. This pseudo-random noise injection is applied at each overlapping window of lesson audio analysis. Test case B is designed to simulate spontaneous noise that may occur in the educational environment and be recorded by the educator's

microphone as they deliver their lesson. Test cases A and B simulate a single microphone setup to capture the educator's voice.

Table 4: Test case scenario specifications used in evaluating the prototypes' performance with different audio pre-processing techniques

Test Case	NI	NR	NC	ASR System
T	FALSE	FALSE	FALSE	Aho-Corasick Search
A	FALSE	FALSE	FALSE	Google, Sphinx
B	TRUE	FALSE	FALSE	
C	TRUE	TRUE	FALSE	
D	TRUE	FALSE	TRUE	
E	TRUE	TRUE	TRUE	

Table 5: Sphinx Average Accuracy and Average True Accuracy per Test Case

Test Case	Average Accuracy	Average True Accuracy
A	0,6333	0,7000
B	0,3907	0,4730
C	0,1660	0,2350
D	0,5300	0,6170
E	0,4600	0,5490
Global Average	0,4360	0,5148

Test cases C, D and E are concerned with noise management interventions and are intended to simulate a dual-microphone equipment setup where one microphone records the educator's voice and the other records environmental classroom noise. These recordings are then aligned, and the recorded environmental noise is used as input to reduce or cancel environmental noise recorded by the educator's microphone. Note that the prototype would be drastically less effective if the microphone recording the environmental noise were also to record the educator's voice, as the noisy sample would include the speech, which would then be cancelled or reduced. Thus, this prototype is not appropriate for a situation where the educator's voice is being amplified by a loud speaker. The educator's microphone in this case is intended only to record their speech and inherently, any environmental noise generated in the classroom which is also captured.

5.4 Prototype Test Results

The prototype was tested one hundred times for each of the test cases defined in Table 4 using Google Cloud Speech Recognition and then using CMUSphinx ASR. Test case T was performed without a recogniser and instead used the Aho-Corasick search to identify keyword terms. The prototype was debugged and tested in a cyclical manner until the results of test case T showed an accuracy of 100% and it had been demonstrated that the logic used to measure each base-line performance metric was accurate. Ultimately, the prototype was able to demonstrate 100% accuracy on test case T and any doubt of inaccuracy of base-line metric measurement was resolved. This prompted the next phase of testing of the prototype using a speech recogniser to transcribe the educator's speech and then to generate a lesson summary. Table 5 shows the performance of the prototype when generating lesson summaries using CMUSphinx cloud-based ASR.

Table 6 shows the performance of the prototype when generating lesson summaries using Google Cloud Speech Speech Recognition.

Table 6: Google Average Accuracy and Average True Accuracy per Test Case

Test Case	Average Accuracy	Average True Accuracy
A	1,0000	1,0000
B	0,8927	0,9390
C	0,4387	0,5360
D	0,9693	0,9900
E	0,8860	0,9400
Global Average	0,8373	0,8810

6 Discussion of Prototype Performance

The discussion of test results will centre around the reported Average True Accuracy across test cases for each recogniser. The results of the prototype testing reported in Table 5 and Table 6 for test case A demonstrate a difference in accuracy of 30% for lesson summary generation between Google Cloud Speech Recognition (CSR) and CMUSphinx ASR. This can be attributed to the design difference of these machine learning models. It has been shown that DNN-based approaches to ASR outperform traditional HMM-based models as they provide more modelling complexity. When noise is injected into the educator's speech in test case B, the performance of both models declines, but Google CSR is reportedly noise robust so the drop in accuracy is far less significant than that of CMUSphinx, where Google had a 6.01% reduction in accuracy, contrasted with a drop in accuracy of 22.7% for CMUSphinx. Both recognisers saw a significant reduction in accuracy with noise injection and then with noise reduction applied in test case C. From a qualitative perspective, the noise reduction algorithm had the effect of dampening the speech and was likely de-emphasising spectral formants, leaving the speech sounding hollow and making it difficult to distinguish utterances by ear. By contrast, when noise cancellation was applied instead in test case D, both speech recognisers saw an improvement in performance; with Google improving in accuracy by 5.1% compared with a greater increase in accuracy for CMUSphinx of 14.4%. This demonstrates that noise cancellation was the ideal noise management technique applied by the prototype, most notably with Google CSR which lost only 1% accuracy of lesson summary

generation when comparing performance without noise injection as opposed to noise injection with noise cancellation applied. The results of test case E are the most surprising where noise reduction and noise injection was applied, a balance in accuracy between noise reduction and noise cancellation was achieved, seemingly diminishing the negative effects of the noise reduction technique, while still removing background noise and retaining the integrity of the speech signal. Nevertheless, the prototype showed satisfactory performance using either recogniser under no noise injection, or while only applying noise cancellation to the noisy speech data. Overall, using Google Cloud Speech Recognition in combination with noise cancellation demonstrated the best performance of the prototype for speech-driven lesson summary generation in a noisy environment.

7 Conclusion and Future Work

This paper has discussed some of the key topic areas involved in incorporating ASR systems for use in educational settings. A model was proposed for the application of an ASR system in a noisy educational environment to automatically generate a lesson summary, driven by the speech of the educator. A prototype system was then developed based on the proposed model and improved and adapted alongside it. The prototype goes beyond the baseline utility of transcribing the speech of the educator with additional analysis on the transcribed text used to identify and associate keyword terms to course content artefacts, summarising this information by monitoring the number of times each keyword is mentioned, providing a reference point for keyword terms and directing students to the underlying course content from which they originate via database bound associations. These relatively simple additions, along with the lesson transcript can allow for the educator to speak at length about the given topics, cross-referencing to related course content artefacts and potentially helping to guide the flow of their lesson, with the peace of mind that relevant sections are made known to the students. Additionally, secondary contact with the lesson transcript after it has been taught helps students with making notes and the additional reference points could help in prioritising certain topics and reaffirming/reinforcing what was communicated during the lesson.

To further the proposed model, future work involves the development, testing and classroom implementation of a more advanced system based on the proposed model and the proof of concept prototype presented and argued herein. The classroom prototype would need to be evaluated using both qualitative and quantitative research methods. For qualitative evaluation, a survey could be designed to gauge both student and educator perspectives on the system's utility, potential for improvement and the overall sentiment of the system implemented in the classroom. Quantitative evaluation could be incorporated into the survey whereby aspects of the prototype's utility could be rated on a Likert scale or Linear Numeric scale to indicate positive or negative sentiment towards particular aspects of the system and the lesson summaries produced. The prototype's database components could also be expanded to cater for multiple lessons across multiple courses and the performance evaluation conducted in this work could then be applied at various levels to measure prototype performance on a lesson and course level and also performance contrasted between courses with varying subject matter. The model could also be improved by automating the definition of keyword terms and their association with course content items so that the educator would no longer need to. This could be done using unsupervised machine learning whereby a machine learning model would have access to course content and automatically extract and assign keyword terms to course content artefacts.

References

[1] P. Blunt, B. Haskins, "A Model for Incorporating an Automatics Speech Recognition System in a Noisy Educational Environment," in 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC), 1–7, IEEE, 2019, doi:10.1109/IMITEC45504.2019.9015907.

[2] G. López, L. Quesada, L. A. Guerrero, "Alexa vs. Siri vs. Cortana vs. Google Assistant: a comparison of speech-based natural user interfaces," in International Conference on Applied Human Factors and Ergonomics, 241–250, Springer, 2017.

[3] J. Meng, J. Zhang, H. Zhao, "An Overview of Speech Recognition Technology," 199–202, IEEE, 2012, doi:10.1109/ICCIS.2012.202.

[4] R. Shadiev, W.-Y. Hwang, N.-S. Chen, Y.-M. Huang, "Review of Speech-to-Text Recognition Technology for Enhancing Learning," Journal of Educational Technology & Society, **17**(4), 65–84, 2014.

[5] A. R. Ahmad, S. M. Halawani, S. Boucetta, "Using Speech Recognition in Learning Primary School Mathematics via Explain, Instruct and Facilitate Techniques," Journal of Software Engineering and Applications, **07**(4), 233–255, 2014, doi:10.4236/jsea.2014.74025.

[6] W.-Y. Hwang, R. Shadiev, C.-T. Kuo, N.-S. Chen, "Effects of Speech-to-Text Recognition Application on Learning Performance in Synchronous Cyber Classrooms," Journal of Educational Technology & Society, **15**(1), 367–380, 2012.

[7] J. Crawford, K. Butler-Henderson, J. Rudolph, B. Malkawi, M. Glowatz, R. Burton, P. Magni, S. Lam, "COVID-19: 20 countries' higher education intra-period digital pedagogy responses," Journal of Applied Learning & Teaching, **3**(1), 1–20, 2020.

[8] T. L. Cox, R. S. Davies, "Using Automatic Speech Recognition Technology with Elicited Oral Response Testing," CALICO Journal, **29**(4), 601–618, 2012, doi:10.11139/cj.29.4.601-618.

[9] N. Das, S. Chakraborty, J. Chaki, N. Padya, N. Dey, "Fundamentals, present and future perspectives of speech enhancement," International Journal of Speech Technology, 1–19, 2020, doi:10.1007/s10772-020-09674-2.

[10] A. Tjandra, S. Sakti, G. Neubig, T. Toda, M. Adriani, S. Nakamura, "Combination of two-dimensional cochleogram and spectrogram features for deep learning-based ASR," in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4525–4529, IEEE, 2015, doi: 10.1109/ICASSP.2015.7178827.

[11] S. Lakshmikanth, K. R. Natraj, K. R. Rekha, "Noise Cancellation in Speech Signal Processing-A Review," International Journal of Advanced Research in Computer and Communication Engineering, **3**(1), 5175–5186, 2014.

[12] K. Gupta, D. Gupta, "An analysis on LPC, RASTA and MFCC techniques in Automatic Speech recognition system," in 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence), 493–497, IEEE, 2016, doi:10.1109/CONFLUENCE.2016.7508170.

[13] U. Shrawankar, V. Thakare, "Feature Extraction for a Speech Recognition System in Noisy Environment: A Study," in 2010 Second International Conference on Computer Engineering and Applications, 358–361, IEEE, 2010, doi:10.1109/ICCEA.2010.76.

[14] A. S. Mukhedkar, J. S. R. Alex, "Robust feature extraction methods for speech recognition in noisy environments," in 2014 First International Conference on Networks Soft Computing (ICNSC2014), 295–299, 2014, doi: 10.1109/CNSC.2014.6906692.

[15] U. Shrawankar, V. Thakare, "Techniques for Feature Extraction in Speech Recognition System: A Comparative Study," International Journal Of Computer Applications In Engineering, Technology and Sciences (IJCAETS), **2**(2), 412–418, 2013.

[16] U. Sharma, S. Maheshkar, A. N. Mishra, "Study of robust feature extraction techniques for speech recognition system," in 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 654–658, 2015, doi:10.1109/ABLAZE.2015.7154944.

[17] M. Sathya, D. S. P. Victor, "Noise Reduction Techniques and Algorithms For Speech Signal Processing," International Journal of Scientific & Engineering Research, **06**(12), 317–322, 2015.

[18] U. Shrawankar, V. Thakare, "Noise Estimation and Noise Removal Techniques for Speech Recognition in Adverse Environment," in Z. S. S. V. A. A. D. Leake, editor, 6th IFIP TC 12 International Conference on Intelligent Information Processing (IIP), volume AICT-340 of *Intelligent Information Processing V*, 336–342, Springer, 2010, doi:10.1007/978-3-642-16327-2_40.

[19] J. Yang, W. Zhenli, "Noise robust speech recognition by combining speech enhancement in the wavelet domain and Lin-log RASTA," 415–418, IEEE, 2009, doi:10.1109/CCCM.2009.5267457.

[20] O. Abdel-Hamid, A. Mohamed, H. Jiang, G. Penn, "Applying Convolutional Neural Networks concepts to hybrid NN-HMM model for speech recognition," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4277–4280, IEEE, 2012, doi:10.1109/ICASSP.2012.6288864.

[21] T. N. Sainath, C. Parada, "Convolutional Neural Networks for Small-Footprint Keyword Spotting," in Proc. Interspeech 2015, 1478–1482, Google, Inc. New York, U.S.A., 2015.

[22] J. Huang, J. Li, Y. Gong, "An analysis of convolutional neural networks for speech recognition," in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4989–4993, IEEE, 2015, doi:10.1109/ICASSP.2015.7178920.

[23] O. Abdel-Hamid, A. Mohamed, H. Jiang, L. Deng, G. Penn, D. Yu, "Convolutional Neural Networks for Speech Recognition," IEEE/ACM Transactions on Audio, Speech, and Language Processing, **22**(10), 1533–1545, 2014, doi:10.1109/TASLP.2014.2339736.

[24] O. Vinyals, S. V. Ravuri, D. Povey, "Revisiting Recurrent Neural Networks for robust ASR," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4085–4088, IEEE, 2012, doi:10.1109/ICASSP.2012.6288816.

[25] C. Weng, D. Yu, S. Watanabe, B. F. Juang, "Recurrent deep neural networks for robust speech recognition," in 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 5532–5536, IEEE, 2014, doi:10.1109/ICASSP.2014.6854661.

[26] M. Ravanelli, P. Brakel, M. Omologo, Y. Bengio, "Light Gated Recurrent Units for Speech Recognition," IEEE Transactions on Emerging Topics in Computational Intelligence, **2**(2), 92–102, 2018, doi:10.1109/TETCI.2017.2762739.

[27] S. . Arık, M. Kliegl, R. Child, J. Hestness, A. Gibiansky, C. Fougner, R. Prenger, A. Coates, "Convolutional Recurrent Neural Networks for Small-Footprint Keyword Spotting," in Proc. Interspeech 2017, 1606–1610, ISCA, 2017, doi:10.21437/Interspeech.2017-1737.

[28] H. Dridi, K. Ouni, "Applying long short-term memory concept to hybrid "CD-NN-HMM" model for keywords spotting in continuous speech," in 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 413–418, IEEE, 2018, doi:10.1109/ATSIP.2018.8364510.

[29] A. Zhang, "SpeechRecognition 3.8.1," Retrieved December 5, 2017, from https://pypi.org/project/SpeechRecognition/, 2017.

[30] C. M. University, "CMUSphinx Open Source Speech Recognition," Retrieved October 23, 2019, from https://cmusphinx.github.io/wiki/, 2019.

[31] Google, "Google Cloud Speech-to-Text API," Retrieved January 17, 2020, from https://cloud.google.com/speech-to-text/docs/apis, 2020.

[32] W. Muła, "pyahocorasick 1.4.0," Retrieved January 14, 2019, from http://www.ldonline.org/article/38655/, 2019.

[33] A. Mitrani, "Evaluating Categorical Models," Retrieved November 28, 2019, from https://towardsdatascience.com/evaluating-categorical-models-e667e17987fd, 2019.

[34] A. Mitrani, "Evaluating Categorical Models II: Sensitivity and Specificity," Retrieved December 6, 2019, from https://towardsdatascience.com/evaluating-categorical-models-ii-sensitivity-and-specificity-e181e573cff8, 2019.

[35] S. Ghoneim, "Accuracy, Recall, Precision, F-Score & Specificity, which to optimize on?" Retrieved April 2, 2019, from https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124, 2019.

[36] N. C. for Technology Innovation, "Speech Recognition for Learning," Retrieved June 12, 2017, from http://www.ldonline.org/article/38655/, 2010.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 541-551 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Complex Order $PI^{a+jb}D^{c+jd}$ Controller Design for a Fractional Order DC Motor System

Pritesh Shah[*,1], Ravi Sekhar [1], Iswanto Iswanto[2], Margi Shah[3]

[1]*Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University) (SIU), Pune, India*
[2]*Department of Electrical Engineering, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia*
[3]*Department of Mechanical Engineering, AISSMS College of Engineering, Pune, India*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 04 March, 2021
Online: 17 March, 2021

Keywords:
Fractional calculus
Fractional modeling
Fractional order controller
DC motor
Complex order controller
Bode diagram
Root locus

ABSTRACT

Industry 4.0 implementation stipulates effective actuator control. In the present work, a complex order $PI^{a+jb}D^{c+jd}$ (COPID) controller was designed for a fractional order model of a direct current (DC) motor system. For comparisons, the DC motor system model was also controlled using the conventional proportional integral (PI), proportional integral derivative (PID), proportional resonant (PR) and fractional order PID controllers (FOPID). Time domain results indicated that the PR controller performed exceedingly well for output signal responses, but fared poorly in case of control signal specifications. The PI controller responses suffered from high time domain characteristics for both control and output signals. The PID controller performed moderately in terms of time domain and peak overshoot metrics. The FOPID controller attained the best time domain characteristics, but was unable to effectively limit the control and output signal peak overshoots. It was only the COPID controller, that successfully minimised / eliminated peak overshoots in control and output signals (0.1 % and 0.0 % respectively). Moreover, the COPID controller was also successful in limiting the rise, peak and settling times. In addition, Bode diagram, root locus plot were obtained and system gain parameters were varied to confirm the robustness of the proposed COPID controller. Thus, COPID controller promises to be an effective solution towards accurate and robust actuator control in modern manufacturing.

1 Introduction

This paper is an extension of the paper presented at the IEEE International Conference on Mechatronics, Robotics and Systems Engineering, MoRSE 2019 [1]. Therein, a direct current (DC) motor system was successfully modeled by the application of fractional calculus to closed loop system identification approach. Of the four identified fractional order models, the best performing model comprised of three parameters and attained R-squared 0.9942, root mean square error (RMSE) 0.0084 and sum of squared estimate of errors (SSE) 0.0711. In the current work, that fractional order model has been utilised to implement PI, PR, PID, fractional order PID (FOPID) and complex order $PI^{a+jb}D^{c+jd}$ (COPID) controller designs with an aim to achieve stable and robust control of the DC motor system with minimum time domain characteristics.

The proportional, integral and derivative (PID) controllers have been implemented in industrial control systems on a large scale owing to their simplicity of design and maintenance [2]. Recently, fractional calculus has been employed to solve various complex problems in engineering and science [3]–[7]. One of these applications is designing controllers using fractional calculus. The fractional order controllers have been shown to exhibit superior control characteristics in numerous applications including the non-minimum phase systems [8]–[12]. This is due to the additional two parameters available for control in the fractional model structure. The complex order $PI^{a+jb}D^{c+jd}$ controller is a logical extension of the fractional calculus based control as it includes two more parameters for control as compared to the FOPID controller [13]–[16].

In [17], the authors implemented a complex order structure to identify hexapod robot movement on the basis of a transfer function for the foot-ground system. In [18], the authors obtained a transfer function to solve a complex order differential equation. In [19], the author employed genetic algorithm to optimize complex order controllers for non linear and linear systems. In [20], the authors investigated robust control of a non linear fuel cell system using complex order architecture. In [21], the authors modeled HIV infection drug resistance using a complex order model. In [16], the authors reviewed time domain, stability and frequency domain in

[*]Corresponding Author: Pritesh Shah, Symbiosis Institute of Technology, Pune, India Contact No. +91 20 28116400 & Email pritesh.ic@gmail.com

complex order modeling of various systems. In [22], the authors tuned complex order controller by numerical optimization for a time delay resonant plant. In [23], the authors tuned complex order controller using genetic algorithm. In [13], the authors compared complex controller performances to integral and fractional order controllers for fractional order systems. They reported superior time domain characteristics attained by the complex order controllers. However, they did not apply complex order controllers to a DC motor system. A complex order controller was designed for DC motors [24]. However, the designed controller in this work [24] was unable to minimise the overshoot in the step response. minimizing peak overshoot in the output signal response is critical to prevent any undesired process variation which may damage the actuated machinery in the long run. Similarly, minimizing control signal overshoots is critical to prevent damage against voltage spikes and enhance the service life of DC motors in industry [25, 26].

The main contribution of the current work lies in the identification of the most suitable DC motor controller design based on minimal output/control signal peak overshoots and settling times among the various controllers considered in the scope of the current study. The following section (2) details upon the controller design methodology adopted in the current work pertaining to the PI, PID, PR, FOPID and COPID controllers. Time and frequency domain results of the investigated controllers are compared and discussed in section 3. Finally, conclusions of the present work are presented in the section 4.

2 Controller Design Methodology

The current work addresses the constant torque region for DC motor modeling and control. The field weakening region (figure 1) is not considered in the present work.

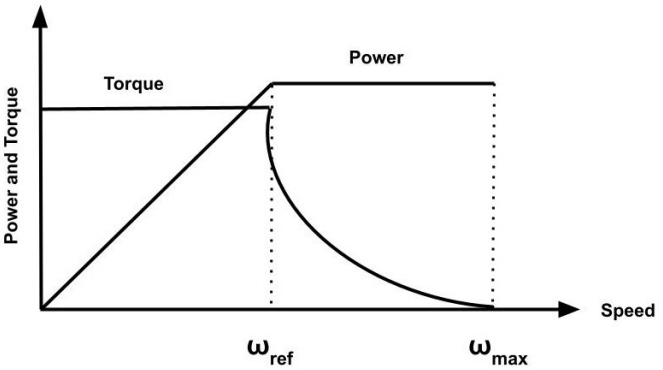


Figure 1: Field weakening region [27]

Following sub-sections describe the PI, PID, PR, COPID and FOPID controllers for the DC motor system considered in the current study.

2.1 PI and PID controller

The PID controller is popular in industry because of its compatibility with the PLCs and its simple structure[2, 28]. The following PID controller structure was used for the DC motor system [29].

$$C(s) = K_P + K_I\left(\frac{1}{s}\right) + K_D\left(\frac{N}{1 + N\frac{1}{s}}\right) \tag{1}$$

where, N is the bandwidth of the low pass filter on the derivative, K_P is proportional gain, K_I is integral gain and K_D is derivative gain. This controller was tuned by the auto tuning capability available in the Matlab Simulink block. Auto-tuning can be done via time and frequency based methods. In this work, frequency based auto tuning method was implemented. In the case of PI controller, K_D is zero and there is no need of derivative filter. It's structure is given as follows.

$$C(s) = K_P + K_I\left(\frac{1}{s}\right) \tag{2}$$

2.2 PR Controller

The PI controller yields infinite gain at s = 0 and generally responds in a slow mode to step inputs without any steady state error. Moreover, it is unable to trace sinusoidal reference input. The PR controller yields infinite gain at zero phase shift and resonant frequency. Its form is given as follows [30]–[32].

$$C(s) = K_P + \frac{K_I * s}{s^2 + \omega^2} \tag{3}$$

2.3 FOPID controller

FOPID controllers are based on fractional calculus wherein the orders of derivation and integration are real numbers [9, 33]. Fractional calculus has been defined in many ways. The following definitions are generally applied in control system applications [34].

2.3.1 Grunwald-Letnikov definition

The GL (Grunwald-Letnikov) definition is very effective for obtaining numerical solutions of the fractional differential equations [35]. This definition is expressed as follows

$${}_aD_t^{\alpha} = \lim_{h\to 0} \frac{1}{h^{\alpha}} \sum_{r=0}^{\left[\frac{t-a}{h}\right]} (-1)^r \binom{n}{r} f(t - rh) \tag{4}$$

where, the integer part is represented by $[\frac{t-a}{h}]$; t and a are the limits of operator; n is an integer subjected to $n - 1 < \alpha < n$. The binomial coefficient is expanded as follows -

$$\binom{n}{r} = \frac{\Gamma(n+1)}{\Gamma(r+1)\Gamma(n-r+1)} \tag{5}$$

Similarly, the Gamma function in the above equation can also be defined as -

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \Re(Z) > 0 \tag{6}$$

2.3.2 Riemann-Liouville definition

The Riemann-Liouville definition is appropriate for the determination of analytical solutions of simple functions like $e^t, t^b, \cos(t)$ [36]. Riemann and Liouville applied fractional operators to derive formulae for the integration of arbitrary numbers and for solving differential equations respectively. The Riemaann-Liouville definition combines the distinct approaches followed by Riemann and Liouville in a composite formula expressed as follows -

$$ {}_aD_t^{\alpha} = D^n J^{n-\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_a^t \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau \qquad (7) $$

where, J is the integral operator, α is a real number, n is an integer subjected to $n-1 < \alpha < n$ and t , a are the limits of integration. For example, if α is 1.8, n will be two because $1 < 1.8 < 2$.

2.3.3 M. Caputo definition

The M. Caputo definition is very popular among engineers because it directly relates the type of fractional derivative to the corresponding type of initial conditions [37]. In this definition, initial conditions like $y(0)$ and $y'(0)$ are allowed; unlike $y^{0.5}(0)$ (fractional conditions) [8]. The Caputo definition of fractional calculus is given as follows -

$$ {}_aD_t^{\alpha} = \frac{1}{\Gamma(n-\alpha)} \int_a^t \frac{f^n(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau \qquad (8) $$

where, α is a real number, n is an integer subjected to $n-1 < \alpha < n$ and t, a are the limits of integration. As stated before, fractional calculus is a branch of mathematics wherein the orders of integration and derivation are real numbers [9, 33, 38]. This real number feature has a significant impact towards improvement of the controller performance [39]–[41]. The classical PID Controllers are particular cases of fractional controllers where λ and μ are equal to one (figure 2). With reference to the PID plane, this implies that instead of moving between four fixed points it is possible for λ and μ to move across the entire plane [42].

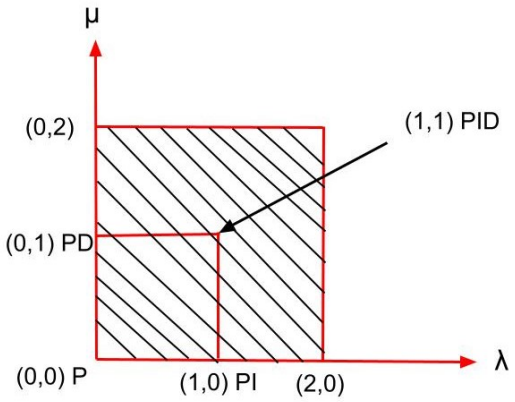


Figure 2: FOPID region in λ - μ plane [9]

The fractional PID controller structure implemented in the current work (figure 3) is given as follows.

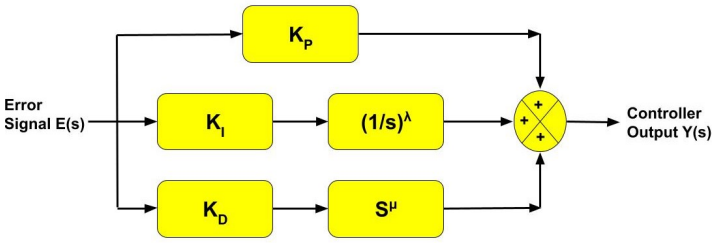


Figure 3: Fractional PID controller structure [43]

$$ C(s) = K_P + K_I \left(\frac{1}{s}\right)^{\lambda} + K_D s^{\mu} \qquad (9) $$

where λ is the order of integration and μ is the order of derivative. In FOPID, there are two more parameters in the controller structure as compared to PID controller. Because of this, two more specifications can be achieved using the FOPID controller. The parameter range of orders is generally accepted 0 to 2 for the stability of the closed loop system because the locations of closed-loop poles lie mostly in the first sheet of Riemann [44] in this order range. The same approach has been followed in the current study as well. The fractional differentiator can be synthesised for FOPID controller application using Oustaloup's recursive approximation method as follows -

$$ s^v \approx K \prod_{k=-N}^{N} \frac{1 + s/\omega_k}{1 + s/\omega_k^{'}} \qquad (10) $$

$$ \omega_u = \sqrt{\omega_h \omega_b} \qquad (11) $$

$$ \omega_0^{'} = \alpha^{-0.5} \omega_u; \omega_0 = \alpha^{0.5} \omega_u; \qquad (12) $$

$$ \frac{\omega_{k+1}^{'}}{\omega_k^{'}} = \frac{\omega_{k+1}}{\omega_k} = \alpha\eta > 1 \qquad (13) $$

$$ \frac{\omega_{k+1}^{'}}{\omega_k} = \eta > 0; \frac{\omega_k}{\omega_k^{'}} = \alpha > 0 \qquad (14) $$

$$ N = \frac{\log(\omega_N/\omega_0)}{\log(\alpha\eta)} \qquad (15) $$

where ω_h, ω_b are the approximation frequency bounds, v is the order of the fractional differentiator, N is the order of fractional differentiator approximation and K is a constant. In the present work, the fractional order controller was implemented using FOMCON toolbox for fractional calculus [45, 46]. In this toolbox, fractional calculus can be approximated using the refined Oustaloup approximation method by varying the order and frequency ranges. This toolbox can be used for time domain, frequency domain and many more analyses for fractional calculus. The fractional order parameters in the current study were tuned based on the approach given in related literature [39].

2.4 COPID controller

Just as the FOPID controller is an extension of the classical PID controller, similarly the COPID controller is an extension of the FOPID controller. In complex order PID controller, integration and

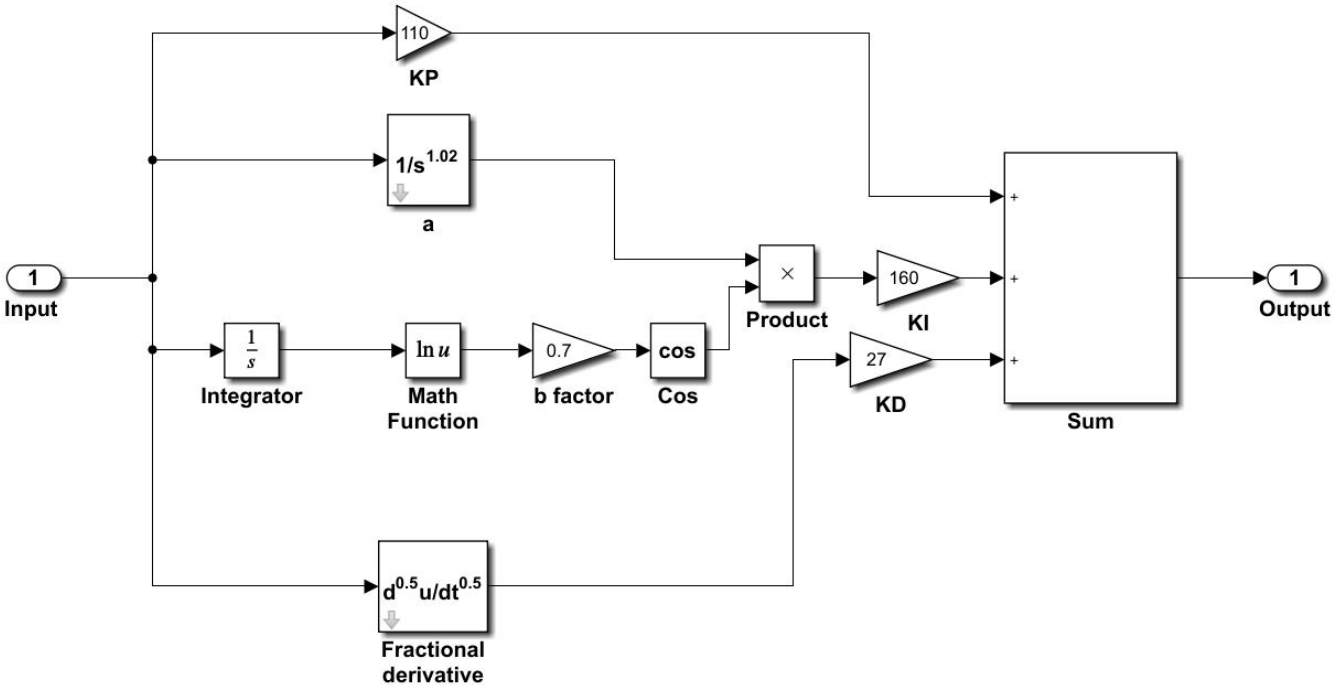


Figure 4: Complex $PI^{a+jb}D^{c+jd}$ controller structure

derivative orders are expressed in the form of a complex number (x+jy). The genesis of COPID controller is related to the CRONE controller (third generation) [47, 48]. The complex order PID controller structure for the DC motor fractional order model in the current study is given as follows [14, 20, 49].

$$C(s) = K_P + K_I\left(\frac{1}{s}\right)^{a+jb} + K_D\,(s)^{c+jd} \tag{16}$$

where, a, b, c, d are the complex order controller orders. The integral gain component of this equation is simplified as follows -

$$K_I\left(\frac{1}{s}\right)^{a+jb} = K_I\left(\frac{1}{s}\right)^{a}\left(\frac{1}{s}\right)^{jb} \tag{17}$$

$$K_I\left(\frac{1}{s}\right)^{a+jb} = K_I\left(\frac{1}{s}\right)^{a} * e^{\ln\left(\frac{1}{s}\right)^{jb}} \tag{18}$$

$$K_I\left(\frac{1}{s}\right)^{a+jb} = K_I\left(\frac{1}{s}\right)^{a} * e^{jb\ln\left(\frac{1}{s}\right)} \tag{19}$$

$$K_I\left(\frac{1}{s}\right)^{a+jb} = K_I\left(\frac{1}{s}\right)^{a} * \left[\cos\left(b\ln\left(\frac{1}{s}\right)\right) + j\sin\left(b\ln\left(\frac{1}{s}\right)\right)\right] \tag{20}$$

It is necessary to omit the imaginary part of the above equation because it cannot be synthesized for time domain implementation of the closed loop system with COPID controller in Simulink [14].

$$K_I\left(\frac{1}{s}\right)^{a+jb} = K_I\left(\frac{1}{s}\right)^{a} * \left[\cos\left(b\ln\left(\frac{1}{s}\right)\right)\right] \tag{21}$$

Similarly, the derivative part of Eq. (16) is simplified as follows -

$$K_D * s^{c+jd} = K_D * s^{c} * s^{jd} \tag{22}$$

$$K_D * s^{c+jd} = K_D * s^{c} * e^{\ln s^{jd}} \tag{23}$$

$$K_D * s^{c+jd} = K_D * s^{c} * e^{jd\ln s} \tag{24}$$

$$K_D * s^{c+jd} = K_D * s^{c} * [\cos(d\ln s) + j\sin(d\ln s)] \tag{25}$$

As followed in case of the integral gain component, it is necessary to omit the imaginary part of the derivative equation as well because it cannot be synthesized for time domain implementation of the closed loop system with COPID controller in Simulink [14].

$$K_D * s^{c+jd} = K_D * s^{c} * [\cos(d\ln(s))] \tag{26}$$

Thus, Eq.(16) may be rewritten for tuning as follows.

$$C(s) = K_P + K_I\left(\frac{1}{s}\right)^{a} * \left[\cos\left(b\ln\left(\frac{1}{s}\right)\right)\right] + K_D * s^{c} * [\cos(d\ln(s))] \tag{27}$$

The tuning of complex order controller is more challenging than FOPID and classical PID controller as there are seven parameters in the complex order controller. In this work, Eq. (27) was utilised for the implementation of complex order controller. This controller was tuned based on the tuning principles provided in literature [39, 50]. Figure 4 shows the structure of the COPID controller implemented in Simulink (Matlab). The initial value of the integer order integration was assumed to be a small number for this implementation. For all controllers, high control signal overshoots were limited by adding saturation block in the closed loop system models.

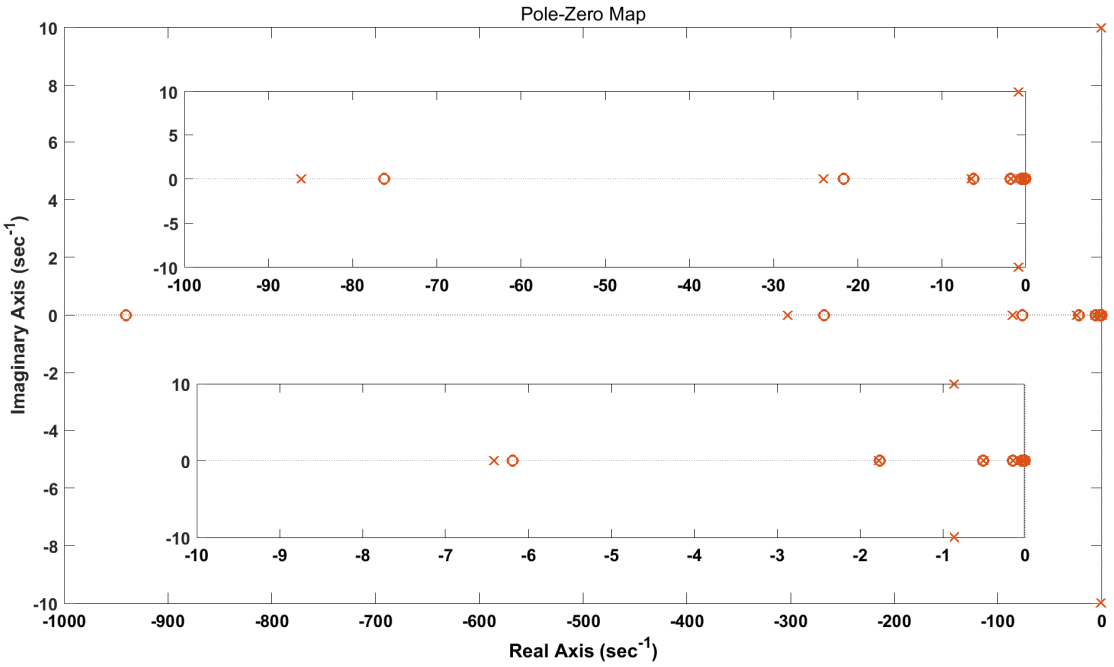


Figure 5: Pole zero map for fractional order DC motor plant model

3 Results and Discussions

The fractional order DC motor system model derived in the work reported in the 2019 IEEE International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE) [1] is given below.

$$G(s) = \frac{0.7992}{s^{1.9018} + 80.0440} \tag{28}$$

It may be noted that this model has a low DC gain. Low DC gains in second / higher integer order system models are expected to cause the system responses to become sluggish due to the presence of the non-dominant pole(s) in the pole zero map. Similar effect is observed in the pole zero map (figure 5) of the fractional order system model considered in the current study as well. However, the presence of a large number of dominant poles keep the system response fast; overcoming the adverse impact of low DC gain. This model consisting three parameters was employed for designing the PI, PR, PID, FOPID and COPID controllers in the current work. The PI controller design for the DC motor system was obtained as follows -

$$C(s) = 10 + 92.6736\frac{1}{s} \tag{29}$$

Figures 6 and 7 show the output and control signal plots for the PI controller respectively.

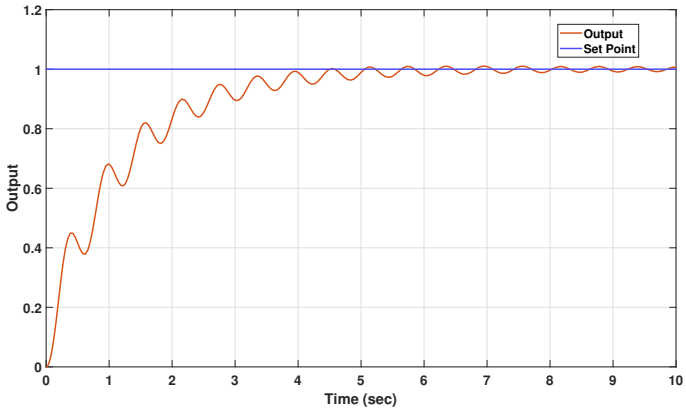


Figure 6: PI controller output for fractional order DC motor plant model

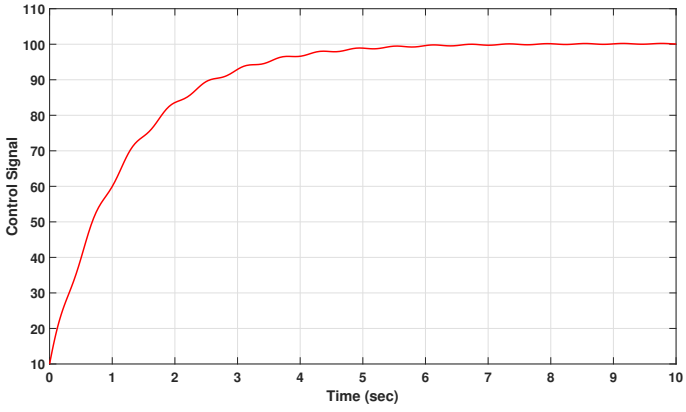


Figure 7: Control signal for PI controller output for fractional order DC motor plant model

Similarly, the following PR controller structure for the DC motor system was generated -

$$C(s) = 1.0469E05 + \frac{-8.2380E08 * s}{s^2 + 10000} \tag{30}$$

The output and control signal responses for the PR controller are depicted in figures 8 and 9 .

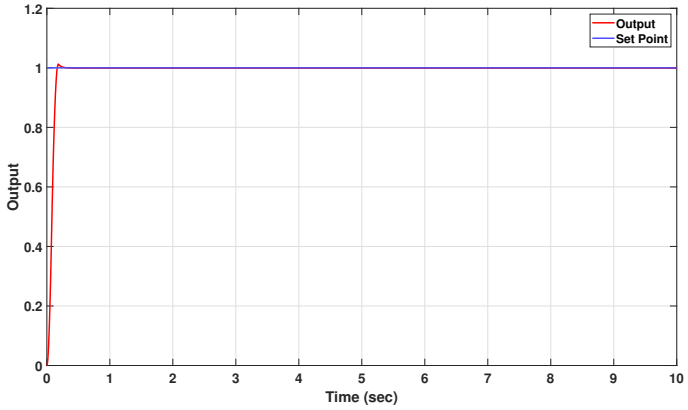


Figure 8: PR controller output for fractional order DC motor plant model

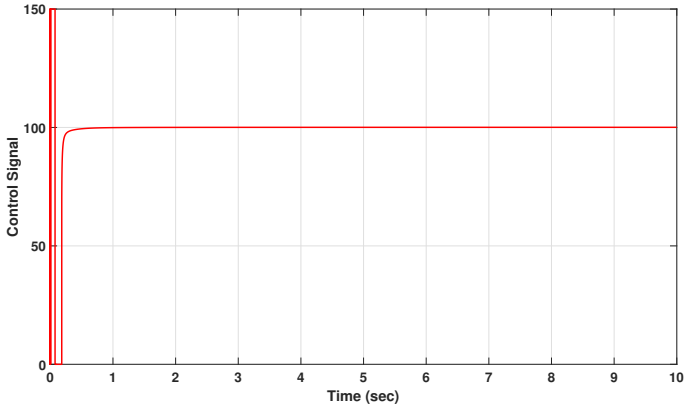


Figure 9: Control signal for PR controller output for fractional order DC motor plant model

The auto tuned PID controller architecture for the fractional order DC motor plant model considered in the current study was obtained as follows -

$$C(s) = 0.00001 + 92.5368\frac{1}{s} + 0.001\frac{100}{1 + 100\frac{1}{s}} \tag{31}$$

Figure 10 shows the output response of the PID controller output, whereas figure 11 shows the control signal plot for the same.

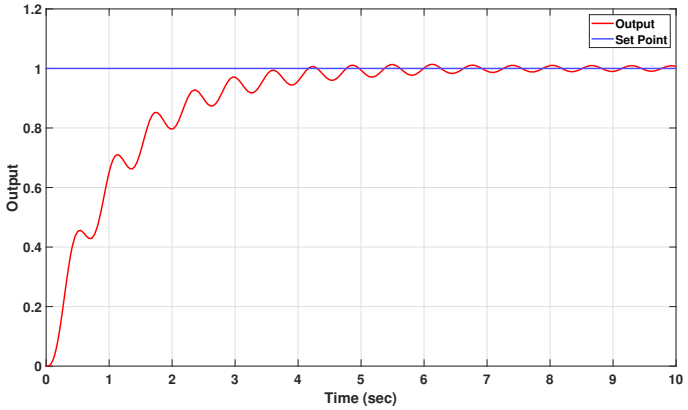


Figure 10: PID controller output for fractional order DC motor plant model

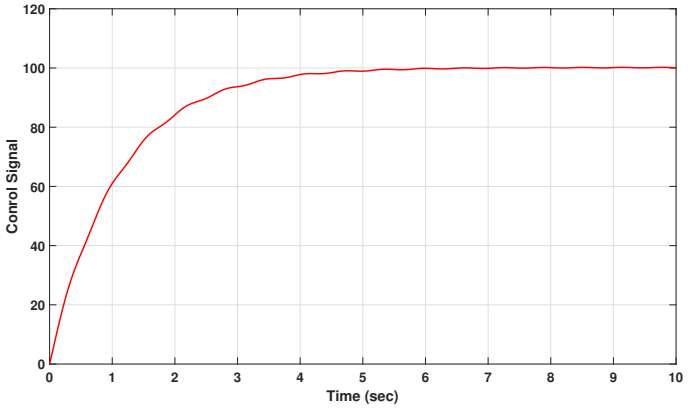


Figure 11: Control signal for PID controller output for fractional order DC motor plant model

The fractional order PID controller structure was obtained as follows -

$$C(s) = 125 + 1125\frac{1}{s^{0.95}} + 51s^{0.77} \tag{32}$$

Figures 12 and 13 depict the output and control signal responses for this FOPID controller, respectively.

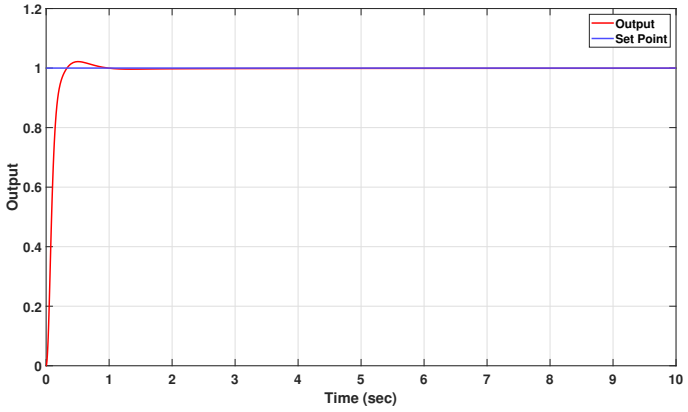


Figure 12: Fractional PID controller output for fractional order DC motor plant model

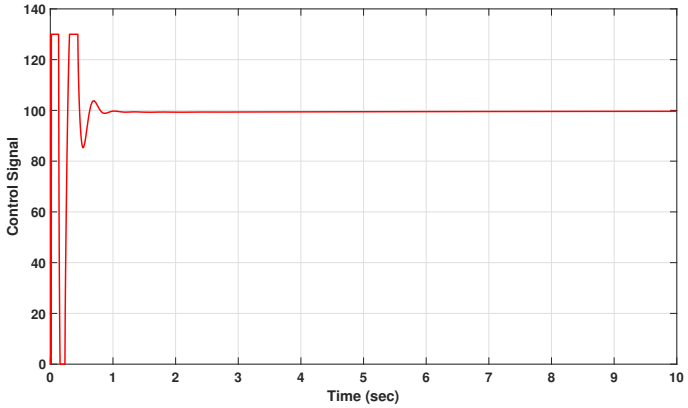


Figure 13: Control signal for Fractional PID controller output for fractional order DC motor plant model

The complex order PID controller structure was obtained as follows -

$$C(s) = 11 + 125\frac{1}{s^{0.99+j0.7}} + 27s^{0.5+j0} \quad (33)$$

Figures 14 and 15 show the output response and the control signal of the complex controller designed for the DC motor system.

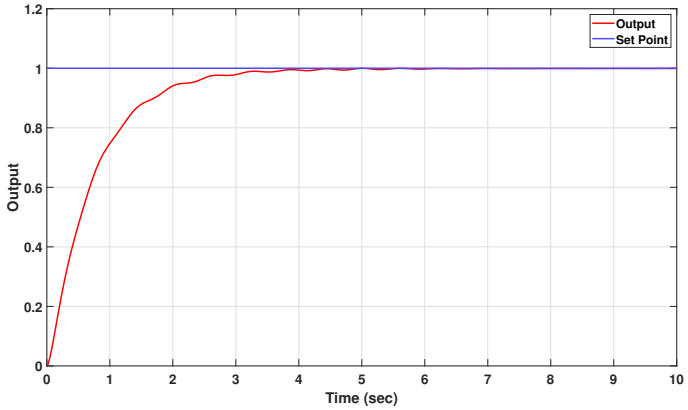


Figure 14: Complex $PI^{a+jb}D^{c+jd}$ controller output for fractional order DC motor plant model

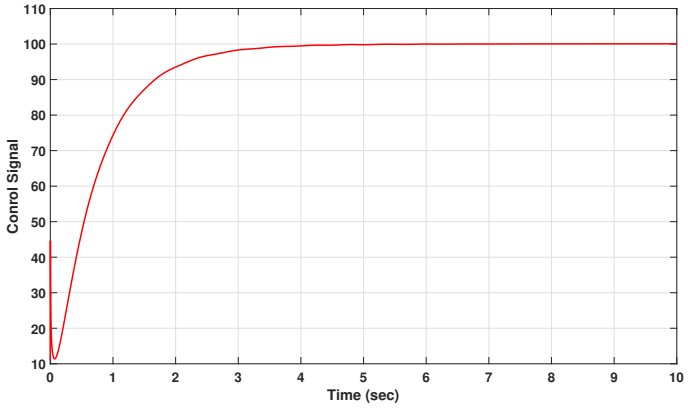


Figure 15: Control signal for complex $PI^{a+jb}D^{c+jd}$ controller output for fractional order DC motor plant model

Table 1 gives the output signal rise and peak time domain specifications of all controllers explored for the DC motor system considered in the current study. The PR controller is successful in attaining the least rise time and peak time among all controllers. The FOPID controller attains the second best peak and rise time characteristics. The COPID controller registers higher rise and peak times as compared to FOPID output signals. The PI and PID controller output signals consume relatively higher rise and peak times. Figures 16 and 17 showcase graphical presentations of output signal overshoots and settling times.

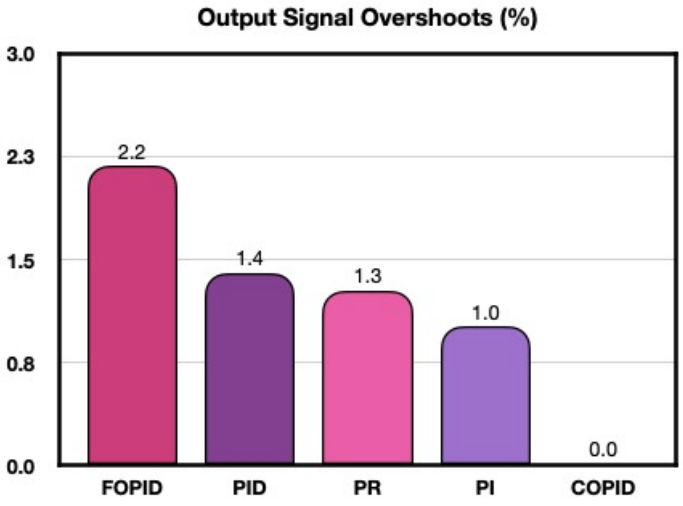


Figure 16: Output signal overshoots (%)

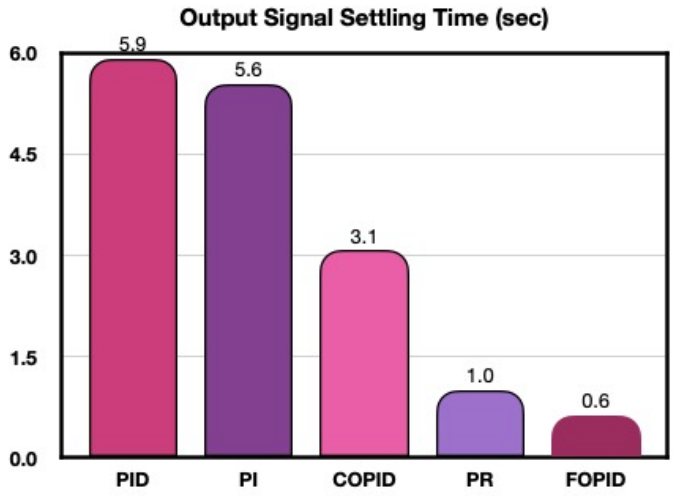


Figure 17: Output signal settling time (sec)

The FOPID controller achieves the least settling time, whereas the COPID controller minimizes peak overshoot percentage to zero. In contrast, the PR, PI, PID and FOPID output signals record overshoots of 1.3 %, 1.0 %, 1.4 % and 2.2 % respectively. This result proves the efficacy of the complex order controller in limiting the output signal overshoot for the DC motor system. The DC motor system output is the output shaft revolutions per minute. In indus-

Table 1: Time domain specifications of output signals

	PI	PR	PID	FOPID	COPID
Rise Time (sec)	3.9	0.0003	2.2473	0.1946	1.7366
Peak Time (sec)	5.7	0.0005	5.4711	0.5011	5.0000

Table 2: Time domain specifications of control signals

	PI	PR	PID	FOPID	COPID
Rise Time (sec)	2.47	-	2.5285	0.0001	1.7365
Peak Time (sec)	9.75	-	6.6835	0.0184	4.7587

trial systems, DC motor outputs actuate machinery for achieving the stipulated processing control. Eliminating peak overshoots in the DC motor output signals is important to avoid undesirable process variations which may be detrimental and even hazardous in some cases.

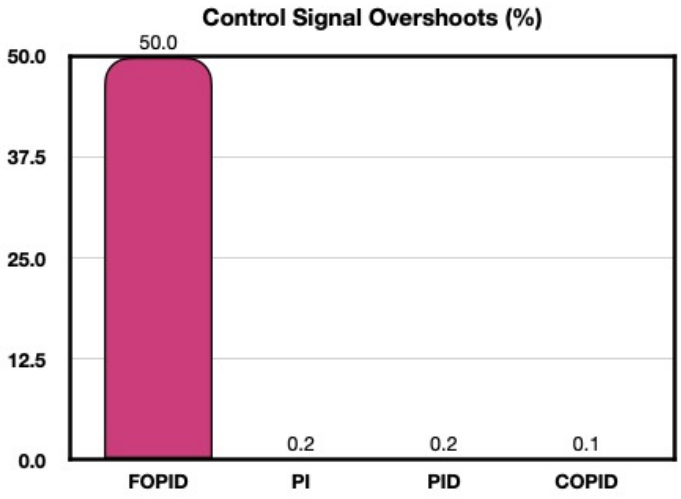


Figure 18: Control signal overshoots (%)

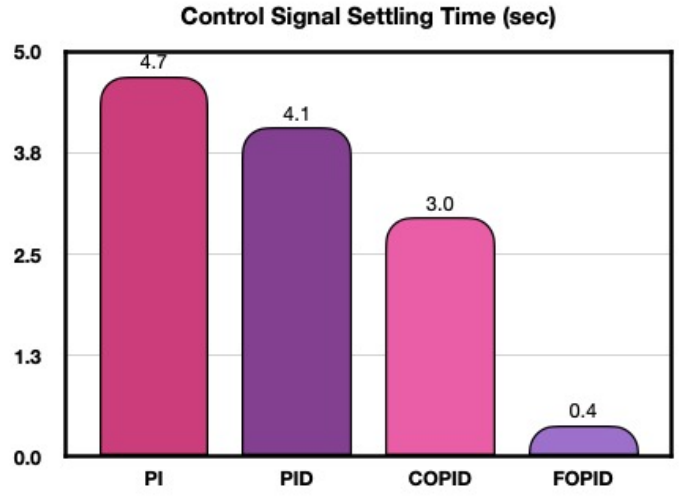


Figure 19: Control signal settling time (sec)

Table 2 shows the control signal time domain characteristics for the five controllers under the scope of the current study. In this case, the PR controller performed very poorly and registered extremely high overshoot (of the order of 10E6); rise, peak and settling times. The FOPID control signals attained minimum rise and peak times, followed by COPID and PID control signals. The PI control signals achieved much higher rise and peak time metrics. Figures 18 and 19 showcase graphical presentations of control signal overshoots and settling times. The PI controller also under performed with respect to the settling time metric. However, it performed much better in case of control signal peak overshoots. Similarly, the PID controller attained very low peak overshoot but suffered from higher settling time. The FOPID controller attained minimum settling time, but was unable to keep the control signal overshoot in check. Thus, it was only the COPID which consistently outperformed all other controllers with minimum peak overshoot and second best settling time characteristics. For DC motor systems, the control signal corresponds to the voltage input. Excessive voltage spikes (as evident in FOPID control signal response) may prove detrimental to the motor's active service life and might even result in premature damage. It is important for the controller design to ensure protection to the DC motor against high voltage shocks. Moreover, faster settling of the input voltage corresponding to the desired output set point is desirable for a highly responsive system. The COPID controller exhibited reliable performances in both aspects.

Figures 20 and 21 show the Bode plot and root locus diagram for the closed loop system with COPID controller. The Bode plot indicates that the gain and phase margin for the closed loop system with COPID controller is ∞. This implies that the system is robust against variations in the system parameters. Similarly, the location of closed loop poles to the left side of the s-plane in root locus diagram also proves that the DC motor system performance is stable and robust under complex order controller. Figure 22 depicts the robustness of the COPID controller against variations in the gain of the DC motor system. The system gain was varied with values of 0.5, 0.8, 2, 5, 10 and 20. The gain variations do not have any significant impact on the settling trends of the COPID controller responses, which is a positive indication.

4 Conclusions

The current work is an extension of the one reported in the 2019 IEEE International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE). Building upon the fractional order

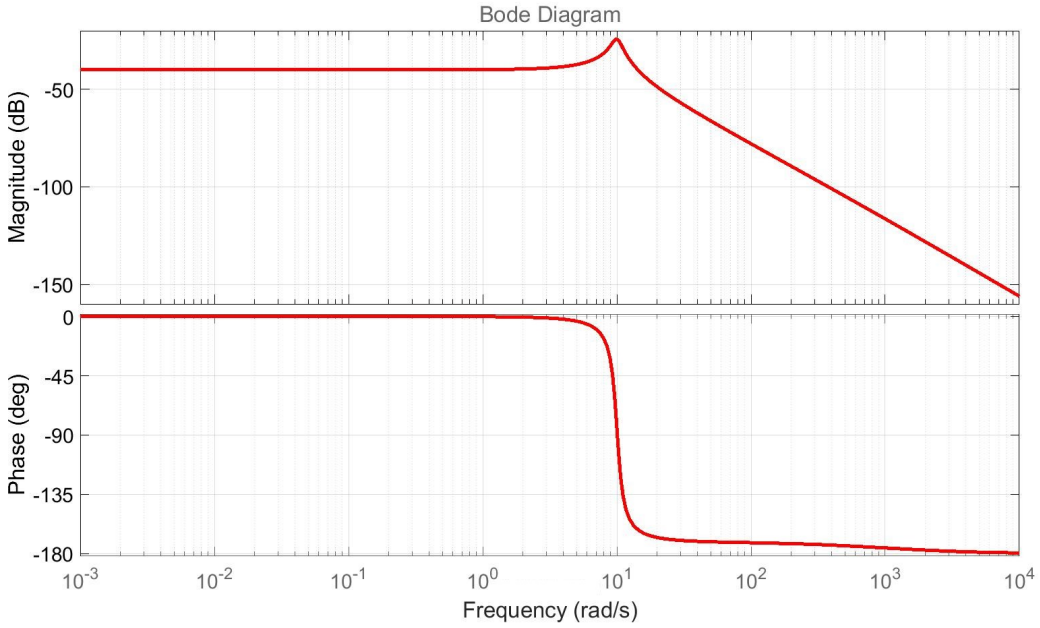


Figure 20: Bode plot for complex order controller for fractional order DC motor plant model

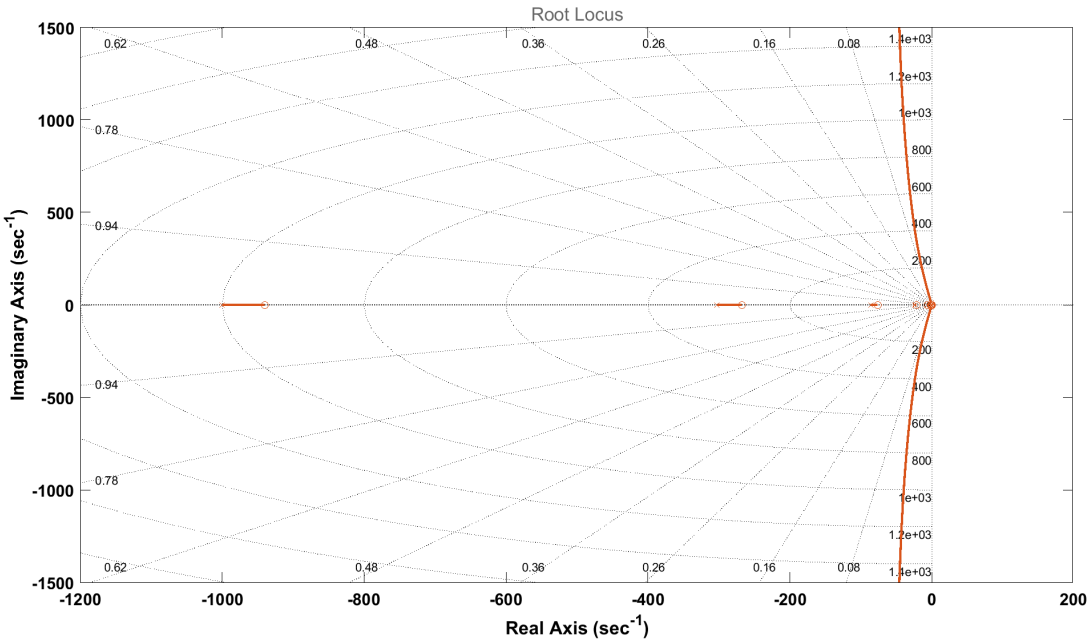


Figure 21: Root locus for complex order controller for fractional order DC motor plant model

model of the DC motor system identified in the previous work; the current work involved design of PI, PR, PID, FOPID and COPID controllers for the same. Primary results indicate that the PR controller performed very well in output signal responses; but fared very poorly in control signal time specifications. The PI controller also performed poorly in terms of output and control signal rise, peak and settling times. However, it performed comparatively better in terms of peak output and control signal overshoots. The PID controller's performance was mediocre in terms of control and output signal peak, rise and settling times. Its performance was also not very impressive with regards to the output signal peak overshoots, but it did much better in minimizing the control signal overshoot percentage. The FOPID controller attained very low rise, peak and settling times for its output signals. However, it suffered from the highest peak overshoots for both output and control signal responses. The COPID controller exhibited excellent performance with regards

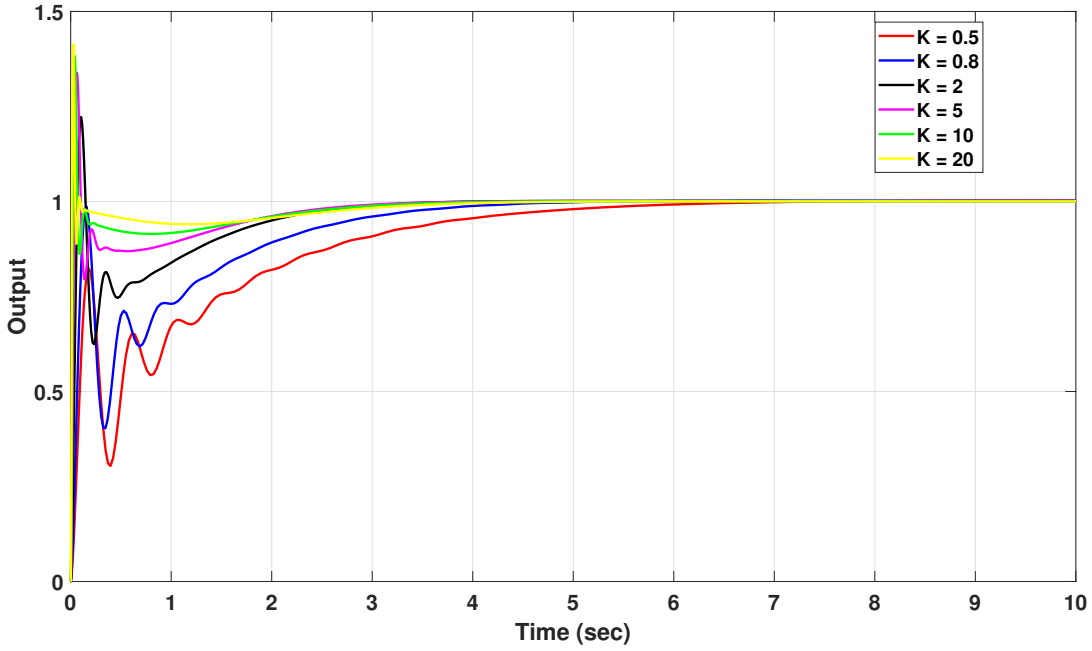


Figure 22: Responses of complex order controller for various values of K

to minimizing / eliminating the peak overshoots in the control (0.1 %) and output signal (0.0 %) responses. COPID controller also attained the second best low range rise, peak and settling time values. Minimisation and possible elimination of the peak overshoots in the control and output signals is vital for ensuring the safety of the controlled process machinery as well as for the long service life of the DC motor based industrial actuators. The COPID controller fulfills this requirement without sacrificing the settling time much. In addition, the Bode plot, root locus diagram and stable controller responses under varying system gain parameter confirm the stability and robustness of the designed complex order controller for the fractional order DC motor plant model considered in the current study. Thus, complex order controllers can be extensively and safely applied towards industry 4.0 oriented implementation of advanced industrial control systems such as smart actuators and collaborative robots.

References

[1] P. Shah, R. Sekhar, "Closed Loop System Identification of a DC Motor using Fractional Order Model," in 2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE), 69–74, IEEE, 2019, doi:https://doi.org/10.1109/morse48060.2019.8998744.

[2] K. J. Åström, P. Eykhoff, "System identification—a survey," Automatica, **7**(2), 123–162, 1971, doi:https://doi.org/10.1016/0005-1098(71)90059-8.

[3] R. Sekhar, T. Singh, P. Shah, "ARX/ARMAX Modeling and Fractional Order Control of Surface Roughness in Turning Nano-Composites," in 2019 International Conference on Mechatronics, Robotics and Systems Engineering (MoRSE), 97–102, IEEE, 2019, doi:https://doi.org/10.1109/MoRSE48060.2019.8998654.

[4] W. Zheng, Y. Luo, Y. Chen, Y. Pi, "Fractional-order modeling of permanent magnet synchronous motor speed servo system," Journal of Vibration and Control, 1–26, 2015, doi:https://doi.org/10.1177/1077546315586504.

[5] P. Shah, S. Agashe, A. P. Singh, "Fractional order modelling using state space theory," International Journal of Engineering and Technology, **5**(3), 2891–2894, 2013, doi:http://www.enggjournals.com/ijet/docs/IJET13-05-03-343.pdf.

[6] R. Bhimte, K. Bhole, P. Shah, "Fractional Order Fuzzy PID Controller for a Rotary Servo System," in 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), 538–542, IEEE, 2018, doi:https://doi.org/10.1109/ICOEI.2018.8553867.

[7] R. Bhimte, K. Bhole-Ingale, P. Shah, R. Sekhar, "Precise Position Control of Quanser Servomotor using Fractional Order Fuzzy PID Controller," in 2020 IEEE Bombay Section Signature Conference (IBSSC), 58–63, IEEE, 2020, doi:https://doi.org/10.1109/IBSSC51096.2020.9332216.

[8] I. Podlubny, "Fractional-order systems and fractional-order controllers," Institute of Experimental Physics, Slovak Academy of Sciences, Kosice, **12**(3), 1–18, 1994, doi:http://people.tuke.sk/igor.podlubny/pspdf/uef0394.pdf.

[9] P. Shah, S. Agashe, "Review of fractional PID controller," Mechatronics, **38**, 29–41, 2016, doi:https://doi.org/10.1016/j.mechatronics.2016.06.005.

[10] M. Zamani, M. Karimi-Ghartemani, N. Sadati, M. Parniani, "Design of a fractional order PID controller for an AVR using particle swarm optimization," Control Engineering Practice, **17**(12), 1380–1387, 2009, doi:https://doi.org/10.1016/j.conengprac.2009.07.005.

[11] P. Shah, R. Sekhar, S. Agashe, "Application of Fractional PID Controller to Single and Multi-Variable Non-Minimum Phase Systems," International Journal of Recent Technology and Engineering, **8**(2), 2801–2811, 2019, doi:https://doi.org/10.35940/ijrte.b2805.078219.

[12] C. A. Monje, Y. Chen, B. M. Vinagre, D. Xue, V. Feliu-Batlle, Fractional-order systems and controls: fundamentals and applications, Springer Science & Business Media, 2010, doi:https://doi.org/10.1007/978-1-84996-335-0.

[13] A. V. Tare, J. A. Jacob, V. A. Vyawahare, V. N. Pande, "Design of novel optimal complex-order controllers for systems with fractional-order dynamics," International Journal of Dynamics and Control, **7**(1), 355–367, 2019, doi:https://doi.org/10.1007/s40435-018-0448-5.

[14] M. Shahiri, A. Ranjbar, M. R. Karami, R. Ghaderi, "New tuning design schemes of fractional complex-order PI controller," Nonlinear Dynamics, **84**(3), 1813–1835, 2016, doi:https://doi.org/10.1007/s11071-016-2608-5.

[15] B. Ross, W. Haven, F. H. Northover, "A use for a derivative of complex order in the fractional calculus," 1978, doi:https://insa.nic.in/writereaddata/UpLoadedFiles/IJPAM/20005a85_400.pdf.

[16] J. A. Jacob, A. V. Tare, V. A. Vyawahare, V. N. Pande, "A review of time domain, frequency domain and stability analysis of linear complex-order systems," in 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), 164–169, IEEE, 2016, doi:https://doi.org/10.1109/WIECON-ECE.2016.8009110.

[17] M. F. Silva, J. T. Machado, R. S. Barbosa, "Complex-order dynamics in hexapod locomotion," Signal processing, **86**(10), 2785–2793, 2006, doi: https://doi.org/10.1016/j.sigpro.2006.02.024.

[18] J. L. Adams, T. T. Hartley, L. I. Adams, "A solution to the fundamental linear complex-order differential equation," Advances in Engineering Software, **41**(1), 70–74, 2010, doi:https://doi.org/10.1016/j.advengsoft.2008.12.014.

[19] J. T. Machado, "Optimal controllers with complex order derivatives," Journal of Optimization Theory and Applications, **156**(1), 2–12, 2013, doi: https://doi.org/10.1007/s10957-012-0169-4.

[20] M. Shahiri, A. Ranjbar, M. R. Karami, R. Ghaderi, "Robust control of nonlinear PEMFC against uncertainty using fractional complex order control," Nonlinear Dynamics, **80**(4), 1785–1800, 2015, doi:https://doi.org/10.1007/s11071-014-1718-1.

[21] C. M. Pinto, A. R. Carvalho, "Effect of drug-resistance in a fractional complex-order model for HIV infection," IFAC-PapersOnLine, **48**(1), 188–189, 2015, doi:https://doi.org/10.1016/j.ifacol.2015.05.162.

[22] A. Guefrachi, S. Najar, M. Amairi, M. Aoun, "Tuning of fractional complex order PID controller," IFAC-PapersOnLine, **50**(1), 14563–14568, 2017, doi: https://doi.org/10.1016/j.ifacol.2017.08.2093.

[23] O. Hanif, G. B. Babu, S. Sharma, "Performance Improvement of PI^{x+iy} D Fractional Complex Order Controller using Genetic Algorithm," in 2018 Fourth International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 1–5, IEEE, 2018, doi: https://doi.org/10.1109/AEEICB.2018.8480981.

[24] K. Khandani, A. A. Jalali, M. R. R. Mehdiabadi, "Robust complex order controller design for DC motors," in 20th Iranian Conference on Electrical Engineering (ICEE2012), 900–903, IEEE, 2012, doi:https://doi.org/10.1109/IranianCEE.2012.6292481.

[25] A. Ma'arif, N. R. Setiawan, "Control of DC Motor Using Integral State Feedback and Comparison with PID: Simulation and Arduino Implementation," Journal of Robotics and Control (JRC), **2**(5), 456–461, 2021, doi: https://doi.org/10.18196/jrc.25122.

[26] M. Khalifa, A. H. Amhedb, M. Al Sharqawi, "Position Control of Real Time DC Motor Using LabVIEW," Journal of Robotics and Control (JRC), **2**(5), 342–348, 2021, doi:https://doi.org/10.18196/jrc.25104.

[27] P.-Y. Lin, Y.-S. Lai, "Novel voltage trajectory control for field-weakening operation of induction motor drives," IEEE Transactions on Industry Applications, **47**(1), 122–127, 2010, doi:https://doi.org/10.1109/TIA.2010.2091092.

[28] D. Somwanshi, M. Bundele, G. Kumar, G. Parashar, "Comparison of fuzzy-PID and PID controller for speed control of DC motor using LabVIEW," Procedia Computer Science, **152**, 252–260, 2019, doi:https://doi.org/10.1016/j.procs.2019.05.019.

[29] C. Sánchez-López, V. Carbajal-Gómez, M. Carrasco-Aguilar, F. Morales-López, "PID controller design based on memductor," AEU-International Journal of Electronics and Communications, **101**, 9–14, 2019, doi:https://doi.org/10.1016/j.aeue.2019.01.019.

[30] N. Zhang, H. Tang, C. Yao, "A systematic method for designing a PR controller and active damping of the LCL filter for single-phase grid-connected PV inverters," Energies, **7**(6), 3934–3954, 2014, doi:https://doi.org/10.3390/en7063934.

[31] J. Zhang, L. Li, D. G. Dorrell, Y. Guo, "Modified PI controller with improved steady-state performance and comparison with PR controller on direct matrix converters," Chinese Journal of Electrical Engineering, **5**(1), 53–66, 2019, doi:https://doi.org/10.23919/CJEE.2019.000006.

[32] J. Gnanavadivel, R. Thangasankaran, N. S. Kumar, K. Krishnaveni, M. Schlenk, "Performance Analysis of PI Controller and PR Controller Based Three-Phase AC-DC Boost Converter with Space Vector PWM," International Journal of Pure and Applied Mathematics, **118**(24), 1–16, 2018, doi:https://acadpubl.eu/hub/2018-118-24/3/501.pdf.

[33] E.-H. Dulf, "Simplified fractional order controller design algorithm," Mathematics, **7**(12), 1166, 2019, doi:https://doi.org/10.3390/math7121166.

[34] K. Oldham, J. Spanier, The fractional calculus : theory and applications of differentiation and integration to arbitrary order, 1^{st} Edition, Elsevier, 1974, doi:https://cutt.ly/RznbhbX.

[35] I. Petras, "Tuning and implementation methods for fractional-order controllers," Fractional Calculus and Applied Analysis, **15**(2), 282–303, 2012, doi:https://doi.org/10.2478/s13540-012-0021-4.

[36] A. Loverro, "Fractional calculus: history, definitions and applications for the engineer," Rapport technique, Univeristy of Notre Dame: Department of Aerospace and Mechanical Engineering, 2004, doi:https://www.academia.edu/download/36579354/FracCalc.pdf.

[37] D. Cafagna, "Fractional calculus: A mathematical tool from the past for present engineers," IEEE Industrial Electronics Magazine, **2**(1), 35–40, 2007, doi:https://doi.org/10.1109/MIE.2007.901479.

[38] R. Sekhar, T. Singh, P. Shah, "Micro and Nano Particle Composite Machining: Fractional Order Control of Surface Roughness," in Third International Conference on Powder, Granule and Bulk Solids: Innovations and Applications, 35–42, 2020, doi:https://cutt.ly/LzvarrI.

[39] P. Shah, S. Agashe, "Experimental Analysis of Fractional PID Controller Parameters on Time Domain Specifications," Progress in Fractional Differentiation and Applications, **3**, 141–154, 2017, doi:http://dx.doi.org/10.18576/pfda/030205.

[40] P. Shah, S. Agashe, A. J. Kulkarni, "Design of a fractional $PI^{\lambda}D^{\mu}$ controller using the cohort intelligence method," Frontiers of Information Technology & Electronic Engineering, **19**(3), 437–445, 2018, doi:https://doi.org/10.1631/FITEE.1601495.

[41] S. Nangrani, S. Bhat, "Fractional order controller for controlling power system dynamic behavior," Asian Journal of Control, **20**(1), 403–414, 2018, doi:https://doi.org/10.1002/asjc.1557.

[42] S. E. Hamamci, "An algorithm for stabilization of fractional-order time delay systems using fractional-order PID controllers," Automatic Control, IEEE Transactions on, **52**(10), 1964–1969, 2007, doi:https://doi.org/10.1109/TAC.2007.906243.

[43] R. Gupta, S. Gairola, "FOPID controller optimization employing PSO and TRSBF function," in 2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS), 1–6, IEEE, 2015, doi: https://doi.org/10.1109/RAECS.2015.7453275.

[44] P. Shah, S. Agashe, V. Vyawahare, "System identification with fractional-order models: A comparative study with different model structures," Prog. Fract. Diff. Appl., **4**(4), 533–552, 2019, doi:https://doi.org/10.18576/pfda/040407.

[45] A. Tepljakov, FOMCON: Fractional-Order Modeling and Control Toolbox, 107–129, Springer International Publishing, Cham, 2017, doi:10.1007/978-3-319-52950-9_6.

[46] A. Tepljakov, E. Petlenkov, J. Belikov, I. Petráš, "FOMCON toolbox for modeling, design and implementation of fractional-order control systems," Applications in Control; Petráš, I., Ed.; De Gruyter: Berlin, Germany, 211–236, 2019, doi:https://doi.org/10.1515/9783110571745-010.

[47] P. Lanusse, A. Oustaloup, B. Mathieu, "Third generation CRONE control," in Proceedings of IEEE Systems Man and Cybernetics Conference-SMC, volume 2, 149–155, IEEE, 1993, doi:https://doi.org/10.1109/ICSMC.1993.384864.

[48] S. Das, Functional fractional calculus, Springer Science & Business Media, 2011, doi:https://doi.org/10.1007/978-3-642-20545-3.

[49] M. Zheng, G. Zhang, T. Huang, "Tuning of fractional complex-order direct current motor controller using frequency domain analysis," Mathematical Methods in the Applied Sciences, 2020, doi:https://doi.org/10.1002/mma.6653.

[50] K. H. Ang, G. Chong, Y. Li, "PID control system analysis, design, and technology," IEEE transactions on control systems technology, **13**(4), 559–576, 2005, doi:https://doi.org/10.1109/TCST.2005.847331.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 580-585 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Homology Modeling of CYP6Z3 Protein of Anopheles Mosquito

Marion Olubunmi Adebiyi*, Oludayo Olufolorunsho Olugbara

ICT and Society Research Group & Luban Workshop, Durban University of Technology, Durban, 4000, South Africa

ARTICLE INFO

Article history:
Received: 21 August, 2020
Accepted: 05 December, 2020
Online: 20 March, 2021

Keywords:
Anopheles Gambiae
Cytochrome P450 Family
Pyrethroid Resistance
Model Structure

ABSTRACT

The Anopheles gambiae's CYP6Z3 protein belongs to the Cytochrome P450 family and functions in oxidation-reduction processes, many studies including our previous work on elucidating insecticide resistance genes of the Anopheles also implicated her in pyrethroid insecticide resistance. Model prediction, functional analysis, and enrichment of the target gene with triplex binding sites may become a useful diagnostic biomarker for the disease subtype, but wrong classification of the model by various existing alignment algorithms is a daunting issue that complicates and misleads in decision making during pathway and functional analysis. The aim of this study is to predict five in-silico model of CYP6Z3 Anopheles protein by homology modeling, evaluate and classify them to elucidate the performance of the sequence alignment algorithm deployed, then characterize the top model that is correctly classified. Template selection from three alignment algorithms with sequence of the target-protein, (Anopheles-CYP6Z3) obtained from UNIPROT served as input, Clustal omega and Clustalw2 algorithms was used to generate alignment files for homologous template search to the target-protein. Best template was sought, and the 3D model built in an-automated-mode. PROCHECK was used to evaluate the best-of-the-five-obtained models. Estimating the quality of all models, the prime model emerged from ClustalW2 alignment algorithm, but was wrongly classified as a homo-tetramer-state. These provided a misleading-information which was revealed during model evaluation and interpretation, that resulted to an inappropriate pathway and functional-analysis, false positive model was then isolated, and the current best model emerged from clustalo alignment algorithm having 87.7% amino residues in the most favorable regions, 0.7% in the disallowed regions at monomer oligo state. Functional analysis of the best Anopheles CYP6Z3 secondary structure showed characteristics that explain the different degrees of genetic regulation translating to resistance mechanism in the malaria vector.

1. Introduction

The *Anopheles gambiae* CYP6Z3 protein is a member of Cytochrome P450 protein family functioning in oxidation-reduction processes. It has the VECTORBASE Annotation ID AGAP008217, primary (citable) accession number Q86LT6 and entry name Q86LT6_ANOGA. It has a sequence length of 492AA and is located on Chromosome 3R: 6,971,669-6,973,290. It has a mass of 56,490(Da) [1], [2]. It has been implicated in pyrethroid resistance [3].

The CYP6Z3 protein is expressed during the mosquito's larval stages [4], this protein belongs to the Cytochrome P450 family and functions in oxidation-reduction processes. The P450 is a large family that plays critical role in xenobiotics detoxification or activation. Instances are insecticide detoxification in the West African *Anopheles gambiae* [5].

Over 30 species of *Anopheles* transmit malaria (http://www.cdc.gov/malaria/about/biology/mosquitoes/), hence identification of resistance mechanisms in other species is a focus of many researches. *Anopheles funestus* is the next as a major malaria vector in Sub-Saharan Africa. When quantitative trait loci (QTL) are considered, several genes are strongly associated with pyrethroid resistance in *A. funestus*, this include our target protein CYP6Z3 [6]. In a recent study, genes transcripts from the four known detoxification genes family, (cytochrome P450s, glutathione transferases, carboxylesterases and UDP glucornyltransferases were reported to be generally enriched in the midgut and malpighian tubules of *A. gambiae*. Recently,

*Corresponding Author: Marion Olubunmi Adebiyi, Email: mariona@dut.ac.za

Cytochrome P450 proteins were reported to have developed different levels of resistance to multiple insecticides [7]. Specifically, the CYP6Z3 family was found to be highly enriched in the malpighian tubules consistently with its role in detoxification [8].

The malpighian tubules therefore display roles similar to liver of vertebrates, kidney and their immune system. Despite the wide distribution of detoxifying enzymes in insects, baseline mechanisms protecting vectors against insecticides is resident mainly in the excretory canal (system) of the insect's, which corresponds to less than 0.1% of its total mass [9].

1.1. Theoretical framework

Homology modelling, also called comparative modelling is the central concept of relative and trans-mutative biology. It is the existence of morphological structures and features in various species or organisms linked by sharing common ancestor. Homology was invented in the nineteenth century and found to be phylogenetic in nature and well rooted in comparative practice even before the invention of evolution theory [10]. It is believed that various developmental mechanisms are responsible for formation of homologous structure, however, varying specie may have anatomical structures with same or different characters, featuring similar shape, internal structure, and function, but may only be closely related species by taxonomic link in different mammals. The idea of homology originated with the recognition that the same structures exist in less closely related species (mammals and birds, or even mammals and fish) and that the sameness of morphological (body) units is independent of their function and form [10], [11]. Figure 1 depicts the flow diagram of the conceptual framework of this research.

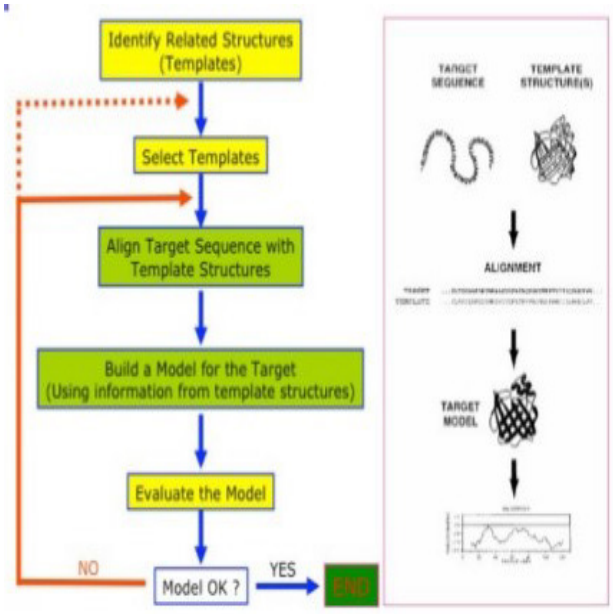


Figure 1: Conceptual framework of homology modelling [12]

The theoretical model is a molecular model premised on the notion that if two proteins possess a high sequence similarity, they will most likely have highly similar three-dimensional structures. To select the best template, a search for homologous Protein that has experimental structure by empirical methods by X-ray crystallography, nuclear magnetic resonance (NMR), or cryo-electron microscopy, existing on Protein Data Bank (PDB) [13].

2. Experimental

The method deployed was adopted from [14]. The FASTA format sequence of the target protein CYP6Z3 of the Anopheles was retrieved from UNIPROT database, (UniProt entry Q86LT6).

2.1. Template Selection

BLAST Program [15], blastp variant was used to the search for homologous sequences within Protein Data Bank (PDB) database [16]. The BLAST search returned many sequences that were homologous with the target sequence but only one was selected as the template.

2.2. Selection Criteria

Sequence identity of 30% or more was considered. Generally, the homologous sequences returned had low sequence identity between 19%-33%. E-value less than 0.001 and query cover above 50% were also considered. Potential templates were filtered based on these criteria and the atomic resolution of their experimentally derived 3D structures as viewed from PDB. Structures that failed to meet one or more of these criteria were excluded and the protein with accession 1TQN_A was selected as the template, with query cover 96%, E-value 2e-68 and sequence identity 31%. It has a medium atomic resolution of 2.05Å and was determined by X-ray crystallography. The 1TQN_A protein is a crystal structure of the human microsomal P450 3a4. It has a sequence length of 486AA. It is also a Cytochrome P450 enzyme functioning in oxidation-reduction processes. The sequence for the template was downloaded directly from the BLAST webpage in the fasta format.

2.3. Sequence Alignment

In [17], [18], the authors were used to perform sequence alignments for the target sequence and template sequence using the default settings. The resulting .clustalw2 and .clustalo files were downloaded to serve as an input files for the next step.

2.4. Building the Model

Swiss Model Server in [19], was deployed for building the 3D structure of the target protein.

3. Results

This study modelled five (3D) secondary structure of the *Anopheles* protein, CYP6Z3 by homology modeling, evaluated, classified, and compared their classified features to gauge how the sequence alignment algorithms rates the features of each of the model. The purpose of comparison was to reveal the best of the five models and elucidate possible errors in the features of their secondary structure, the evaluation result detected a wrong classification error in the supposed best structure, while other features are correctly classified. If the structure with error is the only one modelled and validated, performing further analysis such as functional characterization and pathway analysis based on the result built with error may result in misleading result hypothesis and fatal conclusion.

Clustal omega and Clustalw2 algorithms was used to generate alignment files for homologous template search to the

target-protein. The final correctly classified result of the best CYP6Z3 *Anopheles gambiae* secondary structure was functionally analysed, the result revealed the different degrees of genetic regulation translating to resistance mechanism deployed to pyrethroid class of insecticide by the malaria vector, *Anopheles*.

3.1. Using .clustlaw Alignment File

The "Alignment Mode" was used where the target-template alignment from Clustalw2 was submitted to the server. The server generated a model built with template 1tqn.2.A (Table 1).

Table 1: Features of model from template 1TQN.2.A with Clustal W2 alignment file

Template	Seq ID	Oligo-state	Method	Seq similarity	Coverage
1tqn.2.A	30.30 %	Homo-tetramer	X-ray, 2.05Å	0.36	0.96

3.2. Using .clustalo Alignment File

First, the "Alignment Mode" was used where the target-template alignment from Clustal Omega was submitted to the server (TABLE 2). Next, the "Automated Mode" was used where another entry was submitted into the server with only the target protein as input data, allowing Swiss Model Server to search through databases for its template(s) of choice (TABLE 3).

3.3. Alignment mode

Table 2: Features of model from template 1TQN.1.A with Clustalo alignment file

Template	Seq ID	Oligo-state	Method	Seq similarity	Coverage
1tqn.1.A	30.85%	Monomer	X-ray, 2.05Å	0.36	0.96

3.4. Automated mode

The server found 50 templates and built 3 models

Table 3: Features of the three models from automated mode

Built	Template	Seq ID	Oligo-state	Method	Seq similarity	Coverage
Model 1	1tqn.1.A	30.85 %	Monomer	X-ray, 2.05Å	0.36	0.96
Model 2	4k9w.4.A	30.19 %	Monomer	X-ray, 2.4Å	0.36	0.95
Model 3	3ua1.1.A	30.19 %	Monomer	X-ray, 2.15Å	0.36	0.95

From the models generated, it was observed that using Clustalw2 for sequence alignment, the model was predicted to be a homo-tetramer while for Clustal Omega, all models were predicted to be monomers. Also, both alignment modes (i.e from .clustalw and .clustalo files) generated structures for 1TQN_A proteins (1tqn.2.A and 1tqn.1.A respectively). It was hypothesized that these two proteins were more closely related based on their coverage of 0.96 although they differed in sequence identity compared with a similarly named model (1tqn.1.A) built via automated mode which had a coverage of 0.95.

3.5. Model evaluation

PROCHECK [20], [21] generated the Ramachandran plot, this plot was used to validate the built models (Figures 2, 3, 4, 5, 6).

Table 4: Ramachandran plot interpretation for the modelled structures (result validity table)

Source	Alignment mode	PDB SUM	Most favoured regions	Disallowed regions
Cluster W2 Alignment	1tqn.2.A	d839 (Figure 2)	87.9%	1.2%
Cluster O Alignment	1tqn.1.A	d860 (Figure 3)	87.7%	0.7%
Automated mode	1tqn.1.A	d861 (Figure 4)	80.9%	1.4%
Automated mode	4k9w.4.A	d862 (Figure 5)	83.2%	1.7%

4. Discussion

From the above tables, as hypothesized, the structures derived from Clustalw2 and Clustal Omega sequence alignments were similar, having the highest percentage of residues in the most favored regions than every other model generated via the automated mode (TABLE 4). All the structures built represented the same protein, but clustalw2 'misinterpreted' it to be a homo-tetramer protein. Clustalw2 model would have been the best with 87.9% amino acid residue highly deposited in the most acceptable region and 1.2% deposit in the disallowed region, but since it could not classify the protein appropriately into a monomer oligo state, where the protein family belongs, the result was disregarded and could not be considered for further functional analysis and quality scrutiny.

This shows that the quality of the choice of template sequence used as input data into any modelling server, Swiss model in this case, is very important. For the automated mode, the 3ua1.1.A model was the best with the highest percentage of residues in the most favored/ acceptable regions. Comparing this with those obtained via alignment mode, the model 1tqn.1.A generated from .clustalo (Figure 8) alignment file with Ramachandran plot (Figure 3) respectively was the overall best structure having 87.7% amino residues in the most favorable regions and 0.7% residue deposits in the disallowed regions. All model proteins derived from both alignment and automated modes are in Figures (7, 8, 9, 10, 11).

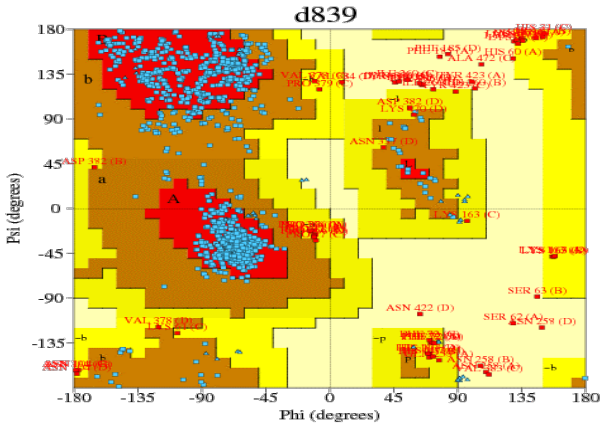


Figure 2: Ramachandran plot of Figure 7

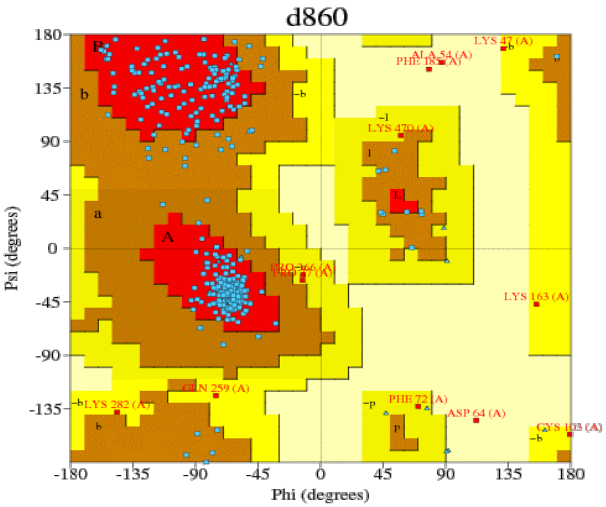


Figure 3: Ramachandran plot of Figure 8

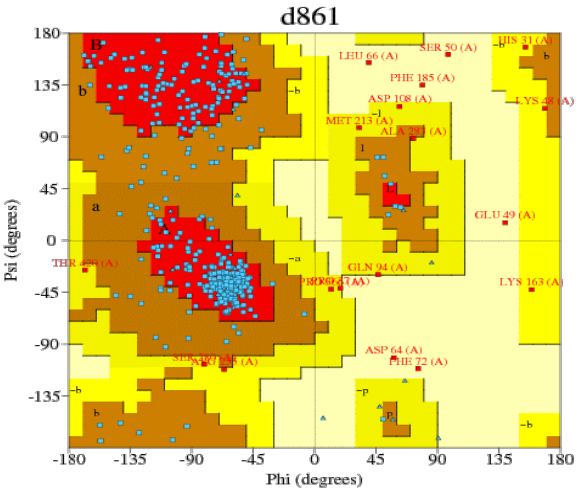


Figure 4: Ramachandran plot of Figure 9

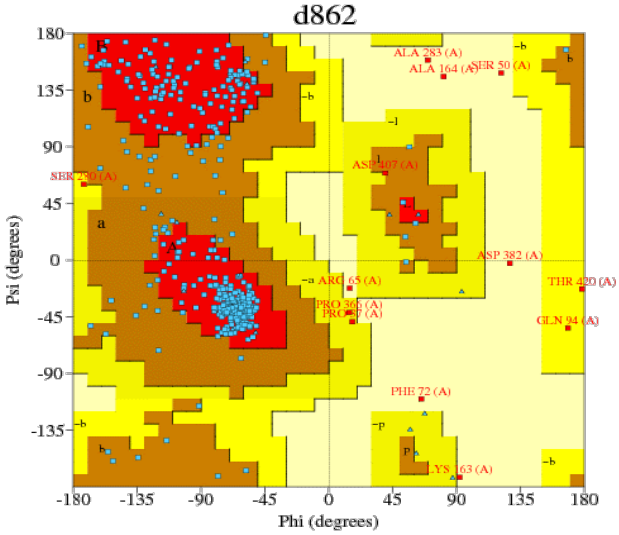


Figure 5: Ramachandran plot of Figure 10

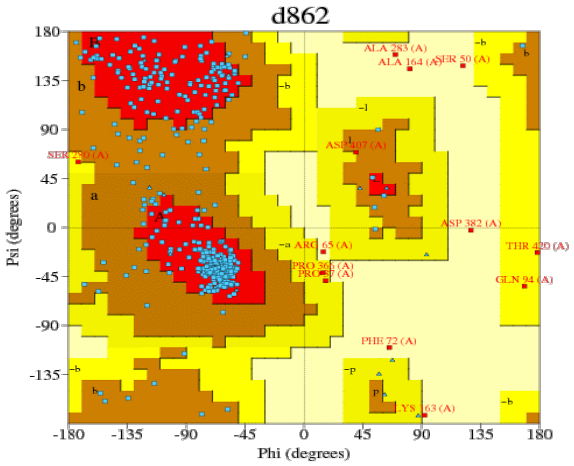


Figure 6: Ramachandran plot of Figure 11

4.1. Alignment mode

In [22], the author used to summarize the statistical details of the prime model (1tqn.1.A from alignment mode), while SaliLab Model Evaluation Server (ModEval) was also used to estimate the quality of the best model. The model from ClustalW2 alignment algorithm came up as the best model, but the alignment algorithm wrongly classified the model to be at homo-tetramer-state. These provided a misleading information which sufficed during result interpretation and model evaluation, these resulted to inappropriate pathway and functional characterization, until model with false positive classification was isolated and the real true positive model emerged the best model and was characterized.



Figure 7: 1tqn.2. A



Figure 8: 1tqn.1. A (Best of all Models)

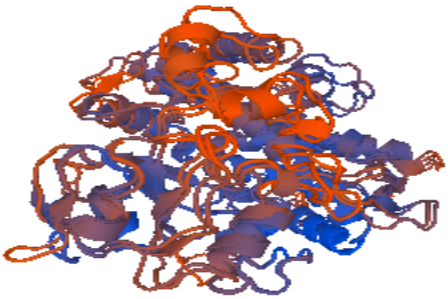


Figure 9: 1tqn.2. A

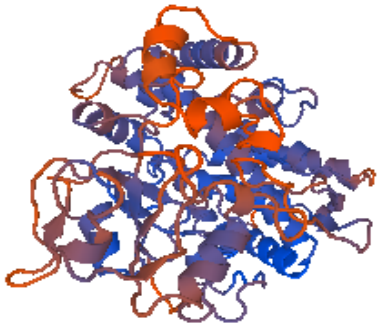


Figure 10: 4k9w.4. A

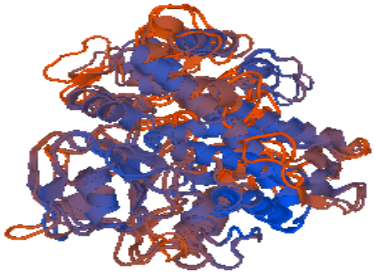


Figure 11: 3ua1.1.A

5. Conclusion

In this work, the best model was obtained by comparing the modeled structures produced by clustalo and .clustalw alignment files, meaning that the 3D-models produced by the two alignment-modes were very-similar, having a high-amino-residue-percentage deposited most favored-regions on the Ramachandran-plot than the model by the SWISS-automated-mode. Meticulosity and carefulness in the selection of procedure for analysis and choice of criteria for template protein selection from the avalanche of selectable criteria is an explorable virtue that can be deployed to avoid performing further analysis on true negative errors and interpreting false positive results. Our result showed that the quality of structure predicted by in-silico modeling cannot be over emphasized. Results showed that the 3D-model generated from template-protein-(1tqn.2.A), by Clustalw2-alignment and 3D-model generated from template-protein (1tqn.1.A) by Clustal Omega alignment were similar, having a high amino residue percentage deposited at regions that are most favored on the Ramachandran plot than the rest of the models generated via the automated mode. MolProbity for statistical analytics-summary and SaliLab-Model-Evaluator (ModEval) estimated the quality of the prime/leading model and evaluated the model from ClustalW2 as the best model, but wrongly classified the protein into a homo-tetramer state. These provided a misleading information which suffice during results interpretation, hypothesis generation and may result to inappropriate pathway and functional analysis. This matter in decision making during interpretation of result. it goes a longer way to determine not just the quality of the outputted target or modelled structure, but also considered aptly in taking decision to elicit the quality of the predicted functions of the target proteins, even in taking hypothesis while interpreting the output model as result, researcher may be misguided if an appropriate criterion was not considered during template selection on BLAST search. The role of CYP6Z3 in physiological processes including hormone and pheromone metabolism as well as insecticide detoxification, especially of pyrethroid, in *Anopheles gambiae* was also deduced.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgment

The appreciation of the first author goes to the Research and Postgraduate Directorate of Durban University of Technology, Durban, South Africa for hosting the Postdoctoral Fellows and to the ICT and Society Research Focus Area for funding this research.

References

[1] P. J. Kersey, J. E. Allen, M. Christensen, P. Davis, L. J. Falin, C. Grabmueller, D. S. Toney Hughes, J. Humphrey, A. Kerhornou, J. Khobova, et al., "Ensembl Genomes 2013: scaling up access to genome-wide data", Nucleic Acids Research, **42**(1), D546–D552, 2014, Doi: 10.1093/nar/gkt979.

[2] UniProt-Consortium, "Ongoing and future developments at the Universal Protein Resource", Nucl. Acids Res. **39**(1): D214-D219, 2011, Doi: 10.1093/nar/gkq1020

[3] P. Müller, M. J. Donnelly, H. Ranson, "Transcription profiling of a recently colonised pyrethroid resistant *Anopheles gambiae* strain from Ghana", BMC Genomics. **8(**36), 2007, Doi: 10.1186/1471-2164-8-36

[4] D. Nikou, H. Ranson, J. Hemingway, "An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistant strain of the malaria vector, *Anopheles gambiae*", Gene, **318**, 91-102, 2003, Doi: 10.1016/s0378-1119(03)00763-7

[5] A. S. Haruna, K. Mavridis, J. Vontas, J. E. Eyo, "First incidence of CYP9K1, CYP6P4 and CYP6Z1 in Anopheles gambiae ss from Nigeria", 2020, Doi: 10.21203/rs.3.rs-19996/v1

[6] H. Irving, J. Riveron, S. Inbrahim, N. Lobo, C. Wondji, "Positional cloning of rp2 QTL associates the P450 genes CYP6Z1, CYP6Z3 and CYP6M7 with pyrethroid resistance in the malaria vector *Anopheles funestus*", Heredity. **109**(6), 383–392, 2012, Doi: 10.1038/hdy.

[7] X. Chen, C. Tang, K. Ma, J. Xia, D. Song, X. W. Gao, "Overexpression of UDP-glycosyltransferase potentially involved in insecticide resistance in *Aphis gossypii Glover* collected from Bt cotton fields in China", Pest Manag Sci. **76**, 1371-1377, 2020, doi: 10.1002/ps.5648. PMID: 31605421.

[8] V. A. Ingham, C. M. Jones, P. Pignatelli, V. Balabanidou, J. Vontas, S. C. Wagstaff, J. D Moore, H. Ranson, "Dissecting the organ specificity of insecticide resistance candidate genes in Anopheles gambiae: known and novel candidate genes", BMC Genomics. **15**(1), 2014, doi: 10.1186/1471-2164-15-1018.

[9] C. Yang, D. J. McCart, S. Woods, K. G. Terhzaz, R. H. Greenwood, ffrench-Constant, and J.A. Dow, "A Drosophila systems approach to xenobiotic metabolism", Physiol. Genomics, **30**(3), 223-231, 2007, doi: 10.1152/physiolgenomics.00018.2007.

[10] I. Brigandt, "Essay: Homology", The Embryo Project Encyclopedia (23 November 2011).

[11] A. L. Panchen, "Homology—history of a concept", Novartis Foundation Symposia, **222,** 5-18, 1999, doi: 10.1002/9780470515655.ch2. ISBN 9780470515655.

[12] S. B. Kumar, "Homology Modeling Lecture Note (ppt)". https://www.slideshare.net/bharathpharmacist/homology-modeling-43813667.

[13] A. Kryshtafovych, T. Schwede, M. Topf, K. Fidelis, J. Moul, "Critical Assessment of Methods of Protein Structure Prediction (CASP), Round XIII. Proteins*"*, 2019, doi: 10.1002/prot.25823.

[14] M. O. Adebiyi, D. Falola, O. Olatunji, R. O. Ogundokun, E. A. Adeniyi, "Computational prediction of corynebacterium matruchotii Protein's 3D structure reveals its capacity to bind to DNA domain site in the Malaria Vector, *Anopheles",* International Journal of Engineering Research and Technology, **12**(11), 1935–1940, 2019.

[15] S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucl. Acids Res. **25**(17), 3389-3402, 1997, Doi: 10.1093/nar/25.17.3389.

[16] H. Berman, K. Henrick, H. Nakamura, J. L. Markley. "The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data", Nucl. Acids Res. **35**(1), D301-D303, 2007, doi: 10.1093/nar/gkl971.

[17] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, et al. "ClustalW and ClustalX version 2". Bioinformatics. **23**(21), 2947-2948, 2007.

[18] F. Sievers, A. Wilm, D. G. Dineen, T. J. Gibson, K. Karplus, W. Li, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega", Mol. Syst. Biology, **7**(539), 2011, doi: 10.1038/msb.2011.75

[19] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information", Nucl. Acids Research, **37**(1), D387-D392, 2014, doi: 10.1093/nar/gkn750.

[20] R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton. "PROCHECK: a program to check the stereochemical quality of protein structures", J. Appl. Cryst. **26**(2), 283-291, 1993, Doi: 10.1107/S0021889892009944

[21] R. A. Laskowski, V. V. Chistyakov, J. M. Thornton. "PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids", Nucl. Acids Research, **33**(1), D266-D268, 2005, Doi: 10.1093/nar/gki001

[22] V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, "MolProbity: all-atom structure validation for macromolecular crystallography", Acta Crystallogr. **66**(1), 12-21, 2010, Doi: 10.1107/S0907444909042073

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 586-592 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Neural Networks and Fuzzy Logic Based Maximum Power Point Tracking Control for Wind Energy Conversion System

Hayat El Aissaoui*, Abdelghani El Ougli[1], Belkassem Tidhaf [2]

[1]*Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, 30050, Morocco*

[2]*National School of Applied Sciences, Mohammed First University, Oujda, 60010, Morocco*

ARTICLE INFO

Article history:
Received: 12 December, 2020
Accepted: 28 February, 2021
Online: 20 March, 2021

Keywords:
Turbine
PMSG
ANN
Fuzzy Logic
MPPT

ABSTRACT

In grid connected wind turbine (WT) systems, the maximum power point tracking (MPPT) approach has a crucial role in optimizing the wind energy efficiency. To search for the maximum power value of the wind turbine, this contribution proposes a new Maximum Power Point Tracking System (MPPT) for wind turbine related to a permanent magnet synchronous generator (PMSG)). The new proposed MPPT combines two techniques: Artificial Neural Network (ANN) and Fuzzy Logic (FL). The ANN is employed to estimate the maximum voltage of the WT, for various values of wind speed, while the control of DC–DC boost converter operation is executed by applying Fuzzy Logic technique. The comparison of our proposed algorithm to P&O technique has shown that it ensures more efficiency, and we used for that a simulation under Matlab/Simulink.

1. Introduction

Nowadays, motivated by the environmental issues caused by the outrageous use of traditional energy resources such as oil, natural gas or coal, many countries in the world tend to an ecological transition. They also pay more attention to the ecological effects generated by this ruinous exploitation such as: global warming, air pollution, etc.

Due to their adaptability and ease of use, renewable energies are currently a crucial form of energy used in different fields. Thus, we are going to concentrate in this contribution on the case of wind turbine as a mechanism of generating renewable energy.

Wind turbine is an unflagging source of renewable energy that is able to produce electricity, using the power of the wind. It is also worth to mention that the wind power plants can be established either on earth or in oceans. [1]

The WT is an instrument that converts the dynamic wind energy into electrical energy, which will be sent to the power grid. In the grid connected systems of wind turbines (WT), the maximum power point tracking algorithm (MPPT) is a pivotal criteria to improve the wind energy productivity.

In last few years, diverse MPPT techniques have been applied, the major controllers that are used extensively are Climb Search (HCS) or Power Signal Feedback (PSF), Perturbation and Observation (P&O), Optimal Torque Control (OTC), Top Speed Ratio (TSR) and methods of soft computing such as Fuzzy Logic Control (FLC) and Artificial Neural Network (ANN).

The Perturb and observe (P&O) or HCS (Hill climb search) control methods are applied when the system's optimal relation is determined. The P&O approach is used to follow the MPP, by changing the maximization factor and measuring the obtained power.

The MPPT controller based on the Fuzzy Logic Controller (FLC) and the (ANN) are designed to overcoming the limitation of all methods listed above. They have a more rapid response, in the case of fast varying wind speeds.

This article exhibits a new *approach* for Maximum Power Point Tracking for Wind Energy Conversion System.

*Corresponding Author: Hayat EL AISSAOUI, Team SEERIA, ENSAO School, Electronics and Systems Laboratory, Faculty of Science, Mohammed First University, Oujda, Morocco, Email: hayatel89@gmail.com

https://dx.doi.org/10.25046/aj060267

In fact, we propose a new MPPT based mainly on fuzzy logic (FL) and Artificial Neural Networks (ANN) and by doing so we will be able to take out maximum power of the wind turbine. The idea is to use the ANN to estimate the maximum voltage of the WT using different values of wind speed, while a fuzzy controller will be responsible of controlling the DC-DC boost converter.

Our system involves wind turbine related to a permanent magnet synchronous generator (PMSG), a rectifier and a DC-DC converter with MPPT control.

Our results have shown that the proposed system guarantee more efficiency in terms of tracking the MPPT and extracting the maximum power of a WT.

The proposed model is presented in figure 1 below:

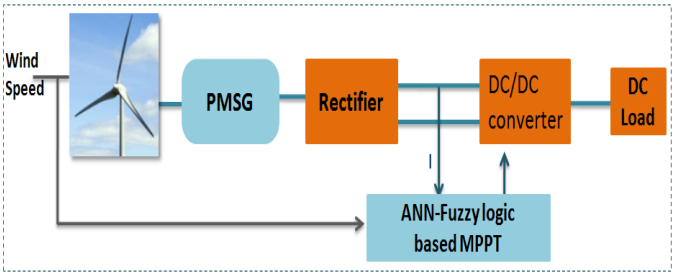


Figure 1: Overall schematic of PMSG based ANN and Fuzzy logic

The manuscript is structured as listed below:

Section 2 describes the designing of a WT wind turbine. The section 3 presents the modeling of the used converter. Section 4 explains the different types of MPPT controllers used in this article: ANN and FLC and later P&O. Section 5 gives the results obtained from the simulation realized in this work with a discussion of the performances of the proposed algorithm.

2. Modeling of A WT

The mathematical expression of the mechanical power generated by the WT is presented by the following equation presents: [2]

$$\text{Par} = \text{Cp}(\lambda, \beta)\frac{\rho * A * Vw^3}{2} \quad (1)$$

Par is the output mechanical power of the turbine (w), ρ is the air density (kg/ m^3), Cp (λ, β) is the performance coefficient of the turbine, , A is the turbine swept area (m^2) , V_W is the wind speed (m/s), λ is the tip speed ratio, and β is the blade pitch angle (deg).

$$\lambda = \frac{R * \omega}{V_W} \quad (2)$$

R means the turbine ratio (m), ω is the turbine angular velocity (rad/s).

The power coefficient is in accordance with β and λ. The coefficient of performance is shown as below:

$$Cp(\lambda, \beta) = 0.22\left(\frac{116}{\lambda i} - 0.4\beta - 5\right)e^{\frac{-21}{\lambda i}} \quad (3)$$

$$\frac{1}{\lambda i} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1} \quad (4)$$

The aerodynamic torque is presented as follows:

$$T = \frac{\text{Par}}{\omega} = \frac{1}{2\lambda}\rho\pi R^3 Vw^2 Cp(\lambda, \beta) \quad (5)$$

Based on equation 1, we can say that the most significant factor to achieve the maximum power produced by a wind turbine is the curve of Cp; in fact, When the Cp is at maximum, it means that the power produced by the wind turbine has reached its maximum. For this work, β is equal to zero.

Every different value of λ we obtain a constant β and an optimum Cp. For every wind speed value, we can define a unique rotor speed that allows achieving a maximal power. Thereby, considering the wind speed constant invariable, the optimum Cp value will be bound only on the wind turbine rotor speed [3].

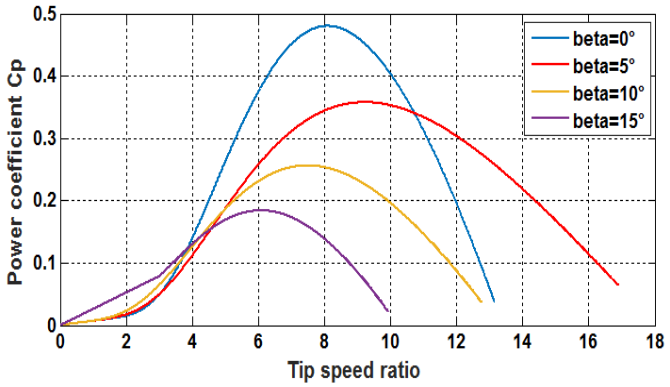


Figure 2: Characteristics of power coefficient

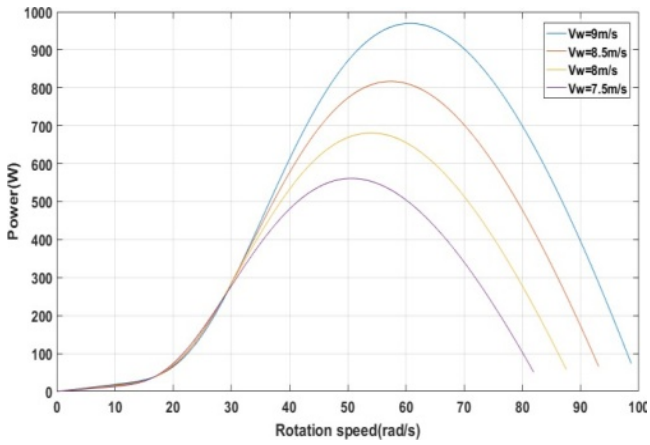


Figure 3: Power graphs under various wind speeds (β=0)

The parameters of the Turbine Generator are detailed as follows:

Table 1: WT generator system characteristics

Characteristics	Values
Rated Voltage	90 V
Rated Power	1000W
Synchronous inductance	1mH
Rated Current	4.8A
Number of poles	8

Characteristics	Values
Synchronous resistance	1.13Ω
Friction coefficient	0.006N.m.s/rad
Magnetic flux	0.16Wb
Moment of inertia	0.005N.m
Blade length	1.2m
Air density	1.2 kg/ m^3

3. Converter Modelling

The boost converter is basically a voltage step-up power converter which accepts a low-voltage input and delivers an output at a much higher voltage. The DC/DC boost converter topology is the most commonly used topology. [4]

The electrical circuit of DC/DC boost converter used in this contribution is shown in the figure below:

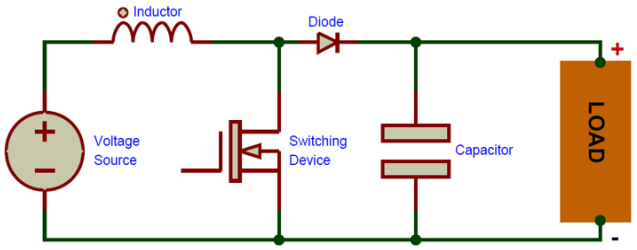


Figure 4: Configuration of Boost converter

The Boost Converter is the conventional basic DC/DC converter, which runs with a single switch.

The boost converter contains a capacitor, a switch, a diode for its operation, and an inductor, as illustrated in Figure 4. [5]

Once the switch is activated, the diode is reverse-biased, the current begins to increase by charging the inductor; and a capacitor holds the charge.

The model of the boost converter is as described below:

$$V_0 = \frac{1}{1-D} V \quad (6)$$

The Vo, V are the output and the input voltage respectively.

The mathematical expression of the parallel chopper is:

In the case of $t \in [0, DT]$:

$$\frac{di}{dt} = \frac{V}{Lf} \quad ; \quad \frac{dV_0}{dt} = \frac{V_0}{RC} \quad (7)$$

In the case of $t \in [DT, T]$:

$$\frac{di}{dt} = \frac{V-V_0}{L} \quad ; \quad \frac{dV_0}{dt} = \frac{i}{C} - \frac{V_0}{RC} \quad (8)$$

The inductor of the boost converter and capacitor are determined by:

$$L = \frac{V.D}{\Delta I_L .f} \quad (9)$$

$$C = \frac{I_0 .D}{\Delta V_0 .f} \quad (10)$$

where ΔI_L and f are the approximate inductor ripple current and the corresponding switching frequency, respectively. ΔV_0 the estimated output ripples voltage.

4. The proposed MPPT method for the Wind Turbine

To search the maximum power produced by the wind turbine, we propose a new power tracking approach based mainly on fuzzy logic (FL) and Artificial Neural Network (ANN), as shown in Figure 5.

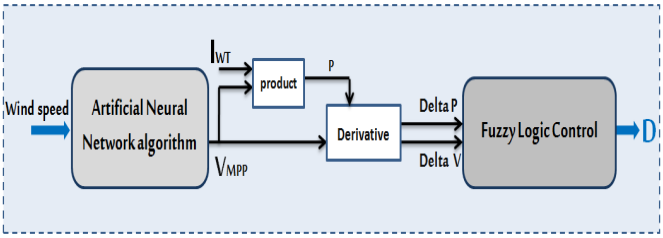


Figure 5: Block diagram developed MPPT

In this paper, two steps are applied to keep track the PPM of the WT.

In the first step, the ANN is used to obtain an Estimated Optimal Voltage Value (V_{MPP}) for each wind speed value.

ANN estimates the maximum value of voltage of the wind turbine associated to each given wind speed. It predicts the value of V_{MPP} based on the training of a database that combines each wind speed with the associated Vout voltage.

During the second step, the Fuzzy Logic method is implemented to give the value of the Duty cycle.

From the maximum voltage (V_{MPP}) and the value of the current I_{WT}, the derivatives deltaP and deltaV are calculated, which will then be the Fuzzy block inputs. The fuzzy logic controller is implemented to indicate the value of Duty cycle based on the inputs delta P (power) and delta V (voltage).

4.1. ANN Based MPPT Algorithm

The design of artificial neural networks is based on the biological neuron structure of the human brain. Artificial neural networks can be described as systems composed of at least two layers of neurons, an input layer and an output layer and usually including hidden layers. If the problem is complex, the artificial neural network must have more layers. Each layer contains a large number of specialized artificial neurons.[6]

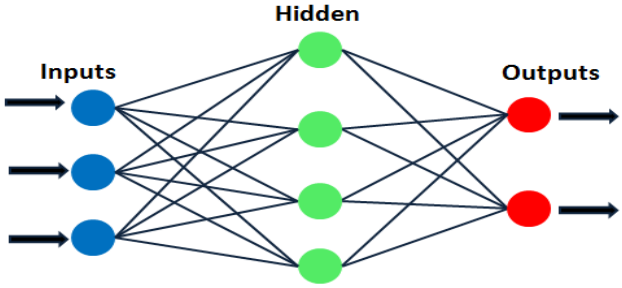


Figure 6: ANN Diagram

Neurons are linked and interchange with each other. Every neuron node has data as an increment and it is able to execute trivial operations on it.

Afterward, the outcome of these actions is delivered to other neurons. The result of each neuron node is called the nodes activation or node value.

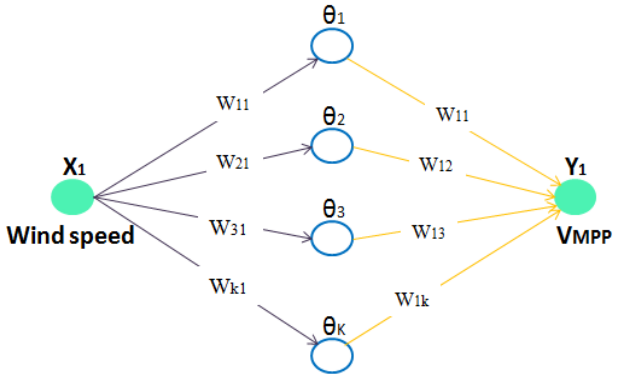


Figure 7: V_{MPP} prediction using ANN

where X_1 represents the wind speed value, it is the input of the ANN, Y_1 represents the output V_{MPP} (optimal voltage).

The artificial neural network is created from a program under Matlab, based on the training of a given database (wind speed, V).

In our case, in the first layer, the optimized number of neurons is forty neurons, the second layer has one purelin neuron, one neuron in the input layer and also one in the output layer is constructed. The model of network is as below:

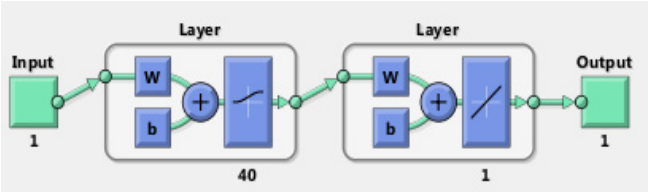


Figure 8: Neural Network Model

4.2. FLC For Wind Turbine

The strong point of the fuzzy logic controller is that it deals with the unspecific inputs and nonlinearity and brisk changes.

FLC is mainly The FLC is mainly structured in three phases that represent the fuzzification, the fuzzy inference engine (rule base table determined by previous instructions) and defuzzification.

Fuzzification: Transform numeric input values to linguistic values using a membership function. [7]

Rules base: Fuzzy rules have been used to characterize the connection that exists between the output and input of the fuzzy control. The number of fuzzy rules depends, among other things, on the partition of the speech universes of the input and output parameters. [8]

Defuzzification: This step consists of converting a linguistic value into a numerical value.

In this method, the fuzzy logic controller is used to indicate the value of D based on the inputs delta P (power) and delta V (voltage).

The process of FLC for the WT subsystem is shown in Figure 10.

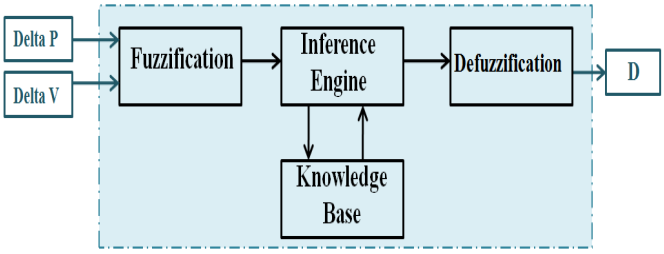


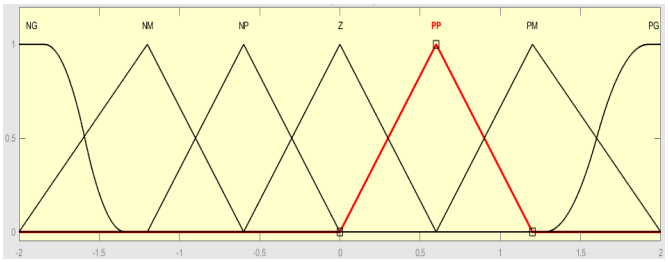
Figure 9: Structure of a fuzzy controller

DeltaP and DeltaV are shown by the equations below:

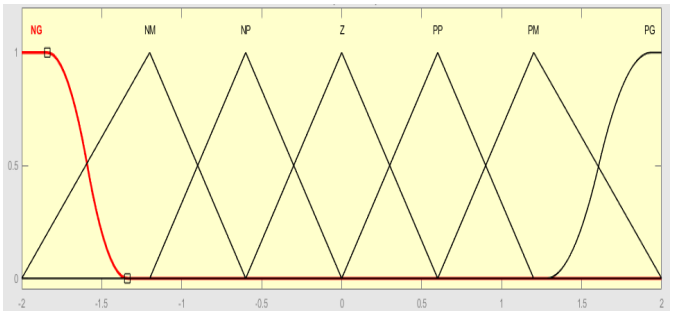
$$DeltaP(k) = P(k) - P(k-1) \quad (11)$$

$$DeltaV(k) = V(k) - V(k-1) \quad (12)$$

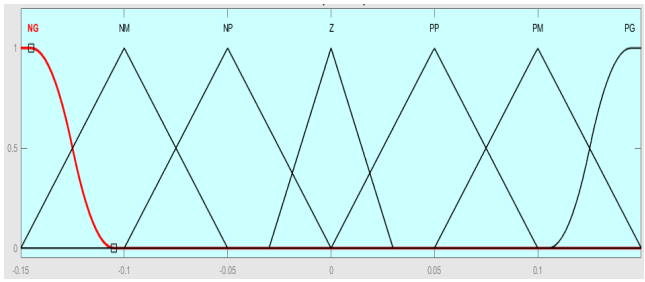
The membership functions of DeltaP, DeltaV and D are respectively shown in Figure 10(a), Figure 10(b) and Figure 10 (c).



(a)



(b)



(c)

Figure 10: Membership functions related to (a) DeltaP, (b) DeltaV, (c) D.

The table 2 shows the inference rules for the various combinations of the variables DeltaP, DeltaV with output D. [9]

Table 2: Inference Rules related to Fuzzy Logic WT

Delta V / Delta P	NG	NM	NP	Z	PP	PM	PG
NG	NG	NG	NG	NM	NM	NP	Z
NM	NG	NG	NM	NM	NP	Z	PP
NP	NG	NM	NM	NP	Z	PP	PM
Z	NM	Z	NP	Z	PP	PM	PM
PP	NM	NP	Z	PP	PM	PM	PG
PM	NP	Z	PP	PM	PM	PG	PG
PG	Z	PP	PM	PM	PG	PG	PG

4.3. Perturb and Observe Based MPPT Algorithm for wind turbine

The Perturb and observe (P&O) method is the most widely employed because of its simple implementation. The concept of this algorithm is illustrated in Figure 11.

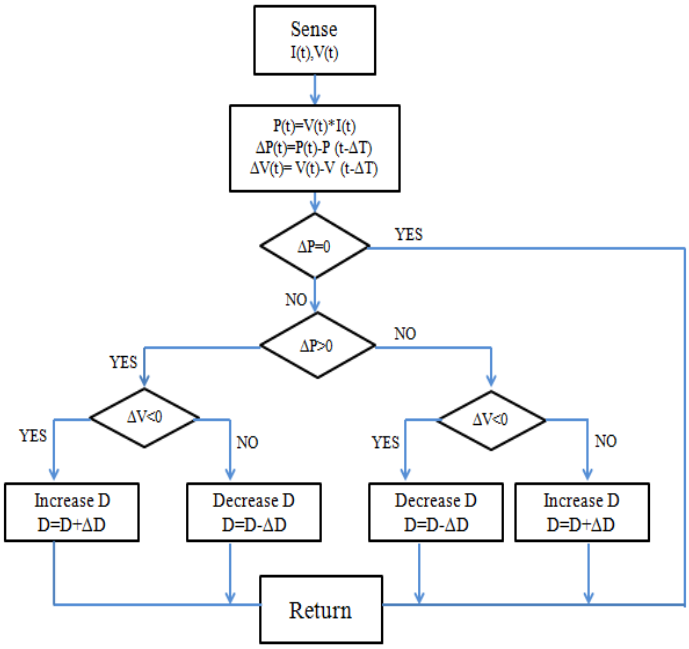


Figure 11. Perturb and Observe Approach (P&O)

5. Results and Discussion

The proposed algorithm is verified with MATLAB SIMULINK, as illustrated in Figure 12:

To implement the new MPPT algorithm and to verify its efficiency, MATLAB / SIMULINK is used as shown in Figure 12.

The system includes the following elements: a wind turbine, a permanent magnet synchronous generator (PMSG), a rectifier, a boost converter, an ANN block and a fuzzy logic block.

Figure 13 shows that the wind speed varied in three stages, it passes from 8.4 m / s to 8.5m / s, then to 9.1 m / s.

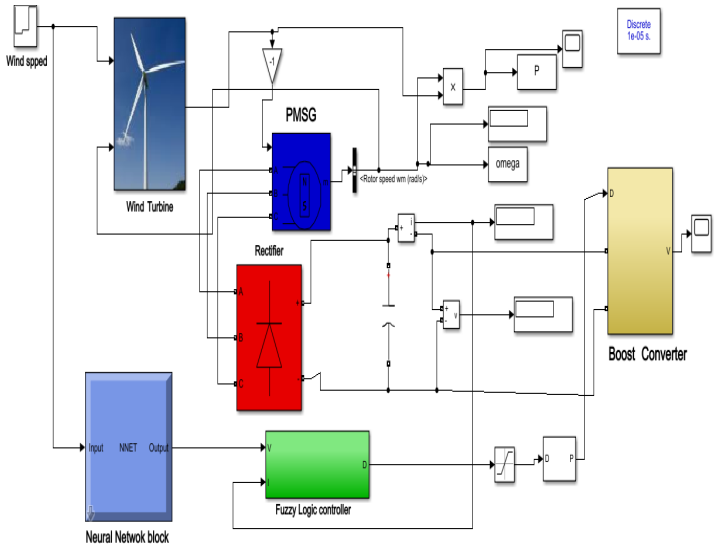


Figure 12: ANN-Fyzzy based MPPT algorithm deployment on Simulink

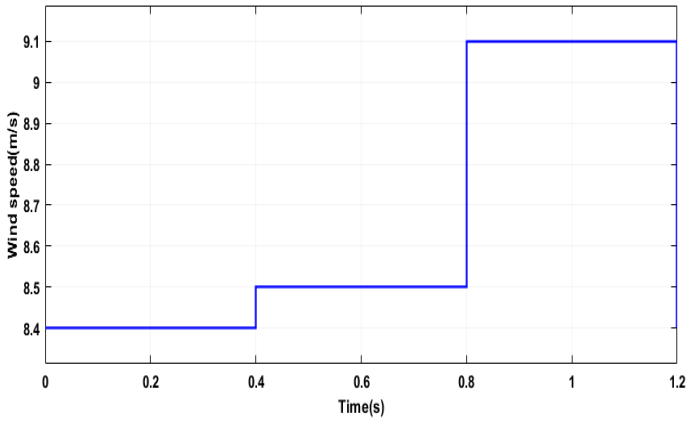


Figure 13: Varying wind speed

Figure 14 shows, the variation of the performance coefficient, for the wind speed which varies between 8.4m/s, 8.5m/s, and 9.1m/s.

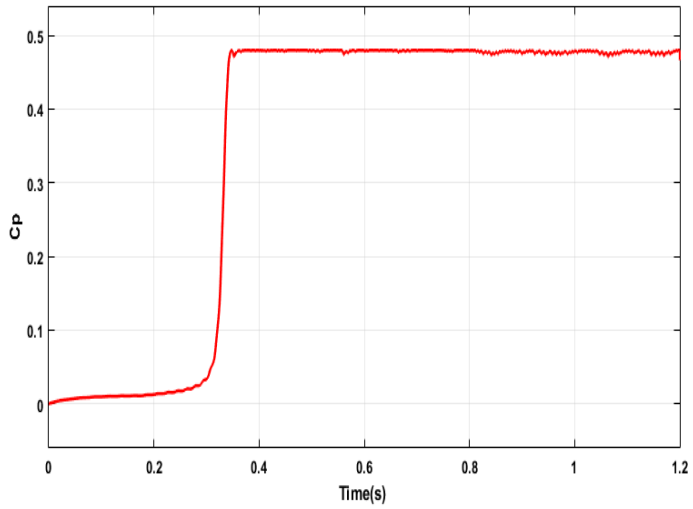


Figure 14: power coefficient for different wind speed.(maximum value is 0.48) using ANN-Fuzzy algorithm

Our results seem to confirm that our system is able to keep track of its MPP (Cp=0.48) considering various wind speed values.

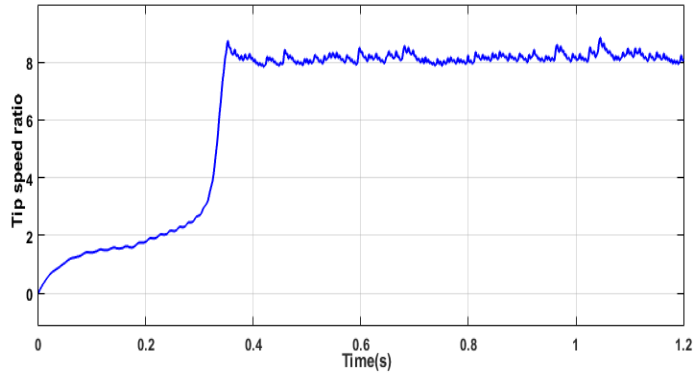


Figure 15: Tip speed ratio curve under a varying wind speed using ANN-Fuzzy logic algorithm

According to the simulation results the power coefficient and tip speed ratio are set at the maximum values.

Figure 16 shows the torque evolution curve related to the proposed approach based on ANN and Fuzzy logic.

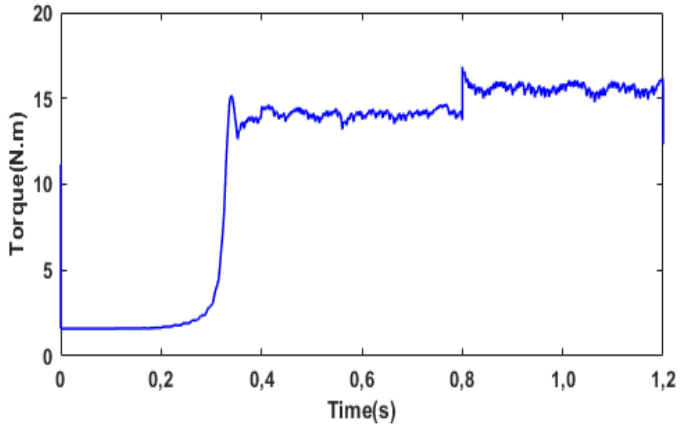


Figure 17 shows the WT's mechanical output power, using the new ANN-FL MPPT approach, for the wind speed which varies between 8.4m/s, 8.5m/s, and 9.1m/s.

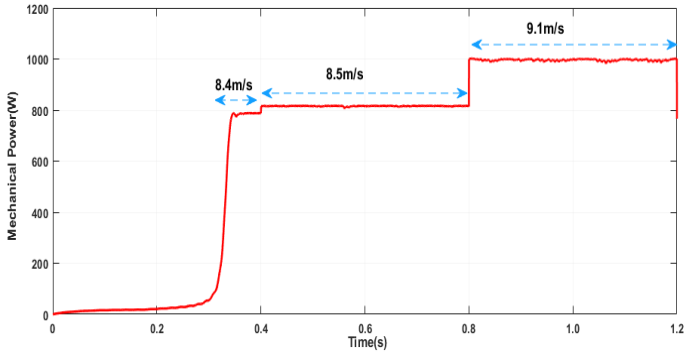


Figure 17: The output mechanical power under different wind speed using ANN-Fuzzy algorithm

By comparing the evolution of the power delivered by WT (figure17), with the results of figure3, under different wind speeds, we can deduce that our system is able to keep track of the maximum mechanical power through new MPPT method based on ANN and Fuzzy logic.

Figure 18 shows the WT's mechanical output power, using the P&O algorithm, for wind speed which varies between 8.4m/s, 8.5m/s, and 9.1m/s.

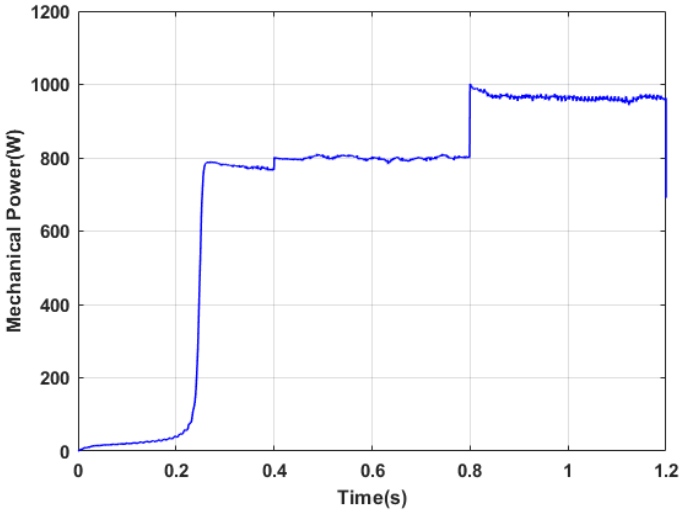


Figure 18: The output mechanical power under different wind speed using P&O algorithm

By comparing the evolution of the power delivered by WT (figure18), with the results of figure3, under different wind speeds, we can deduce that the system can't reach the maximum mechanical power through MPPT based on P&O.

By comparing the power curves provided by WT using the ANN and Fuzzy Logic algorithm and the one using P&O, it seems that the ANN technique reaches the maximum power with more stability than the P&O technique.

The P&O technique is not able to achieve MPP, because the power delivered using P&O is not maximum and it represents a significant oscillation, as shown in the power curve.

6. Conclusion

As a way to improving the power produced by the wind turbine it is necessary to track and extract the maximum power point.

This work presents a new MPPT method based on the ANN and fuzzy logic method. The idea is to combine these two methods, to provide an efficient system that outstrip the P&O based one in terms of MPPT and extracting maximum power from the WT. In the proposed approach, the ANN estimates the value of V_{MPP} for various wind speed values while Fuzzy logic controller gives the value of Duty cycle D for controlling the DC-DC boost converter.

In order to show the efficiency of the proposed new approach based on ANN and fuzzy logic, it has been compared with the famous P&O algorithm.

The simulation realized in this work demonstrates that the maximum power point tracking based on the ANN and FLC algorithm can track and maintain the maximum power delivered by the wind energy production for every wind speed value. According to these results, it is undeniable that the new MPPT

based on ANN and fuzzy logic is more efficient than the P&O method.

Based on the results of the simulation, it appears that the validity of the proposed MPPT controller has been confirmed for wind speed variations, using MATLAB /SIMIULINK.

References

[1] A. Choumane, O. Boukhari, "Wind energy in Algeria: potential and achievements," in International conference on renewable energy strategies and its role in achieving sustainable development.

[2] M. Zerouali, M. Boutouba, A. El Ougli and B. Tidhaf, " Control of variable speed wind energy conversion systems by fuzzy logic and conventional P&O," in 2019 International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS), 1-5, IEEE, 2019, doi: 10.1109/ISACS48493.2019.9068866.

[3] P. Suresh, B.C. Sujatha, "Maximum Power Point Tracking Of A Hybrid Solar-Wind Power Generation System For A Smart Grid Using Fuzzy Logic Control," International Journal of Scientific & Engineering Research, **7**(5), 165-170, 2016.

[4] S.S. Dheeban, N.B. Muthu Selvan, C. Senthil Kumar, "Design Of Standalone Pv System," International Journal Of Scientific & Technology Research, **8**(11), 2019.

[5] I. Duka, C. Noble. " High frequency DC/DC boost converter," submitted to the Faculty of the Worcester Polytechnic Institute in partial fulfillment of the requirements for the Degree of Bachelor of Science in Electrical and Computer Engineering , 2011.

[6] J.K. Gowri, M.C. Popuri, "Neutral Networks Based Maximum Power Point Tracking (Mppt) Of Wind Power Generation," in The International Journal of Scientific Development and Research (IJSDR), **5**(2), 2020.

[7] K.H. Lee, "First course on fuzzy theory and applications," in Springer Science & Business Media, 2004.

[8] C.C. Lee, "Fuzzy logic in control systems: fuzzy logic controller. I," IEEE Transactions on systems, man, and cybernetics, **20**(2), 404-418, 1990, doi: 10.1109/21.52551.

[9] H. Othmani, H. Chaouali, D. Mezghani, A. Mami. " Optimisation de la Technique de Perturbation et Observation par la logique floue," International Journal of Scientific Research & Engineering Technology (IJSET), 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 593-601 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A framework for the alignment of ICT with Green IT

Manuel Landum[1,*], Maria Margarida Madeira e Moura[1], Leonilde Reis[2]

[1]*Universidade do Algarve, DEEI/FCT, Campus de Gambelas, Faro, Faro, 8005-139, Portugal*

[2]*Polytechnic Institute of Setúbal, Campus do IPS, Estefanilha, Setúbal, Setúbal. 2910-761, Portugal*

ARTICLE INFO

Article history:
Received: 12 December, 2020
Accepted: 28 February, 2021
Online: 20 March, 2021

Keywords:
Green IT
Environment
Information and Communication Technologies

ABSTRACT

The Public Administration is forced to transform itself by taking advantage of the contribution of ICT to in the process of reducing bureaucracy and increase transparency, promoting the dematerialization of processes, increasing the quality of online services, allowing greater ubiquity of access, reducing response times, in the search for improvement of the quality of life of its citizens. Decision-making should consider objectives not only technical but also financial, environmental, and social objectives, ideally aligning with Green IT. The objective of the paper is to present a framework, supported by international standards and frameworks, that allows measuring and guiding the alignment of ICT with Green IT for the optimization of practices instituted in organizations, namely in Local Government. This framework includes a qualitative component with several phases and quantitative component with a metric that allows the evaluation of alternative strategies for a given goal. The phases considered are Problem Identification; Problem assessment; Study and planning; Project; Telecommunications and Printing; Information Security; Innovation; Improvement of Citizen Service; Evaluation / opinion. The complete metric includes four valences: IT, financial, environmental, and social. The IT valence, its indicators and corresponding weights are illustrated in a practical example. The proposed framework is an innovative contribution to the area, clearly integrating the support of the perspective of Green IT and thus actively contributing to the implementation of sustainability policies and alignment with Green IT best practices in Local Government, as well as presenting with greater level of detail the components of the structure that emphasize Green IT concerns. The main expected results of the application of the framework are contributing to the implementation, in local government, of sustainable policies and good practices aligned with Green IT, whist targeting cost reduction and optimization, ubiquity of access, increased productivity and ensuring safety standards.

1. Introduction

This paper is an extended version of the article originally presented in 2020 at the 15th Iberian Conference on Information Systems and Technologies (CISTI), where the themes address the alignment of public administration (PA) good practices, namely the alignment with Green IT of local administration (LA) practices [1].

The pervasive way in which Information and Communication Technologies (ICT) are currently part of our daily lives, associated with environmental concerns, motivated this work of integrating environmental sustainability aspects into information technology (IT) governance. Environmental sustainability should be translated to strategic thinking within organizations and is based on three objectives, pollution prevention, product management and the use of clean technologies; these objectives incorporate underlying concerns such as carbon footprint, global warming, fossil fuels, volume of material recycling and economic sustainability [2]. The three objectives can be applied at three different levels: individual, organizational, and social. Its combination can be used to identify IS or IT deployment opportunities to improve sustainability [3].

The objective of this communication is reporting a framework that includes ICT best practices in technology governance, aligned with Green IT, where the focus is on the phases that are more

*Corresponding Author: Manuel Landum, Email: A55265@ualg.pt

aligned. The design of this framework was supported by international standards and frameworks, and aims at optimizing governance in organizations, namely in LA. The framework incorporates PA specific constraints into implementing best practices and how these best practices can be defined. The main contributions of this framework are based on the metrics presented, sustainability and the improvement of the ICT ecosystem.

This article is organized into four sections: firstly, in this introduction, the contextualization of the problem and the objective of the article are presented; in the second section, the literature review is presented, which includes the theoretical foundations of the supporting thematic; following, in section three, the framework that includes ICT best practices, aligned with Green IT and supported by international standards and frameworks in order to optimize practices instituted in organizations, mainly in the LA under study, is presented; finally, in the fourth section, some conclusions highlighting the consequences from the use of the framework and its contributions to the community are presented.

2. Literature review

The literature review, based on the collection of articles on the subject, was supported by references and citations of papers, together with keywords directed to the theme, as well as in the analysis of sustainability and the pursuit of good practices in alignment with Green IT.

1.1. Digital Transformation

The process of digital transformation arises in the daily life of organizations, particularly in PA and more specifically in LA, driven by the need felt to modernize internally, externalizing more and better services for their customers, in a ubiquitous way, allowing for greater mobility and freedom in the way these customers relate to the organization. And ubiquity is a good practice, aligned with Green IT.

It should be noted that "computer technology and the increasing availability of digital data are radically changing forms of research and production of knowledge" [4] in our areas. Customers seek solutions that lead to a streamlining of processes, where requests and deliveries of services are made remotely, so that they can avoid face-to-face calls, that would otherwise cause loss of time and costs. With this demand, customers press services to ensure that processes suffer digital transformation, conducing to prompt, fast and click-away responses.

The process of "digital transformation is not just about computerizing a productive system. It involves exploring these systems to develop, maintain or improve skills that a company needs or dominates. Having this idea in mind is important when a company thinks about taking the step towards digital transformation" [5]. From this, it can be inferred that digital transformation is not just about digitization of processes, and it is much more than that. From the perspective of [4], digital copies of printed theories do not transform these theories into digital theories, just as literature is much more than a mere transliteration of oral discourse.

Moreover [6], the digital transformation of society does not follow directly from digital technology. The latter is a necessary condition, but not enough. For digital transformation to take root, digital technologies must be imbued with social significance. Digital transformation implies process reengineering, creation, and optimization of flows, incorporating the acquisition of new skills and business practices. Organizations are thus driven to adhere to digital transformation and automation, maintaining or strengthening their competitive position, further enhancing the human-machine relationship, and increasing the degree of automation of processes.

1.2. Virtualization

Virtualization will be a way to implement the digital transformation process more quickly, being an important technique for the increasingly digital world of our daily life. One of the achievements of this technique is the creation of several servers or virtual desktops, from one or more physical machines; this allows, among other factors, to reduce the number of physical machines, leading to the reduction simultaneously of the amount of power supplies, the reduction of the time until readiness by facilitating the management of virtual machines, the reduction of energy consumption and environmental improvement and consequently leads to a decrease in the need to purchase new equipment. Virtualized data centers are designed to provide better management flexibility, lower cost, scalability, better resource utilization and energy efficiency, and thus, consequently, reducing costs with installation, maintenance, cooling, and physical space, thus contributing for an alignment with Green IT [7].

Cloud computing, or cloud computing, turns out to be one of the factors responsible for boosting virtualization, given the fact that it is defined as a shared set of virtualized resources, where in fact users access cloud resources using a unit, independent, called Virtual Machine (VM) [8]. These VMs are housed on physical servers, within large infrastructures, of data centers. Commercial offers facilitate the access to these virtualized pools of resources, providing users with nearly infinite computing resources in the form of processing power, memory, and disk space. Thus, and according to [9], the concept of virtualization is applied in cloud computing systems to help both users and owners get better use and efficient management of the cloud at the lowest cost.

Generically, virtualization is a technique that allows the reduction of resource use and contributes to environmental improvement, fostering the alignment with Green IT.

1.3. Green IT

Based on the alignment mentioned above, a definition was sought and a set of good practices that demonstrate it. If there is no single definition of what Green IT is, it is then considered a set of good practices about the use of computers and IT resources, in a sustainable, more efficient, and responsible way with the environment.

Green IT, as a concept, surfaces as a worldwide movement from the concern of organizations and/or individuals in preserving the environment, seeking the promotion of social and environmental responsibility, in this technological world. This movement focuses its ideals on "preserving the environment, in the search for the reduction of environmental impacts, based on sustainable practices and, in this particular case, interconnected

with ICT" [10]. Currently, the theme is of particular interest given the problems of climate change and sustainability, considering the daily reality, where these issues need to be included in the decision-making process to contribute in a timely manner to the environmental non-aggravation.

It should be seen that as the population increases, developing or underdeveloped economies end up committing to the rapid growth of infrastructure, leading to a great pressure from the consumption of natural resources. The realization that natural resources are not inexhaustible and that it is not possible to continue with economic growth without considering the variable environment and society, opens the way to the search for new alternative solutions for the productive system, such as the concept of Sustainable Development, which aims to maintain the balance between economic growth, social equity, and the natural environment [11].

Currently, with growing concern about the impacts of environmental strategies at local and global level, there is a need to achieve development in a more environmentally responsible manner, by balancing the choices between the environmental, economic, and social aspects of sustainability [12]. It is therefore necessary for organizations, particularly those responsible for hardware development, to look at the concept of Green IT and develop policies of social and environmental responsibility. Some organizations, in response to customer pressure, already exhibit an awareness and an effort to reduce operational and technological costs, through energy efficiency, in alignment with the good practices of Green IT.

The concerns regarding ICT consider the negative impact on the environment, not only by the electricity consumed, but also by the materials used in the manufacture of hardware, such as lead, arsenic, antimony trioxide, selenium, cadmium, chromium, cobalt, and mercury. This problem of society, in particular ICT, "includes electronic waste, municipal waste, the toxicity and hazard stemming from this equipment and whose solution is to comply with the directives on electronic waste, not allowing it to be deposited in bins, but rather collected by suppliers, treated and recycled, where even some restrictions are placed on manufacturing materials and legislation to control the transfer of electronic waste between borders" [10].

It is generally accepted that in organizations the use of some resources in some processes is conducive to waste, such as energy waste, paper waste, or of other materials, time, and financial resources. In view of this fact, Green IT practices can generate financial returns, sustainability, and environmental improvement. It is necessary to raise awareness among organizations and the PA, to consider the environment, but effectively, not only with good intentions. In [13], the author defends that the perspectives of sustainability transitions emphasize the need of the public sector to adapt policies towards technological innovation, and to be institutionally redesigned. Ecological concerns that allow environmental sustainability, such as reducing the carbon footprint or reducing the use of fossil fuels, are effective if actions involving e-waste management and waste management are considered, adopting an ecological stance in the life cycle of products that allows the reduction of the quantities of equipment and materials to be recycled and the costs inherent. In this way, Haley considers the urgent need to make the transition to a low-carbon economy and has reintroduced the importance of technology-specific policy approaches [13].

In view of this scenario, more efficient resource use practices should also be considered and allow for the reduction of energy consumption, which leads to a reduction in CO2 emissions and reduction of energy costs in the case of virtualization. Actions that lead to paper reduction, such as dematerialization and digital transformation, are also important.

It should be noted that the implementation of sustainability policies and practices in public administration has been adopted at different paces, depending on the country, the level of administration (e.g., local versus national or central) or the activities and objectives of each organization [14]. All this depends on the organizations being willing to adopt a posture of social and environmental responsibility. The adoption of environmental practices and tools in general management is a growing reality in local public organizations as these institutions are shifting their management to sustainability, although local governments move faster than other levels of the public sector in terms of integrating environmental and sustainability aspects into operations and strategies [15]. It should also be considered that in the 2030 Sustainable Development Goals Agenda, in addition to the limited obligations of governments, there is an emphasis on impact, and the means for implementing the targets should be significantly updated [16]. Therefore, "local governments are faced with the need to be part of the change from the perspective of sustainability and integrate good practices to improve their own performance" [15].

Faced with this reality and considering the need for organizations, particularly in LA, to find sustainable ICT good practice models that require less of the environment, aligning them with Green IT, a framework is presented.

3. The framework

This framework for the alignment of ICT with Green IT encompasses two complementary approaches: a qualitative methodology, supported by international standards and frameworks, to provide the optimization of practices in organizations, namely in LA; a quantitative methodology for the assessment of strategies. Given the approach developed, and which can be implemented in other municipalities of the country, it is considered that the results can constitute an added value. Generally, considering the set of skills of the organization, the Chief Information Officer (CIO) should validate the suitability and the weighting factors used in the assessment of alternative strategies.

In Sections 3.1 and 3.2 an overview of the qualitative and quantitative components of the framework are presented and, where appropriate, illustrated with small instances from the use in a real case scenario.

1.4. Qualitative components of the framework

This component includes practices, deemed good, which are added value because they contribute simultaneously to increased productivity, increased information security, and management optimization. Different contributions towards the definition of

these good practices were taken into consideration. Considered relevant were all those practices that contribute to economic and environmental sustainability, process optimization through dematerialization and digital transformation, combat waste, reduce materials to be recycled, reduce costs, reduce CO2 emissions, fostering greater alignment with Green IT. These good practices were formalized into the model. This component also incorporates sustainability factors, aligning with Green IT and includes the generality of the objectives and constraints, identified in the municipality under study, at the level of ICT.

Figure 1 represents the phases of the model, starting in phase zero, with the Identification of the Problem, followed by eight more phases, in a governance model that allows identifying good practices that align with the objective of the activity.

- Phase 0, Problem Identification.
- Phase 1, Evaluation and Continuous Improvement.
- Phase 2, Study and Planning.
- Phase 3, Project.
- Phase 4, Telecommunications and Printing.
- Phase 5, Information Security.
- Phase 6, Innovation.
- Phase 7, Improvement of Citizen Services.
- Phase 8, Evaluation/Opinion.

The model will be iterative from phase 2 to phase 6, going through all phases up to phase 8, where, based on the collected elements, evaluation and opinion, there may be a need to go through the entire cycle again, optimizing processes, alluding to the Plan, Do, Check and Act (PDCA) cycle, but this with more phases. It should be noted that when applying the model, choices are made and, in the corresponding qualitative component, greater weights are given to greater sustainability.

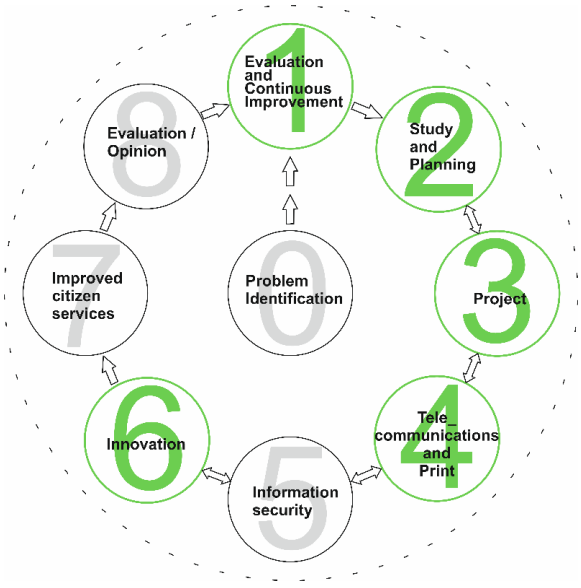


Figure 1 - The Framework

Complementing the context of the main objective of this article, it is intended to deepen the phases where the alignment of good practices with Green IT; consequently, the phases 1, 2, 3, 4 and 6, marked in green color, are more focused.

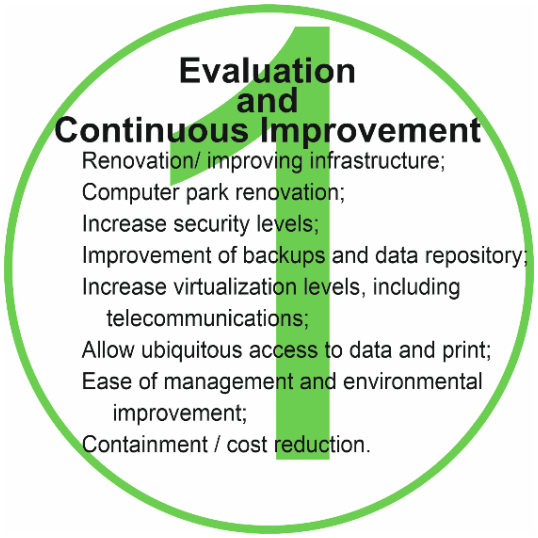


Figure 2 - Phase 1, Evaluation and Continuous Improvement

In Phase 1, as illustrated in Figure 2, the goal is to study the possibility of moving from a traditional model to a virtualized infrastructure model, including telecommunications. This virtualized model can be a way to reduce costs with maintenance contracts, enhance increased information security and internal know-how in the maintenance and configuration tasks of the virtual plant. In the context of this evaluation phase, it is also intended to ensure the ubiquity of access, between the various buildings, decentralized throughout the county, providing the data on impressions to ensure greater mobility of people and autonomy daily [17]. Particularly relevant are the two final tasks in this Phase 1: the evaluation of ease of management and environmental improvement, always present throughout this project the environmental preservation and the use of technologies that allow this preservation, in alignment with Green IT; and the containment equation or whenever possible, cost reduction without compromising the project [17].

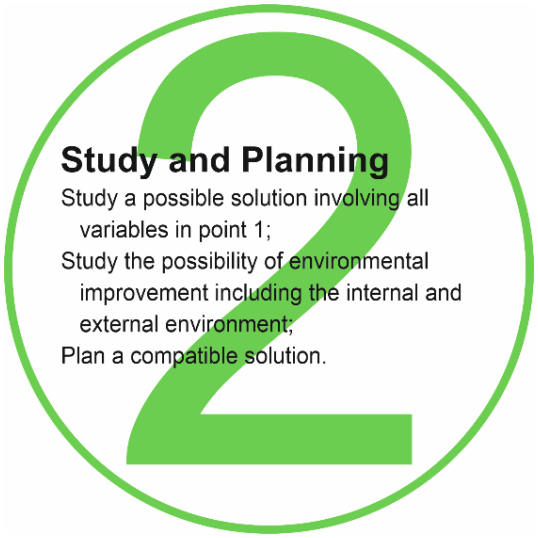


Figure 3 - Phase 2, Study and Planning

Phase 2, as summarized in Figure 3, begins with the Study and Planning, having been concluded the assessment of the constraints encountered, the needs for improvement, and upgrade of the solution.

This phase should be considered as one of the most relevant and decisive in the conduct of the whole approach, although we are facing an iterating process between this phase and phase 6, as it will be based on the results of the study and planning carried out, that all subsequent phases will be supported, and where the need to change the areas outlined may bring unexpected costs to the organization or unforeseen results [17]. Thus, it is intended to study an approach that allows the involvement of all variables described in the previous phase, and that simultaneously includes the possibility of environmental improvement, encompassing both the internal environment and the external environment., aligning with Green IT. Consequently, there is the intention of planning a solution that can be compatible and retro compatible with the existing one, used, at the level of ICT, both hardware and software, also minimizing the costs of the final solution.

After the completion of the Study and Planning, Phase 3 begins (see Figure 4) where the architecture will be developed in design.

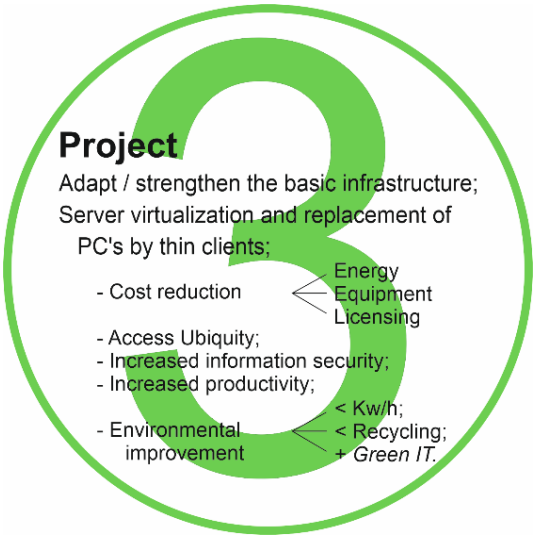


Figure 4 - Phase 3, Project

The input of this phase will be based on the output of the previous phases, where it is intended to adapt and strengthen the basic infrastructure as fundamental support of the project.

The aim is to build a data center internal to the municipality, which can serve to housing the active and passive equipment necessary for the interconnection of all buildings of the municipality, through routing and switching, internet connection and data storage, to allow ubiquity of access to data and printing, increasing productivity increase and without compromising information security, preferably aligned with Green IT. This infrastructure should be in a virtualized environment, as described in phase 1 and all equipment must be supported and connected to a UPS capable of maintaining the infrastructure for some time in the event of a power outage. In addition to server virtualization, with the creation of several VMs, in a distributed environment, the replacement of most PC's, in the end of life, by thin clients that will connect to the servers, that after the cost/benefit assessment and considering the requirements, one can opt for connections by Remote Desktop Protocol (RDP), or connections to Virtual Desktop Infrastructure (VDI). Costs must be evaluated together with the benefits of the adopted solution [17]. Also, in the field of virtualization, it is noted that the replacement of end-of-life PC's by thin clients is conducting to reduction of costs, particularly energy as will be proven later, not undermining the alignment with Green IT.



Figure 5 – Phase 4, Telecommunications and Print

In phase 4, one targets the solutions that allow the reformulation of telecommunications and printing, as can be seen in Figure 5 to optimize these two areas of resource expenditure and equate how contributions to environmental improvement and containment or cost reduction can be achieved. Although distinct, it is considered that they should be treated in an integrated way per se, because they group a set of features and valences related to digital transformation and that allow, in this context, the clear expression of added value for the organization.

Specifically, in telecommunications, it was identified a critical situation with the continued use of physical telephone centers in old, deteriorated buildings, where humidity prevails, implying constant breakdowns and obviously the breakdown of telecommunications between buildings and the exterior, with the consequent constraints and intrinsic costs. Another problem is the difficulty of daily management of IT technicians, hindered by the need to go to the different locations and locally deal with issues; an integrative solution should be achieved in the future and one that allows facilitated management, without increased costs [17]. Even if the option is virtual telephone centers, we will also be contributing to environmental improvement, with the reduction of the number of materials to be recycled and, after checking the numbers, one may state that a reduction in energy consumption occurred, as less movement will translate to a reduction of CO2 emissions to the atmosphere, which will help in reducing the ozone hole, clearly in alignment with Green IT.

Regarding printing, one of the concerns of the municipality, or of any municipality or organization, is the high costs, of the equipment and space occupied, of the printing itself, in terms of paper and ink or tonner, as well as of keeping stock of printing consumables, worsened by the proliferation of various models,

requiring the increase of stocks, contributing all this to the increase of costs. It should be noted that there is also the stock management process: adding to the need of a large investment due to the number of different models, it may happen that cartridge stocks are renewed by the terminus of their shelf life, even if the corresponding printers have already been disposed of and information not propagated [17].

Thus, solutions should be sought to reduce the number of equipment and achieve homogeneity of models, contributing to scale savings, simultaneously aiming to reduce or end stocks, and so contributing to process optimization, faster response and containment or cost reduction. The ubiquity of access, in a safe way, to printed documents is a target allowing better productivity and improvement of processes in everyday life, to those who use it in the different buildings. As far as printing is concerned, the decision to be taken should also be aligned with environmental improvement, where the acquisition of new equipment, and therefore, the amount and type of materials to be recycled in the future, is also in alignment with Green IT is to be reduced.



Figure 6 - Innovation

Phase 6 will address the innovative aspects that the PA, in this case and more properly LA, can and should provide its citizens, customers of its organization, with more and better services preferably innovative and strategic, at the distance of a click, thus saving travel costs and loss of time. To this effect, the LA must take the path of digital transformation, adopting the digitization of processes based on a document management system, unique and transversal to the whole organization, that allows integration with online services, providing added value services to citizens.

Referring to innovation, it was stated [18] that the word innovate is widely used when we approach transformation and strategy. This is the path that has been outlined for the municipality under study. Realistic and well-founded transformation should be seen as a future perspective and seen as digital transformation; as such it will be necessary that on the part of organizations and the PA, there is a concerted strategy so that a new, faster, collaborative mode of action among several actors can implement responses to the needs and perspectives of citizens and citizens in general [17]. There is an increasing interest in the form of stimulating innovation in the public sector, being noted the mention [19] of recent research points of collaboration between actors as a factor of superior innovation. In this sense, the public sector or other non-governmental organization consider in their daily agenda the incentive to innovation, by political, economic weight and quality of services. Similarly, there is some agreement [20] with the previous statement and the recognition that the economic weight of these entities leads to the growing political interest in stimulating innovation in PA, which may culminate in a marked improvement and increased efficiency in the use of resources, adding quality of public services and preparing to face the most varied social challenges.

Innovation can be another catalyst for improving services, increasing productivity, and, containing or reducing costs, aligned with Green IT good practices, can bring benefits at various levels.

1.5. Quantitative component of the framework

The qualitative component of the framework aims at evaluating different alternatives. It was understood that alignment with Green IT can be measured through four dimensions [21]:

IT valences: Hardware, software, information systems, security, governance.

Financial valences: Direct and/or indirect costs of IT.

Environmental valences: Environmental impact including CO2 emissions, use of natural resources, contribution to global warming, recycling.

Social valences: satisfaction of the citizen, speed of response, number of trips, time on the move, image of the municipality.

Thus, the Green IT alignment (1) metric includes the IT, financial (Fin), environmental (Env), and social (Soc) dimensions.

$$\text{Green IT alignment} = 0.3\,\text{IT} + 0.2\,\text{Fin} + 0.3\,\text{Env} + 0.2\,\text{Soc} \quad (1)$$

In each dimension, the evaluation is achieved by a set of weighted indicators. The indicators correspond to the valences identified. The direction - positive (+) or negative (-) value - of the indicators will convey if it furthers or hinders the green alignment, respectively. The relative relevance of an indicator is conveyed by a weight (in %).

Greater sustainability, in one case, may be obtained by the virtualization of the infrastructure in the data center of the municipality, which will favor the exchange of end-of-life computers by thin clients. As an example of this component of the framework we will consider two strategies: maintain PCs or replace them by thin clients. It must be noted that some indicators are not applicable to the current case study and are presented but not included in the comparison; a one-year timeframe was considered; costs and prices refer to the time of writing; to illustrate the application, we will instantiate the financial dimension.

Table 1 lists some of the IT indicators that shape operational objectives, especially those aimed at optimizing reliability, service availability, management times, productivity, breakdown reduction, access ubiquity and information security; in short, the relevant IT indicators when in the search for sustainability.

Table 1: IT indicators

Indicators	Direction	Weight factor
reliability	+	20 %
management times	-	5 %
service availability	+	10 %
productivity	+	15 %
quality management	+	5 %
number of malfunctions	-	10 %
information security	+	20 %
ubiquity of access	+	10 %
reduction of file space on digital support	+	5 %

The PA should pay particular attention to its cost management, as on the one hand it strives to remain competitive about the technology at its disposal, on the other hand it should be moderate and contain the costs to keep the budget balanced. Nowadays the IT sector is a strategic asset within each organization and digital transformation assumes significant role in an organization although the study and adoption of technological solutions may raise costs in the annual budget, informed choices reflect good strategic planning preferably aligning with Green IT.

Table 2 illustrates the measurement of some financial indicators demonstrating infrastructure and energy costs, energy consumption and equipment lifetime in the municipality.

Table 2: Financial indicators and instantiated values

Indicators	Direction	PC	Thin client	Weight factor
cost of infrastructure	-	€ 650	€ 250	45 %
equipment lifetime	+	4 Years	7 Years	25 %
energy consumption	-	377 Wh	5 Wh	30 %

After normalizing the financial indicators, one would reach the values -0,61 and + 0,08, for PC and thin client, respectively, indicating that the latter will better approach Green IT.

Society today is already governed by the maximum environmental improvement or preservation of nature, as such, organizations should also be aware that it needs to be properly aligned with the business, but not to be able to improve a good alignment of IT with Green IT, as it will certainly become more expensive to remedy than to prevent environmental problems.

Table 3 presents some environmental indicators which involve CO2 emissions, care for nature preservation, sustainability and recycling, and concerns about global warming. These indicators also demonstrate the existing concern with environmental quality, also aligned with the financial indicators contained in Table 2.

Table 3: Environmental indicators

Indicators	Direction	Weight factor
CO2 emissions	-	50 %
preservation of natural resources	+	20 %
suitability for recycling	+	10 %
global warming	-	20 %

Achieving excellence in the valences of IT, Financial and Environmental, implies, achieving the satisfaction of citizens, improving the image of the municipality, the well-being of employees, increasing the speed of response to processes, avoiding travel to the service services, which contributes to the reduction of costs to the citizens themselves, therefore, a social improvement that should not be bleached. Reducing calls to care also contributes to reducing CO2 emissions, once again in alignment with Green IT and environmental improvement.

The mirrored indicators in Table 4 corroborate the opinion already demonstrated by the authors and are proof of the above, and should be taken into account, as an example of social responsibility in PA, which can also be understood as environmental social responsibility, in a clear allusion to the respect shown by people and human values and simultaneously by environmental causes [21].

Table 4: Social indicators

Indicators	Direction	Weight factor
satisfaction of the citizen	+	10 %
image of the municipality	+	10 %
well-being of employees	-	10 %
time in the circulation of information	-	15 %
speed in the response to the citizen	+	10 %
number of trips to the service	-	10 %
loss of time on travel	-	10 %
travel costs to citizens	-	15 %
service times	-	10 %

Finally, the evaluation of a new process or the global activity be qualitatively resumed as presented in Table 5.

Table 5: Green IT Alignment

Nominal score	Qualitative score
<= 20 %	Very little aligned
>20% ^ <= 45%	Little lined up
>45% ^ <= 75%	Aligned
>75% ^ <= 90%	Very aligned
>90% ^100%	Strongly aligned

This model can be used to support strategic decisions and informed choice on changes such as substituting desktop computers by thin clients or making services available online can be facilitated.

4. PC vs Thin Client in detail

The reasoning between the implementation of PCs or Thin Clients regarding consumption, costs, and quantity of materials to be recycled, is shown in Table 6. The implementation of thin clients allows the reduction of energy consumption, contributing to the reduction of the carbon footprint and environmental improvement. As they are small equipment, they also allow the reduction of the number of materials to be recycled. They simultaneously allow the reduction of management times, which means increased productivity, and licensing costs are also lower, not in any way decorating the quality of service and information security.

Table 6: Comparison PC - Thin Client

	PC	Thin Client
Power supply	600 W	30 W
Average consumption	377 Wh	5 Wh
Weight	7,900 kg	0,240 kg

In the calculation of annual consumption, it was considered that the equipment of individual use is connected 7 hours a day, 5 days a week for 48 weeks. Whereas a PC consumes 377 Wh [22], if connected 7 hours a day you will have an annual consumption of about 633.40 kWh. Likewise, considering that a thin client consumes 5 Wh [23], if you are connected the same 7 hours a day, you will have an annual consumption of about 8.40 kWh. Comparing the values of the two equipment, there is a reduction in energy consumption in the order of 625kWh year. If we use 400 thin clients then we will have a reduction in annual energy consumption of 262 MWh, which at an average cost of € 0.165 per kWh, translates into a cost reduction of around € 41,250.00 per year.

Regarding the number of materials to recycle, it is noted that a PC has about 7,900 kg of material and the thin client has only 0.240 kg, and there is a gain in the reduction of materials to be recycled in the exchange of the PC for thin client. It should also be noted that the lack of treatment of these wastes, through the recycling of equipment, may lead to contamination of soil, water, and air, considering heavy metals and toxic substances, existing in computer components, such as cadmium, lead, bromine, copper, and nickel present in small quantities in equipment and devices, especially boards and that affect both people and animals and plants.

5. Conclusions

It has been assumed that the implementation of ICT measures, which address sustainability concerns, is crucial. Thus, it was possible to identify and quantify options to optimize established practices, thus contributing assertively to sustainability. About the Green IT, this article has several contributions, namely on the presentation of indicators for the reduction of energy consumption, contributing to the reduction of costs and the carbon footprint and consequent environmental improvement. Another aspect of Green IT focuses on reducing the volume of materials to be recycled and its pollutants, considering the replacement of computers by thin clients. In addition to the IT aspect, the financial aspect has also been exemplified, where it has been shown that the alignment of IT with Green IT can also be a boost in reducing costs, both in equipment costs, in increasing the life of the same, by reducing consumption and the number of materials for recycling, obviously aligning with Green IT.

Environmental indicators have shown that it is possible in the PA to contribute to an improvement in environmental quality, without increasing costs and technological developments, in reducing the carbon footprint. Finally, the four strands presented here, it was also possible to demonstrate that the social aspect is no less important and that it is possible in the PA, in this case in LA, to contribute to the satisfaction of citizens, by increasing the speed in response, by reducing the number of visits to services, which in itself end up contributing to the reduction of time losses and the reduction of travel costs , which eventually contribute to the improvement of environmental quality, increased confidence in THE and the improvement of the image of the municipality, still allied to the improvement of workers' working conditions, all aligned with Green IT.

The main contribution of this work is the example of a model to focus efforts to optimize ICT management practices that integrate environmental improvement and quality in a sustained manner. The main conclusions of the exercise presented here illustrate how the model increases productivity by reducing management times and containing or reducing costs in ICT. Based on the proposed framework and as a possible line of research to be developed, it is considered that the inclusion of weights in the various stages of the cycle allows their use in various real environments, such as local authorities. An aggregating instrument of all the information inherent in the framework may be another line of future research and may culminate in its development and improvement.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. Landum, M.M.M. Moura, L. Reis, 'ICT Good Practices in alignment with Green IT', in Iberian Conference on Information Systems and Technologies, CISTI, IEEE, Sevilha, 2020, doi:10.23919/CISTI49556.2020.9141166.

[2] S.L. Hart, 'Beyond greening: strategies for a sustainable world', Harvard Business Review, **75**(1), 66–77, 1997.

[3] M. Boudreau, Marie-Claude and Chen, Adela and Huber, 'Green IS: Building Sustainable Business Practices', Information Systems: A Global Text, 1--7, 2008.

[4] S. Roth, H.F. Dahms, F. Welz, S. Cattacin, 'Print theories of computer societies. Introduction to the digital transformation of social theory', Technological Forecasting and Social Change, **149**, 119778, 2019, doi:10.1016/j.techfore.2019.119778.

[5] L. Moreira, 'A quarta revolução industrial no setor metalomecânico português', 2019.

[6] J.S. Guy, 'Digital technology, digital culture and the metric/nonmetric distinction', Technological Forecasting and Social Change, **145**, 55–61, 2019, doi:10.1016/j.techfore.2019.05.005.

[7] M.F. Bari, R. Boutaba, R. Esteves, L.Z. Granville, M. Podlesny, M.G. Rabbani, Q. Zhang, M.F. Zhani, 'Data Center Network Virtualization: A Survey', IEEE Communications Surveys & Tutorials, **15**(2), 909–928, 2013, doi:10.1109/SURV.2012.090512.00043.

[8] A. Satpathy, S.K. Addya, A.K. Turuk, B. Majhi, G. Sahoo, 'Crow search based virtual machine placement strategy in cloud data centers with live migration', Computers & Electrical Engineering, **69**, 334–350, 2018, doi:10.1016/j.compeleceng.2017.12.032.

[9] M. Noshy, A. Ibrahim, H.A. Ali, 'Optimization of live virtual machine migration in cloud computing: A survey and future directions', Journal of Network and Computer Applications, **110**, 1–10, 2018, doi:10.1016/j.jnca.2018.03.002.

[10] M. Landum, Cloud na administração local: estudo de caso, Escola Superior de Ciências Empresariais, 2013.

[11] E. Coral, 'Modelo de planejamento estratégico para a sustentabilidade empresarial', Universidade Federal de Santa Catarina, 282, 2002.

[12] A. Leoneti, A. Nirazawa, S. Oliveira, 'Proposta de índice de sustentabilidade como instrumento de autoavaliação para micro e pequenas empresas (MPEs)', REGE - Revista de Gestão, **23**(4), 349–361, 2016, doi:10.1016/j.rege.2016.09.003.

[13] B. Haley, 'Designing the public sector to promote sustainability transitions: Institutional principles and a case study of ARPA-E', Environmental Innovation and Societal Transitions, **25**, 107–121, 2017, doi:10.1016/j.eist.2017.01.002.

[14] I. Figueira, A.R. Domingues, S. Caeiro, M. Painho, P. Antunes, R. Santos, N. Videira, R.M. Walker, D. Huisingh, T.B. Ramos, 'Sustainability policies and practices in public sector organisations: The case of the Portuguese Central Public Administration', Journal of Cleaner Production, **202**, 616–630, 2018, doi:10.1016/j.jclepro.2018.07.244.

[15] L. Nogueiro, T.B. Ramos, 'The integration of environmental practices and tools in the Portuguese local public administration', Journal of Cleaner Production, **76**, 20–31, 2014, doi:10.1016/j.jclepro.2014.03.096.

[16] J.H. Spangenberg, 'Hot Air or Comprehensive Progress? A Critical Assessment of the SDGs', Sustainable Development, **25**(4), 311–321, 2017, doi:10.1002/sd.1657.

[17] M.A. dos S. Landum, L. Reis, M.M.M. Moura, Concept of Approach to Optimize ICT Management Practices, 174–199, 2021, doi:10.4018/978-1-7998-4099-2.ch009.

[18] M. Camacho, 'Christian Bason: Design for Public Service', She Ji: The Journal of Design, Economics, and Innovation, **2**(3), 256–268, 2016, doi:10.1016/j.sheji.2017.02.002.

[19] J. Torfing, 'Collaborative innovation in the public sector: the argument', Public Management Review, **21**(1), 1–11, 2019, doi:10.1080/14719037.2018.1430248.

[20] A. Arundel, C. Bloch, B. Ferguson, 'Advancing innovation in the public sector: Aligning innovation measurement with policy goals', Research Policy, **48**(3), 789–798, 2019, doi:10.1016/j.respol.2018.12.001.

[21] M. Landum, M.M.M. Moura, L. Reis, 'Evaluation of Green IT and Local Administration', ITEMA 2020 Conference Proceedings (pre-draft), Serbia, 2020.

[22] Extreme, Power Supply Calculator - PSU Calculator | OuterVision, Feb. 2021.

[23] DELL, Wyse 3040 Thin Client for Virtual Desktop Experience | Dell USA, Feb. 2021.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 601-606 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

CNN-LSTM Based Model for ECG Arrhythmias and Myocardial Infarction Classification

Lana Abdulrazaq Abdullah[*, 1, 2, 3], Muzhir Shaban Al-Ani[1]

[1]*Department of Information Technology, College of Science and Technology, University of Human Development, Sulaymaniyah,46001, Iraq*

[2]*Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah, KRG, 46001, Iraq*

[3]*Department of Applied Computer, Health and Applied Science, Charmo University, Sulaymaniyah, KRG, 46001, Iraq*

ARTICLE INFO

Article history:
Received: 11 July, 2020
Accepted: 15 September, 2020
Online: 05 October, 2020

Keywords:
ECG classification arrhythmia
Convolutional neural network
Long Short Term Memory
Myocardial Infarction

ABSTRACT

ECG analysis is commonly used by medical practitioners and cardiologists for monitoring cardiac health. A high-performance automatic ECG classification system is a challenging area because there is difficulty in detecting and clustering various waveforms in the signal, especially in the manual analysis of electrocardiogram (ECG) signals. In this paper, an accurate (ECG) classification and monitoring system are proposed using the implementation of 1D Convolutional Neural Networks (CNNs) and Long Short Term Memory (LSTM). The learned features are captured from the CNN model, and then fed to the LSTM model. No handcraft features are required for the model for the ECG classification. The result of the CNN-LSTM model has demonstrated superior performance than several state-of-the-arts that cited in the result section. The proposed models are evaluated on MIT-BIH arrhythmia and PTB Diagnostics datasets. Based on the obtained results, the CNN-LSTM method can improve the accuracy rate, such that 98.1 % and 98.66 % on Myocardial Infarction (MI) and arrhythmia classification, respectively.

1. Introduction

One of the most crucial organs in the human body is the heart, and monitoring of the heart has become a necessary diagnosis to the health of the human being. The electrical activity of the heart is well known as ECG, and Abnormality in ECG may cause symptoms of heart disease, such as persistent ventricular tachycardia, low blood pressure, rapid atrial fibrillation, and persistent ventricular tachycardia. These mentioned diseases are dangerous for human life and desire an urgent treatment [1]. The most significant information about the state of the heart can be observed from the ECG analysis, which can be monitored manually and automatically [2]. Due to the fact that there are various morphologies in the ECG signal, manual diagnosis is found difficult. Therefore, an automatic system of the ECG diagnosis has been interested in. Both feature extraction and classification approaches are the critical successes of any automatic ECG classification. There are plenty of applications that ECG analysis and classification involve, such as ischemic heart disease, arrhythmia, and myocardial infarction [3].

Feature extraction process is a viral area in ECG classification, and many approaches have been developed for this purpose as the feature has a significant impact on any machine learning system. Nowadays, there are numerous approaches for extracting features from ECG signals such as Discrete Wavelet Transform (DWT), Wavelet Transform (WT), and Mel Frequency Coefficient Cepstrum (MFCC).The DWT has been widely adopted in ECG classification as an effective feature, for instance, Desai at el. in [4], proposed system-based approach for computer-assisted detection of five classes of ECG arrhythmia beats by adopting DWT as a feature to train Support Vector Machine (SVM). In [5] developed a comprehensive model based on random forest techniques and discrete wavelet for arrhythmia classification. The authors in [6]used the DWT to capture features from the ECG signal based on ST-segment elevation and inverted T wave logic.

The MFCC was used as a handcrafted, which was extracted from the ECG signal directly, then the artificial neural network (ANN) was fed by the MFCC features [7]. The authors in [8] developed a hybrid features, which was merged MFCC and DWT after that the proposed feature was classified by k Nearest

[*]Corresponding Author: Lana Abdulrazaq Abullah, Sulaymaniyah, KRG, Iraq: lana.abdulrazaq@gmail.com

Neighbor (kNN). In addition, The WT was conducted to take features out from ECG signal for classifying four cardiac types such as healthy (N), Paced beats (P), right bundle branch block (RBBB), and left bundle branch block (LBBB). Then ANN and SVM were used to classify the extracted features [9].

The handcrafted feature is not always the best candidate feature for the ECG application, as it is a time-consuming and boring task [10]. Therefore, a featureless model like CNN has been interested in classifying the ECG analysis. CNN has been conducted into two forms, such as one dimensional and two-dimensional forms, for instance, Zubair at el. in [11] proposed the raw data fed a model-based 1D-CNN for classifying ECG signals into some classes that were suggested by the Association for Advancement of Medical Instrumentation (AAMI) and the model. Wang in [12] developed a model based on 1D CNN and modified Elman neural network (MENN),which was consisted of the 11-layers neural network, and the non-transform ECG signal also fed the model. While some researchers proposed their CNN model for ECG analysis based on two-dimensional forms, for instance, Zheng at el in [13] converted the one dimensional ECG signal into two-dimensional images and then fed to the CNN model. Yildirim at el. developed a model for detecting a diabetic subject by adopting pre-trained 2D-CNN model with frequency spectrum images, which were obtained from heartbeat signals [14].

In this paper, a hybrid model, which is composed by CNN and LSTM, is developed. No handcrafted feature requires the model as the CNN part is responsible for obtaining the feature. The learned features are given to the LSTM model to learn a deep pattern in the learned feature and classify into healthy and abnormal ECG signals, where the health ECG signal is chosen as the absolute normal ECG pattern template. According to our experimental result, LSTM can improve the classification rate, as shown in the result section. The rest of the paper is organized as follows: Section 2 explained a concise background area of the proposed model, Section 3 methodology. Results and discussion are addressed in Section 4. Finally, the conclusion section is given in Section 5.

2. Background

A representation of the cardiac activity of the heart is known as the ECG signal. The automatic monitoring system of the ECG signal has been focused area as it is related to the human health condition. Feature extraction and classification are the primary processes in any automatic monitoring system. In this paper, the suggested model is a fusion model, which consists of CNN and the LSTM model, as shown in Figure 1. A brief background of CNN and LSTM is addressed in the below section as CNN is used to extract the learned features, and LSTM is conducted to classify the learned features.

2.1 Convolution Neural Networks

The CNNs also known as the feature learner has excellent potential to extract useful features automatically from input raw data. The CNNs are normally comprised of two parts and each part has different function. Feature extractor is the first part of the CNNs and which is normally responsible to extract features automatically. The second part is classifier which is called a fully connected multi-layer perceptron (MLP). The fully connected part performs classification depend on the learned features from the first part. The feature extraction's part includes some layers namely; convolution layer and pooling layer. Any convolutional layer can take feature maps out from the previous layer. convolutional layer contains several convolution kernels (filters), added by bias, and then put throughout the activation function to generate a feature map for the next layer [15].

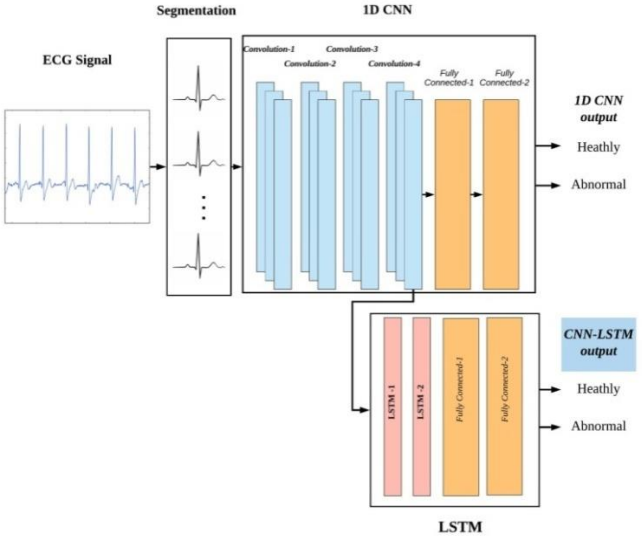


Figure 1: proposed CNN_LSTM scheme.

One of the common sub layers in feature extraction's part is called pooling layer. There are some types of the pooling layer which are Max, Min and Average pooling. Pooling layer can shrink the resolution of the feature maps. Max-pooling operation is used in the proposed model, which calculates the maximum value in a set of nearby inputs [11].There is difference between the both parts in term of complexity. The feature extraction's part, which is primary layer, performs more computations including feature extraction and feature selection compare to the fully connection layers [16].

Many similarities can be seen between the CNN structure the ANN with an input, hidden, and output layers. On the other hand, unlike NN, CNN is an enhanced version of the ANN, which is both translational and shift invariant. Typically, the CNNs comprise of various types of layers such as input, convolution, max pooling, average pooling, drop layer, and Softmax layer, etc. and play a specific role in the model [17].

2.2 Long Short-Term Memory

RNN is one of the types of network architectures, which is specifically designed to deal with the sequential problem, and it is commonly used in sequence classification. A well-known improvement of RNN was introduced by Hochreiter and Schmidhuber in 1997 [18], which is called Long Short Term Memory (LSTM). Plenty of applications of LSTM has been published recently [19]–[23]. The LSTM unit comprises a cell; a forget gate, an output gate, and an input gate[19]. The cell unit is in charge of remembering values at every time interval. The rest unit gates manage the flow of information into and out of the unit. In the memory block structure, a simple one-layer neural network

controls the forget gate. The functionality of this gate is formulated as Eq (1) [24].

$$f_t = \delta(w[x_t, h_{t-1}, c_{t-1}] + b_f) \quad (1)$$

where, x_t is the sequence of input; c_{t-1} is the previous LSTM block memory; h_{t-1} is the earlier block output; b_f is the bias vector; δ is the logistic sigmoid function, and W symbolizes separate weight vectors for every input. An input gate is a unit where the new memory is formed by the previous memory block effect and a simple NN with an activation function which is tanh. These operations are calculated by Eq (2) and (3)[25].

$$i_t = \delta(w[x_t, h_{t-1}, c_{t-1}] + b_i) \quad (2)$$

$$c_t = f_t . c_{t-1} + i_t . \tanh(w[x_t, h_{t-1}, c_{t-1}] + b_i) \quad (3)$$

The output gate is an output of the current LSTM block and can be formulated using Eqs (4) and (5) [19].

$$\delta_t = \delta(w[x_t, h_{t-1}, c_t] + b_0) \quad (4)$$

$$h_t = \delta_t . \tanh(c_t) \quad (5)$$

Where, b_i and b_i are the outputs of previous memory block;These mentioned units are linked to each other, as illustrated in Figure 2, which permit information to cycle between steps of adjacent time and also construct an inner feedback state which the network to the temporal feature in the given data [26].

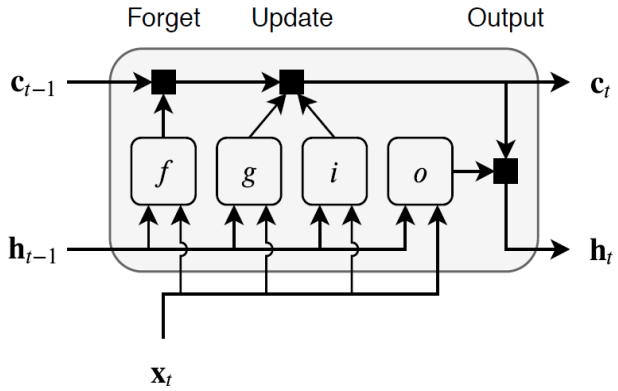


Figure 2: Connection LSTM units, where (i) is Input gate, (g) is Cell candidate, (f) is Forget gate, and (o) is Output gate [26].

3. Methodology

This section is divided into; data description, preprocessing, and experimental design.

3.1 Data description

3.1.1 MIT-BIH Arrhythmia Database

In this study, one of the examined database is a gathering of annotated ECG recordings that achieved by the Arrhythmia Laboratory of Boston's Beth Israel Hospital, known as the MIT-BIH Arrhythmia Database [27]. It comprises 48 half-hour excerpts of two-channel ambulatory ECG recordings, which were obtained from 47 subjects, and the dataset was recorded in the BIH Arrhythmia Laboratory between 1975 and 1979. The recordings are obtained with a frequency of 360 Hz; this means each record approximately consists of 30 *360 = 10800 samples and contains recordings from two leads: one is ML2, and the other is one of $V1, V2, V3, V4, V5,$ or $V6$. Two or more cardiologists independently annotated each record for a total of 110,000 annotations covered with the database: it is worth to mention that each annotation is placed in correspondence to the R peak of a single beat so that we have totally solved for this particular case the beat discovery problem. The MIT-BIH Arrhythmia database is split into two classes healthy and abnormal.

3.1.2 The Physionet PTB Diagnostic ECG database

The second examined database, which is used in this paper, was recorded by the Department of Cardiology at Benjamin Franklin University in Berlin, Germany. The dataset consists of 290 subjects and include 549 records (aged 17 to 87, mean 57.2; and 81 women, mean age 61.6 and 209 men, mean age 55.5). One to five records were observed from each subject. The dataset was split into two sections namely, healthy signals and different cardiac diseases such as myocardial infarction, bundle branch block, arrhythmia, cardiomyopathy/heart failure, myocardial hypertrophy, etc. Each record includes 15 simultaneously measured signals: the conventional 12 leads $(i, ii, iii, avr, avl, avf, V1, V2, V3, V4, V5, V6)$ together with the 3 Frank lead ECGs (vx, vy, vz)[8].

3.2 Preprocessing

In the MIT-BIH database, the beats are labeled based on the R position. The authors segment each record sample intoa separated beat and take samples from both sides of R waves. Each heartbeat consists of 130 sample points (65 samples before the R-peak and 64 samples after the peak), which is the smallest length ever been used for ECG classification. As showed in Figure. 3, the balanced dataset was created; the number of healthy beats is 86456 and 11230 beats of the abnormal beat.

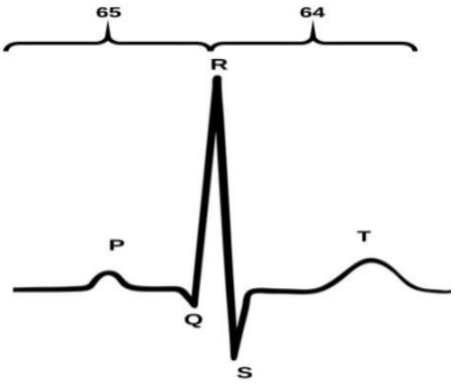


Figure 3: ECG signal.

Feature from the raw signal and eliminate noise from the ECG signals. The LSTM section can deal with the characteristics of ECG related to time series properly and able to predict the future QRS complexes using the previous QRS complexes. The 1D CNN structure contains 19 layers, including four convolution layers, two Dropout layers with half percent drop, and two fully-connected layers.

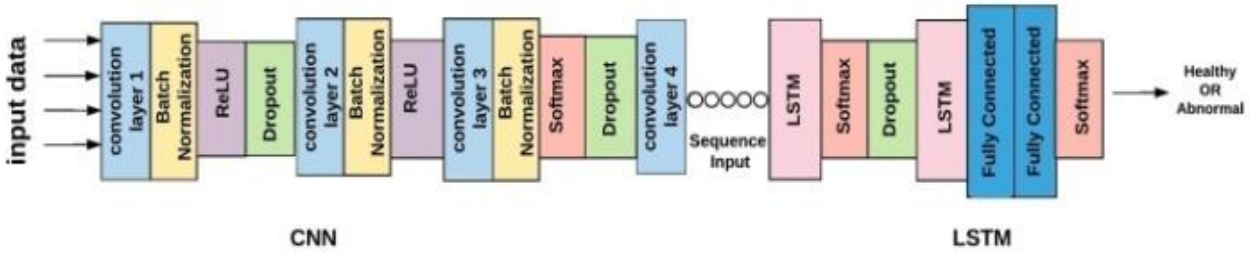


Figure 4: Structure of CNN-LSTM model.

Some parameters of the model such as number of convolution layer, filters and epochs are tuned based on cross validation methods. The description of the CNN-layers,are set out in Table 1.The spatial and local feature map is well extracted from the thirteenth layers of the CNN model, which is the convolutional layer with ten filters. The filters act to identify different features present in an ECG signal like edges, vertical and horizontal lines. Then, the features are fed to the LSTM model. For MIT-BIH Arrhythmia Dataset, the length of the learned features is ten times of the original signal thus the length of the feature is equal to 1300.Then; the learned feature is divided into ten separated beats and later merged them into a cell to be input to the LSTM model. In terms of the PTB Diagnostic ECG Dataset, the length of the ECG signal and the length learned features is 187.

Table 1: Description of 1D CNN model.

No. layer	Name	Description
1.	Input	Depend on the input with 'zero center' normalization
2.	Convolution Layer 1	20 7x7 convolutions with stride [1 x 1] and padding “same”
3.	Batch Normalization1	Batch normalization
4.	Clipped ReLU	Clipped ReLU with ceiling 5
5.	Dropout	50% dropout
6.	Convolution Layer 2	20 9x9 convolutions with stride [1 x 1] and padding “same”
7.	Batch Normalization2	Batch normalization 2
8.	Leaky ReLU	Leaky ReLU with scale 0.01
9.	Convolution Layer 3	30 5x5 convolutions with stride [1 x 1] and padding “same”
10.	Batch Normalization3	Batch normalization
11.	Softmax	Softmax
12.	Dropout	50% dropout
13.	Convolution Layer 4	10 3x3 convolutions with stride [1 x 1] and padding “same”
14.	Batch Normalization4	Batch normalization
15.	ReLU	ReLU
16.	Fully Connected1	60 fully connected layer
17.	Fully Connected2	Two fully connected layer
18.	Softmax	Activation function for the last Fully Connected layer
19.	Classification	Output cross-entropy

3.3 Experimental design

In the current study, a cross-learning method based on deep learning is proposed for automatically classifying healthy and abnormal of the ECG signal. The proposed model is developed by integrating CNN and LSTM, which consists of 19 layers,as illustrated in Figure 4,where the CNN section is responsible for capturing the locally

The implemented LSTM structure contains two lstmlayers, dropout layers, and two fully-Connected layers, as it is shown in Table 2. The output of the proposed model is two classes, which are healthy and abnormal Arrhythmia for MIT-BIH Arrhythmia Database and healthy and Myocardial Infarction for the PTB Diagnostic ECG Database.

Table 2: Description of LSTM model.

No. layer	Name	Description
1.	Input	Sequence input with input dimensions
2.	LSTM Layer 1	LSTM with 120 hidden units
3.	Softmax	Softmax
4.	Dropout	30% dropout

5.	LSTM Layer 2	LSTM with 120 hidden units
6.	Fully Connected1	Ten fully connected layer
7.	Fully Connected2	Two fully connected layer
8.	Softmax	Softmax
9.	Classification	Output cross-entropy

4. Result and discussion

2.3 MIT-BIH arrhythmia database

Nowadays, machine learning has been adopted for many applications, such as arrhythmia classification. One of the well-known datasets is the MIT-BIH arrhythmia dataset that has been adopted by a mounting number of researchers in ECG research. Table 3 shows the result of the proposed methods and some state of arts who have studied a model for detecting ECG arrhythmias. Inspire of these state of arts, ten-fold cross-validation is utilized to evaluate the proposed model. Although the proposed CNN-LSTM model is fed by the shortest length of the ECG signal ever, the result demonstrates higher classification accuracy than other related studies. However, the result of 1D CNN does not exceed the state of arts. Moreover, the performance of our proposed model exceed the performance of a CNN-LSTM model which was conducted by [28]. However, the structure of their model (11 layers) is simplest than our model (19 layers).

Table 3: Result of arrhythmia classification.

Study	Number of Classes	Classifier	Signal Length	Accuracy (%)
Andersen et al. 2019 [25]	2	CNN +RNN	360 , 250	97.80%
Wang 2020[12]	2	CNN+MENN	360	97.40%
Lih Oh[28]	5	CNN+LSTM	1000	98.48%,
Proposed method	2	CNN	130 sample	94.70%
	2	CNN+LSTM	130 sample	**98.66%**

2.4 MI Classification

Another well-known datasets are the PTB Diagnostics dataset that is used to evaluate the proposed methods. MI classification is an application of using the PTB Diagnostics dataset, and the MI can be treated as two classes' problem with infracted and non-infracted classes. The length of the signal 14552 samples and divided by 4046 infracted and 10506 non-infracted. Based on relevant researches in the literature, ten-fold cross-validation is adopted to weigh the proposed method. The experimental result in Table 4 shows a significant improvement in the classification accuracy. Moreover, the percentage of precision and recall are improved compared to the existing approaches available in the literature.

Table 4: MI classification result.

Study	Percentage (%)		
	Accuracy	Precision	Recall
Kachuee et al. [29]	95.90	95.20	95.10
Achary el al. [30]	93.5	92.80	93.70
Safdarian et al.[31]	94.70	-	-
Scratched 1D-CNN proposed	97.55	96.00	97.51
CNN-LSTM proposed	98.13	96.80	98.00

5. Conclusion

Detecting heart abnormalities using ECG signals is very challenging area. An accurate model to detect the abnormal ECG signals leads to offer correct treatment to the patients in the earlier stage of disease. In this paper, an effective arrhythmias and myocardial classification method that combines 1-D CNN and LSTM model. The proposed CNN-LSTM model is applied to the computerized recognition of abnormal ECG signals. According to our experimental result, the learned features from CNN can be useful features for a time series approach like LSTM.

Moreover, the LSTM can find out a better pattern in the learned feature compare to the fully-connection layer itself. Consequently, the classification result by LSTM has outperformed the states of arts. The authors recommend to use the proposed model to classify multiclass of arrhythmias and myocardial instead of binary class.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank assistant professor Mr, Zrar Kh Abdul for his support during the preparation of this work. Many thanks for the reviewers for their valuable comments and suggestions on the manuscript which improve the presentation of the paper.

References

[1] E. J. Benjamin et al., Heart Disease and Stroke Statistics—2017 Update, **135**(10), 2017.
[2] S. Saechia, J. Koseeyaporn, and P. Wardkein, "Human identification system based ECG signal," IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON, 2007, 2–5, 2005, doi: 10.1109/TENCON.2005.300986.
[3] M. A. Serhani, H. T. El Kassabi, H. Ismail, and A. N. Navaz, "ECG Monitoring Systems: Review, Architecture, Processes, and Key Challenges," Sensors, 2020, doi: 10.3390/s20061796.
[4] U. Desai, R. J. Martis, C. G. Nayak, K. Sarika, and G. Seshikala, "Machine

intelligent diagnosis of ECG for arrhythmia classification using DWT, ICA and SVM techniques," 12th IEEE Int. Conf. Electron. Energy, Environ. Commun. Comput. Control (E3-C3), INDICON **2015**, 2–5, 2016, doi: 10.1109/INDICON.2015.7443220.

[5] G. Pan, Z. Xin, S. Shi, and D. Jin, "Arrhythmia classification based on wavelet transformation and random forests," Multimed. Tools Appl., **77**(17), 21905–21922, 2018, doi: 10.1007/s11042-017-5225-5.

[6] P. Kora, C. U. Kumari, K. Swaraja, and K. Meenakshi, "Atrial Fibrillation detection using Discrete Wavelet Transform," in Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019, 1–3, 2019. DOI: 10.1109/ICECCT.2019.8869498.

[7] M. Boussaa, I. Atouf, M. Atibi, and A. Bennis, "ECG signals classification using MFCC coefficients and ANN classifier," Proc. 2016 Int. Conf. Electr. Inf. Technol. ICEIT, 2016, 480–484, 2016. DOI: 10.1109/EITech.2016.7519646.

[8] S. A. A. Yusuf and R. Hidayat, "MFCC feature extraction and KNN classification in ECG signals," 2019 6th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE, 2019, 1–5, 2019. DOI: 10.1109/ICITACEE.2019.8904285.

[9] S. Sahoo, B. Kanungo, S. Behera, and S. Sabut, "Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities," 55–66, 2017. DOI: 10.1016/j.measurement.2017.05.022.

[10] Ö. Yıldırım, U. B. Baloglu, and U. R. Acharya, "A deep convolutional neural network model for automated identification of abnormal EEG signals," Neural Comput. Appl., 10–20, 2018. doi: 10.1007/s00521-018-3889-z.

[11] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," in 2016 6th International Conference on IT Convergence and Security, ICITCS 2016. doi: 10.1109/ICITCS.2016.7740310.

[12] J. Wang, "A deep learning approach for atrial fibrillation signals classification based on convolutional and modified Elman neural network," Futur. Gener. Comput. Syst., **102**, 670–679, 2020, doi: 10.1016/j.future.2019.09.012.

[13] Z. Zheng, Z. Chen, F. Hu, J. Zhu, Q. Tang, and Y. Liang, "An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology," Electron., **9**(1), 1–15, 2020, doi: 10.3390/electronics9010121.

[14] O. Yildirim, M. Talo, B. Ay, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals," Comput. Biol. Med., **113**, 103387, Oct. 2019, doi: 10.1016/j.compbiomed.2019.103387.

[15] W. Yin, X. Yang, L. Zhang, and E. Oki, "ECG Monitoring System Integrated with IR-UWB Radar Based on CNN," IEEE Access, **4**, 6344–6351, 2016, doi: 10.1109/ACCESS.2016.2608777.

[16] J. Li, Y. Si, T. Xu, and S. Jiang, "Deep Convolutional Neural Network Based ECG Classification System Using Information Fusion and One-Hot Encoding Techniques," Math. Probl. Eng., 2018. doi: 10.1155/2018/7354081.

[17] U. R. Acharya et al., "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network," Futur. Gener. Comput. Syst., **79**, 952–959, 2018. doi: 10.1016/j.future.2017.08.039.

[18] S. Hochreiter and J. Urgen Schmidhuber, "Long Shortterm Memory," Neural Comput., **9**(8), p. 17351780, 1997.

[19] C. Chen, Z. Hua, R. Zhang, G. Liu, and W. Wen, "Automated arrhythmia classification based on a combination network of CNN and LSTM," Biomed. Signal Process. Control, **57**, 101819, 2020, doi: 10.1016/j.bspc.2019.101819.

[20] X. H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of Long Short-Term Memory (LSTM) neural network for flood forecasting," Water (Switzerland), **11**(7), 2019, doi: 10.3390/w11071387.

[21] E. Balouji, I. Y. H. Gu, M. H. J. Bollen, A. Bagheri, and M. Nazari, "A LSTM-based deep learning method with application to voltage dip classification," Proc. Int. Conf. Harmon. Qual. Power, ICHQP, 2018-May, 1–5, 2018, doi: 10.1109/ICHQP.2018.8378893.

[22] F. Weninger et al., "Speech enhancement with LSTM recurrent neural networks and its application to noise-robust ASR," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 9237, 91–99, 2015, doi: 10.1007/978-3-319-22482-4_11.

[23] F. Liu et al., "An Attention-based Hybrid LSTM-CNN Model for Arrhythmias Classification," Proc. Int. Jt. Conf. Neural Networks, 2019-July, 1–8, 2019, doi: 10.1109/IJCNN.2019.8852037.

[24] Ö. Yildirim, "A novel wavelet sequences based on deep bidirectional LSTM network model for ECG signal classification," Comput. Biol. Med., **96**(March), 189–202, 2018, doi: 10.1016/j.compbiomed.2018.03.016.

[25] R. S. Andersen, A. Peimankar, and S. Puthusserypady, "A deep learning approach for real-time detection of atrial fibrillation," Expert Syst. Appl., **115**, 465–473, 2019, doi: 10.1016/j.eswa.2018.08.011.

[26] J. Cheng, L. Dong, and M. Lapata, "Long short-term memory-networks for machine reading," EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc., 551–561, 2016.

[27] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," IEEE Engineering in Medicine and Biology Magazine, **20**(3), 45–50, 2001, doi: 10.1109/51.932724.

[28] S. L. Oh, E. Y. K. Ng, R. S. Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," Comput. Biol. Med., **102**, 278–287, 2018, doi: 10.1016/j.compbiomed.2018.06.002.

[29] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG heartbeat classification: A deep transferable representation," Proc. - 2018 IEEE Int. Conf. Healthc. Informatics, ICHI 2018, 443–444, 2018, doi: 10.1109/ICHI.2018.00092.

[30] U. R. Acharya et al., "A deep convolutional neural network model to classify heartbeats," Comput. Biol. Med., **89**, 389–396, 2017, doi: 10.1016/j.compbiomed.2017.08.022.

[31] N. Safdarian, N. J. Dabanloo, and G. Attarodi, "A New Pattern Recognition Method for Detection and Localization of Myocardial Infarction Using T-Wave Integral and Total Integral as Extracted Features from One Cycle of ECG Signal," J. Biomed. Sci. Eng., **07**(10), 818–824, 2014, doi: 10.4236/jbise.2014.710081.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 616-621 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Improved Approach for QoS Based Web Services Selection Using Clustering

Mourad Fariss*, Naoufal El Allali, Hakima Asaidi, Mohamed Bellouki

MASI Laboratory, FPN, UMP, Nador, 62000, Morocco

ARTICLE INFO

Article history:
Received: 22 December, 2020
Accepted: 11 March, 2021
Online: 20 March, 2021

Keywords:
Web service selection
K-Means clustering
Prefiltering
Skyline techniques
BBS algorithm

ABSTRACT

With the rising number of web services created to build complex business processes, selecting the appropriate web service from a large number of web services respond to the same client request with the same functionality are developed independently but with different quality of service (QoS) attributes. From this point, there are many approaches to web service selection. Nevertheless, this is still deficient due to a considerable number of discovered web services. The prefiltering is a solution to reduce the number of web services candidates. In this paper, the K-means clustering is applied to determine similar services based on QoS information. The results of this prefiltering are considered at the selection task using the Branch and Bound Skyline (BBS) algorithm. The experimental evaluation performed on real Dataset proves that our approach presents efficient results for web service selection.

1. Introduction

This paper is an extension of [1], where an advanced mechanism of prefiltering and selection of web services based on QoS is proposed.

Over the past decade, many researchers have developed a strong interest in web services, an important standard of Service Oriented Architecture (SOA). It is a novel paradigm to build the large-scale of distributed applications. Web service is defined as a software-system and identified using an URI, where its public interface and binding description use the XML language, can be discovered and invoked by other web services. This invocation requires a prescribed of resources using XML messages via such protocols of the Internet.

WSDL, SOAP, and UDDI are the series of technology criteria for web services [2], on which other technologies closer to the application problem can be specified and implemented. It presents standard web service protocols to implement /develop the interaction between applications (services) among diverse platforms. The web service architectures are based on the following three entities; (i) service provider, (ii) service registry, and (iii) service customer. The service provider corresponds to the proprietor of the service. It is required to depict the web service and publish it in the service registry (a central entity). The service registry possesses the technical details of web service and the service provider information to facilitate and find services for customers. The customer is the application that is going to search for and invoke a service. The client application can itself be a web service.

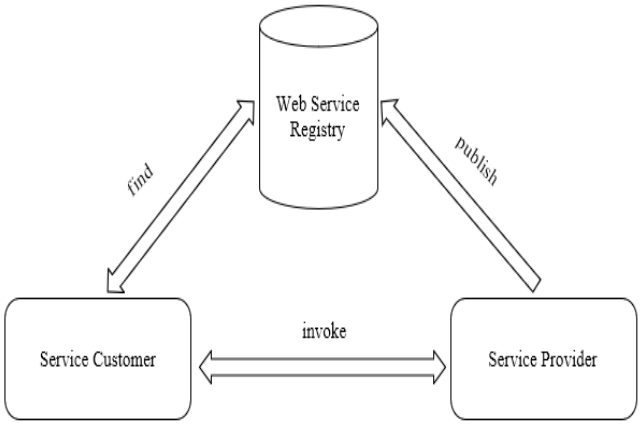


Figure 1: Web Service Model

The increased web applications usage for different fields, making service providers to respond to customers by releasing an enormous number of web services ; the customer finds a problem in choosing the web service that meets his request with this large number of published web services. QoS appears as a solution to help customers select an adequate web service that meets the

*Corresponding Author: Mourad Fariss, m.fariss@ump.ac.ma

customers' requirements. These requirements seek to benefit from more web service performances as cost, response-time, and other QoS properties [3].

One of the immense challenges of SOA is to attribute QoS to the description of web services to facilitate the choice for customers according to their requirements as well as to dynamically select the most efficient web services for each customer according to the criteria of the requirements given by the client or just the best web service among the web services found in the web services registry that have similar functionalities.

The rest of the paper is presented as follows: Section 2 gives the related works. Section 3 explains the background used in this paper as the QoS in web service, the K-Means clustering and the BBS algorithm. Section 4 contains our proposed approach. Section 5 discusses the experimental results. The final section concludes the paper.

2. Related works and motivation

The selection of web service is a hard process because various web services offer similar features. Applications in their consumption of services struggle to employ the optimal QoS; however, the selection phase faces many hardships since the QoS is at the same time influences by several inconsistent QoS features. Present solutions have a shortage in performance for the reason that they take in consideration the potential web services to find QoS features. In case, customer needs are taking place in the selection process. We can use this bit of information in order to distinguish between web services that possess similar QoS as end-user QoS features. For the sake of gathering similar web services together, it is useful to use cluster technology with reference to QoS properties. The selection considers only web services, which is abided by the customer's QoS requirements.

Recently, to solve the selection web service problem, the skyline algorithms have been introduced by selecting service as the optimal candidate services [4–9]. The BBS approach is the most famous skyline algorithm suggested by [10]. As far as the large data spaces are concerned, it is the most efficacious algorithm.

The authors of [11] compared two algorithms: the BBS and the SFS algorithms on the service web selection. The service selection system was performed efficaciously and reliably by the BBS algorithm as the experimental outcomes show.

Authors in [12] were the first to propose the filtering of web services system named "*F-WebS system"*. The system builds the performance on the description and discovery area of web services [11,13–15] as semantics-based web service filtering and utilizes a variety of matching algorithms such as those in the discovery task.

In [7], the authors proposed a framework named "*KRSWS*" to reduce the web services candidates based on QoS attributes and the customer requirements, the proposed method use the Fuzzy AHP method and a new version of Promethee [16].

The clustering mechanism for generating services clusters according to the same QoS is useful for determining the relevant web service [17]. Nonetheless, to cluster the generic type of QoS properties in one group can negatively effect the efficiency of the selection of web service [18].

The authors of [19] proposed a cluster and filtering system architecture model, and have demonstrated that the use of the clustering technique does not effect on the system's accuracy and pertinence, yet it increases the speed of the process of simple service processing.

The proposed method for selecting web services uses a cluster approach based on QoS parameters to pre-filter web services, once we filter the web services with a pre-filter based on K-Means clustering, we obtain a decreased set that contains the web services filtered, after that we select the dominate web service with the skyline technique. The proposed solution presents a significant precision and performance of selecting web services regarding other approaches cited in the literature.

3. Preliminaries

3.1. QoS in Web Service

Several works that have been conducted in web service discovery focus only on the functional features (content requirements) of a web service, but this phase remains insufficient to meet the customer's requirements because of the many similar web services that offer the same functionalities for the customer. However, the web service selection phase introduces the notion of non-functional features (context requirements), namely QoS, to determine the most efficient web service that meets customer requirements.

QoS is a set of features and characteristics of an entity or a service that gives it the ability to meet stated or implicit needs. The needs can be linked to parameters such as accessibility, availability, response time, reliability, cost, etc. The parameters can help to select from the candidate web services and reduce the consumed time.

We can distinguish between customer-independent QoS attributes such as (cost, reputation, accuracy) and customer-dependent QoS attributes such as (throughput, scalability, availability) [2], [3].

QoS has the capacity to satisfy its significance by:

- Defining the operational measurements for the web service.
- Distinguishing between providers and services.
- Filtering and ranking the web services.
- Selecting the efficient and appropriate service that achieves the whole customer needs.

Those QoS attributes can be classified into six categories as shown in Table 1.

Table 1: QoS Categories and Attributes [20].

QoS categories and attributes	1. Service Provider	Service provider reputation, Accountability, Throughput, Scalability, Availability.
	2. Service Customer	Response time, Reliability,

		Usability, Cost, Discoverability.
	3. Service Developer	Maintainability, Interoperability, Composability, Reusability, Stability, Traceability, Testability.
	4. Service runtime management	Exception handling, Completeness, Robustness.
	5. Security	Authentication, Confidentiality, Authorization, Non repudiation, Auditability, Encryption, Integrity.
	6. Network infrastructure	Server failure, Guaranteed messaging, Bandwith, Delay time, Packet loss ratio.
	7. Network infrastructure	Server failure, Guaranteed messaging, Bandwith, Delay time, Packet loss ratio.

3.2. K-Means Clustering

The clustering algorithms are classified into hierarchical clustering, exclusive clustering, probabilistic clustering, and superimposed clustering [21]. The K-Means clustering can be considered the most known to resolve many clustering problems. Our use for this algorithm is to rank web service applications according to the QoS attributes.

The advantages of the K-Means clustering are as follows [22]:

- The larger number of the variables are, the smaller the number of the clusters, and the smaller the speed of calculation than the hierarchical clustering algorithms.
- If the clusters are globular, K-Means will produce tighter clusters, which will be tighter than hierarchical clusters.

Despite all of these advantages, K-Means has also some limitations, but these later ones do not influence our approach to study the QoS between queries and published web services.

The goal of applying the K-Means clustering algorithm is to classify the database of QoS attributes offered by the list of discovered web services. This is done in multiple steps. The first is to randomly initialize the number K of centroids. The centroids represent the centers of the clusters. The following process takes place in two stages called expectation and maximization; these two stages involve assigning each data element to its nearest centroid. Next, the algorithm calculates the new centroid of all the points of each cluster and define the new centroid. The following algorithm describes the K-Means algorithm' steps:

Algorithm 1: K-Means algorithm

1: Define *K // The number of clusters*
2: Initialize *K* centroids (Randomly)
3: repeat
4: **expectation:** Assign the points to their closest centroid
5: **maximization:** Calculate the new centroid of clusters
6: until: The position of centroid does not change

3.3. Branch and Bound Skyline Algorithm

The BBS algorithm is considered as an enhancement of the K Nearest Neighbors (KNN) algorithm with a difference that the BBS algorithm crosses the R-tree only once. The algorithm uses a priority queue, where data points are organized according to their minimum distances (*mindist*) or minimum bounding rectangles (*MBR*) from an origin point. A minimal bounding rectangle is used to evaluate a complex shape. It is a rectangle with parallel sides to the x and y-axis and minimally surrounds the utmost complex shape [23].

The algorithm chooses at each step, among all the unvisited points, the closest tree points to the origin. In addition, it keeps these discovered points in a set *S* for the validation step of dominance.

The description of the BBS algorithm is given as follows:

Algorithm 2: BBS Algorithm

```
01: P=ø // P is a set of dominant points
02: fill the root R by all entries in the heap
03: while R not empty yet do
04:     c is removed // c is the top entry from R
05:     if c is dominated by other points in P remove c
06:     else
07:         if c is an intermediate entry
08:             for each child c_i of c
09:                 if c_i is not dominated by some point in P
                    insert c_i into heap
10:                 else // c is a data point
11:                     insert c_i into P
12:                 end
13:             end
14:         end
15:     end
16: end
```

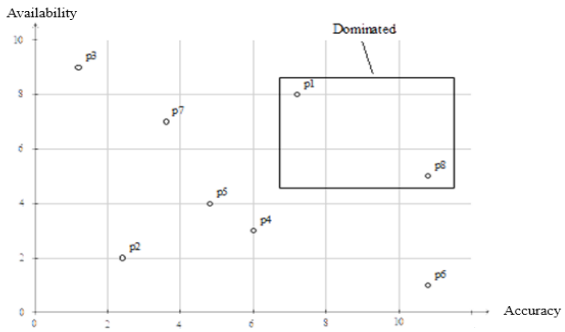


Figure 2: Example of BBS Algorithm

The BBS algorithm performs better than other skyline algorithms, ensuring a minimum cost of input/output, and the number of R-tree node access, and processing time. However, since the number of attributes is increased, the number of points in the skyline is increased substantially [24]. Hence, the idea of using a filter is to decrease the number of candidate points.

Figure 2 provides an example of domination point using two attributes of QoS.

4. Proposed Approach

The QoS-based selection consists to choose the best web service from the candidate (discovered) web services to satisfy the customer's non-functional requirements as QoS needs. This selection depends on the specification adopted when defining the QoS criteria and the QoS profile of the web service.

For solving this problem, we propose to add a filter with K-Means clustering as a first step to generate the clusters of web services based on the QoS properties assigned to each discovered web service. This technique determines the number *K* of classes as an input and generates the *K* clusters. When the process starts, it chooses centers randomly. Then at each step, it recalculates the new cluster centers as the mean of the QoS for this cluster. The criterion function used in this step is expressed in equation (1) [25] below:

$$E = \sum_{i-1}^{k} \sum_{ws \epsilon c_i} |ws - Mi|^2 \qquad (1)$$

where *E* is the error value between the consumer 'constraints and the candidate web services, *ws* refers to the candidate web service, and *Mi* is the mean of the cluster *Ci* which contains *ws*. Each cluster created by the K-Means algorithm contains web services with similar QoS attributes. The algorithm serves as a pre-filter for the discovered web services. The cluster obtained from this phase has the most efficient centroid to meet customer requirements while all web services in this cluster can also meet these requirements, and they are affected to step two.

The second step is to exploit the filter results and apply the BBS algorithm on the filtered cluster web services. The objective of this step is to determine the dominant web services among the obtained cluster web services using their QoS properties. Eliminating inappropriate web services in the pre-filter phase with K-Means clustering makes it easier for the BBS algorithm to find the most appropriate web service and meets customer requirements.

Figure 3 summarizes the model of our proposed approach. The customer submits a request for a service that meets their needs and requirements. In the discovery stage, the web service registry determines a set of candidate web services that can meet customer needs. In the web service selection stage, we propose to add a pre-filter using K-Means algorithm which minimizes the candidate web services. This step takes place by creating clusters that contain web services including similar QoS properties and determining which cluster meets the customer requirements, thereby the BBS algorithm is applied to find the appropriate web service for customer needs and requirements.

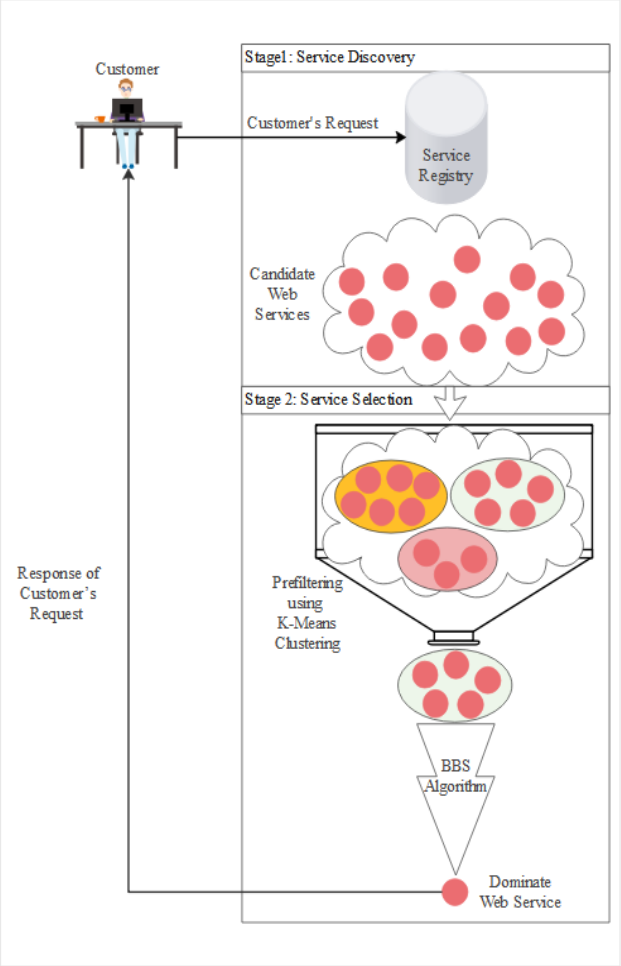


Figure 3: Web Service Selection Model

5. Experimental Results

The evaluation of the proposed approach aims to show the interest of adding a pre-filter to a web service selection system. The pre-filter proposed in our approach is based on K-Means clustering.

A real-world dataset of QoS attributes named QWSDataset [26] is used for experimentation. The dataset contains 9 QoS attributes per service, downloaded by a web service Crawler Engine [27]. The used version contains 2507 web services. This dataset contains 11 segments representing 9 QoS attributes, the URL of WSDL file and the service name. The database is used in experiments to prove the efficiency and performance of our proposed approach. Table 2 represents some values of QWSDataset used in the experimentation.

To reduce the search domain, we apply the K-Means algorithm to determine the web services that have similar QoS parameters. The selection process is conducted using the BBS algorithm. The performance and efficiency of our proposed approach are verified using the evaluation metrics: Success rate and execution time for a selection process with different approaches.

5.1. Success Rate

The success rate (SR) of all selected web services is the proportion of customers' QoS requirements (C_i) to the QoS values

Table 2: QoS Values of QWSDataset.

Service Name	Response Time	Availability	Throughput	Successability	Reliability	Compliance	Best Practices	Latency	Documentatio
MAPPMatching	302.75	89	7.1	90	73	78	80	187.75	32
Compound2	482	85	16	95	73	100	84	1	2
USDAData	3321.4	89	1.4	96	73	78	80	2.6	96
GBNIRHoliday Dates	126.17	98	12	100	67	78	82	22.77	89
CasUsers	107	87	1.9	95	73	89	62	58.33	93

of these web services ($\overline{U_i}(S)$). The success rate of a web service (*SRn)* equals 1 if its value is greater than a threshold value (*ts*).

$$SR = \frac{SRn}{n} \times 100\% \quad (2)$$

and

$$SRn = \sum_{i=1}^{n} \begin{cases} 1, \bigcap_{i=1}^{n} \frac{C_i}{\overline{U_i}(S)} \geq ts \\ 0, otherwise \end{cases} \quad (3)$$

where *SRn* is the success rate for n web services. Based on [11], the values of parameters used in the experimentation are *ts*=0.86 and *n*=200.

Figure 4 presents the success rate of our approach compared to other approaches depending on the number of candidate web services. The success rate is increased for our approach compared to other approaches. Furthermore, the more the number of candidate web services is increased, the more the success rate of our approach is increased, which means that the proposed approach is scalable for the large dataset of web services. However, the other approaches have a stable success rate or a decreased success rate when the number of candidate web services is increased.

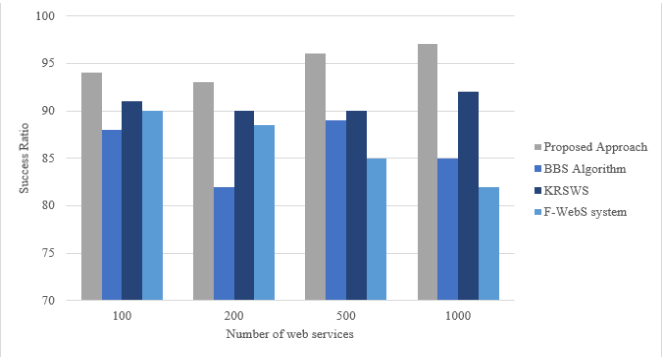


Figure 4: Success Rate Comparison with QWSDataset.

5.2. Computation Time

To evaluate the efficiency of our approach in comparison with other ones, the execution time is computed as shown in Figure 5. The execution time for our approach is minimized than other approaches depending on the number of web services. As a consequence, our approach has a high performance considering the execution time even we deal with a large dataset.

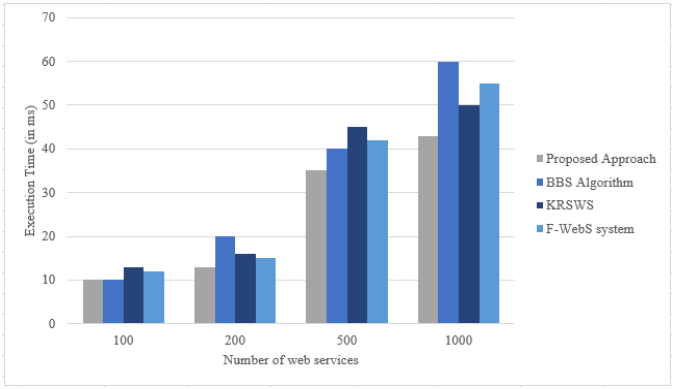


Figure 5: Time Execution Comparison.

After that, we have calculated the execution time by modifying the number of QoS attributes from 3 to 9. This modification is performed to examine the impact of QoS attributes' number to select the adequate web service. Moreover, the comparison between the proposed approach with other approaches is performed in terms of the execution time. As illustrated in Table 3, the execution time for the compared approaches shows a high performance of our approach, which uses the K-Means clustering as a pre-filter. The effectiveness of our approach is proved by increasing the number of QoS attributes.

Table 3: Execution time according to QoS attributes.

Number of QoS attributes	Execution Time (*ms*)			
	BBS Algorithm	F-WebS system	KRSWS	Proposed Approach
3	5	7	6	4
4	7	7	8	6
5	12	10	11	10
6	15	16	17	13
7	17	19	17	15
8	34	30	32	28
9	60	55	53	43

Adding a pre-filter to the web service selection process has allowed us to improve the success rate up to 97% and minimize the execution time of web service selection compared to other approaches. This is due to the elimination of the inappropriate

web services using a new pre-filer mechanism that is based on the K-Means clustering.

6. Conclusion

The selection web service problem consists to find the most adequate web service from a large dataset with a short period of time based on the QoS proprieties. To resolve this problem, we proposed a new mechanism using the K-Means clustering as a pre-filter to eliminate the inappropriate web services, which leads to minimizing the search space. Then the BBS algorithm is applied to select the dominant web service for increasing the precision rate. The experimentation results show a high performance of our approach compared to other ones in terms of the scalability, the execution time, and the precision rate.

The current work will provide many benefits and advantages to end-users, practitioners, and researchers who deal with a large data set of web services. It allows to prefilter an enormous number of web services that are generated from different systems such as smart health, smart agriculture, smart city, etc.

In the future, we can use this approach to resolve the web service composition problems and minimized the composition time, and we try to integrate uncertain QoS parameters in our approach.

References

[1] M. Fariss, N. El Allali, H. Asaidi, M. Bellouki, "Prefiltering Approach for Web Service Selection Based on QoS," Proceedings - 2019 4th International Conference on Systems of Collaboration, Big Data, Internet of Things and Security, SysCoBIoTS 2019, 1–5, 2019, doi:10.1109/SysCoBIoTS48768.2019.9028043.

[2] F. Curbera, M. Duftler, R. Khalaf, W. Nagy, N. Mukhi, S. Weerawarana, "Unraveling the Web services Web: An introduction to SOAP, WSDL, and UDDI," IEEE Internet Computing, **6**(2), 86–93, 2002, doi:10.1109/4236.991449.

[3] S.Y. Hwang, C.C. Hsu, C.H. Lee, "Service selection for web services with probabilistic QoS," IEEE Transactions on Services Computing, **8**(3), 467–480, 2015, doi:10.1109/TSC.2014.2338851.

[4] M. Alrifai, D. Skoutas, T. Risse, "Selecting skyline services for QoS-based web service composition," in Proceedings of the 19th international conference on World wide web - WWW '10, ACM Press, New York, New York, USA: 11, 2010, doi:10.1145/1772690.1772693.

[5] H.T. Kung, F. Luccio, F.P. Preparata, "On Finding the Maxima of a Set of Vectors," Journal of the ACM (JACM), **22**(4), 469–476, 1975, doi:10.1145/321906.321910.

[6] C.H. Papadimitriou, M. Yannakakis, "Multiobjective query optimization," in Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, Santa Barbara, California, USA: 52–58, 2001, doi:10.1145/375551.375560.

[7] A. Ouadah, A. Hadjali, F. Nader, "A Hybrid MCDM Framework for Efficient Web Services Selection Based on QoS," in 2018 International Conference on Applied Smart Systems (ICASS), IEEE: 1–6, 2018, doi:10.1109/ICASS.2018.8652037.

[8] S.S. Joshi, O.B.V. Ramanaiah, "An integrated QoE and QoS based approach for web service selection," in 2016 International Conference on ICT in Business Industry & Government (ICTBIG), IEEE: 1–7, 2016, doi:10.1109/ICTBIG.2016.7892671.

[9] R. Karim, C. Ding, C.-H. Chi, "An Enhanced PROMETHEE Model for QoS-Based Web Service Selection," in 2011 IEEE International Conference on Services Computing, IEEE: 536–543, 2011, doi:10.1109/SCC.2011.81.

[10] Marios Kokkodis., "Implementation Of Skyline Query Algorithms," 2003.

[11] M. Fariss, H. Asaidi, M. Bellouki, "Comparative study of skyline algorithms for selecting Web Services based on QoS," Procedia Computer Science, **127**, 408–415, 2018, doi:10.1016/j.procs.2018.01.138.

[12] W. Abramowicz, A. Godlewska, J. Gwizdala, M. Kaczmarek, D. Zyskowski, "Application-Oriented Web Services Filtering," in: IEEE, ed., in International Conference on Next Generation Web Services Practices (NWeSP'05), IEEE, Los Alamitos: 63–68, 2005, doi:10.1109/NWESP.2005.17.

[13] L. Li, I. Horrocks, "A Software Framework for Matchmaking Based on Semantic Web Technology," International Journal of Electronic Commerce, **8**(4), 39–60, 2004, doi:10.1080/10864415.2004.11044307.

[14] K. Lynch, D., Keeney, J., Lewis, D., O'Sullivan, "A Proactive Approach to Semantically-Driven Service Discovery," in the Proceedings of 2nd Workshop on Innovations in Web Infrastructure, Edinburgh, 2006.

[15] N. Srinivasan, M. Paolucci, K. Sycara, "Semantic Web Service Discovery in the OWL-S IDE," in Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06), IEEE: 109b-109b, 2006, doi:10.1109/HICSS.2006.431.

[16] E. Al-Masri, Q.H. Mahmoud, "QoS-based discovery and ranking of Web services," Proceedings - International Conference on Computer Communications and Networks, ICCCN, 529–534, 2007, doi:10.1109/ICCCN.2007.4317873.

[17] Y. Xia, P. Chen, L. Bao, M. Wang, "A QoS-Aware Web Service Selection Algorithm Based On Clustering," in 2011 IEEE International Conference on Web Services, 428–435, 2011, doi:10.1109/ICWS.2011.36.

[18] M. Ahmadi Oskooei, S. Mohd Daud, "Quality of service (QoS) model for web service selection," in 2014 International Conference on Computer, Communications, and Control Technology (I4CT), IEEE: 266–270, 2014, doi:10.1109/I4CT.2014.6914187.

[19] W. Abramowicz, K. Haniewicz, M. Kaczmarek, D. Zyskowski, "Architecture for Web Services Filtering and Clustering," in Second International Conference on Internet and Web Applications and Services (ICIW'07), IEEE: 18–18, 2007, doi:10.1109/ICIW.2007.19.

[20] K. Kaewbanjong, S. Intakosum, "QoS Attributes of Web Services : A Systematic Review and Classification," Journal of Advanced Management Science, **3**(3), 194–202, 2015, doi:10.12720/joams.3.3.194-202.

[21] J. Han, M. Kamber, "Data mining concepts and techniques San Francisco Moraga Kaufman," 2001.

[22] N. Arunachalam, A. Amuthan, C. Kavya, M. Sharmilla, K. Ushanandhini, M Shanmughapriya, "A SURVEY ON WEB SERVICE CLUSTERING," in 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC), IEEE, Melmaruvathur, India: 247–252, 2017, doi:10.1109/ICCPEIC.2017.8290371.

[23] J. Wood, Minimum Bounding Rectangle, Springer US, Boston, MA: 660–661, 2008, doi:10.1007/978-0-387-35973-1_783.

[24] K. Beyer, J. Goldstein, R. Ramakrishnan, U. Shaft, When Is "Nearest Neighbor" Meaningful?, 217–235, 1999, doi:10.1007/3-540-49257-7_15.

[25] J. Han, M. Kamber, J. Pei, Data Mining : Concepts and Techniques : Concepts and Techniques (3rd Edition), 3rd ed., San Francisco, CA, USA: Morgan Inc, Kaufmann Publishers, 2012.

[26] E. Al-Masri, Q.H. Mahmoud, "Investigating web services on the world wide web," in Proceeding of the 17th international conference on World Wide Web - WWW '08, ACM Press, New York, New York, USA: 795, 2008, doi:10.1145/1367497.1367605.

[27] E. Al-Masri, Q.H. Mahmoud, "Crawling multiple UDDI business registries," 16th International World Wide Web Conference, WWW2007, 1255–1256, 2007, doi:10.1145/1242572.1242794.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 619-625 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Particle Swarm Optimization, Genetic Algorithm and Grey Wolf Optimizer Algorithms Performance Comparative for a DC-DC Boost Converter PID Controller

Jesus Aguila-Leon[*,1,2], Cristian Chiñas-Palacios[1,2], Carlos Vargas-Salgado[2,3], Elias Hurtado-Perez[2,3], Edith Xio Mara Garcia[1]

[1]*Departamento de Estudios del Agua y de la Energía, Centro Universitario de Tonalá, Universidad de Guadalajara, Tonalá, 45425, Mexico*

[2]*Instituto Universitario de Ingeniería Energética, Universitat Politècnica de València, València, 46022, Spain*

[3]*Departamento de Ingeniería Eléctrica, Universitat Politècnica de València, València, 46022, Spain*

ARTICLE INFO

Article history:
Received: 16 November, 2020
Accepted: 21 January, 2021
Online: 30 January, 2021

Keywords:
PID Tunning
Grey Wolf Optimizer
Particle Swarm Optimization
Genetic Algorithm
Boost converter

ABSTRACT

Power converters are electronic devices widely applied in industry, and in recent years, for renewable energy electronic systems, they can regulate voltage levels and actuate as interfaces, however, to do so, is needed a controller. Proportional-Integral-Derivative (PID) are applied to power converters comparing output voltage versus a reference voltage to reduce and anticipate error. Using PID controllers may be complicated since must be previously tuned prior to their use. Many methods for PID controllers tunning have been proposed, from classical to metaheuristic approaches. Between the metaheuristic approaches, bio-inspired algorithms are a feasible solution; Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) are often used; however, they need many initial parameters to be specified, this can lead to local solutions, and not necessarily the global optimum. In recent years, new generation metaheuristic algorithms with fewer initial parameters had been proposed. The Grey Wolf Optimizer (GWO) algorithm is based on wolves' herds chasing habits. In this work, a comparison between PID controllers tunning using GWO, PSO, and GA algorithms for a Boost Converter is made. The converter is modeled by state-space equations, and then the optimization of the related PID controller is made using MATLAB/Simulink software. The algorithm's performance is evaluated using the Root Mean Squared Error (RMSE). Results show that the proposed GWO algorithm is a feasible solution for the PID controller tunning problem for power converters since its overall performance is better than the obtained by the PSO and GA.

1. Introduction

Power converters have an important role in Industry applications; their main purpose is to regulate power in electronic appliances and adjusting current and voltage signals to desired levels by a high frequency switching control device [1,2]. One of the most significant modern applications for power converters is related to Renewable Energy Sources (RES), since the power generated for these sources highly depends on environmental conditions power converters are a reliable solution to stabilize output voltage and current of RES. A Boost Converter is an electronic device whose main purpose is to raise an input voltage and stabilize it to a desired highest level [3,4]. Since voltage regulation on power converters, and therefore, boost converters, depends on a switching signal, is needed a controller to generate a proper pulse width modulation (PWM) to modify the duty cycle of the switching signal. PID controllers are commonly used for this kind of applications, however, PID controllers need to be tunned prior its use [5]. The PID controller tunning can be a challenging task, many methods had been proposed for this purpose, from classical modeling and analysis based on system response [6] to modern techniques based on metaheuristic algorithms [7]. Nature had been an inspiration for modern metaheuristic algorithms, most of them based on animal behavior, for system design and control optimal parameters finding [8]. Bio-inspired algorithms imitate animal collective intelligence to explore, find, and exploit food and resources. Collective intelligence is the sum of individual behavior based on simple rules, and these behaviors and strategies can be translated into computational optimization algorithms. The most common bio-inspired algorithms are Genetic Algorithms (GA) and

[*]Corresponding Author: Jesus Aguila Leon, Centro Universitario de Tonalá, Universidad de Guadajara, Av. Nuevo Periférico 555, Tonalá, México.
Email: jesus.aguila@academicos.udg.mx

Particle Swarm Optimization (PSO) algorithms. Among these bio-inspired optimization algorithms applications are control of load frequency [9], systems optimal sizing and design [10], power flow applications [11], predictive control for microgrids based on renewable energies [12], power converters optimal design [13] and regulation for voltage [14]. In [15], an evolutionary approach using Symbiotic Organisms Search (SOS) algorithm, which is based on PSO algorithms and the trophic chain of the ecosystems, and a two-round fuzzy inference engine were presented for energy management in microgrids. Since PSO is general modeling for animal behavior, there had been developed and studied variants of PSO based on particular animal species [16,17]. Authors in [18] applied a Whale Optimization Algorithm for an optimal design of a PID controller for a DC-DC converter using a transfer function and step response approach to analise the system performance; they found beter transciente response in comparison to a compared GA algorithm, however, they did not considered fluctutanting input voltages or load changes. In 2014, the author presented the Grey Wolf Optimizer (GWO), based on hunting strategies of wolves herds [19]. The GWO had been studied in comparison to other optimization algorithms as well as some engineering applications. Authors in [20] applied GWO to a PID controller to optimize the system response of a steam pressure system, they found an improvement in system stability and response. In [21] the GWO was applied to a modeled levitation system, improving time domain response and reducing the system error in comparison to the PID tunning MATLAB tool. Also, GWO had been applied to power systems.

In this paper, an optimal tunning for a PID controller using the GWO algorithm applied to a boost converter is presented. The boost converter is modeled unsung state-space equations and then simulated in MATLAB/Simulink. The proposed GWO-PID tunned algorithm performance is compared with PSO and GA algorithms in terms of the RMSE and the system response for variable load and variable input voltage. The paper is organized as follows: Section 2 depicts the boost converter model and description. Section 3 presents the GWO, PSO, and GA optimization algorithm basis. Section 4 shows the optimization methodology for the boost converter. Section 5 summarizes the results and, finally, Section 6 are the conclusions of this study.

2. Mathematical model of the Boost Converter

The boost converter is an electronic device that rises input voltage to the desired highest output voltage. Voltage regulation is made by a PWM signal, applied to an inductor (L) and capacitor (C) arrangement, carried out by a fast-switching transistor that according to the control signal. Changes in the PWM signal modifies the L-C charge and discharge cycles and change the output signal [22]. The boost converter configuration is shown in Figure 1.

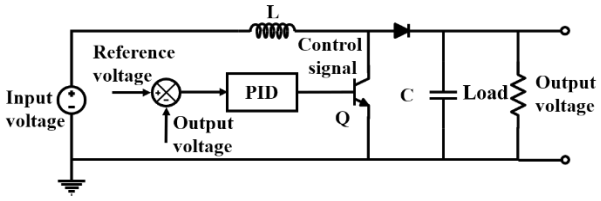


Figure 1: Electric diagram for the Boost converter.

Modeling of the boost converter was carried out using state-space equations, then MATLAB/Simulink software was used for simulation.

State-space equations were obtained employing Kirchhoff's voltage and current laws analysis for each system switching state determined by the u control signal value. Since u the signal can only adopt a 1 o 0 value, there are two possible Boost converter electric configurations, that correspond to charge and discharge cycles for the L and C elements. The two possible circuit configurations are condensed in a single matrix form of state-space equations as showed in (1).

$$\begin{bmatrix} \frac{dv_o(t)}{dt} \\ \frac{di_L(t)}{dt} \end{bmatrix} = \begin{bmatrix} \frac{1-u}{C} & -\frac{1}{CR} \\ 0 & -\frac{1-u}{L} \end{bmatrix} \begin{bmatrix} v_o(t) \\ i_L(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} v_i(t) \quad (1)$$

In (1) L is the inductor, C is the capacitor, $i_L(t)$ is the current in the inductor, $v_o(t)$ is the output voltage, $v_i(t)$ is the voltage ant input of the boost converter, and u is the PWM control signal generated by the PID controller.

3. Optimization Algorithms

A performance comparison between three optimization algorithms: PSO, GA, and GWO is presented in this paper, the main objective of this research is to get insights on the best optimization algorithm for PID controllers tunning applied to a power converter. A brief description of each of the optimization algorithms is presented in the following subsections.

3.1. Grey Wolf Optimizer

Wolves had a strong hierarchy inside their herds, they are organized in a group of 5 to 12 wolves. Each herd had a leading wolf, called alpha α wolf; secondary wolves called beta β wolves; wolves subordinated to α and β wolves are delta wolves δ ; and finally, follower wolves are omega ω wolves. The optimization using the GWO is carried out in three main stages that mimic the hunting process of grey wolves' herds in nature: encircling, hunting, and attacking.

Encircling stage

In this stage, each wolf updates its position in the search space according to the relative best position to prey, dictated for α wolf. The encircling and corral behavior of the prey is modeled mathematically by (2-5).

$$\vec{X}(t+1) = \vec{X}_{pos}(t) - \vec{A}\vec{D} \quad (2)$$

$$\vec{A} = 2ar_1 - a \quad (3)$$

$$\vec{C} = 2r_2 \quad (4)$$

$$\vec{D} = \left| \vec{C}\vec{X}_{pos}(t) - \vec{X}(t) \right| \quad (5)$$

where $\vec{A}$ $\vec{C}$ $\vec{X}_{pos}$ $\vec{X}$ vector is the position of the wolf, t is

current iteration, r_1 and r_2 values randomly generated between zero and one, and, vector a value decreases linearly according to iterations.

Hunting stage

The position of wolves is rearranged according to their proximity to the prey. The wolf with the closest distance to the prey is assigned to be the alpha α wolf, β and δ wolves are assigned according to their position to prey. Equations (6-8) describe the wolve position updating.

$$\vec{X}_i(t+1) = \frac{\vec{X}_{i1} + \vec{X}_{i2} + \vec{X}_{i3}}{3} \tag{6}$$

$$\left.\begin{aligned} \vec{X}_{i1} &= \vec{X}_\alpha(t) - \vec{A}_1 \vec{D}_\alpha \\ \vec{X}_{i2} &= \vec{X}_\beta(t) - \vec{A}_2 \vec{D}_\beta \\ \vec{X}_{i3} &= \vec{X}_\delta(t) - \vec{A}_3 \vec{D}_\delta \end{aligned}\right\} \tag{7}$$

$$\left.\begin{aligned} \vec{D}_\alpha &= \left|C_1 \vec{X}_\alpha(t) - \vec{X}_i(t)\right| \\ \vec{D}_\beta &= \left|C_2 \vec{X}_\beta(t) - \vec{X}_i(t)\right| \\ \vec{D}_\delta &= \left|C_3 \vec{X}_\delta(t) - \vec{X}_i(t)\right| \end{aligned}\right\} \tag{8}$$

where $\vec{X}_i(t+1)$ is the wolf that has the best position to the prey and i is the current iteration number of the GWO algorithm.

Attack stage

Attack to prey occurs when the heard is upon the prey, before this, it is necessary to minimize the distance between wolves and the prey. The prey is the optimization problem's best global solution. Must be defined as a vector A, shown in(3), to make a decreasing coefficient dependent on the iteration number. The vector A value is reduced as the value index a value is reduced according to (9).

$$a = 2 - t\left(\frac{2}{T}\right) \tag{9}$$

The GWO pseudo code is as follows,

GWO: pseudo code.

Result: The best set of particles for the fitness function.
X generation; creation of an initial population of wolves.
Parameters initialization (a,A,C);
Evaluation of position X(0);
Selection of new (a , β and δ);
Selection of new position X(0);
for *e = 1 to MaxIteration* **do**
  for *each wolf*$_i$ *in* ω set **do**;
    for *i = 0 to* DIMENSION **do**;
      Position(*i*,*j*) updating;
    end for
  Change (a,A,c) factors;
  Calculate *X(t+1)* position
  end for
end for

3.2. Particle Swarm Optimization

PSO algorithms are based upon the swarm behavior that some animal species show when they search for resources in their environment. Since PSO is based on collectiveness, each search agent must be modeled. For this purpose, agents are modeled as particles of a swarm, having their position relative to exploration space, velocity, and acceleration rates. At the beginning of the PSO algorithm, several particles are set up in a search space where somewhere in is the global solution for the optimization problem. Finding a global solution depends on the evaluation and minimization of the defined objective function. Iteration by iteration positions, velocities, and acceleration of particles are updated to converge to the global solution. For each particle, a fitness function is numerically evaluated. The best value obtained for all the particle's fitness function is called to be the best global g_{best}. During the iteration process, each best particle fitness function value is called to be the personal best p_{best}. During iterations, the speeds of the particles are accelerated toward the best global solution and the best personal according to (10).

$$v_n = w * v_n + c_1 rand() * (g_{best,n} - x_n) + c_2 rand() * (p_{best,n} - x_n) \tag{10}$$

Where v_n is the speed update of particles, w is a factor of inertia whose value is decreased from 0.9 to 0.4 over time, c_1 and c_2 are coefficients of acceleration pointing to the best global and the best personal.

3.3. Genetic Algorithm

Genetic Algorithms are based on the genetic evolution process, imitating the genes mutation and crossover to create the best-adapted organisms to the environment. Mathematical modeling of these mechanisms allows the algorithm to refine solutions carrying out artificial genes crossover, while the mutation mechanism adds uncertainty to experiment with not expected genes, this makes the algorithm avoid local solutions where it could be trapped. During iterations, the best set of artificial genes is obtained, and therefore, a best-adapted species to the environment. The best group of genes for i generations are said to be the found best solution for the optimization problem.

GA can be expressed as a four-step algorithm,

Step 1. An initial population is created. Crossover rate and mutation rate are randomly generated; generation number is setup.

Step 2. Is evaluated the defined fitness function for each set of artificial genes.

Step 3. Starts the crossover process and mutation process to set the next generation of artificial genes.

Step 4. Return to Step 2 until the stop criterion is reached.

4. Tunning of a PID controller for a Boost converter using optimization algorithms

PID controllers allow the system to operate near desired output values, ensuring good response in the face to possible disturbances. The proposed Boost Converter includes a PID controller for voltage regulation for load changes. The PID controller carries out proportional, integrative, and derivative actions to reduce and prevent error $e(t)$ between output and reference signals of the system. The controller signal $u(t)$ is mathematically modeled according to (11).

$$u(t) = K_p e(t) + \frac{K_p}{T_i}\int_0^t e(t)dt + K_p \frac{T_d de(t)}{dt}] \quad (11)$$

In (11) K_p , T_i and T_d are the proportional, integration time and derivation time constants, respectively. By adjusting the values of these constants, the error in the system can be reduced. The integration and derivative time can be expressed in terms of K_p according to (12-15).

$$K_p = k \quad (12)$$

$$K_i = {K_p}/{T_i} \quad (13)$$

$$K_d = K_p T_d \quad (14)$$

where k is the constant of proportional gain K_p , K_i is the constant of integrative gain, and K_d is the constant of derivative gain for the PID controller. The goal of the GWO, PSO, and GA algorithms is to find the best K_p , K_i , K_d values so the error $e(t)$ is minimized. A vector X is defined to include the constant values according to (15). To minimize the Error $e(t)$ the Root Mean Squared Error (RMSE) is defined to be the objective function for the optimization problem according to (16).

$$\overrightarrow{X} = [K_p, K_i, K_d] \quad (15)$$

$$RMSE = \sqrt{\frac{\sum_{t=0}^{T_{sim}} \left(v_{out}(t) - v_{reference}(t)\right)^2}{T_{sim}}} \quad (16)$$

Where T_{sim} is simulation time, $v_{out}(t)$ is the output voltage of the power converter, and $v_{reference}(t)$ is the reference voltage signal. The Figure 2 illustrates the overall flowchart for the optimization to find the values for the PID controller constant gains using the GWO, PSO, and GA, the first step of all three algorithms evaluated is to generate a random search agent population, that according to the algorithm is the number of wolves, particles, or gene population, then for each agent, the objective function is evaluated in an iterative loop until the best solution is found.

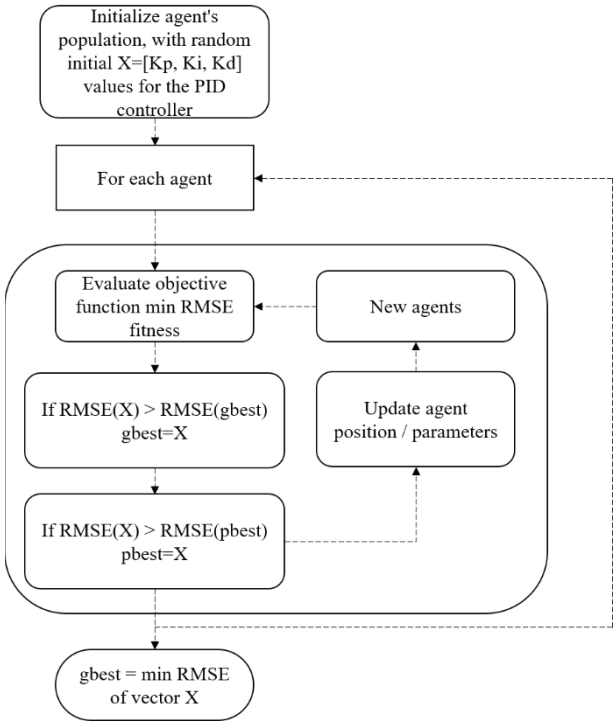


Figure 2: Overall flowchart for the optimization to find the optimal values PID controller constant gains for the proposed optimization algorithms.

5. Simulation, results, and discussion

The optimization algorithms of GWO, PSO, and GA were implemented using MATLAB and Simulink software. Several simulations run for each of these algorithms were performed to find the best PID controller gains values. The difference between each run is the variation of the limits of the search space for the optimization variables, that is, different minimum and maximum values for each gain constant of the controller. A scan was made with different values to determine the search space with the best possible solution for the algorithms evaluated for this particular application. The Boost Converter configuration parameters are shown in Table 1.

Table 1: Boost Converter Configuration

Description	Value	Units
Capacitor	250	μF
Inductor	1.5	mH
Input voltage	12	V
Reference voltage	24	V
Load	3-15	Ω

Since optimization algorithms require some constant parameters these values must be specified at the beginning of each test. One of the most significant benefits of using GWO is that fewer initial parameters are required in comparison to PSO and GA. The parameters initial values for each algorithm used in this work are shown in Table 2. For all three optimization algorithms, the search

K_p, K_i and K_d were varied from 0 to 100 for each PID controller gain.

Table 2: Algorithm parameters initial values

Algorithm	Initial Values	
	Description	***Value***
GWO	Wolves number	12
	Maximum Iterations	7
PSO	Particles	50
	Factor of inertia	0.4-0.9
	Weight of self-adjustment	3
	Weight of social adjustment	1
GA	Size of populations	50
	Rate of crossover	0.9
	Rate of mutation	0.6
	Maximum generations	120

The Table 3 summarizes the PID controller constant gains obtained after the simulation process for each algorithm. The performance is evaluated using RMSE.

Table 3: Best PID gains values and RMSE.

Parameter	Algorithm		
	GWO	***PSO***	***GA***
K_p	0.0532E-5	1.7223E-5	0.1269E-5
K_i	1.6048	2.8391	4.0815
K_d	4.8572E-5	7.2477E-5	4.1911E-5
T_i	3.3151E-7	6.0664E-6	3.1092E-7
T_d	91.3008	4.2082	33.0268
RMSE	1.0683	2.1755	2.2367
Simulation time (s)	1286	1997	1175

As observed in Table 3, the obtained PID controller gains for all the three algorithms had no significant value for de K_d constant, in this sense, the proposed controller built by the metaheuristic algorithms is a PI controller. Using (13) and (14) it is possible to calculate T_i and T_d for each PID controller tunned by the three algorithms, as shown in Table 3. For a greater T_i than K_p value a significative integrative action of the controller is observed according to (11) and (13); and when K_p value is much smaller than T_d the PID controller will have a small derivative action over the plant, according to (11) and (14). The controller performance evaluated by the RMSE shows that the GWO has the lowest error since its RMSE is about 50.89% lower than the one obtained by the PSO and 44.70% lower than the one obtained using the GA. Since the three evaluated algorithms differ in their search mechanisms, different convergence curves to the best solution found by GWO, PSO and GA algorithms were found. The Figure 3 shows the convergence curve for each of the evaluated algorithms for the Boost converter PID controller tunning optimization problem.

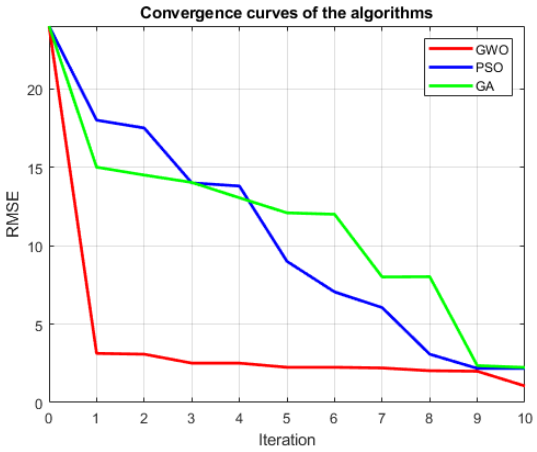


Figure 3: Convergence curves to the best solution found by GWO, PSO, and GA algorithms.

As can be observed in Figure 3, the GWO algorithm is faster to get closer to the best solution set of K_p, K_i and K_d in comparison to PSO and GA whose convergence curves are slower to decrease the resulting RMSE value.

Once the GWO was chosen as the best algorithm for this application, more simulation runs were made varying the search space limits for the optimization variables and with a different number of wolves to ensure a more refined PID gains constants optimal solution for the Boost Converter. The main results of this second round of tests made with the GWO are shown in Table 4.

Table 4: Second round of test for GWO algorithm

Run	Obtained optimal values					
	Search space limits [K_p, Ki, K_d]	***K_p***	***K_i***	***K_d***	***RMSE***	***T***
1	[0-10,0-10,0-10]	0.0012	0.8681	0.0502 E-3	1.1436	948.43
2	[0-0.5,0-5,0-0.5]	0.0025	0.7626	0.0263 E-5	1.1099	968.40
3	[0-0.25,0-2.5,0-0.25]	0.0191	0.0167	0.0098 E-3	8.6105	953.45
4	[0-0.01,0-0.5,0-0.01]	0.0033	0.3118	2.9451 E-5	0.7768	9030.77
5	[0-5,0-32,0-0.005]	0.0532 E-5	1.6048	4.8572 E-5	1.0683	1286.00

As observed in Table 4, the best PID controller gains were obtained in run number four, where the RMSE is minimum with a value of 0.7768. Different RMSE results were obtained for different search space limits; the best results were obtained for the search space limit of [0 to 0.01, 0 to 0.5, 0 to 0.01] corresponding to the $\vec{X} = [K_p, K_i, K_d]$ optimization variables vector. However, when plotting the system, see Figure 4, response for each run, it is found that the system response of the fifth run, with an RMSE of

1.0683, has a faster response to achieve the reference voltage depicts the higher oscillation rates obtained. Best PID controller gains values must be selected according to the needs: a faster response with oscillations, or a slower response with fewer oscillations. The fifth run is chosen as the best solution since a faster response is desirable for this application.

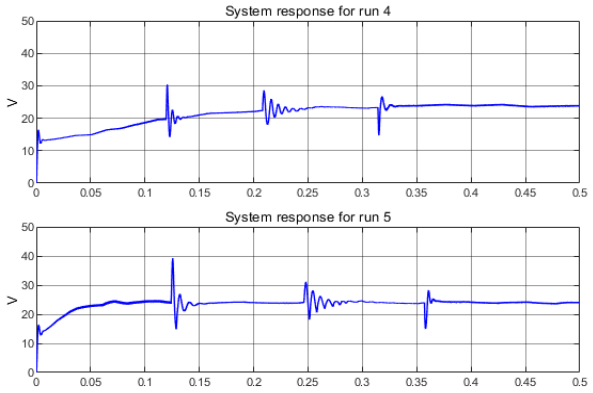


Figure 4: Comparison between run 4 (a) and run 5 (b)for the best GWO simulation results

Once the best solutions for the three evaluated algorithms were chosen, a test of the performance to varying load conditions was carried out to evaluate the system response for the GWO, PSO, and GA tunned PID controller. In Figure 5 the system response for variable load and variable input voltage condition comparison is shown for each algorithm best solution.

As observed in Figure 5, the PID controllers tuned by PSO and GA have similar system responses, presenting important oscillations for load R=15 Ohm. The input voltage to the converter is a signal that varies over time between 11.7 and 12.3 V for the time scale used. The PID controller tunned using the GWO has a better performance to stabilize system response under load and voltage changes in comparison to GA and PSO, despite having a slightly slower response than the other algorithms.

6. Conclusion

In this paper, the tunning of a PID controller for a boost converter was presented. Three algorithms were implemented and compared for this purpose using MATLAB/Simulink: PSO, GA, and GWO. The power converter is modeled using state-space equations. Several simulations are performed to find optimal values. Results are evaluated using the RMSE and the system response for variable conditions of input voltage and load at the output of the power converter. The GWO-PID tunned controller had the best performance with an RMSE about 50.89% lower than PSO and 44.70% lower than GA. The differences between the obtained RMSE using the three algorithms showed in Table 3, gives insights on the greater susceptibility of PSO and GA to be trapped in local optimum solutions, while the lower RMSE value obtained using the GWO algorithm for the same search space limits indicates that this algorithm manages in a better way to circumvent a greater number of local optimum solutions to find a better optimal solution than the PSO and GA algorithms for this application. After the controller was tunned using each algorithm and the best gain constants were found, the system response was evaluated under input voltage and load changes. The system response for variable load also showed a better performance for the GWO with fewer oscillations for load changes in comparison to PSO and GA. Also, the GWO algorithm had the advantage over the PSO and GA that GWO requires fewer configuration parameters for the optimization process. However, the PID tuned through the GWO was a little slower to reach the reference voltage than the other algorithms. This work gives insights into the GWO algorithm for controller design and control applied to power converters

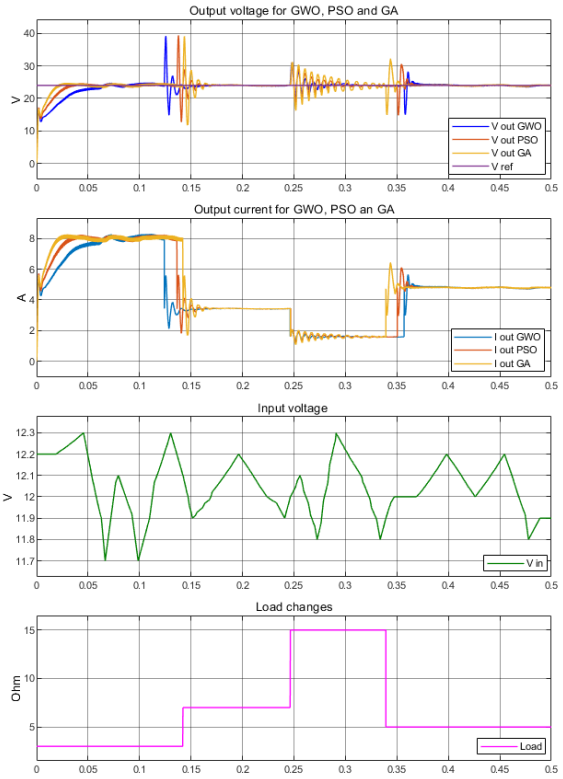


Figure 5: Comparison of system responsefor (a) output voltage, (b) output current under variable (c) input voltage and (d) load

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors wish to thank the Institute of Energy Engineering of the Polytechnic University of Valencia, Spain, and the Department of Water and Energy Studies of the University of Guadalajara, Mexico, for all their support and collaboration.

7. References

[1] J. Aguila-Leon, C.D. Chinas-Palacios, C. Vargas-Salgado, E. Hurtado-Perez, E.X.M. Garcia, "Optimal PID Parameters Tunning for a DC-DC Boost Converter: A Performance Comparative Using Grey Wolf Optimizer, Particle Swarm Optimization and Genetic Algorithms," in 2020 IEEE Conference on Technologies for Sustainability, SusTech 2020, 2020, doi:10.1109/SusTech47890.2020.9150507.

[2] H. Sira-Ramírez, R. Silva-Ortigoza, Control Design Techniques in Power Electronic Devices, 2013, doi:10.1017/CBO9781107415324.004.

[3] G.A. Raiker, S.R. B, P.C. Ramamurthy, L. Umanand, S.G. Abines, S.G. Vasisht, "Solar PV interface to Grid-Tie Inverter with Current Referenced Boost Converter," in 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), IEEE: 343–348, 2018, doi:10.1109/ICIINFS.2018.8721313.

[4] S.E. Babaa, G. El Murr, F. Mohamed, S. Pamuri, "Overview of Boost Converters for Photovoltaic Systems," Journal of Power and Energy Engineering, **06**(04), 16–31, 2018, doi:10.4236/jpee.2018.64002.
[5] J. Berner, K. Soltesz, T. Hägglund, K.J. Åström, "An experimental comparison of PID autotuners," Control Engineering Practice, **73**, 124–133, 2018, doi:10.1016/J.CONENGPRAC.2018.01.006.
[6] K. Ogata, Modern Control Engineering, 5th ed., Prentice Hall, 2010.
[7] K. Nisi, B. Nagaraj, A. Agalya, "Tuning of a PID controller using evolutionary multi objective optimization methodologies and application to the pulp and paper industry," International Journal of Machine Learning and Cybernetics, **10**(8), 2015–2025, 2019, doi:10.1007/s13042-018-0831-8.
[8] M.T. Özdemir, D. Öztürk, "Comparative performance analysis of optimal PID parameters tuning based on the optics inspired optimization methods for automatic generation control," Energies, **10**(12), 2017, doi:10.3390/en10122134.
[9] G.-Q. Zeng, X.-Q. Xie, M.-R. Chen, "An Adaptive Model Predictive Load Frequency Control Method for Multi-Area Interconnected Power Systems with Photovoltaic Generations," Energies, **10**(11), 1840, 2017, doi:10.3390/en10111840.
[10] Y. Sawle, S.C. Gupta, A.K. Bohre, "Optimal sizing of standalone PV/Wind/Biomass hybrid energy system using GA and PSO optimization technique," Energy Procedia, **117**, 690–698, 2017, doi:10.1016/j.egypro.2017.05.183.
[11] S. Surender Reddy, C. Srinivasa Rathnam, "Optimal Power Flow using Glowworm Swarm Optimization," International Journal of Electrical Power & Energy Systems, **80**, 128–139, 2016, doi:10.1016/J.IJEPES.2016.01.036.
[12] C.Y. Acevedo-arenas, A. Correcher, C. Sánchez-díaz, E. Ariza, D. Alfonso-solar, C. Vargas-salgado, J.F. Petit-suárez, "MPC for optimal dispatch of an AC-linked hybrid PV / wind / biomass / H2 system incorporating demand response," Energy Conversion and Management, **186**(February), 241–257, 2019, doi:10.1016/j.enconman.2019.02.044.
[13] M. Çelebi, "Efficiency optimization of a conventional boost DC/DC converter," Electrical Engineering, **100**(2), 803–809, 2018, doi:10.1007/s00202-017-0552-0.
[14] Q.Y. Lu, W. Hu, L. Zheng, Y. Min, M. Li, X.P. Li, W.C. Ge, Z.M. Wang, "Integrated coordinated optimization control of automatic generation control and automatic voltage control in regional power grids," Energies, **5**(10), 3817–3834, 2012, doi:10.3390/en5103817.
[15] J. Aguila-Leon, C. Chiñas-Palacios, E.X.M. Garcia, C. Vargas-Salgado, "A multimicrogrid energy management model implementing an evolutionary game-theoretic approach," International Transactions on Electrical Energy Systems, **30**(11), 2020, doi:10.1002/2050-7038.12617.
[16] Ovat Friday Aje, Anyandi Adie Josephat, "The particle swarm optimization (PSO) algorithm application – A review," Global Journal of Engineering and Technology Advances, **3**(3), 001–006, 2020, doi:10.30574/gjeta.2020.3.3.0033.
[17] N.K. Jain, U. Nangia, J. Jain, A Review of Particle Swarm Optimization, Journal of The Institution of Engineers (India): Series B, **99**(4), 407–411, 2018, doi:10.1007/s40031-018-0323-y.
[18] B. Hekimoğlu, S. Ekinci, S. Kaya, "Optimal PID Controller Design of DC-DC Buck Converter using Whale Optimization Algorithm," in 2018 International Conference on Artificial Intelligence and Data Processing, IDAP 2018, Institute of Electrical and Electronics Engineers Inc., 2019, doi:10.1109/IDAP.2018.8620833.
[19] S. Mirjalili, S.M. Mirjalili, A. Lewis, "Grey Wolf Optimizer," Advances in Engineering Software, **69**, 46–61, 2014, doi:10.1016/j.advengsoft.2013.12.007.
[20] S.-X. Li, J.-S. Wang, "Dynamic Modeling of Steam Condenser and Design of PI Controller Based on Grey Wolf Optimizer," Mathematical Problems in Engineering, **2015**, 1–9, 2015, doi:10.1155/2015/120975.
[21] S. Yadav, S.K. Verma, S.K. Nagar, "Optimized PID Controller for Magnetic Levitation System," IFAC-PapersOnLine, **49**(1), 778–782, 2016, doi:10.1016/J.IFACOL.2016.03.151.
[22] R.H.G. Tan, L.Y.H. Hoo, "DC-DC converter modeling and simulation using state space approach," 2015 IEEE Conference on Energy Conversion (CENCON), (2), 42–47, 2015, doi:10.1109/CENCON.2015.7409511.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 622-628 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Node-Node Data Exchange in IoT Devices Using Twofish and DHE

Bismark Tei Asare[*,1,2], Kester Quist-Aphetsi[2], Laurent Nana[1]

[1]*Univ Brest, Lab-STICC, CNRS, UMRS 6285, F-29200 Brest, France*

[2]*Ghana Communication Technology University, Computer Science Department, Ghana*

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 08 March, 2021
Online: 20 March, 2021

Keywords:
Secure Cloud Communication
IoT Security
Twofish
Diffie-Hellman Key Exchange

ABSTRACT

Internet of Things provides the support for devices, people and things to collaborate in collecting, analyzing and sharing sensitive information from one device onto the other through the internet. The internet of things is thriving largely due to access, connectivity, artificial intelligence and machine learning approaches that it supports. The stability and enhanced speed of the internet is also attributable to the huge adoption rate that IoT continues to enjoy from Governments, industry and academia in recent times. The increased incidences of cyber-attacks on connected systems in recent times, has inspired the heightened efforts from Governments, industry practitioners and the research world towards improving existing approaches and the engineering of new innovative schemes of securing devices, the software or the platforms for the deployment of IoT. Security solution for Internet of things includes the use of secure ciphers and key exchange algorithms that ensures the provisioning of a security layer for the: devices or hardware, communication channels, cloud, and the life cycle management constituting the Internet of things. The use of key exchange algorithms in resilient cryptographic solution that have less computational requirements without compromising the security efficiency in the encryption of messages for IoT continues to be the preferred approach in securing messages in a node-node exchange of data. This paper aims at providing a cryptographic solution that uses a key exchange cryptographic primitive and a strong cipher in encrypting messages for exchange between nodes in an IoT. Towards achieving this goal, the Diffie-Hellman key exchange (DHE) protocol was used to provide a secure key exchange between the communicating nodes, whiles the Twofish block cipher was used in the encryption and decryption of messages, assuring the security, privacy and integrity of messages in a node-node IoT data exchange. The cryptographic solution has a high throughput.

1. Introduction

The use of sensors, actuators, radars and other diodes in the collection, intelligent analysis and communication of sensitive data in connected systems have driven the user specific needs for homes, businesses, public, private institutions as well as security agencies in expanding access to support the scaling of network resources to drive new value propositions. From surveillance to the aggregation of sensitive information such as humidity, pressure, temperature and in some case depth measurement in restricted environments like mining, construction and naval systems, IoT has been the preferred solution in achieving these desired goals.

This paper is an extension of work originally presented in the International Conference on Communications, Signal Processing and Networks/International Conference on Cyber Security and Internet-of-Things ICCSPN/ICSIoT. The relatively complex key schedule that Twofish cryptographic primitive supported together with the DHE provided an efficient and tamper-resistant security scheme for data exchange between nodes in IoT [1].

The security of communication in connected devices has understandably been of great concern to both governments, industry and academia due to the rather recent increase in the number of high profile cybersecurity attacks on these networks where several accounts with sensitive information have been tempered with and in some cases, holders of these accounts have lost huge sums of monies. The cases of cyber-attacks in the past

[*]Corresponding Author: Bismark Tei Asare, Ghana Communication Technology University, Ghana, +33638556158, btasare@gctu.edu.gh

five years have motivated several public and private organisations to increase their budgetary allocations and investments into providing a stronger security infrastructure to secure their networks [2]. The devices serving as the main actors for the internet-of-things could introduce additional sources and points for attacks that hackers could use in exploiting the vulnerabilities in IoT because most of these devices are not secure by design and hence lack the appropriate robust security configuration to ensure secure communication of sensitive data, yet these devices end up forming essential components for systems that aggregate sensitive data for communication across several devices through the internet. These devices are mostly susceptible to several cybersecurity challenges including the backdoor vulnerability where hackers could easily exploit to successfully manipulate these IoT systems to their skewed benefits [3]-[6].

There exist various techniques for supporting secure node-to-node authentication that assure integrity and availability of sensitive data. Most of these techniques demand higher resource requirements for effective and efficient implementation. In the particular context of IoT devices, due to the resource constraint nature with regards to computational power, storage and power consumption, these encryption techniques and approaches are not adoptable for them, in most of the cases. In addition, since IoT devices automate their processes by eliminating human intervention in their operations, the use of some of the off the shelves solutions that exists are not practicably suited for them. Symmetric encryption involves the use of a shared key between two participating nodes in a communication. It allows these participating nodes to generate key pairs - a private key and a corresponding public key using a key generator - a trustee component in a key exchange protocol application that allows two nodes that have no prior knowledge of each other to jointly share secret for encryption and decryption of message. The public keys are common keys that are known by the participating nodes before any communication is even established. The public key is also known as a shared key. Both private (secret) key and public (shared or symmetric) keys are used in encrypting and decrypting data in a communication session between two participating nodes or parties. There are several key exchange protocols. The Diffie-Hellman key exchange (DHE) protocol allows two nodes to securely exchange session key (the public or shared key) prior to communicating messages between them [7]. The Diffie-Hellman key exchange protocol ensures the creation of a new session key for each message to be communicated, this provides another layer of security. Session keys are generated using random number generator schemes or algorithms. The key exchange protocol only supports secure sharing of session key and does not provide authentication of the source node for creating the message. Replay attacks, Man-in-the-middle attacks, Dictionary attacks, Key compromise impersonation attacks and ephemeral key compromise attacks are some of the vulnerabilities that DHE protocol suffers [1]. Cryptographic algorithms provide mechanisms for authentication using various approaches like digital signature schemes and public-key certificates.

The scalable nature of IoT and its ability to support the rapid generation volumes of traffic and their associated processing and temporary storage, places additional overload on the edge nodes used in the network. The nodes are originally constrained by design as well as operational capabilities particularly for large computational activities. Hence, complex security infractions and exploitations are constantly deployed on these IoT nodes using modern and sophisticated tools by hackers [8]-[11].

A key exchange algorithm like the DHE by itself is limited in use alone unless, it is fed into another protocol for undertaking encryption and decryption of messages. The use of the DHE in the Twofish encryption algorithm increases the confusing of the encryption key to enhance the diffusion of the distribution of the ciphertext to avoid redundancy. This improves the level of security in the cipher. The use of DHE within the Twofish cryptographic symmetric key block cipher for message encryption assured privacy and confidentiality of IoT node data.

The rest of the paper is structured into five sections. Section 2 presents key concepts needed for a better understanding of the article and related works. Section 3 is dedicated to the proposed methodology. Section 4 discusses results and Section 5 concludes the paper.

2. Background and Related Works

2.1. Background

A description of key notations and terminologies used in the paper is outlined as follows:

- Key – Is any random number as an output from a random number generator. It is used in encryption or decryption of messages. It is mainly referred to as a cryptography key.
- Public Key – It is also known as a shared or symmetric key. This key is a common key shared by two participating nodes prior to initiating a session for message communication. Two keys are generated for every participating node the public key and its corresponding private key.
- Private Key – The private key is a unique key for each participating node. The private key is used in combination with the public key of the sender for encryption of messages. During decryption of messages also, the private key of the receiver is used.
- Encryption – It is the process of encoding or hiding the content or message details by changing them from plaintext to a ciphertext using cryptographic keys. In a symmetric encryption, the private key is used to encrypt and decrypt messages.
- Decryption – It involves the use of cryptographic keys and approaches in changing ciphertexts to their original plain texts. In symmetric encryption algorithms, the private key of the destination node is used together with the public key of the source node for the decryption.
- Plaintext – It describes an unencrypted message in which the original message format and content are maintained.
- Ciphertext – It is a description for an encrypted message. Thus, the output or transformed text for any plaintext message on an encryption algorithm.
- Cipher – It is a cryptographic algorithm for transforming plaintext to ciphertext. It is used for encryption and decryption of messages.

- Algorithm – It involves a standard outline of sequential procedure or steps for undertaking encryption or decryption of messages.

The next section presents related works.

2.2. Related Work

2.2.1. Secure Cloud Communication

In our previous conference paper, we pointed out the security weakness of the DHE protocol in the provision of security solution against threats such as man-in-the-middle attacks due to vulnerabilities in the protocol to authenticate participating nodes using their unique identifying features in block stream cipher [1].

Cloud computing infrastructure services where enterprises manage and share resources between the local devices that are stationed in their premises and remote servers offer some convenience but have security challenges including data loss and tampering [12].

The absence of a trusted and robust security framework in the cloud to protect against security threats including access control management, authentication challenges, integrity of stored information that cloud services suffer is a concern that attracts equal attention from industry and academia to explore approaches to address the security vulnerabilities in these services to improve the security of the cloud to support secure remote processing and storage of information [13]-[16].

2.2.2. Internet-of-Things Security

In [17] an extensive survey on the security, privacy and authentication challenges in internet of things confirmed the fact that although there have been. improvement in the design and development of cryptographic solutions tailored for the individual elements within IoT, there is equally an increase in the number of exploitation of attacks in recent times. Therefore, authentication, confidentiality, and data integrity challenges exist in IoT domain making it demand a cryptographic solution that adopts an algorithm that provides efficient privacy and security.

In [18], the authors delved into the material agents used in the design and development of smart objects and internet of things systems. Majority of the devices used in an IoT system have sensing capabilities to independently collect data from its environment for transport to other devices within the network. A defective device by design is a threat by itself and introduces a weak link for attack exploitation to a system that has such a device as component. Compatibility challenges in intelligence sensing, interoperability, distributed intelligence and flexibility to adapt to a universal authentication solution for an IoT ecosystem have become fundamental in implementing IoT security solution.

An efficient encryption scheme that involves less computational overheads and runs on adequate power for constrained devices without compromising on the security, privacy and integrity of data in IoT is needed to in modern IoT security architectures [19]. The algorithm developed in [20] combined machine learning approaches to detect and control network congestion in IoT using the fitness function that is based on the grey wolf optimization algorithm (GWO), since network congestion could increase the execution overheads for cryptographic algorithms to run efficiently.

The intensity and size for IoT investments by enterprises keep increasing to improve the security of their network, as the number and the scale in the incidences of security attack on IoT keep rising [21], [22]. Modern and more sophisticated malware that targets IoT devices and the various components of IoT systems to exploit the security vulnerabilities in these devices to the skewed benefits of those attackers is on the increase [8].

2.2.3. Twofish

The authors in [23] adopted the Twofish encryption algorithm in securing data and maintaining confidentiality of the communication network and implemented the solution using the chilkat encryption activeX. Figure 1[24] illustrates the working of the Twofish encryption algorithm. The encryption involved splitting the input data into four halves of 128 bits, where the XOR operations were conducted on the bits input with a key. The key whitening and the XOR operation provided the needed security to the data to assure privacy of the data in the communication network.

The Twofish encryption algorithm was used in addition to other cryptographic primitives to enhance the security in Bluetooth encryption. The combined encryption approach supported the protection of data in Bluetooth transmission. The use of the Twofish cryptographic algorithm improved the security and efficiency of the encryption scheme in the Bluetooth communication [25].

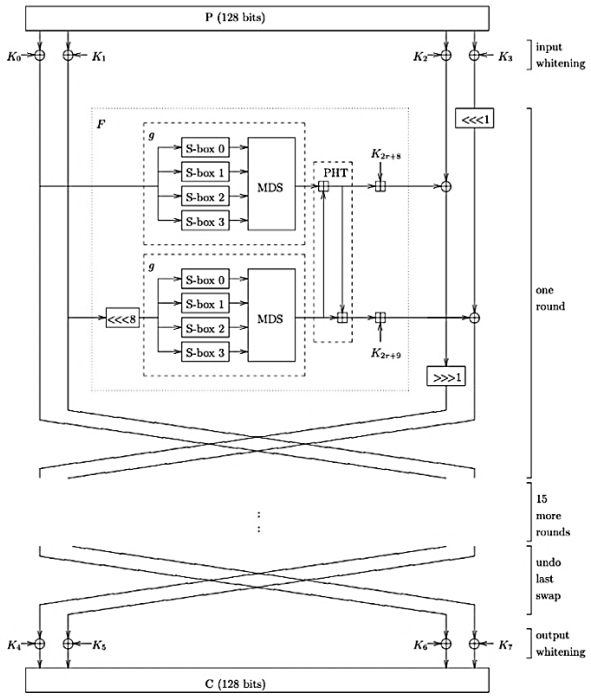


Figure 1: Twofish [24]

The Twofish cryptographic algorithm provided an enhanced security for a software to support secure communication of information over the internet. In their paper, the Twofish algorithm with 192-bits key space was implemented in encrypting messages.

The algorithm provided efficient and secure encryption of data for communication [26], [27].

In [28], a variant of the Twofish algorithm with an increased complexity for multi-level keys adoptable for dynamic bit size inputs to improve its cryptographic strength that made it resistant against the differential attacks was adopted for encrypting images.

The standard Feistel encryption algorithm involving several rounds of encryption with key whitening techniques to improve the security of the encryption is what the Twofish encryption algorithm offers. The input and output data are XOR-ed with eight sub-keys. The encryption in the Twofish encryption algorithm involves encrypting a message over 16-rounds in a Feistel network that uses a bijective function comprising four-byte wide pre-computed key dependent substitution (S) boxes; a matrix; a key schedule, and a Pseudo Hadamard Transform of bitwise rotation. After each round of encryption, the ciphertext generated is swapped and fed as input into the next round such that the left encrypted text is interchanged with the right counterpart and vice versa. The final round of encryption produces a ciphertext in two halves-the right and left halves of cipher texts. The resultant ciphertext for the Twofish encryption is achieved by interchanging the positions of the ciphertexts and combining both as the ciphertext for the message [23], [24]

2.2.4 Diffie-Hellman Key Exchange (DHE)

The communicating nodes through a key generator acquire two sets of key pairs (a private key and a public key) to enable them authenticate for communicating data. The DHE is susceptible to several attacks including key compromise impersonation attacks [23], [29].

The authors in [30] adopted an algorithm that is based on the Linear Feedback Shift Register (LFSR)-dependent correlation technique to supplement the Diffie-Hellman key exchange protocol to maintain the privacy of message communication between the cloud and the local device. The LFSR correlation algorithm detected and verified digital signatures from a digital signature pool between the devices and the cloud by using the correctional framework for digital signature analysis through the calculation of linear complexities between the cloud and the local devices. This technique helped with the detection of errors with digital signatures of messages thereby maintaining the privacy of the information shared between the cloud and the local devices, eliminating the possibility of key compromise or impersonation and the related attacks.

In [31], cryptographic-based public key infrastructure approaches provided at the device levels were adopted to support secure communication of message in an end-to-end encryption that ensured privacy, confidentiality and device integrity. The Elliptic Curve Diffie-Hellman key exchange (ECDH) guaranteed a secure key sharing between communicating devices on a chat application that operated on the android environment. The RC4 encryption scheme was used to encrypt the multimedia messages for communication between the devices used on the chatting application.

The use of authentication scheme based on the Diffie-Hellman model supported the efficient key exchange and management for communication devices with more than one identity. The authors in [32] used a key agreement and management protocol adaptable for IoT communicating devices with more than one identity. The key agreement and management protocol supported the efficient selection and initialization of session key pairs from devices with multiple identities and helped authenticate the communicating devices in the IoT system.

The next section deals with the proposed methodology.

3. Methodology

The proposed methodology is a combined cryptographic primitive consisting of a key exchange protocol or key agreement protocol such as the Diffie-Hellman Key Exchange (DHE) protocol illustrated in figure 2, and a cryptographic cipher for encryption and decryption such as the Twofish cryptographic algorithm illustrated in figure 1.

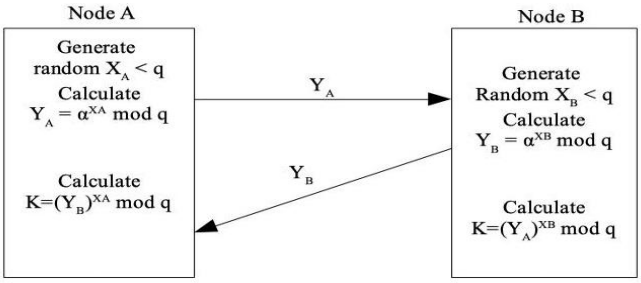


Figure 2: Diffie-Hellman Key Exchange between Node A and Node B.

The Diffie-Hellman Key Exchange (DHE) protocol is a public-key cryptographic primitive with an implementation involving two unknown nodes or devices to securely establish communication by generating and exchanging shared secrets between them. The securely shared secret is used in performing a symmetric key encryption and decryption. The DHE is structured on the discrete logarithm problem which is based on a one-way function; finding the factorization of the product of two prime numbers p and g.

Such that p is a prime number and g is a primitive root mod (p).

where: p and g (generator) are two large integer number.

$1 < n \leqslant (p-1)$; where $n = g^k \bmod p$

Thus, for every number n between 1 and p-1 inclusive, there is a power k of g such that $n=g^k \bmod p$.

Publicly available numbers of p and g are to be used by any two untrusting and unknown devices or nodes to generate their key pairs: namely the public key (Pu) and private key (Pv).

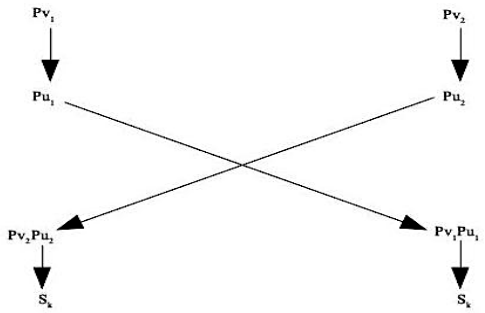


Figure 3: The generator for the shared key in DHE

As shown in Figure 3, the notations for Pv_1 , Pu_1 and Pv_2 , Pu_2 represent the Private key and shared key for the source node and the receiver node respectively. The DHE protocol used four numbers in all. The first two numbers are to generate the shared key (of same value) $Pu_1 = Pu_2$. Thus, both sender and receiver have the same value for the shared key Pu_1.

The DHE and Twofish Algorithm:

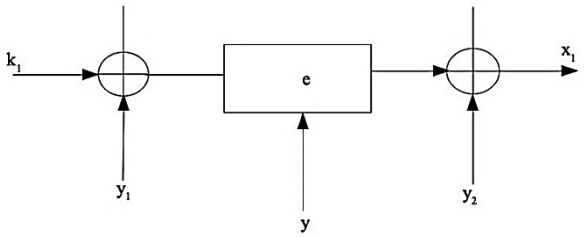


Figure 4: Twofish Encryption Algorithm based on the DHE Protocol

As shown in Figure 4, (k_1, y_1) represents the private key and public key respectively for the source node.

(e), represents the Twofish cryptographic function.

(x_1, y_2) represents the private and shared keys respectively for the receiver node.

The shared key for the communicating nodes is $(y) = y_1 = y_2$.

Plaintext of arbitrary size data up to 128 bits. The input bits are grouped into four sections, thus four parts of each constituting a 32-bit part. The four sections are divided into halves. The first two sections of 32 bits each forms the right component of the input bits and the other two sections of two 32 bits, form the left component [23].

The four keys for whitening the encryption are applied on the Bit-XOR input.

$R_{0,i} = P \oplus K_i \; ; i = 0, \dots, 3$

Where:

R denotes the rounds of encryption in the Feistel network.

K denotes the key (Ki) where i represents the sub-key for whitening. There are four sub-keys. 0, 1, 2, 3.

Encryption: $e_k, k_1, y_1(M), = e_k(M \oplus k_1) \oplus y_1 = C_T$

Decryption: $e_k^{-1}, x_1, y_2(C_T), = e_k^{-1}(C_T \oplus x_1) \oplus y_2 = M$

where:

M represents the plaintext.

C_T denotes the ciphertext or encrypted message.

e_k represents the Twofish encryption function.

e_k^{-1} represents the Twofish decryption function.

k_1 denotes the private key for the source node.

x_1 represents the private key for the receiver node.

y_1 denotes the public key of the source node.

y_2 denotes the public key of the receiver node.

$y_1 = y_2 = y$, denotes the shared secret for encryption and decryption of the DHE-Twofish cryptographic solution.

Encryption at the source node is implemented using the shared key with the Twofish algorithm encryption function.

Algorithm 1: Secure Node-Node Data Exchange - *Encryption* - Twofish (M, S_k)

Input : M, S_k
Output: C_T
Begin
  Split M into four 32-bit parts $M_{i=0,1,2,3}$
  Split S_k into four 32-bit partial keys key $K_{i=0,1,2,3}$
  For i=0..1 **loop**
    $R_{i,0}$ = M XOR K_{2*i}
    $L_{i,0}$ = M XOR K_{2*i+1}
    For j=1..3 **loop**
      $L_{i,j} = R_{i,j-1}$
      $R_{i,j} = L_{i,j-1}$ XOR (((2 $R_{i,j-1}$ * K_j)^ x) % (2^32-1))
    End loop
  End loop
  Combine L_{13}, R_{13}, L_{03} and R_{03} into C_T
  Return C_T
End

Algorithm 2: Secure Node-Node Data Exchange - *Decryption* - Twofish (CT, S_k)

Input : C_T, S_k
Output: M
Begin
  Split C_T into four 32-bit parts $C_{T\ i=0,1,2,3}$
  Split S_k into four 32-bit partial keys key $K_{i=0,1,2,3}$
  For i=0..1 **loop**
    $R_{i,3}$ = C_T XOR K_{2*i}
    $L_{i,3}$ = C_T XOR K_{2*i+1}
    For j=2..0 **loop**
      $R_{i,j} = L_{i,j+1}$
      $L_{i,j} = R_{i,j+1}$ XOR (((2 $L_{i,j+1}$ * K_j)^ x) % (2^32-1))
    End loop
  End loop
  Combine L_{10}, R_{10}, L_{00} and R_{00} into M
  Return M
End

The Algorithm 1 above is the encryption procedure for the Twofish cipher at the source node.

The Algorithm 2 above is used for the decryption of the ciphertext at the receiving node using the Twofish cipher.

In Algorithms 1 and 2:

M represents the plaintext.

S_k shared key, represents the Key K

C_T represents the ciphertext.

K_i represents the partial key or the subkeys where i represents the indexes from (0,1,2,3)

Li and Ri are Left and Right partitioned 64-bits block size
f represents the Twofish cipher.

% represents the Modulo operator.

Li represents the Leftmost partitioned part of the Feistel structure

Ri represents the Rightmost partitioned part of the Feistel structure

x represents the total number of rounds, which is 16 in this case.

4. Results and Discussion

At the source node, using the DHE through a handshake request with the destination node established a shared key for encryption and decryption.

The Twofish encryption uses the data, and the shared key to produce the ciphertext. Decryption of ciphertext at the destination node is also implemented using the Twofish on the shared key and ciphertext.

Table 1: Displayed results

ID	SK	Data	Ciphertext	Recovered
1	262262262	2345	OXEVxUcVBXZk lzbf0F1D+A==	2345
2	232232232	1291	7efDJnVQrVwEV qBD3rguLA==	1291
3	452452452	4672	+gJXYsjBJP/Y/UE djUMIWQ==	4672
4	232232232	1456	RXec4NLFwt1nmx eWqQkm3g==	1456
5	232232232	2121	IeTtTlkRgQuvjD0h AFQFIg==	2121

The ID denotes the unique node identity number which represented the order of arriving data for the encryption. The 1st, 2nd, 3rd, 4th, 5th showed the first data, second data, third data, fourth data, and firth data to arrive in that order.

The SK represents the shared key by the DHE for message encryption.

Data denotes the plaintext.

Ciphertext represents the output of the plaintext with the Twofish cryptographic algorithm.

Recovered Data represents the decrypted data.

The Diffie Hellman key exchange protocol that took the unique ID of the communicating devices for establishing the shared key using the key generator helped for encryption of node data.

Table 2: Average Encryption Throughput (MiB) in the ARMv7-a of Samsung and Xiaomi Devices [33]

Algorithm/ Pack Size	1 MiB	5 MiB	10 MiB
AES	77.539	78.058	77.586
RC6	51.84	53.556	53.1065
Twofish	47.7135	48.6145	47.1835

As presented in Table 2 [33], the average encryption throughput for data packet sizes 1MiB, 5MiB and 10 MiB shows an improved throughput for Twofish algorithm. The management information base (MiB) values catalogued the properties and data objects from the encryption of implementation of AES, RC6 and Twofish respectively for the block ciphers. Across all the pack sizes of the MiB values, the encryption throughput for Twofish was higher than the AES and the RC6.

Table 3: Average Decryption Throughput (MiB) in the ARMv7-a of Samsung and Xiaomi Devices [33]

Algorithm/ Pack Size	1 MiB	5 MiB	10 MiB
AES	66.671	68.541	69.026
RC6	52.0065	53.3635	52.877
Twofish	47.785	48.6215	47.0225

As seen in Table 3 [33], the average decryption throughput for data packet sizes of 1MiB, 5MiB and 10 MiB shows an improved throughput for Twofish algorithm. The management information base (MiB) values catalogued the properties and data objects from the encryption of implementation of AES, RC6 and Twofish respectively for the block ciphers. Across all the pack sizes of the MiB values, the encryption throughput for Twofish was higher than the AES and the RC6.

Table 4: Speed of AES Candidates for different Key Lengths [34]

Algorithm Name	Key Setup	Encryption
MARS [BCD+98]	Constant	Constant
RC6 [BCD+98]	Constant	Constant
Rijndael [DR98]	Increasing	128: 10 rounds 192: 20% slower 256: 40% slower
SERPENT [ABK98]	Constant	Constant
Twofish [SKW+98, SKW+99a]	Increasing	Constant

As shown in Table 4 [34], the speed of the AES candidates for different key lengths for encryption is presented. The Twofish algorithm adopts an increasing key setup but encrypts and decrypts at a speed independent of the length. The performance and security of an encryption is as a function of the key space or key length.

5. Conclusion

The combined cryptographic scheme consisted of the DHE, an algorithm based on mathematical approaches for exchanging a shared secret between the communicating nodes and the Twofish. The Twofish encryption algorithm provided a relatively better encryption time than the AES and RC6. The use of the DHE increased the confusion and diffusion of the key for the encryption algorithm to improve the strength of the cryptographic algorithm of Twofish in eliminating the possibilities of relative key attacks that could result in man-in-the-middle and its associated attacks.

The Twofish cryptographic algorithm complemented by the pre-shared secret key protocol in the DHE key exchange provided authentication for the nodes as well as the creation of a secure channel between the communicating nodes to guarantee a secure

node-to-node exchange of IoT messages. It helped in assuring the integrity of the content of messages between the communicating nodes since the DHE, in generating the shared key, only included and engaged the intended receiver node prior to the actual message communication. The complex key exchange logic provided in DHE algorithm improved the key diffusion in the Twofish cryptographic symmetric key cipher assuring an improved security with a relatively high throughput for end-to-end encryption for secure communication within IoT systems. As presented in Tables 2, 3 and 4 [33], [34], the Twofish encryption algorithm produced an encryption and decryption with a high throughput [32], [33], [35].

An analysis on the throughput, battery drain, key space and its impact on the security of related symmetric key ciphers for IoT constrained devices would be explored for future works.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] B.T. Asare, K. Quist-Aphetsi, L. Nana, "Secure data exchange between nodes in IoT using TwoFish and DHE," Proceedings - 2019 International Conference on Cyber Security and Internet of Things, ICSIoT 2019, 101–104, 2019, doi:10.1109/ICSIoT47925.2019.00024.

[2] G. Strategy, L. Council, "Internet of Things : Where Your Competitors Are Investing Overview," Gartner, 2020.

[3] S. Siboni, V. Sachidananda, Y. Meidan, M. Bohadana, Y. Mathov, S. Bhairav, A. Shabtai, Y. Elovici, "Security Testbed for Internet-of-Things Devices," IEEE Transactions on Reliability, 2019, doi:10.1109/TR.2018.2864536.

[4] C. DeWitt, J. Ellis, "Control system cyber security: Staying ahead of the evolving threats," in Proceedings of the Annual Offshore Technology Conference, 2013, doi:10.4043/24393-ms.

[5] S.M. Saeed, S.S. Ali, O. Sinanoglu, Scan design: Basics, advancements, and vulnerabilities, 2017, doi:10.1007/978-3-319-44318-8_6.

[6] P. Chan, T. Barnett, A.-H. Badawy, P.W. Jungwirth, "Cyber defense through hardware security," 2018, doi:10.1117/12.2302805.

[7] M.R. Mishra, J. Kar, "A Study on Diffie-Hellman Key Exchange Protocols," International Journal of Pure and Apllied Mathematics, **114**(2), 2017, doi:10.12732/ijpam.v114i2.2.

[8] C. Kolias, G. Kambourakis, A. Stavrou, J. Voas, "DDoS in the IoT: Mirai and other botnets," Computer, **50**(7), 80–84, 2017, doi:10.1109/MC.2017.201.

[9] P. Wang, S. Sparks, C.C. Zou, "An advanced hybrid peer-to-peer botnet," IEEE Transactions on Dependable and Secure Computing, 2010, doi:10.1109/TDSC.2008.35.

[10] E.B. Beigi, H.H. Jazi, N. Stakhanova, A.A. Ghorbani, "Towards effective feature selection in machine learning-based botnet detection approaches," in 2014 IEEE Conference on Communications and Network Security, CNS 2014, 2014, doi:10.1109/CNS.2014.6997492.

[11] D. Zhao, I. Traore, B. Sayed, W. Lu, S. Saad, A. Ghorbani, D. Garant, "Botnet detection based on traffic behavior analysis and flow intervals," Computers and Security, 2013, doi:10.1016/j.cose.2013.04.007.

[12] F. Sabahi, "Cloud computing security threats and responses," 2011 IEEE 3rd International Conference on Communication Software and Networks, ICCSN 2011, 245–249, 2011, doi:10.1109/ICCSN.2011.6014715.

[13] H. Sato, A. Kanai, S. Tanimoto, "A cloud trust model in a security aware cloud," Proceedings - 2010 10th Annual International Symposium on Applications and the Internet, SAINT 2010, 121–124, 2010, doi:10.1109/SAINT.2010.13.

[14] A. Behl, "Emerging security challenges in cloud computing: An insight to cloud security challenges and their mitigation," Proceedings of the 2011 World Congress on Information and Communication Technologies, WICT 2011, 217–222, 2011, doi:10.1109/WICT.2011.6141247.

[15] M.D. Ryan, "Cloud computing security: The scientific challenge, and a survey of solutions," Journal of Systems and Software, **86**(9), 2263–2268, 2013, doi:10.1016/j.jss.2012.12.025.

[16] C. Rong, S.T. Nguyen, M.G. Jaatun, "Beyond lightning: A survey on security challenges in cloud computing," Computers and Electrical Engineering, **39**(1), 47–54, 2013, doi:10.1016/j.compeleceng.2012.04.015.

[17] P.M. Chanal, M.S. Kakkasageri, "Security and Privacy in IoT: A Survey," Wireless Personal Communications, **115**(2), 1667–1693, 2020, doi:10.1007/s11277-020-07649-9.

[18] C. Savaglio, M. Ganzha, M. Paprzycki, C. Bădică, M. Ivanović, G. Fortino, "Agent-based Internet of Things : State-of-the-art and research challenges," Future Generation Computer Systems, **102**, 1038–1053, 2020, doi:10.1016/j.future.2019.09.016.

[19] P. Sudhakaran, C. Malathy, "Energy efficient distributed lightweight authentication and encryption technique for IoT security," International Journal of Communication Systems, (August), 1–10, 2019, doi:10.1002/dac.4198.

[20] M.S. Manshahia, "Grey Wolf Algorithm based Energy-Efficient Data Transmission in Grey Wolf Algorithm based Energy-Efficient Data Transmission in Internet of Things Internet of Things," Procedia Computer Science, **160**, 604–609, 2019, doi:10.1016/j.procs.2019.11.040.

[21] I. Lee, K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," Business Horizons, 2015, doi:10.1016/j.bushor.2015.03.008.

[22] A. Mayuri, Bhabad, B. Sudhir, T., "Internet of Things: Architecture, Security Issues and Countermeasures," Future Generation Computer Systems, **10**(6), 1–4, 2015, doi:10.5120/ijca2015906251.

[23] M.A. Muslim, B. Prasetiyo, Alamsyah, "Implementation twofish algorithm for data security in a communication network using library chilkat encryption activex," Journal of Theoretical and Applied Information Technology, **84**(3), 370–375, 2016.

[24] S. Bruce, J. Kelsey, D. Whiting, D. Wagner, C. Hall, N. Ferguson, Twofish : A 128-Bit Block Cipher, John Wiley & Sons Inc, 1999.

[25] M.A. Albahar, O. Olawumi, K. Haataja, P. Toivanen, "Novel Hybrid Encryption Algorithm Based on Aes, RSA, and Twofish for Bluetooth Encryption," Journal of Information Security, **09**(02), 168–176, 2018, doi:10.4236/jis.2018.92012.

[26] P. Gehlot, S. R. Biradar, B. P. Singh, "Implementation of Modified Twofish Algorithm using 128 and 192-bit keys on VHDL," International Journal of Computer Applications, 2013, doi:10.5120/12024-8087.

[27] D. Smekal, J. Hajny, Z. Martinasek, "Hardware-Accelerated Twofish Core for FPGA," in 2018 41st International Conference on Telecommunications and Signal Processing, TSP 2018, 2018, doi:10.1109/TSP.2018.8441386.

[28] S.M. Kareem, A.M.S. Rahma, "A novel approach for the development of the Twofish algorithm based on multi-level key space," Journal of Information Security and Applications, 2020, doi:10.1016/j.jisa.2019.102410.

[29] N. Li, "Research on diffie-hellman key exchange protocol," ICCET 2010 - 2010 International Conference on Computer Engineering and Technology, Proceedings, **4**(4), V4-634-V4-637, 2010, doi:10.1109/ICCET.2010.5485276.

[30] M.P. Rewagad, M.Y. Pawar, "Use of digital signature with diffie hellman key exchange and aes encryption algorithm to enhance data security in cloud computing," Proceedings - 2013 International Conference on Communication Systems and Network Technologies, CSNT 2013, 437–439, 2013, doi:10.1109/CSNT.2013.97.

[31] M. Begum, M.A. Waheed, "Protecting Data Privacy in Cloud," International Journal of Engineering and Advanced Technology (IJEAT), (9), 2249–8958, 2020, doi:10.35940/ijeat.C6096.029320.

[32] A. Ali, A. Sagheer, "Design of Secure Chatting Application with End to End Encryption for Android Platform," Iraqi Journal for Computers and Informatics, 2017, doi:10.25195/ijci.v43i1.73.

[33] D.A.F. Saraiva, V.R.Q. Leithardt, D. de Paula, A.S. Mendes, G.V. González, P. Crocker, "PRISEC: Comparison of symmetric key algorithms for IoT devices," Sensors (Switzerland), **19**(19), 1–23, 2019, doi:10.3390/s19194312.

[34] B. Schneier, J. Kelsey, D. Whiting, D. Wagner, C. Hall, N. Ferguson, "On the twofish key schedule," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **1556**, 27–42, 1999, doi:10.1007/3-540-48892-8_3.

[35] A. Ghosh, "Comparison of Encryption Algorithms : AES , Blowfish and Twofish for Security of Wireless Networks," International Research Journal of Engineering and Technology, (June), 4656–4659, 2020, doi:10.13140/RG.2.2.31024.38401.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 626-630 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Cardiovascular Risk in Patients who go to the Medical Office of a Private Health Center in North Lima

Jairo Zegarra-Apaza[1], Sara Oliveros-Huerta[1], Santiago Vilela-Cruz[1], Rosita Chero-Benites[1], Gissett Marcelo-Ruiz[1], Leslie Yelina Herrera-Nolasco[1], Brian Meneses-Claudio[2], Hernan Matta-Solis[1,*], Eduardo Matta-Solis[3]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[3]*Health Sciences, Instituto Peruano de Salud Familiar, 15304, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 12 January, 2021
Online: 30 January, 2021

Keywords:
Adolescents
Quality of life
Students
Health
Adolescent psychology

ABSTRACT

Cardiovascular diseases are the group of conditions produced in the heart or blood vessels. This is one of the main causes of death in Peru and the world, produced mostly by non-communicable diseases and harmful habits, which makes it an extremely predictable disease. These factors include body mass index, smoking, diabetes, age, blood pressure, total cholesterol, and high-density lipoproteins. Therefore, this study aims to identify patients who go to the medical office of a private health center in North Lima who do not have a prior history of a cardiovascular accident, using the cardiovascular risk calculator provided by the Organization World Health. The present research work had a quantitative, non-experimental, descriptive, and cross-sectional approach, in a population of 99 adult and elderly patients. Regarding the results, it was found that 46.5% presented a low cardiovascular risk, 37.4% a moderate risk, 11.1% a high risk and 5.1% an extremely high risk. The information found contrasts with the number of deaths caused by this disease and may be an indicator of greater prevention by populations with higher economic income. Finally, it is concluded that diabetes, smoking and the age group are predisposing factors to an increased cardiovascular risk.

1. Introduction

Cardiovascular risk is specified as the probability of a heart or blood vessel condition occurring [1]. The origin of cardiovascular diseases (CVD) is closely linked to harmful habits or non-communicable diseases, from which derives its progressive and highly predictable nature [2].

According to the Framingham Heart Study, conducted from 1948 to the present, the main predictors of a CVD are age, body mass index, HDL and total cholesterol, smoking, diabetes, and treated high systolic blood pressure or untreated [3], [4]. Therefore, these predictors are also connected to illnesses and harmful habits, present and increasing mainly in developed and developing countries.

In Peru, according to data collected by the Instituto Nacional de Estadística e Informática (INEI), from 2015 to 2018 the Peruvian population has had an annual increase in cases with high blood pressure (14.8%), diabetes mellitus (3.9%) and obesity (23%), which are mainly concentrated in Lima, while cases of overweight (37%), cigarette consumption (19%) and daily cigarette consumption (1.7%) make up almost constant in the years of study [5]. Likewise, its importance in relation to cardiovascular risk (CVR), death attributed to diabetes also increased [6], which gives an indication of inefficient control of the disease. Consequently, as mentioned above, each year the cardiovascular disease (CVD) threat is amplified and the people with the longest longevity are affected.

On the other hand, in [7], a study was carried out in Peruvian youth and adults between 20 and 59 years old who were users of popular dining rooms, these premises being focused on food supply in areas of poverty and extreme poverty. The research found a prevalence of abdominal obesity (51.6%), low HD-C (42.2%), hypertriglyceridemia (35.3%), overweight (35.8%), high blood pressure (21.0%) and hyperglycemia (14.1%), thus conforming an overall metabolic syndrome of 40.1%. This not only allows to infer that the state of this type of population makes them prone to CVD, but also without a preventive measure that takes an active role in the search for these patients, the risk will

*Corresponding Author: Hernan Matta-Solis, Mr., +51 1 999751065 & hmatta@uch.edu.pe

probably not be reduced until they are close to suffering a condition of this type, because despite having insurance focused on low-income populations, regular visits in search of prevention are not usually integrated in Peru.

In the study [8], carried out in a public hospital for internal medicine inpatients, vascular age and cardiovascular risk were evaluated in 238 participants who were between 30 and 74 years old without previous cardiovascular events. The objective was to determine these variables as well as the predominant risk factor. This study concluded with a medium and high cardiovascular risk in the highest percentage of the population, a difference between chronological and vascular age of 6.9 years old, and finally, diabetes and male sex were identified as predominant factors.

According [9], a research carried out in Spain, they were able to note that heart disease is the main cause of morbidity in a tertiary hospital and men between 50 and 60 years old are the ones with the highest morbidity in this pathology, causing high rates of incapacity for work, in addition the problem goes beyond the hospital setting because these patients will require cardiac rehabilitation. They also point out that emphasis should be placed on prevention, since studies have documented that of the patients who were discharged, 75% of patients re-entered for any cardiac complication.

In [10], a study was carried out in an educational institution of 100 schoolchildren from 10 to 19 years old, with the aim of determining cardiovascular risk. The study found a prevalence of cardiovascular risk of 84.6% in adolescents (15-18 years old), as well as that the consumption of alcoholic beverages and unhealthy eating habits are related to the increase in cardiovascular risk. So, it is suggested that prevention from an early age is essential to reduce its appearance in more advanced stages of life.

The objective of this research work is to identify cardiovascular risk in patients with no prior history of CVD, excluding hypertension in untreated cases, and who go to a cardiology office in a private health center, using a calculator provided by the Organization. Pan American Health (PAHO). Because of having an alteration in the cardiovascular system that is not a serious alteration, patients of the cardiology service are one of the groups to be suspected of having a cardiovascular condition that may cause irreversible damage or death of the patient.

The instrument used in this study is the CVR calculator provided by PAHO, based on a form proposed by the World Health Organization (WHO) to evaluate CVR in Latin American and Caribbean countries [11]. This instrument uses a risk score that is based on the Framingham study, which also is adapted to the characteristics of the region and these data will be processed in matrix form by the IBM SPSS statistics 25 program. This application deals with to assess CVR quickly and easily in patients, which is mainly focused as an aid to medical health personnel, but due to the simplicity of its use, this can be a great option for the recruitment of risk patients who are in areas that are difficult to access or that lack medical personnel.

The organization of the work carried out is structured in the following way: in Section II, the development carried out for the identification of cardiovascular risk in patients who go to the medical office will be found. In section III, it will find the results of the data collected for the measurement, as well as the most relevant values of the CVR found. In section IV, the discussion of the research work and finally in section V, the conclusions, as well as the expectations that are had for the use of this research work.

2. Methodology

This section is made up of the process followed for the cardiovascular risk assessment, which are: data acquisition, data processing and finally the results with their corresponding risk values. The steps in carrying out the cardiovascular risk assessment are represented in a flowchart in Figure 1. All this was passed through the ethics committee, to give the authorization for the research work, followed by the ethical principles in part for the study in patients, such as beneficence, autonomy, non-maleficence, and justice.

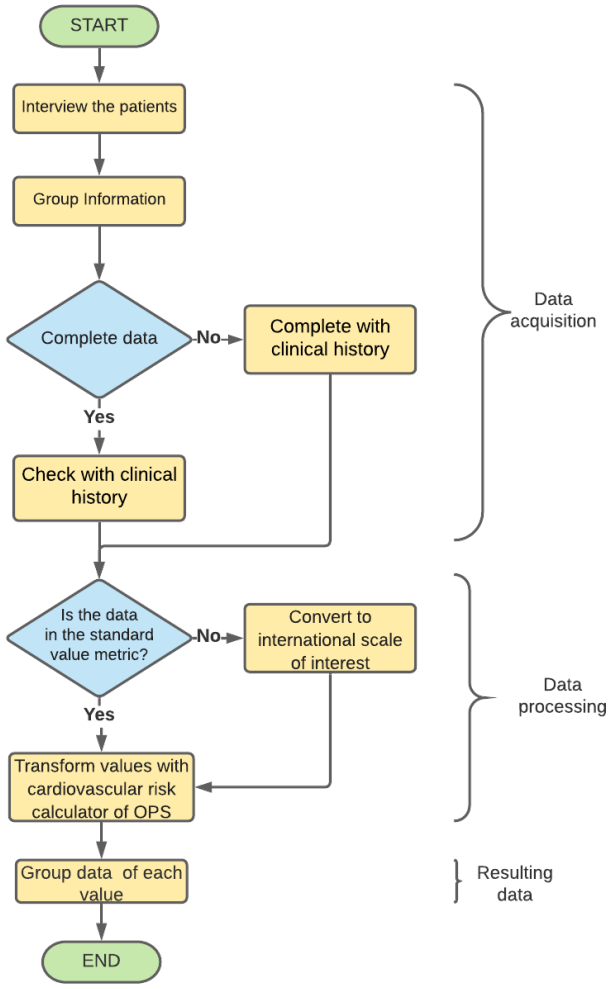


Figure 1: Flowchart for cardiovascular risk data collection

2.1. Data Acquisition

This study has a quantitative, non-experimental, descriptive, and cross-sectional approach. For the acquisition of the data, coordination was required with a private polyclinic located in the jurisdiction of North Lima, in which the patients of the cardiology service were evaluated in the first quarter of the year. Likewise, the data collected in the research work was cataloged in

sociodemographic data of the patient and descriptive parameters of cardiovascular risk, the subdivisions are shown in Table 1.

Table 1: Sociodemographic data and descriptive parameters of cardiovascular risk of people from a cardiology clinic of a Private Polyclinic in North Lima

Sociodemographic Data	Gender
	Degree of instruction
	Occupation
Descriptive parameters of cardiovascular risk	Maximum / Systolic Pressure
	Smoking
	Total cholesterol
	Diabetes
	Age

Population and Sample

The population is made up of 110 adult patients (49-59) and older adults (60 or more) from the jurisdiction of the health center, but the sample size was made reaching 99 since the remaining patients did not want to participate voluntarily in the study.

Inclusion and Exclusion Criteria

The inclusion criteria that were considered at the time of data collection were: Patients who are between the ages of 49 and 79, who have high blood pressure, who are continuing or go to medical office more than once, who belong to the jurisdiction of the polyclinic and agree to participate voluntarily in the study.

In addition, the exclusion criteria corresponding to the study were: patients who depended on a wheelchair for mobilization and those who suffered from cognitive or hearing impairment.

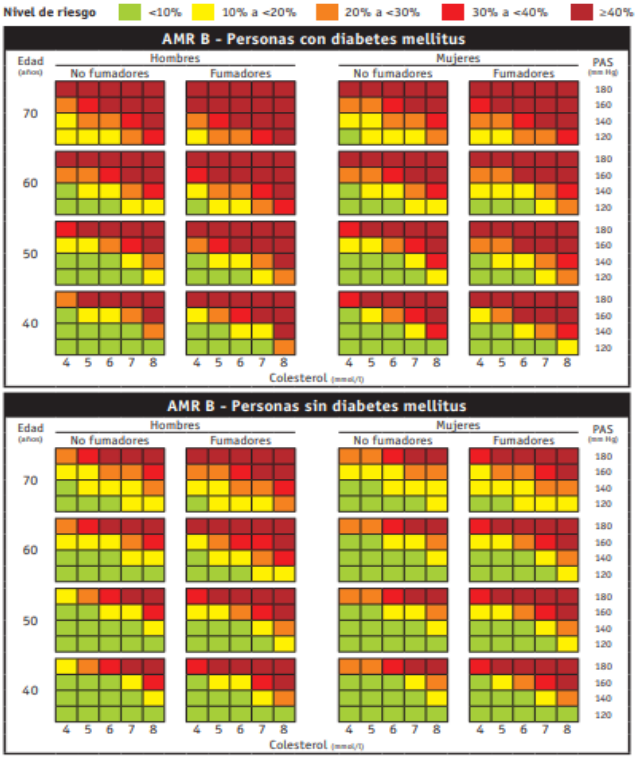


Figure 2: WHO/ISH AMR B risk prediction table, for settings in which blood cholesterol can be measured

2.2. Data Processing

The data corresponding to the descriptive parameters of cardiovascular risk were transformed into the international system of measurement units that were adequate or preferable for the measurement of each parameter and subsequently entered the cardiovascular risk calculator provided by PAHO that takes place during the month of November. These data were cataloged by the calculator according to the percentage of CVR of the patient suffering from a cardiovascular condition in the subsequent 10 years corresponding to the table displayed in Figure 2.

The table displayed in Figure 2 subdivides cardiovascular risk into 5 subsectors, relating them to one color according to the range of cardiovascular risk that occurs, these being: green, for a CVR <10%, yellow for a CVR from 10% to <20%, orange from 20% to <30%, red from 30% to <40% and dark red for values> 40%.

Likewise, the acquired data were entered in matrix form to the IBM SPSS Statistics 25 software, for subsequent management and analysis. The variables considered for data analysis are ID, gender, degree of institution, family type, occupation, smoking, diabetes, maximum systolic pressure, total cholesterol, cardiovascular risk, age, and Age 2, these two being the last ages corresponding to the adult and older adult age groups.

The participation of health personnel in the process of obtaining data is significant due to the complications that may arise due to the patient's lack of knowledge regarding the topic corresponding to each variable, the patient's mistrust of a person who do not work in the public health service, as well as greater efficiency in the collection of data in the clinical history and in the reliability when measuring the value of blood pressure.

3. Results

99 adult and elderly patients from the cardiology service were evaluated in the first quarter of the year. Sociodemographic data and descriptive parameters of cardiovascular risk were identified in each of them, as shown in Tables 2, 3 and 4.

Table 2: Sociodemographic data of patients from a cardiology clinic of a Private Polyclinic in North Lima

Sociodemographic data	**N**	**%**
Age [1/]		
Adult	57	57.6%
Older adult	42	42.4%
Gender		
Female	35	35.4%
Male	64	64.6%
Degree of instruction		
Illiterate	2	2.0%
Primary	16	16.2%
Secondary	73	73.7%
Technical	5	5.1%
University	3	3.0%
Occupation		
Does not work	17	17.2%
Retired	10	10.1%
Laborer	12	12.1%

Employee	15	15.2%
Informal work	45	45.5%
Total	**99**	**100.0%**

[1/]Mean = 57.7; Standard Deviation (SD) = 11.08

These results indicate that the mean age of the patients was 57.7 years, the majority being adults representing 57.6%, likewise, 64.6% were men, 73.7% had a secondary level, 45.5% worked informally and 35.4% have an extended family as shown in Table 2.

Table 3: Cardiovascular risk parameters in patients of a cardiology office of a Private Polyclinic in North Lima

Cardiovascular risk parameters	□ ± SD [1/]	
Maximum Systolic Pressure (mmHg)	132.9	11.6
Total Cholesterol (mg/dL)	225.6	22.4
Smoking	**N**	**%**
Yes	52	52.5%
No	47	47.5%
Diabetes		
Yes	84	84.8%
No	15	15.2%

[1/] Mean (□); Standard Deviation (SD)

Table 3 shows that the mean maximum systolic pressure was 132.9 mmHg, and the mean total cholesterol was 225.6 mg/d. On the other hand, 52.5% were smokers and 84.8% had diabetes.

Table 4: Cardiovascular risk and descriptive parameters of cardiovascular risk

			Cardiovascular Risk				Total
			Low	Moderate	High	Very High	
Age group	Adult	N	34	20	3	0	57
		%	34.3 %	20.2%	3.0%	0.0%	57.6%
	Older adult	N	12	17	8	5	42
		%	12.1 %	17.2%	8.1%	5.1%	42.4%
Smoking	Yes	N	15	24	10	3	52
		%	15.2 %	24.2%	10.1%	3.0%	52.5%
	No	N	31	13	1	2	47
		%	31.3 %	13.1%	1.0%	2.0%	47.5%
Diabetes	Yes	N	34	36	10	4	84
		%	34.3 %	36.4%	10.1%	4.0%	84.8%
	No	N	12	1	1	1	15
		%	12.1 %	1.0%	1.0%	1.0%	15.2%

Table 4 shows that: adults are 57.6%, with 34.3% belonging to a low CVR, while older adults are 42.4%, and their CVR is mostly moderate with 17.2%; those who smoke make up 52.5% of the population, their CVR being mostly moderate with 24.2%, while those who do not smoke make up 47.5% and their CVR is mostly low with 31.3%; finally, those with diabetes make up 84.8% of the population and their CVR tends to be moderate with 36.4% or low being 34.3%, while those without diabetes make up 15.2% and most of them have a low CVR representing the 12.1%. These data help to deduce the trend of cardiovascular risk in relation to its risk factors in private health center.

On the other hand, the chi square results were 0.001 for the age group and cardiovascular risk, 0.027 for diabetes and cardiovascular risk, and 0.001 for smoking and cardiovascular risk for P <0.05, confirming that the differences in the proportions observed between the two groups are statistically significant. These results allow to identify the relationship of risk factors with cardiovascular risk in populations that go to a private health center, as well as to differentiate their situation from the current situation in Peru.

Table 5: Cardiovascular risk in patients of a cardiology office of a Private Polyclinic in North Lima

Cardiovascular Risk	**N**	**%**
Extremely high Risk	5	5.1%
High Risk	11	11.1%
Moderate Risk	37	37.4%
Low Risk	46	46.5%
Total	**99**	**100.0%**

Table 5 shows that 46.5% of adult and elderly patients presented a low cardiovascular risk, 37.4% a moderate risk, 11.1% a high risk and 5.1% an extremely high risk.

The knowledge of these results is important because it will allow to identify cardiovascular risk in patients with no prior history of CVD, and it also helps to know the risk factors to which many people, such as smokers, are exposed. This information allows to realize that the greater control there is over risk factors, whether through a change in lifestyle (healthy eating, playing sports, etc.), the less likely there is to have cardiovascular risk.

4. Discussion

The results confirm that there is a significant relationship between smoking, age group and diabetes with cardiovascular risk, thus agreeing with the findings of the Framingham heart study.

On the other hand, the prevalence of CVR found is mostly belonging to a low-risk one, in other words, a population that has a less than 10% risk of suffering cardiovascular disease in the next 10 years, which contrasts significantly with the analysis carried out in [12], which lists circulatory diseases as the third cause of death for large groups and cardiac ischemia as the second cause of death. As it is a private health center, it can indicate a greater emphasis on prevention by a sector of the population with higher economic incomes.

Peru is a developing country where child malnutrition and respiratory diseases are still considered highly relevant problems and their financing is a high priority. For this reason, this research is not only focused on identifying the status of patients entering

cardiology offices, but also on the scientific contribution, because its use is also expected to promote the use of this instrument for the early recruitment of patients, both by health personnel and residents through self-evaluation, because due to its easy access and simplicity it can mean a great advance for the prevention of cardiovascular diseases, in a country where, as declared in [13], self-knowledge, adherence or control of hypertension are still below to other Latin American countries.

5. Conclusions

It is concluded that in patients who go to the medical office of the private health center in North Lima, diabetes, smoking and the age group are predisposing factors to increased cardiovascular risk in patients.

It is also concluded that the majority were adults, with a secondary institution degree and who work informally. As well as 5 out of 10 patients were smokers and 8 out of 10 had diabetes.

Finally, it is concluded that cardiovascular diseases are made up of a series of characteristics that can be regulated to prevent a cardiovascular risk, such as smoking, diabetes, total cholesterol, and maximum systolic pressure. As well as that the problem increases in relation to the old age of the population and therefore measures should be taken focused on these age groups.

This research work was studied to contribute to the lack of scientific information with the use of the PAHO cardiovascular risk calculator, in addition to promoting its use in the Peruvian community nursing setting for the recruitment of risk patients, because despite having problems such as malnutrition in the child population, the obstacle that CVDs pose to life expectancy in adults and older adults cannot be ignored.

The Strength in our research work is teamwork to be able to carry out this research work

The limitation in the research work is access to people to be in the study, since many of them mentioned that they preferred to enter their consultations than to be in the study.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] M. Arboleda, A. García, "Riesgo cardiovascular: análisis basado en las tablas de Framingham en pacientes asistidos en la unidad ambulatoria 309, IESS – Sucúa," Revista Med, **25**(1), 20–30, 2017, doi:10.18359/rmed.1949.

[2] J. Álvarez, A. Álvarez, W. Carvajal, M. González, J. Duque, O. Nieto, "Determinación del riesgo cardiovascular en una población.," Revista Colombiana de Cardiologia, **24**(4), 334–341, 2017, doi:10.1016/j.rccar.2016.08.002.

[3] R. D'Agostino, R. Vasan, M. Pencina, P. Wolf, M. Cobain, J. Massaro, W. Kannel, "General cardiovascular risk profile for use in primary care: The Framingham heart study," Circulation, **117**(6), 743–753, 2008, doi:10.1161/CIRCULATIONAHA.107.699579.

[4] J. Alvarez, V. Bello, G. Pérez, O. Antomarchi, M. Bolívar, "Factores de riesgo coronarios asociados al infarto agudo del miocardio en el adulto mayor.," Medisan, **17**(1), 54–60, 2013.

[5] Instituto Nacional de Estadistica e Informática, Peru, Enfermedades No Transmisibles Y Transmisibles, 2018.

[6] N. Atamari, M. Ccorahua, A. Taype, C. Mejia, "Mortalidad atribuida a diabetes mellitus registrada en el Ministerio de Salud de Perú, 2005-2014," Revista Panamericana de Salud Pública, **42**, 1–7, 2018, doi:10.26633/rpsp.2018.50.

[7] K. Adams, J. Chirinos, "Prevalencia de factores de riesgo para Síndrome Metabólico y sus componentes en usuarios de comedores populares en un distrito de Lima, Perú," Revista Peruana de Medicina Experimental y Salud Publica, **35**(1), 39–45, 2018, doi:10.17843/rpmesp.2018.351.3598.

[8] J. Mayta, A. Morales, A. Cárdenas, J. Mogollón, V. Armas, L. Neyra, C. Ruíz, "Determinación de riesgo cardiovascular y edad vascular según el score de Framingham en pacientes del Hospital Nacional Arzobispo Loayza," Horizonte Med, **15**(2), 27–34, 2015, doi:10.1520/stp30190s.

[9] A. Sánchez, M. Bobadilla, B. Dimas, M. Gómez, G. González, "Enfermedad cardiovascular: primera causa de morbilidad en un hospital de tercer nivel.," Revista Médica de Cardiologia, **27**(3), 98–102, 2016.

[10] C. Torres, D. Illera, D. Acevedo, M. Cadena, L. Meneses, P. Ordoñez, L. Pantoja, M. Pastás, "Riesgo cardiovascular en una población adolescente de Timbío, Colombia," Revista de La Universidad Industrial de Santander. Salud, **50**(1), 59–66, 2018, doi:10.18273/revsal.v50n1-2018006.

[11] OPS/OMS, Calculadora de riesgo cardiovascular, Organización Panamericana de La Salud - Uruguay,.

[12] Ministerio de la Salud, ASIS-R: Análisis de Situación de Salud de los Gobiernos Regionales, 120, 2019.

[13] P. Herrera, J. Pacheco, G. Valenzuela., G. Málaga, "Autoconocimiento, Adherencia al tratamiento y control de la Hipertensión Arterial en el Perú: una versión narrativa.," Revista Peruana de Medicina Experimental y Salud Publica, **34**(3), 497–504, 2017, doi:10.17843/rpmesp.2017.343.2622.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1378-1383 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A-MnasNet and Image Classification on NXP Bluebox 2.0

Prasham Shah*, Mohamed El-Sharkawy

IoT Collaboratory at IUPUI, Department of Electrical and Computer Engineering, Purdue School of Engineering and Technology, Indianapolis, 46202, USA

ARTICLE INFO

Article history:
Received: 28 October, 2020
Accepted: 08 February, 2021
Online: 28 February, 2021

Keywords:
Deep Learning
A-MnasNet
NXP Bluebox 2.0
Cifar-10

ABSTRACT

Computer Vision is a domain which deals with the challenge of enabling technology with vision capabilities. This goal is accomplished with the use of Convolutional Neural Networks. They are the backbone of implementing vision applications on embedded systems. They are complex but highly efficient in extracting features, thus, enabling embedded systems to perform computer vision applications. After AlexNet won the ImageNet Large Scale Visual Recognition Challenge in 2012, there was a drastic increase in research on Convolutional Neural Networks. The convolutional neural networks were made deeper and wider, in order to make them more efficient. They were able to extract features efficiently, but the computational complexity and the computational cost of those networks also increased. It became very challenging to deploy such networks on embedded hardware. Since embedded systems have limited resources like power, speed and computational capabilities, researchers got more inclined towards the goal of making convolutional neural networks more compact, with efficiency of extracting features similar to that of the novel architectures. This research has a similar goal of proposing a convolutional neural network with enhanced efficiency and further using it for a vision application like Image Classification on NXP Bluebox 2.0, an autonomous driving platform by NXP Semiconductors. This paper gives an insight on the Design Space Exploration technique used to propose A-MnasNet (Augmented MnasNet) architecture, with enhanced capabilities, from MnasNet architecture. Furthermore, it explains the implementation of A-MnasNet on Bluebox 2.0 for Image Classification.

1 Introduction

Computer Vision is becoming an essential application in this modern world. With advances in technology autonomous cars, drones and UAVs, robots etc have been enabled with vision capabilities. These technologies use Convolutional Neural Networks (CNNs) to process the images or video input. They are used for applications like Image Classification, Object Detection and Semantic Segmentation etc. CNNs are a part of Deep Learning, which is a subset of Machine Learning (ML). Due to the advances in AI, ML capabilities increased and this enabled a whole new field of Deep Learning. This field deals with creating, optimizing and implementing algorithms which enables technology to become self-reliant and gain human level precision. The prime goal is automation of these technologies in a way that they are able to operate perfectly without any human intervention.

Convolutional Neural Networks are used in developing healthcare technologies, which have almost human level precision and are used to save lives in hospitals. Medical imaging devices which, detect even minute particles, are used for diagnosis of various diseases. Autonomous cars which will make roads safer and will reduce fatal life-threatening accidents, Autonomous Drones which will revolutionize the logistics and will deliver packages more efficiently, UAVs which will aid militaries with surveillance and help prevent wars, smart cameras which will be able to recognize people and track their activities to reduce crimes. Smart imaging of atmosphere for weather prediction and warnings for natural calamities like tsunami, tornadoes etc. These are a few examples of how CNNs are making a major impact in our lives.

These technologies require computational power, speed, accuracy and precision. In order to make them efficient, the algorithms have to be fast having low latency, working efficiently on low power, being more accurate and consuming less memory. CNNs have so many layers so as the architectures become deeper and wider, giving more accuracy, their computational cost increases. These CNN models have to be more compact and should work as efficiently as the novel architectures. With new research and advent of new algorithms, now it is possible to make CNNs more efficient having

*Corresponding Author: Prasham Shah, Email: pashah@purdue.edu

competitive size and accuracy. This research aims to contribute towards this same goal, making CNNs compact in terms of model size and increasing its accuracy so that they can be used for such applications[1].

A-MnasNet [2] is derived by Design Space Exploration of MnasNet. New algorithms were implemented for more efficiency with competitive size and accuracy. Furthermore, it is deployed on NXP Bluebox 2.0 for Image Classification [3]. RTMaps Studio is used to deploy the model on the hardware by establishing a TCP/IP connection.

In this paper, section 2 discusses MnasNet architecture. Section 3 introduces A-MnasNet [2] and discusses the new algorithms implemented to achieve better results. Section 4 demonstrates deployment of A-MnasNet [2] on NXP Bluebox 2.0 for Image classification [3]. Section 5 states the hardware and software requirements to conduct this research. Section 6 illuminates the results that demonstrate the model comparison, model size and accuracy trade-offs and deployment results. Section 7 is conclusion which gives the gist of this research.

2 Prior Work

This section gives an insight on MnasNet [4] architecture. It explains the features of this architecture and what can be improved in order to make it more efficient.

2.1 MnasNet Architecture

MnasNet [4] was introduced in 2019 by Google Brain Team with a goal of proposing an approach of neural architecture search which would optimize the overall accuracy of the architecture and reduce the latency of the model on mobile devices. MnasNet outperformed all existing novel architectures on ImageNet dataset with enhanced accuracy. This architecture extracts features by using depthwise separable convolutional layers with 3*3 and 5*5 filter dimensions and mobile inverted bottleneck layers. It has ReLU [5] activation layers which introduces non-linearity to the network. It uses Batch normalization to adjust and scale the output of activation layers. Dropout regularization is used to reduce the overfitting in the network.

Figure 1 shows MnasNet architecture and illuminates its convolutional layers. This architecture was designed for ImageNet dataset so the input dimensions were set to 224*224*3. It had a validation accuracy of 75.2% and outperformed all other existing state-of-the-art architectures. For this research, MnasNet was used with CIFAR-10 [6] dataset so the input dimensions were modified to 32*32*3. Table 1 shows various layers of the architecture.

An important aspect to observe about MnasNet is that it only explores the channel dimensions of the feature maps throughout its architecture. The convolutional layers fail to explore the spatial feature scale in the network. The activation layers used in the architecture have become outdated and today there are better algorithms which work more efficiently then the used activations.

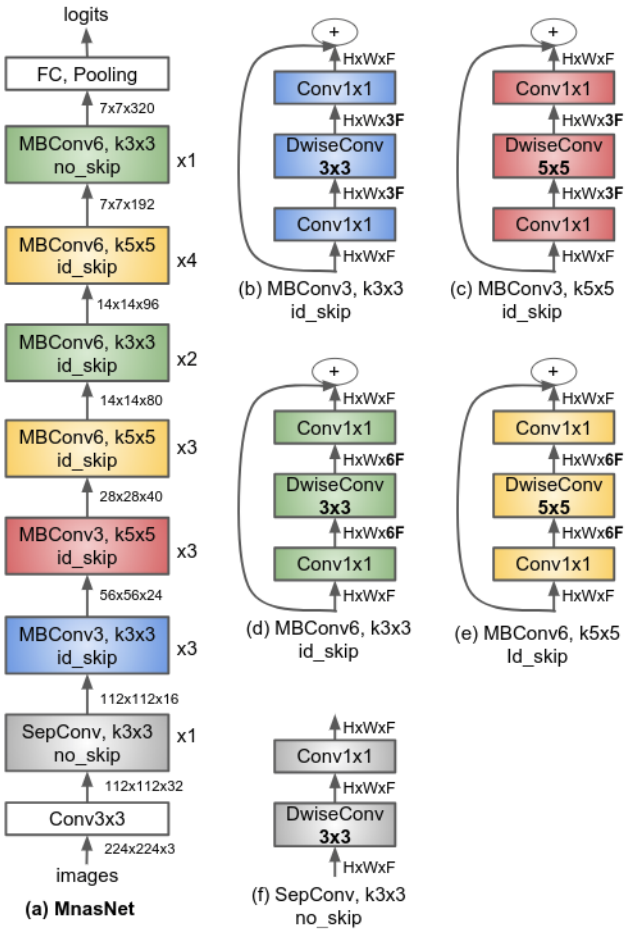


Figure 1: MnasNet architecture

Table 1: MnasNet Architecture

MnasNet Architecture					
Layers	Convolutions	t	c	n	s
$32^2 \times 3$	Conv2d 3×3	-	32	1	1
$112^2 \times 32$	SepConv 3×3	1	16	1	2
$112^2 \times 16$	MBConv3 3×3	3	24	3	2
$56^2 \times 24$	MBConv3 5×5	3	40	3	2
$28^2 \times 40$	MBConv6 5×5	6	80	3	2
$14^2 \times 80$	MBConv6 3×3	6	96	2	1
$14^2 \times 96$	MBConv6 5×5	6	192	4	1
$7^2 \times 192$	MBConv6 3×3	6	320	1	1
$7^2 \times 320$	FC,Pooling			10	

t: expansion factor, c: number of output channels, n: number of blocks and s: stride

3 A-MnasNet Architecture

This section gives an insight on the A-MnasNet [2] CNN architecture and its features.

3.1 Convolution Layers

Harmonious Bottleneck Layers [7] were introduced in the architecture. This special convolutional layers extract features across the spatial and channel dimensions. They are more efficient then the depthwise separable convolutional layers. Harmonious layers consist of spatial contraction-expansion keeping the number of channels constant and channel expansion-contraction keeping the spatial dimensions constant. This down-scaling reduces the number of parameters of the model. Hence, the model size is reduced. Since it extracts features more efficiently, the overall model accuracy is enhanced.

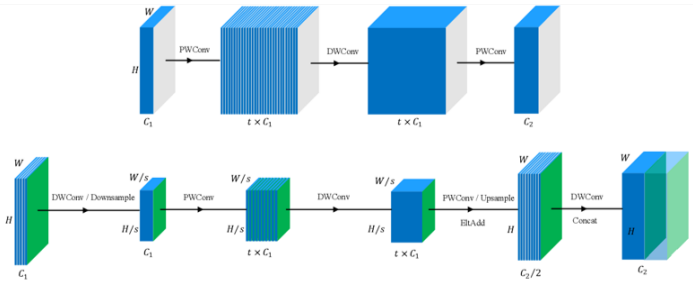


Figure 2: Comparison of Depthwise Separable Convolution Layer and Harmonious Bottleneck Layer.[7]

A comparison between Harmonious bottleneck layers [7] and depthwise separable layers is demonstrated in Figure 2. Here, Input size and output size is represented by H and W respectively. C1 and C2 denote the input and output feature channels where K is the size of the filter and s is the stride value.
The total cost of depthwise separable convolution is

$$(H \times W \times C1 \times K \times K) + (H \times W \times C1 \times C2) \quad (1)$$

The total cost of harmonious bottleneck layer is

$$B/s^2 + (H/s \times W/s \times C1 + H \times W \times C2) \times K^2 \quad (2)$$

where, B is the computational cost of the blocks inserted between the spatial contraction and expansion operations.

Dowmsampling the channel dimensions and extracting features from the spatial dimensions results in reduced spatial size of wide feature maps. The parameters of the model are reduced by a factor of the stride value. As a result the model size reduces and the accuracy of the model increases.

Hence, these layers were added to enhance the performance of the model. Table 2 shows various layers of A-MnasNet [2].

Table 2: A-MnasNet Architecture

A-MnasNet Architecture					
Layers	Convolutions	t	c	n	s
$32^2 \times 3$	Conv2d 3×3	-	32	1	1
$112^2 \times 32$	SepConv 3×3	1	16	1	2
$112^2 \times 16$	MBConv3 3×3	3	24	3	2
$112^2 \times 24$	Harmonious Bottleneck	2	36	1	1
$56^2 \times 36$	MBConv3 5×5	3	40	3	2
$112^2 \times 40$	Harmonious Bottleneck	2	72	1	2
$28^2 \times 72$	MBConv6 5×5	6	80	3	2
$112^2 \times 80$	Harmonious Bottleneck	2	96	4	2
$14^2 \times 96$	MBConv6 3×3	6	96	2	1
$14^2 \times 96$	MBConv6 5×5	6	192	4	1
$7^2 \times 192$	MBConv6 3×3	6	320	1	1
$7^2 \times 320$	FC,Pooling			10	

t: expansion factor, c: number of output channels, n: number of blocks and s: stride

3.2 Learning Rate Annealing or Scheduling

This parameter defines the rate at which the network learns the parameters. It sets the speed of learning features. It a large value is set then the speed of training would increase but with that, the accuracy would decrease as the efficiency of learning more features would reduce. This parameter is tuned throughout the training process in a descending manner. The learning rates are scheduled so that the step size in each iteration gives out the best output for efficient training. There are various methods to set learning rates. Some them have been demonstrated in Figure 3.

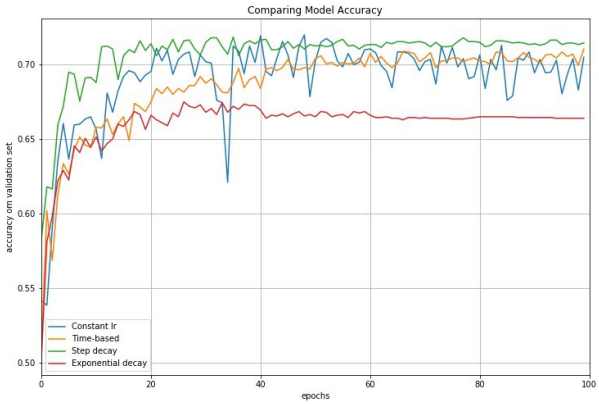


Figure 3: Comparision of different LR scheduling methods.

3.3 Optimization

Optimization is an essential aspect of enhancing the efficiency of the CNN models. Basically, it is a method of correcting the calculated error. It might sound simple but it just as complicated. There are various optimizers which are used to do this optimization. There are constant learning rate algorithms and adaptive learning rate algorithms. For this research, Stochastic Gradient Descent optimizer

[8] was used because of its enhanced performance on optimizing A-MnasNet [2]. It was used with varying values of learning rates during training with a momentum set to 0.9.

3.4 Data Augmentation

This step is used to preprocess the dataset. Various transformations are applied on the images. The purpose of data augmentation is the enhance the accuracy of the model. In other words, it enables the CNN model to learn features more efficiently. AutoAugment [9] is implemented to preprocess the dataset. It uses reinforcement learning to automatically select the transformations that are most efficient for the CNN model. It defines two parameters for the process. First is the transformation and the other is the magnitude of applying it. These two together are defined as a Policy. A Policy consists of sub-policies which have those two parameters. Different sub-policies are selected by the controller which best suit the training. This process is demonstrated in Figure 4. It was evident that by using AutoAugment [9] the accuracy of A-MnasNet increased to 96.89% from 92.97%.

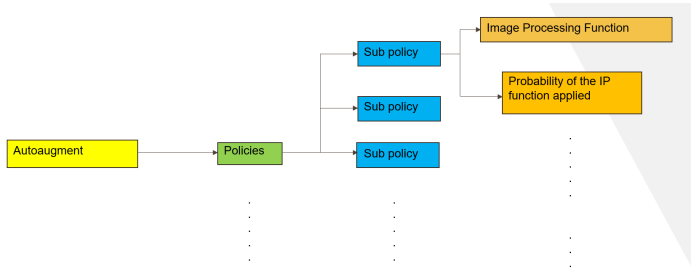


Figure 4: AutoAugment

4 Image Classification on NXP Bluebox 2.0

A-MnasNet [2] was deployed on NXP Bluebox 2.0 for Image Classification [3]. BlueBox 2.0 by NXP is a real-time embedded system which serves with the required performance, functional safety and automotive reliability to develop the self-driving cars. It is a ASIL-B and ASIL-D compliant hardware system which provides a platform to create autonomous applications such as ADAS systems, driver assistance systems.

There are three processing units in Bluebox 2.0. There is a vision processor, S32V234, a radar processor, S32R274, and a compute processor, LS2048A. This enables to develop autonomous applications. The vision processor and the compute processors are used to perform Image Classification [3] using A-MnasNet [2]. Figure 5 shows the methodology to deploy A-MnasNet [2] on Bluebox 2.0 for Image Classification.

After training A-MnasNet [2], the model is imported in the RTMaps Studio using its python component. A TCP/IP connection is used for data transmission between RTMaps Design Studio and NXP Bluebox 2.0. The model was successfully implemented on the hardware and was able to predict the object in the input images accurately.

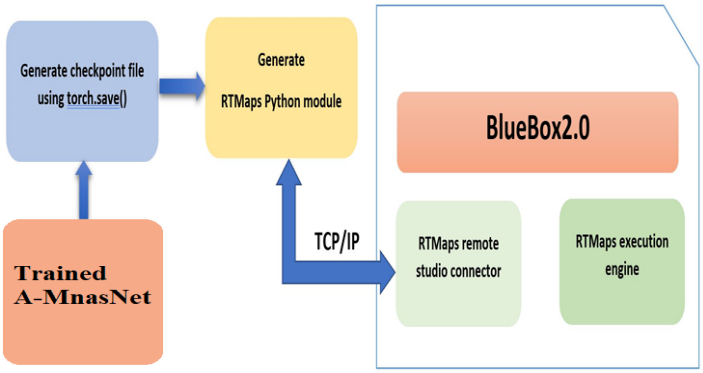


Figure 5: Deployment of A-MnasNet on NXP BlueBox 2.0

RTMaps Studio was used for real-time implementation of A-MnasNet[2] on NXP Bluebox 2.0. It is an asynchronous high performance platform which has an advantage of an efficient and easy-to-use framework for fast and robust developments. It provides a modular toolkit for multimodal applications like ADAS, Autonomous vehicles, Robotics, UGVs, UAVs, etc. It provides a platform to develop, test, validate, benchmark and execute applications.



Figure 6: RTMaps connection with Bluebox 2.0

Figure 6 shows the role of RTMaps in the implementation. It provides an interface between the hardware and the embedded platform. It uses a TCP/IP connection by assigning a static IP address to enable data transfer between them.

5 Hardware and Software used

- NXP Bluebox 2.0
- Aorus Geforce RTX 2080Ti GPU
- Python version 3.6.7.
- Pytorch version 1.0.

- Spyder version 3.6.
- RTMaps Studio
- Tera Term
- Livelossplot

6 Results

MnasNet [4] achieved 80.8% accuracy with CIFAR-10 [6] dataset achieving 12.7 MB model size without data augmentation. New convolutional blocks resulted in an increase of accuracy to 92.97%. Furthermore, AutoAugment [9] was implemented for enhancing overall accuracy which increased from 92.97% to 96.89%. The final model is of 11.6 MB and a validation accuracy of 96.89%.

Implementation of new algorithms resulted in enhanced accuracy. It is demonstrated in Table 3. The overall accuracy increases by 16.09% with a slight decrease in model size.

Table 3: Comparison with baseline architecture

Comparison of models		
Architecture	Model Accuracy	Model size (in MB)
MnasNet	80.8%	12.7
A-MnasNet	96.89%	11.6

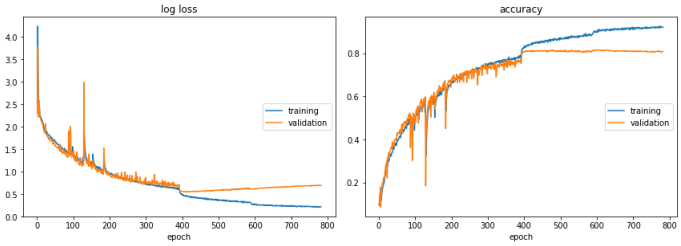


Figure 7: MnasNet (Baseline Architecture)

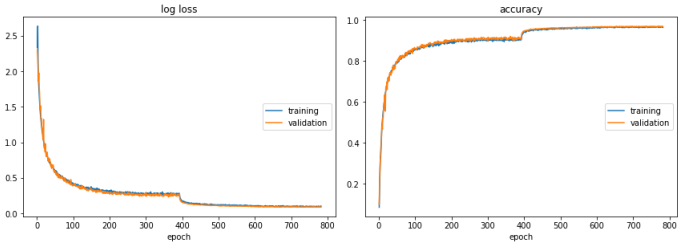


Figure 8: A-MnasNet (Proposed Architecture)

Figure 7 and figure 8 demonstrate the training curves. It is evident that by implementing new algorithms resulted in reduced overfitting and better accuracy.

For resource constrained hardware, desired model size and accuracy can be obtained by tuning width multiplier. Table 4 demonstrates the idea of having multiple size and accuracy by scaling width multiplier.

Table 4: varying model size by width multiplier

Scaling A-MnasNet with width multiplier		
Width Multiplier	Model Accuracy	Model size (in MB)
1.4	97.16%	22
1.0	96.89%	11.6
0.75	96.64%	6.8
0.5	95.74%	3.3
0.35	93.36%	1.8

Figure 9 shows the console output for Image Classification on NXP Bluebox 2.0. The model predicts that the object in the input image is a car. The window in the left corner shows the console output in Tera Term.

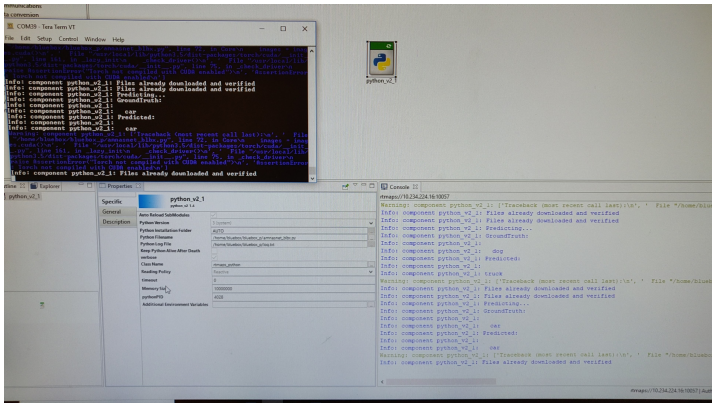


Figure 9: Image Classification using A-MnasNet

7 Conclusion

Design Space Exploration of MnasNet [4] architecture resulted in enhanced accuracy and a decrease in model size. The proposed architecture, A-MnasNet [2], has an accuracy of 96.89% and a size of 11.6 MB. It surpasses MnasNet [4] in terms of model size and accuracy. New convolutional layers were added to make A-MnasNet [2] more efficient in extracting features. AutoAugment[9] was used to preprocess the dataset. The model was scaled with width multiplier to get various trade-offs between model size and accuracy. The architectures were trained and tested on Cifar-10 [6] dataset. Furthermore, they were implemented on NXP Bluebox 2.0 for Image Classification. Results show that the new architecture successfully detects the object and successfully performs Image Classification.

Conflict of Interest The authors declare no conflict of interest.

References

[1] P. Shah, "DESIGN SPACE EXPLORATION OF CONVOLUTIONAL NEURAL NETWORKS FOR IMAGE CLASSIFICATION," 2020, doi:https://doi.org/10.25394/PGS.13350125.v1.

[2] P. Shah, M. El-Sharkawy, "A-MnasNet: Augmented MnasNet for Computer Vision," in 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS), volume 2020-Augus, 1044–1047, IEEE, 2020, doi: 10.1109/MWSCAS48704.2020.9184619.

[3] C. Wang, Y. Xi, "Convolutional Neural Network for Image Classification," Johns Hopkins University, 2015.

[4] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, Q. V. Le, "MnasNet: Platform-Aware Neural Architecture Search for Mobile," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2815–2823, IEEE, 2019, doi:10.1109/CVPR.2019.00293.

[5] A. F. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," arXiv, 2018.

[6] "CIFAR-10 and CIFAR-100 datasets," doi:https://www.cs.toronto.edu/{\~{}}kriz/cifar.html.

[7] D. Li, A. Zhou, A. Yao, "HBONet: Harmonious Bottleneck on Two Orthogonal Dimensions," in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 3315–3324, IEEE, 2019, doi:10.1109/ICCV.2019.00341.

[8] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv, 2016.

[9] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, Q. V. Le, "AutoAugment: Learning Augmentation Policies from Data," arXiv, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1384-1402 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

BrcLightning - Risk Analysis and Scaling for Protection against Atmospheric Discharge - Extender

Biagione Rangel De Araújo*

BRC – Biagione Rangel Consulting, Natal, 59064-490, Brazil

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 09 February, 2021
Online: 28 February, 2021

Keywords:
Risk Analysis
Lps Sizing
BrcLightning

ABSTRACT

This manuscript intending to publicize the improvements incorporated in the BrcLightning application, including the Risk Analysis module with the help of a pop-up, which provides the result and assists in the identification of mitigating measures by the professional, which must be defined to reduce the calculated risks. Other points addressed in this extension are the improvements added to the database to meet the corporate demands of companies, referring to the Risk Analysis module. It also incorporates flexibilities to perform the sizing separately, in the design, evaluation and scaling modules of the LPS – Lightning Protection System that using rolling sphere method and Angle Method, incorporating, in some of the modules, the issue of opinions or alerts. These modules use the mathematical approach methodology. In addition to these improvements, this review included the reporting module of facilities in the filter system, which allows the use of the database more selectively for the emission of these documents. This filter has a structure for issuing corporate demands of reports. The results can be obtained quickly and easily, on-screen or printing several reports. The reliability and safety of the results can be assessed through the check with the examples of the standards that define the criteria and methodology, that must be followed to carry out for cases of Risk Analysis or through graphic drawings on AutoCad platform or similar for the sizing modules. Other improvements in this extension are the addition of topics for new modules for which we already have the equations modeled in Excel, although we have not yet coded in the programming language.

1. Introduction

This article is an extension of the work originally presented at the Symposium SIPDA XV [1]. In the event, representatives of a large company attended our explanation and reported that the great difficulty they faced was to have the control and traceability of the various Risk Analyses that are performed for each building or structures of the various units of the company. So, considering that our system is structured in the form of a database, we evaluated that, by introducing structural changes in this database, it will be possible to meet this demand.

Concomitantly with this observation, we also identified that other improvements would be required, even to use the various modules of the system in isolation (before having to register a project), either in the database or through an app which may be via mobile phones. These improvements aim to meet specific demands, such as those to evaluate the effectiveness of the volume of protection of a given SPDA concerning the various structures or equipment and facilities that, in principle, are or would be under protection against direct impacts of atmospheric discharges. In this case, the data is not embedded in the database but can be consolidated into a report, with options to be printed or archived.

Regarding the Reporting Module, this extension was less extensive, as it was composed by the improvements of the filters defined to facilitate the selective issuance of Database Reports, with limited coverage of the specificity of each module that make up the current system. However, because it is a database, we intended (in a new version) to provide tools in this module, to enable the data export specifically to the Excel environment, to allow better tools for analyzing historical data.

2. Definitions

- Coverage radius (rc): The distance between the point of the cover margin, of a determined envelopment of an LPS. This distance determines the size of the horizontal projection of the fictitious plane, given by equation (1):

$$rc = [h_1(2R - h_1)]^{1/2} - [h_2(2R - h_2)]^{1/2} \quad (1)$$

*Corresponding Author: BRC – Biagione Rangel Consulting, contato.brc@biagione.com.br, +55-84-98723-8753

where:

rc	horizontal protected distance
h1	height of the higher mast
R	rolling sphere radius
h2	height of the lower mast:

- *Cover margin height (hc):* The dimensions of the height of the nearest point on the envelope over the structure under the protection of the LPS;
- *Distance from the critical point (d1):* The distance between the critical point and one metallic element such as rods or wire air-terminations
- *Envelopment:* Geometric shape that limits the protected volume according to the rolling sphere method;
- *Coverage margin (cm):* The shortest distance between a point of the structure under the protection of the LPS and the envelopment of the protective volume. The dimension of the margin corresponds to the perpendicular tangent measurement of the envelopment to the nearest point of the structure under protection. The graphical representations of "cm" are as shown in Figure 1a, Figure 1b, and Figure 1c (all contained in Figure 1):

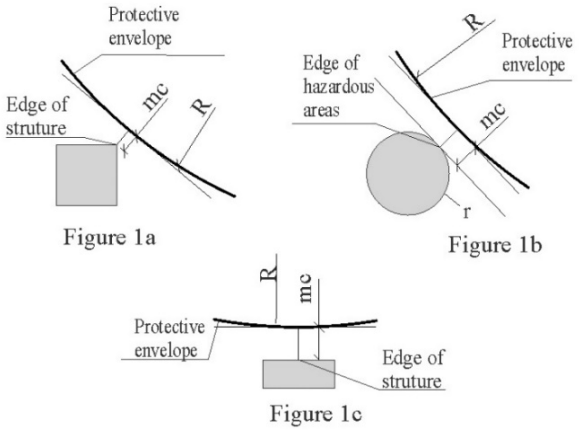


Figure 1: Examples of coverage margins

- *Fictitious plane (Fp):* Assumed horizontal plane that provides coverage of protection at a given height;
- *Hazardous areas:* Area surrounding storage facilities or transfer station of flammable liquids or gases, due to the possibility of containing flammable or explosive mixtures. They are defined in: Zone 0 (when the explosive/inflammable mixture is still there or will be for long periods); Zone 1 (likely to occur in normal operating conditions) and Zone 2 (it's unlikely – an abnormal condition of operation);
- *Sizing module marks*: The marks indicated in the graphs of the sizing and limit variation modules correspond to the extreme points of the structure that should be protected or evaluated on the effectiveness of the coverage by the SPDA protection wrap, as shown in figures 2 and 3 below:

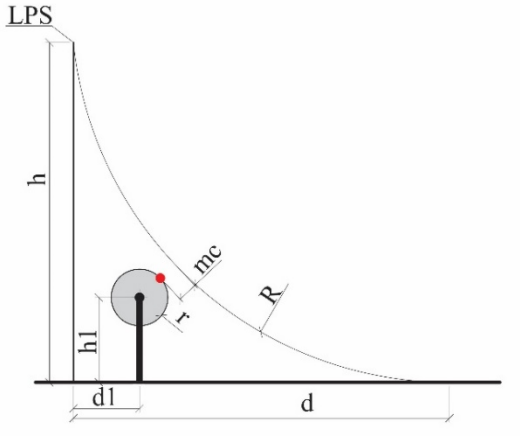


Figure 2: Hazardous areas limiting point figure (see Figure 1b)

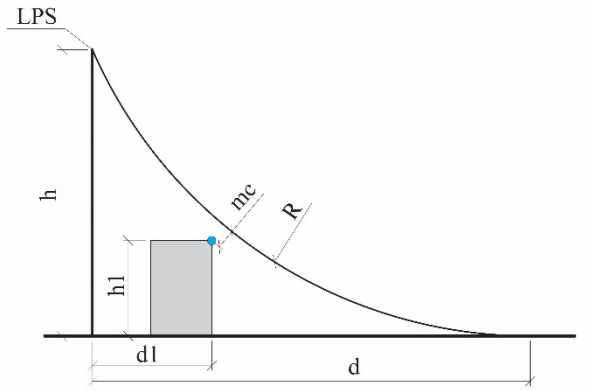


Figure 3: Structure limiting point figure (see Figure 1a)

3. Software Development

The application is developed in Access, in VBA language, to allow it to run on Windows and on the Internet. We had a project to carry out the migration to an Internet language; however, due to the Covid 19 pandemic, we had to cancel it. Nevertheless, we hope that, considering the problems due to the current situation, we can resume the project in the year 2021.

In this version, the system continues with the same configuration of modules and sections of the previous publication, from which this paper is an extension of (SIPDA XV [1]). Thus, we will focus on detailing the improvements that have been introduced, as described in sequence (items 4 to 7).

4. Risk Analysis Module

Risk Analysis consists of 5 sections: 1. Registration Module; 2. Zone Registration Module; 3. The configuration data of the connected lines and Zone.; 4. Risk Analysis Factors Module (typical loss value) and 5. Results Module.

4.1. Registration Section

In this section, there are specific data related to design registration that can be common to all modules and all sections, but also, there are important data specific to the Risk Analysis Module. Due to this, we have identified that it will be more effective for these data input to be dismembered in two modules: one for general project registration and another for registration and

configuration of the Risk Analysis Module. However, this should only be performed in a new version.

The improvements that we introduced in this section included the fields of input to meet the demand to perform the registration of organizational structures, containing up to 5 levels: Level 1 - Hold Company; Level 2 – The Business Unit, which may or may not cover several other Operating Units; Level 3 - The Operational Unit; Level 4 - The Operational Sector and Level 5 - The installation subject to risk analysis or the one containing Air Terminal. This improvement is only incorporated into the Registration Form of this Section. Therefore, in this version, it has not yet been possible to integrate this organizational structure registration with the other modules, including the Risk Analysis Module itself, regarding the reports. This requires numerous changes in the reports already formatted. The reported modifications are highlighted in green as shown in Figure: 4.

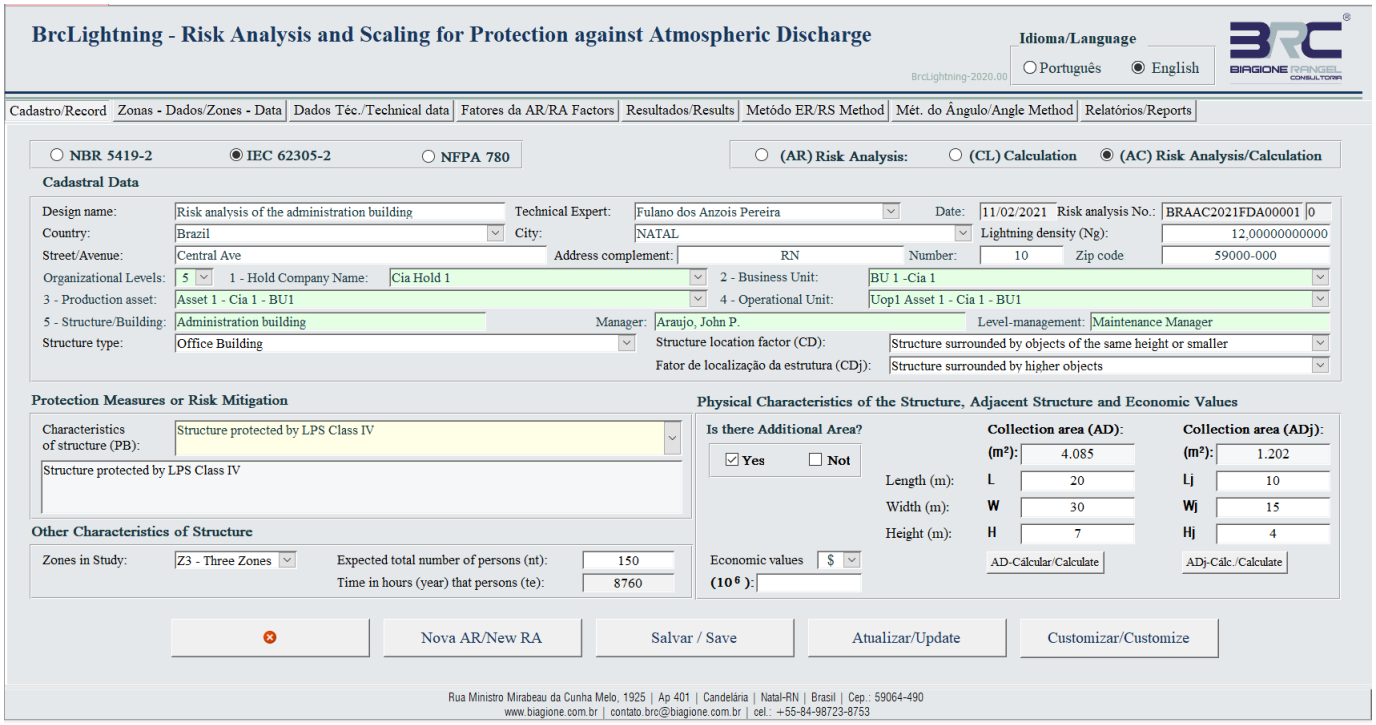


Figure 4: Main screen of data input

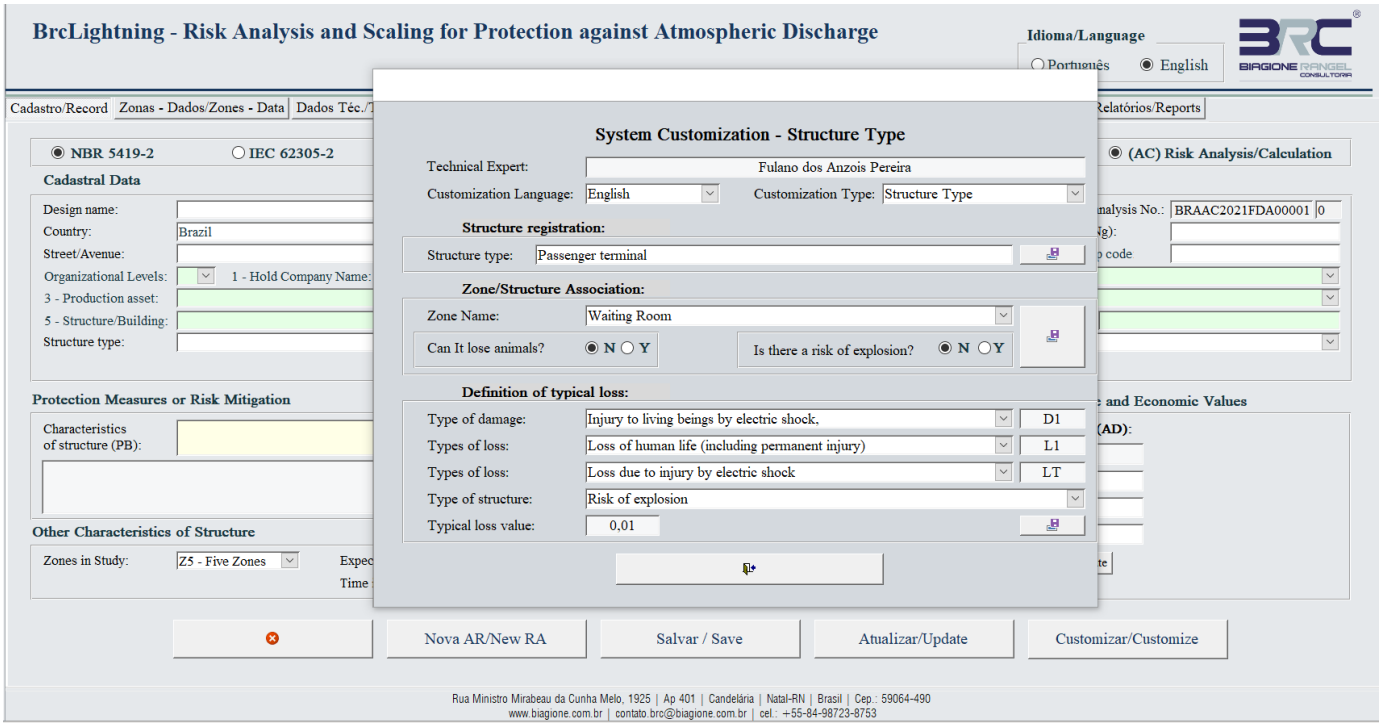


Figure 5: Screen for customizing structure and zones)

4.2. Zone registration module

The System was engineered to cover up to 5 Zones (NBR-5419-2 [2] and IEC 62305-2 [3] do not establish limits, but provide examples of up to 5 Zones), as we observe that this is a practice used and that already surpasses the examples provided by the standards cited. The number of Zones is selected as shown in Figure 4 (Zone in Study).

The System already makes several types of structures available, which were understood as necessary and sufficient to meet the range of facilities and buildings that must be studied and related to each of them, the different types of Zones. However, the Designer can insert new structures and new zones, but he or she must proceed to register them. To do so, it will be necessary to customize them and make the association with the loss factors, as shown in Figure 5.

In this Section, the designer inserts the data for each zone that is defined for Risk Analysis, according to Section One of the Register, as shown in Figure 4. In this section, no improvement has been introduced, as show in Figure 6. For more details see SIPDA XV [1].

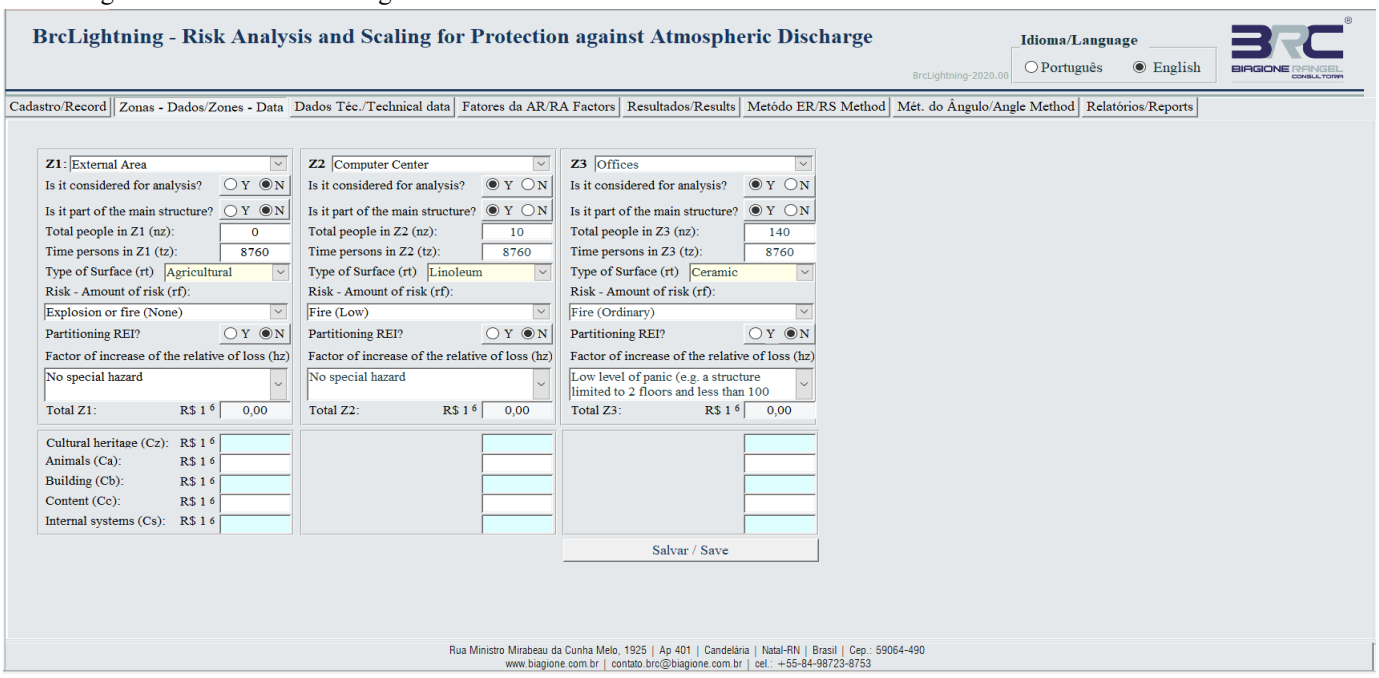


Figure 6: Screen of the setting module of the Zones (structure with all 3 Zones)

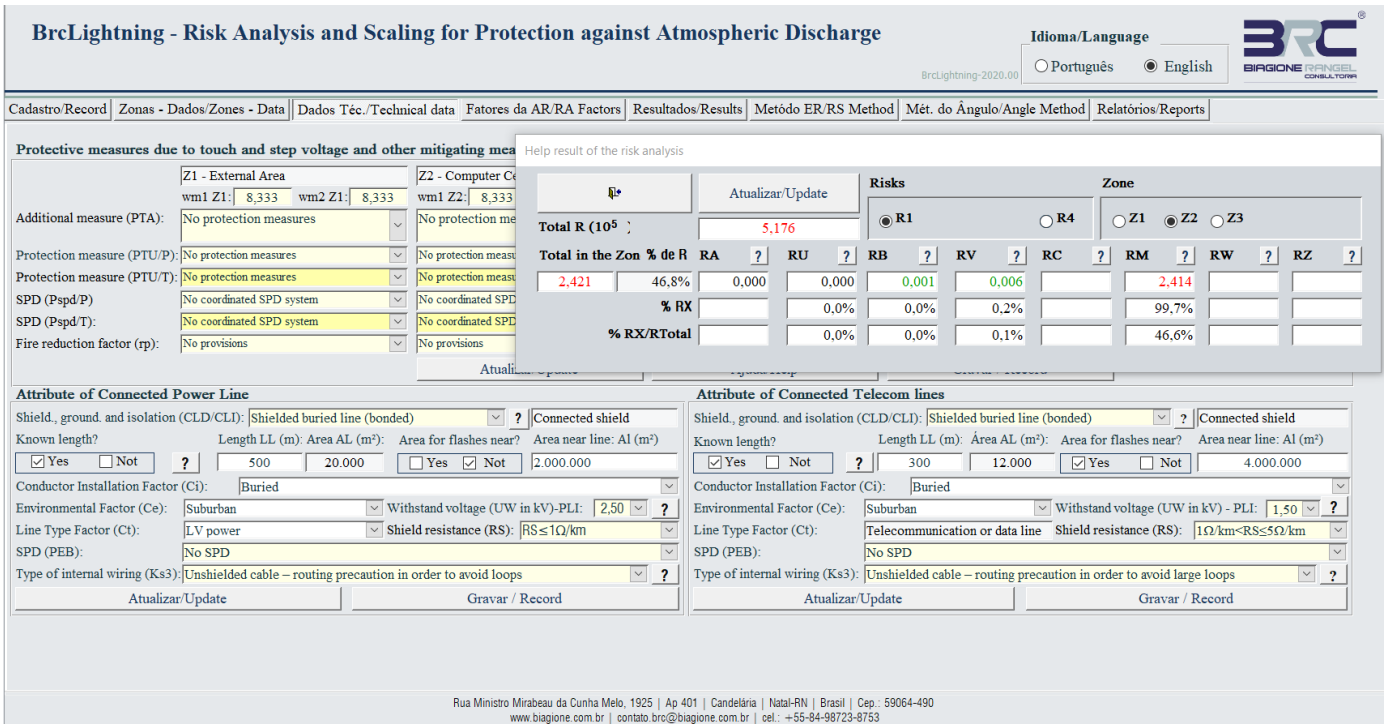


Figure 7: Setting screen of the connected lines and other zone data

4.3. The setting data of the connected lines and Zone

This Section provides the space to enter information about power lines and telecommunications lines: how they connect to the site (building; installations; etc.). The designer can also enter protective measures due to touch and step voltage and other risk-mitigating measures, individually for each defined Zone. In this section, it was introduced improvements to the Help pop-up form with the result of the risk analysis, as shown in Figure 7. Thus, the designer can monitor the effectiveness of each mitigating measures in real-time. This pop-up, once visible, can be viewed in all sections of the risk analysis module.

Also, there is a pop-up to help define line parameters and other alert parameters as shown in Figure 8.

4.4. Risk Analysis Factors Module

This section presents the typical loss values for each zone (as shown in Figure 9), based on the definitions established in section one (as shown in Figure 4) and two (as shown in Figure 5).

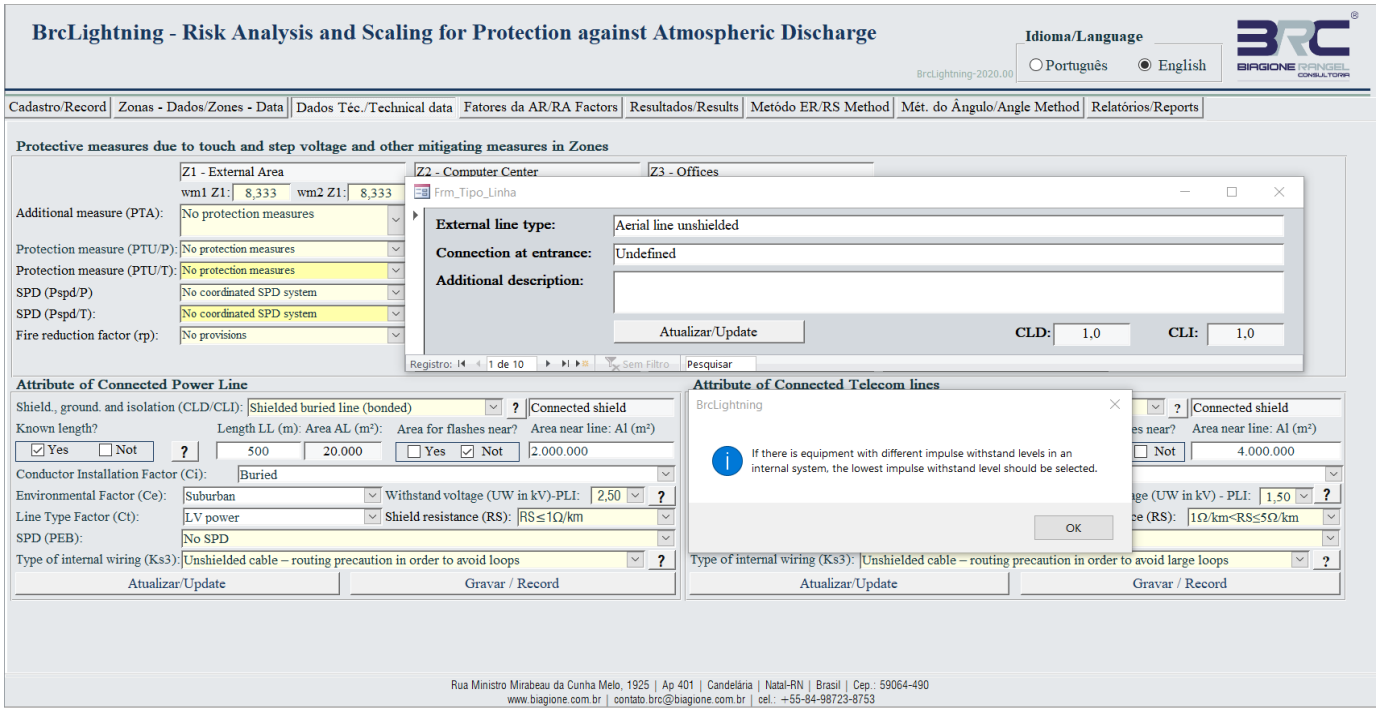


Figure 8: The pop-up helps with line parameters and other information

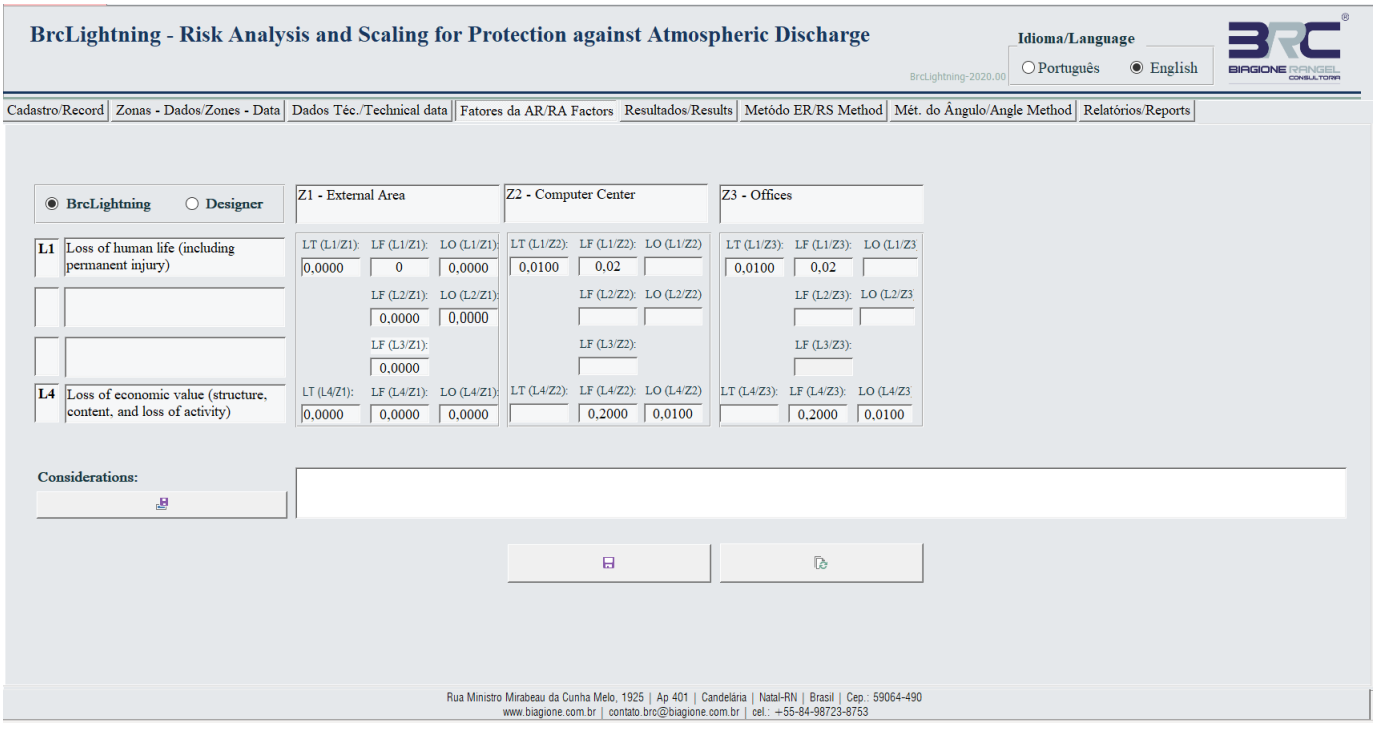


Figure 9: Loss Type Factors Module Screen

In the version presented in SIPDA XV [1], it was already possible for designers to change the typical loss value, if they had another interpretation regarding the one previously defined in the system (these factors defined in the system were based on the definitions established in NBR 5419 -2 [2] and IEC 62305-2 [3] for typical loss values). In this case, in the previous version, the designers had to standardize their basic values for each type of installation in advance. The improvement introduced allows that, in case they don't have a database parameterized with the values according to their interpretation criteria, they can do so by simply selecting the designer parameters (in the input field) and applying a double click on each factor that they want to parameterize. This way, a pop-up (as shown in Figure 10) form will appear for each factor and they will define it at their discretion. The new defined typical loss value, after being updated, will be highlighted in green (as shown in Figure 10), but there are security restrictions and other requirements (see item 9.4).

4.5. Results Module

In this section, the final result is as shown in Figure 11. Mitigating measures can also be inserted (the description of these measures), if the results indicate that this is necessary. Mitigation measures, as needed, can be included (these should be inserted descriptively). A technical note with the analysis design also can and should be inserted. Another improvement is the possibility of issuing, in a summarized format, the risk analysis report, as shown in Figures: 12 (a) and (b). This report contains the basic information of the risk analysis study and the result.

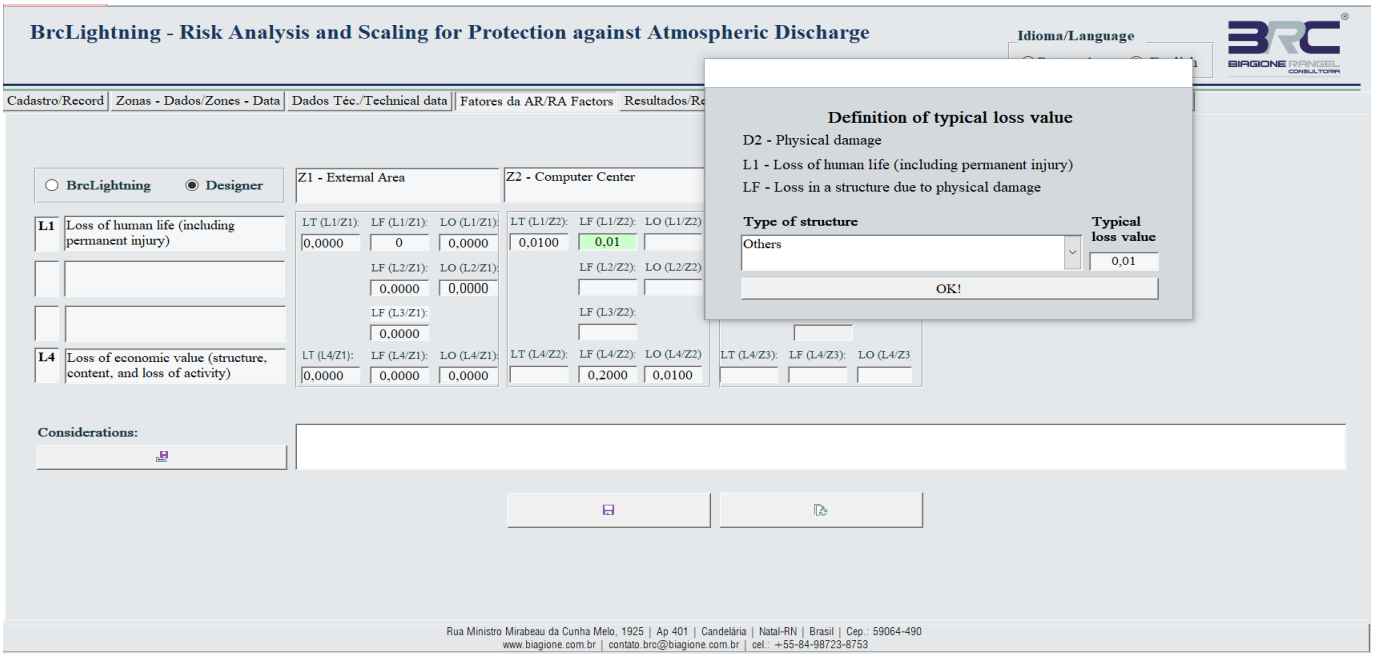


Figure 10: Loss Type Factors Module Screen with individualization pop-up

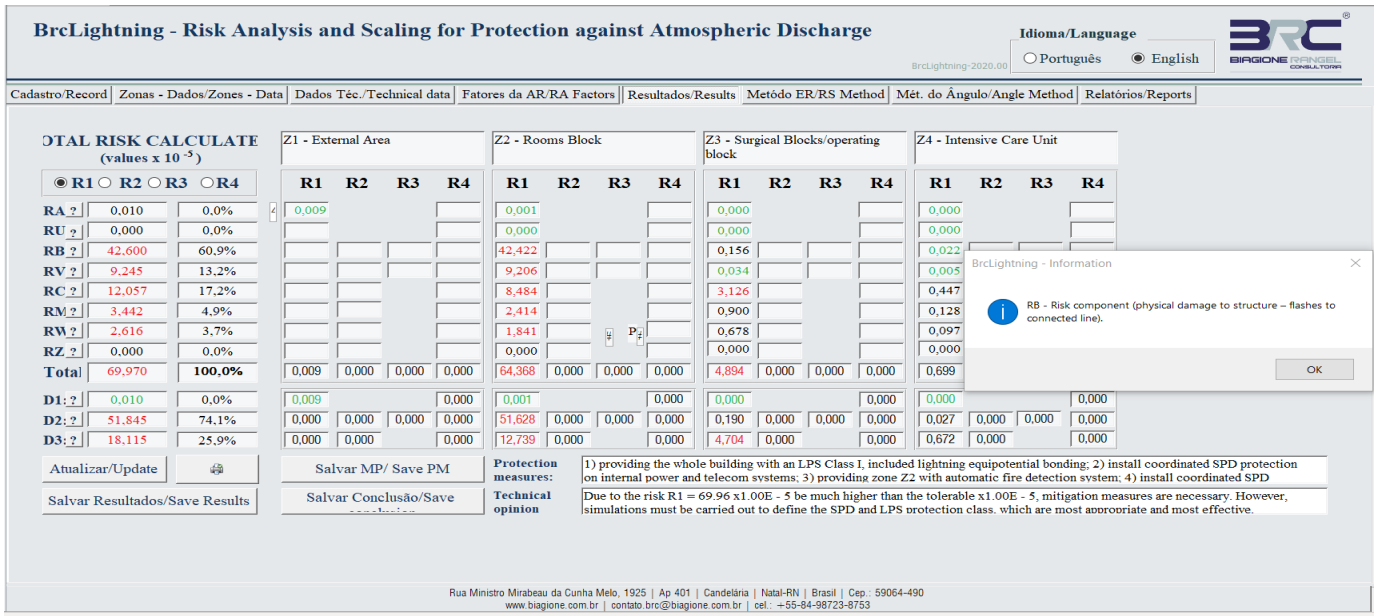
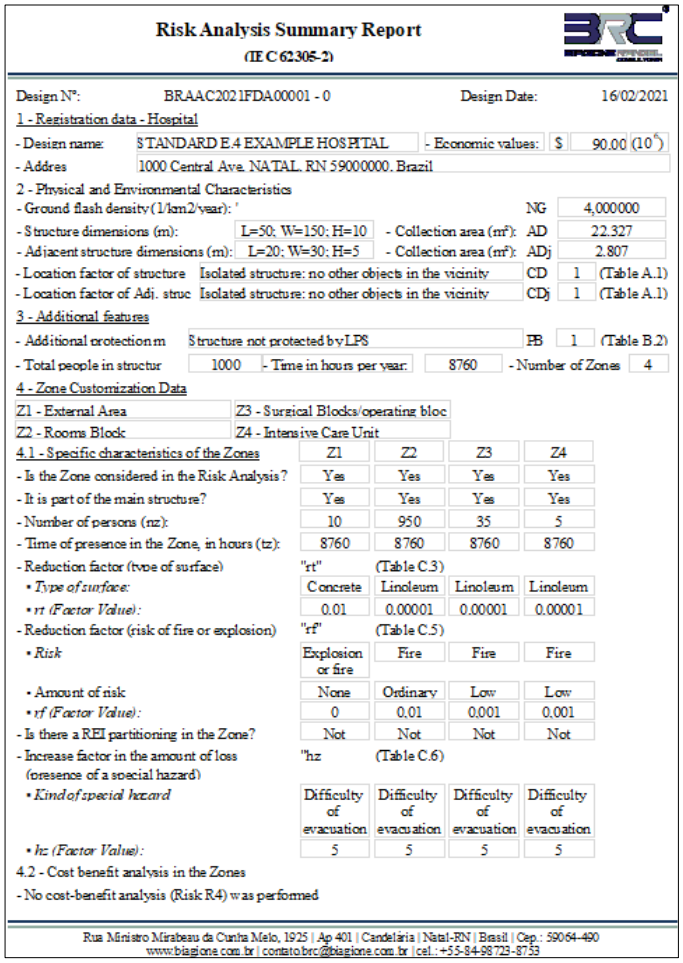


Figure 11: Result Module Screen and the button for printing the summarized report



Risk Analysis Summary Report
(IEC 62305-2)

Design N°: BRAAC2021FDA00001 - 0 Design Date: 16/02/2021

1 - Registration data - Hospital

- Design name: STANDARD E.4 EXAMPLE HOSPITAL - Economic values: $ 90.00 (10^6)
- Addres 1000 Central Ave. NATAL. RN 59000000. Brazil

2 - Physical and Environmental Characteristics

- Ground flash density (1/km2/year):			NG	4,000000
- Structure dimensions (m):	L=50; W=150; H=10	- Collection area (m²):	AD	22.327
- Adjacent structure dimensions (m):	L=20; W=30; H=5	- Collection area (m²):	ADj	2.807
- Location factor of structure	Isolated structure: no other objects in the vicinity		CD	1 (Table A.1)
- Location factor of Adj. struc	Isolated structure: no other objects in the vicinity		CDj	1 (Table A.1)

3 - Additional features

- Additional protection m Structure not protected by LPS PB 1 (Table B.2)
- Total people in structur 1000 - Time in hours per year: 8760 - Number of Zones 4

4 - Zone Customization Data

Z1 - External Area | Z3 - Surgical Blocks/operating bloc
Z2 - Rooms Block | Z4 - Intensive Care Unit

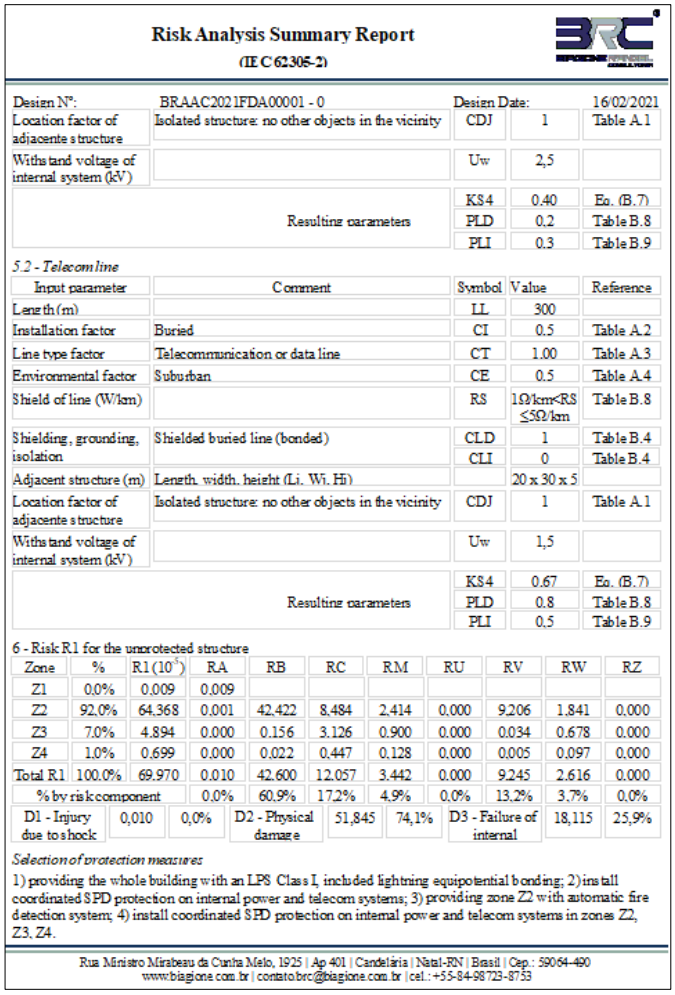
4.1 - Specific characteristics of the Zones	Z1	Z2	Z3	Z4
- Is the Zone considered in the Risk Analysis?	Yes	Yes	Yes	Yes
- It is part of the main structure?	Yes	Yes	Yes	Yes
- Number of persons (nz):	10	950	35	5
- Time of presence in the Zone, in hours (tz):	8760	8760	8760	8760
- Reduction factor (type of surface) "rt" (Table C.3)				
• *Type of surface:*	Concrete	Linoleum	Linoleum	Linoleum
• *rt (Factor Value):*	0.01	0.00001	0.00001	0.00001
- Reduction factor (risk of fire or explosion) "rf" (Table C.5)				
• *Risk*	Explosion or fire	Fire	Fire	Fire
• Amount of risk	None	Ordinary	Low	Low
• *rf (Factor Value):*	0	0.01	0.001	0.001
- Is there a REI partitioning in the Zone?	Not	Not	Not	Not
- Increase factor in the amount of loss (presence of a special hazard) "hz (Table C.6)				
• *Kind of special hazard*	Difficulty of evacuation	Difficulty of evacuation	Difficulty of evacuation	Difficulty of evacuation
• *hz (Factor Value):*	5	5	5	5

4.2 - Cost benefit analysis in the Zones

- No cost-benefit analysis (Risk R4) was performed

Rua Ministro Mirabeau da Cunha Melo, 1925 | Ap 401 | Candelária | Natal-RN | Brasil | Cep.: 59064-490
www.biagione.com.br | contato.brc@biagione.com.br | cel.: +55-84-98723-8753

(a)



Risk Analysis Summary Report
(IEC 62305-2)

Design N°: BRAAC2021FDA00001 - 0 Design Date: 16/02/2021

Location factor of adjacente structure	Isolated structure: no other objects in the vicinity	CDJ	1	Table A.1
Withstand voltage of internal system (kV)		Uw	2,5	
Resulting parameters		KS4	0.40	Eq. (B.7)
		PLD	0,2	Table B.8
		PLI	0.3	Table B.9

5.2 - Telecom line

Input parameter	Comment	Symbol	Value	Reference
Length (m)		LL	300	
Installation factor	Buried	CI	0.5	Table A.2
Line type factor	Telecommunication or data line	CT	1.00	Table A.3
Environmental factor	Suburban	CE	0.5	Table A.4
Shield of line (W/km)		RS	1Ω/km<RS ≤5Ω/km	Table B.8
Shielding, grounding, isolation	Shielded buried line (bonded)	CLD	1	Table B.4
		CLI	0	Table B.4
Adjacent structure (m)	Length, width, height (Li, Wi, Hi)		20 x 30 x 5	
Location factor of adjacente structure	Isolated structure: no other objects in the vicinity	CDJ	1	Table A.1
Withstand voltage of internal system (kV)		Uw	1,5	
Resulting parameters		KS4	0.67	Eq. (B.7)
		PLD	0.8	Table B.8
		PLI	0.5	Table B.9

6 - Risk R1 for the unprotected structure

Zone	%	R1 (10^{-5})	RA	RB	RC	RM	RU	RV	RW	RZ
Z1	0,0%	0.009	0.009							
Z2	92,0%	64.368	0.001	42.422	8.484	2.414	0.000	9.206	1.841	0.000
Z3	7,0%	4.894	0.000	0.156	3.126	0.900	0.000	0.034	0.678	0.000
Z4	1,0%	0.699	0.000	0.022	0.447	0.128	0.000	0.005	0.097	0.000
Total R1	100,0%	69.970	0.010	42.600	12.057	3.442	0.000	9.245	2.616	0.000
% by risk component			0,0%	60.9%	17.2%	4.9%	0,0%	13.2%	3,7%	0,0%

D1 - Injury due to shock: 0,010 | 0,0% | D2 - Physical damage: 51,845 | 74,1% | D3 - Failure of internal: 18,115 | 25,9%

Selection of protection measures

1) providing the whole building with an LPS Class I, included lightning equipotential bonding; 2) install coordinated SPD protection on internal power and telecom systems; 3) providing zone Z2 with automatic fire detection system; 4) install coordinated SPD protection on internal power and telecom systems in zones Z2, Z3, Z4.

Rua Ministro Mirabeau da Cunha Melo, 1925 | Ap 401 | Candelária | Natal-RN | Brasil | Cep.: 59064-490
www.biagione.com.br | contato.brc@biagione.com.br | cel.: +55-84-98723-8753

(b)

Figure 12 (a): Summary of risk analysis report (first page) (b): Summary of risk analysis report (last but one page)

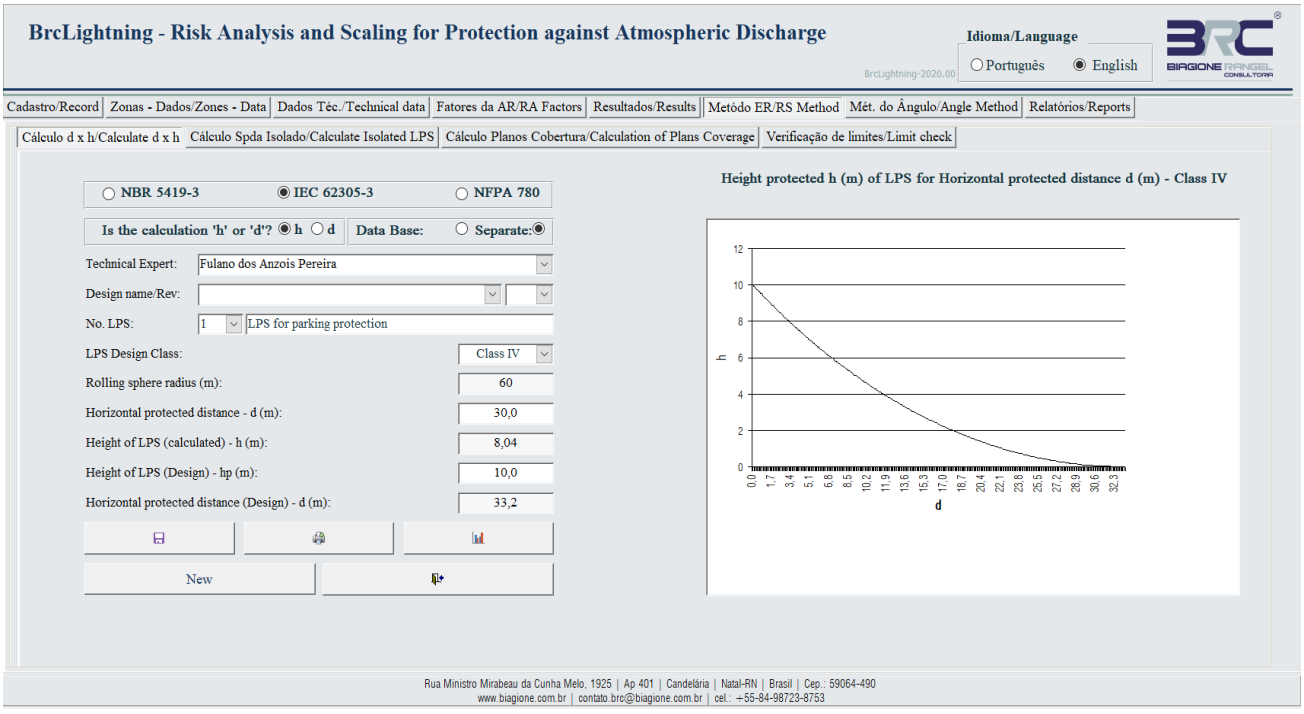


Figure 13: Example of calculation of h x d for Class IV

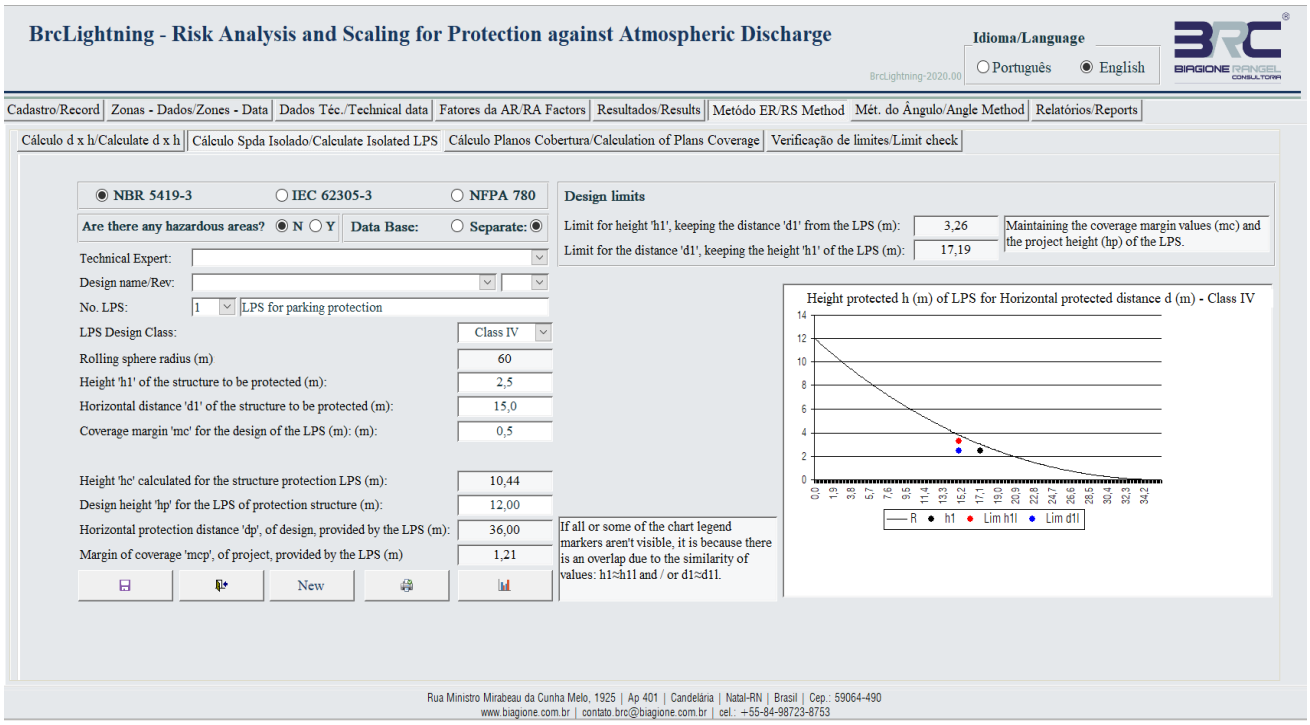


Figure 14: Sizing of the Isolated LPS, without hazardous area

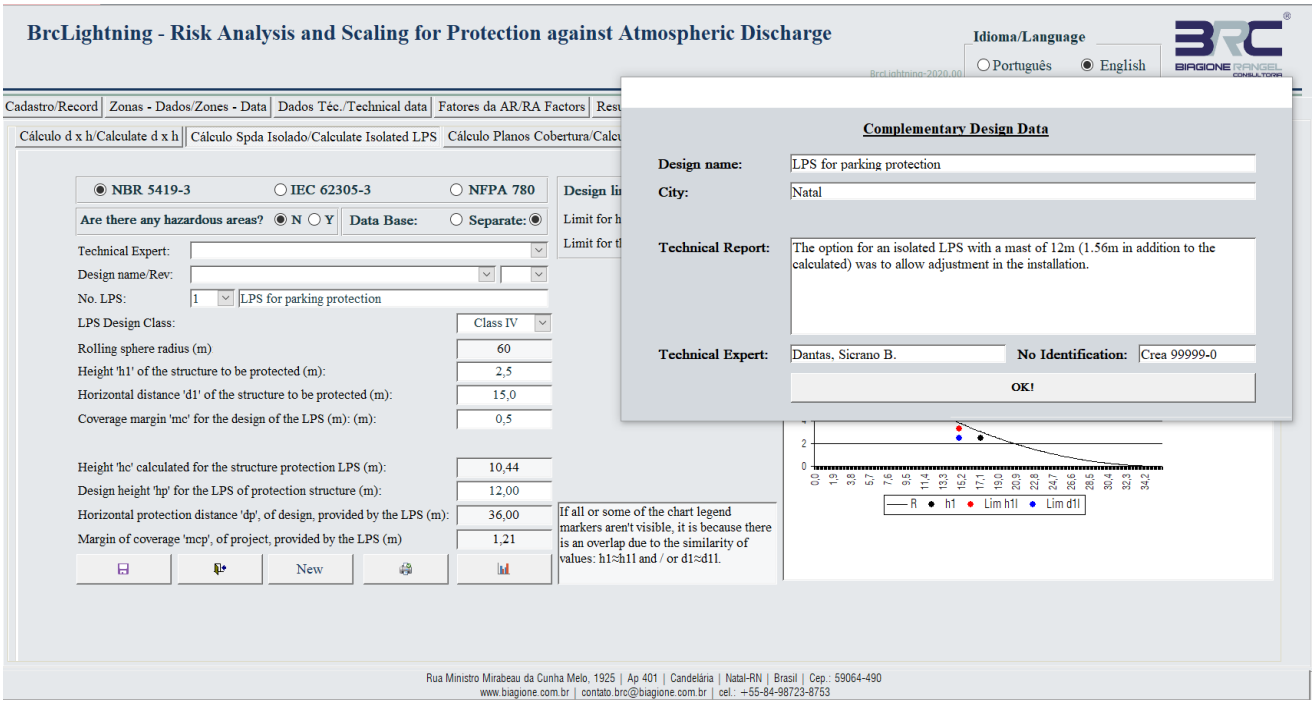


Figure 15: Pop-up for supplementary data

5. Rolling sphere module

This Module maintains the 4 Sections of the original version presented in SIPDA XV [1] and is in accordance with NBR 5419-3 [4]; IEC 62035-3 [5] and NFPA 780 [6]. However, the possibility of performing the sizing separately and the printing of the report of this dimensioning was introduced in all four sections. This can be seen in the sequence. Other improvements have been introduced and are detailed in each of the sections.

5.1. Section for sizing 'h' or 'd'

In the version presented in SIPDA XV [1], the system included one graph related to the calculation of the protected horizontal distance at ground level or on a reference surface when the height of the LPS is known and another for the calculation of the height, required to provide protection of a horizontal distance at ground level or a reference surface. Now, these graphs have been integrated into one (as shown in Figure 13). When calculating the height, it is a non-standard commercial value, this version provided

an input field to adjust the calculated height. In addition, it was added the option for issuing the report of this dimensioning.

In this section, the height of the insulated LPS is scaled (mast or air termination system of the insulated (only one) overhead wire to protect a specific structure) based on the required LPS Class and the determined coverage margin, including the characterization of a structure with or without hazardous area, as shown in Figure 14. Improvements were also incorporated to print the report of the sizing (the button with the printer icon) and the option to scale the LPS of a database design or a separate design for the issuance of the report, it is necessary to provide additional data, as shown in Figure 15 (in the case of a separate project). The pop-up appears when the print button is pressed. The issued report is shown in Figure 16. Separated projects are not saved in the database.

When the design definitions of the structure to be protected ('h1' and 'd1', besides 'mc') do not allow the sizing of the LPS to the defined protection Class, the System issues an alert of the impossibility of sizing and indicates the limit of the maximum distance and height the structure can be protected by LPS, as shown in Figure 17.

BrcLightning - Risk Analysis and Scaling for Protection against Atmospheric Discharge

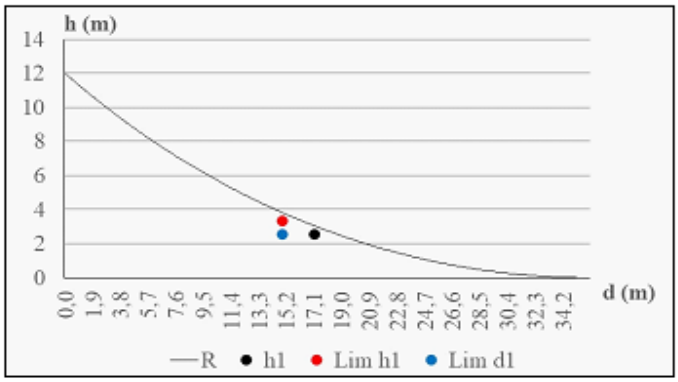
Height Sizing Report for a LPS - Rolling Sphere Method

(Structure without Hazardous Areas)

Data

Standard: IEC 62305-3

Design name:	Parking protection LPS
LPS Number:	1
LPS description:	Central LPS
Class of LPS:	Classe IV (60m)
Structure height to protect - h1 (m):	2.50
Horizontal distance to be protected - d1 (m):	15.00
Design coverage margin - mc (m):	0.50
Calculated height for LPS - hc (m):	10.44
Design height for LPS - h (m):	12.00
Horizontal protected distance, design - d (m):	36.00
Coverage margin, design - mcp (m)	1.21
Limit for height 'h1', keeping the 'd1' from the LPS (m):	3.26
Limit for the distance 'd1', maintaining the 'h1' of the LPS (m):	17.19



Nota - 1: If all or some of the chart legend markers aren't visible, it is because there is an overlap due to the similarity of values: h1≈h11 and / or

Technical Report:

The option of an isolated LPS with a mast of 12 m (1.56 m in addition to the calculated one) will be to allow adjustment in installation.

Araujo, Sivrano P.

Crea 9999-0

Rua Ministro Mirabeau da Cunha Melo, 1925 | Ap 401 | Candelária | Natal-RN | Brasil | Cep.: 59064-490
www.biagione.com.br | contato.brc@biagione.com.br | cel.: +55-84-98723-8753

Figure 16: Example report of sizing the isolated LPS

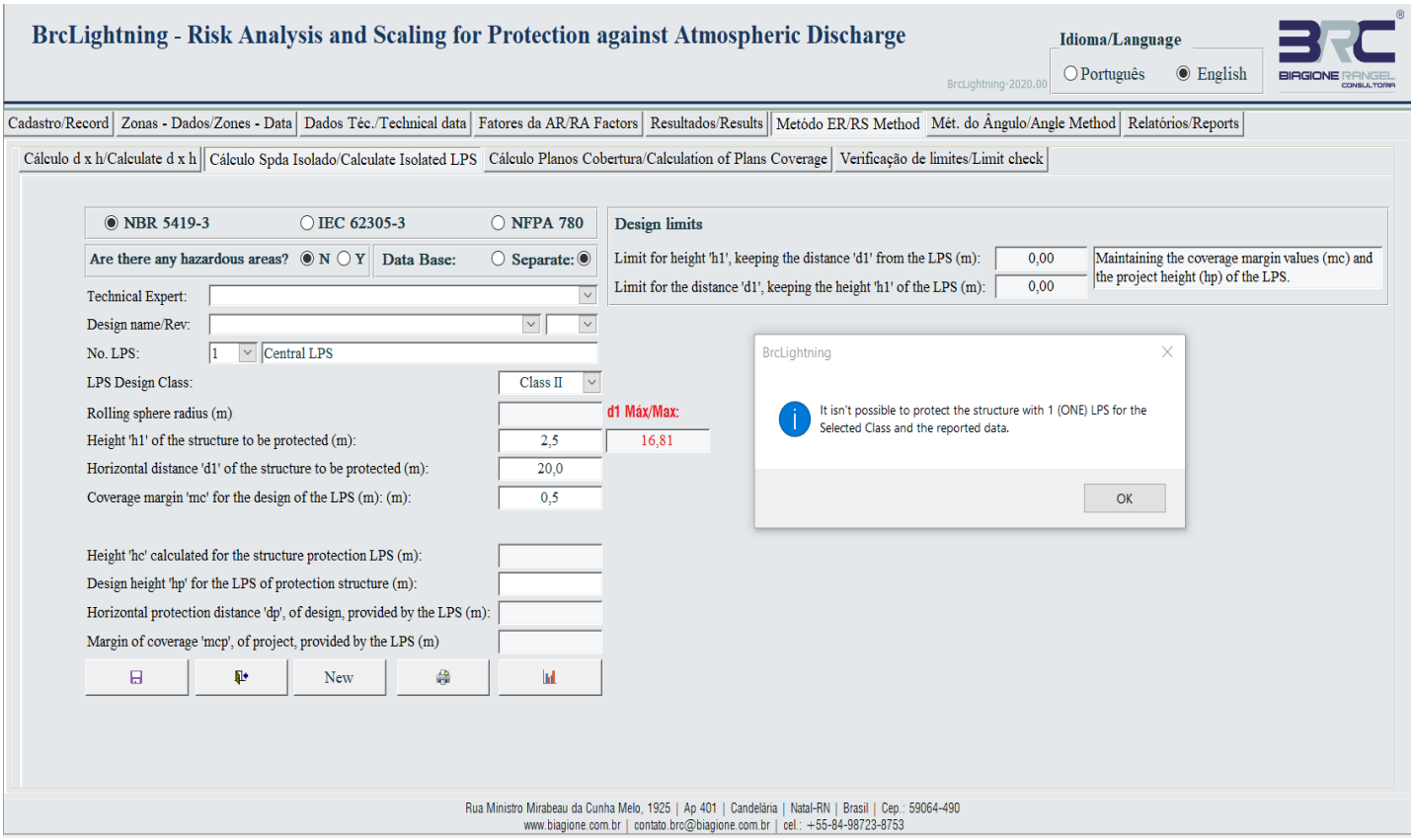


Figure 17: Alert for impossibility of isolated LPS sizing

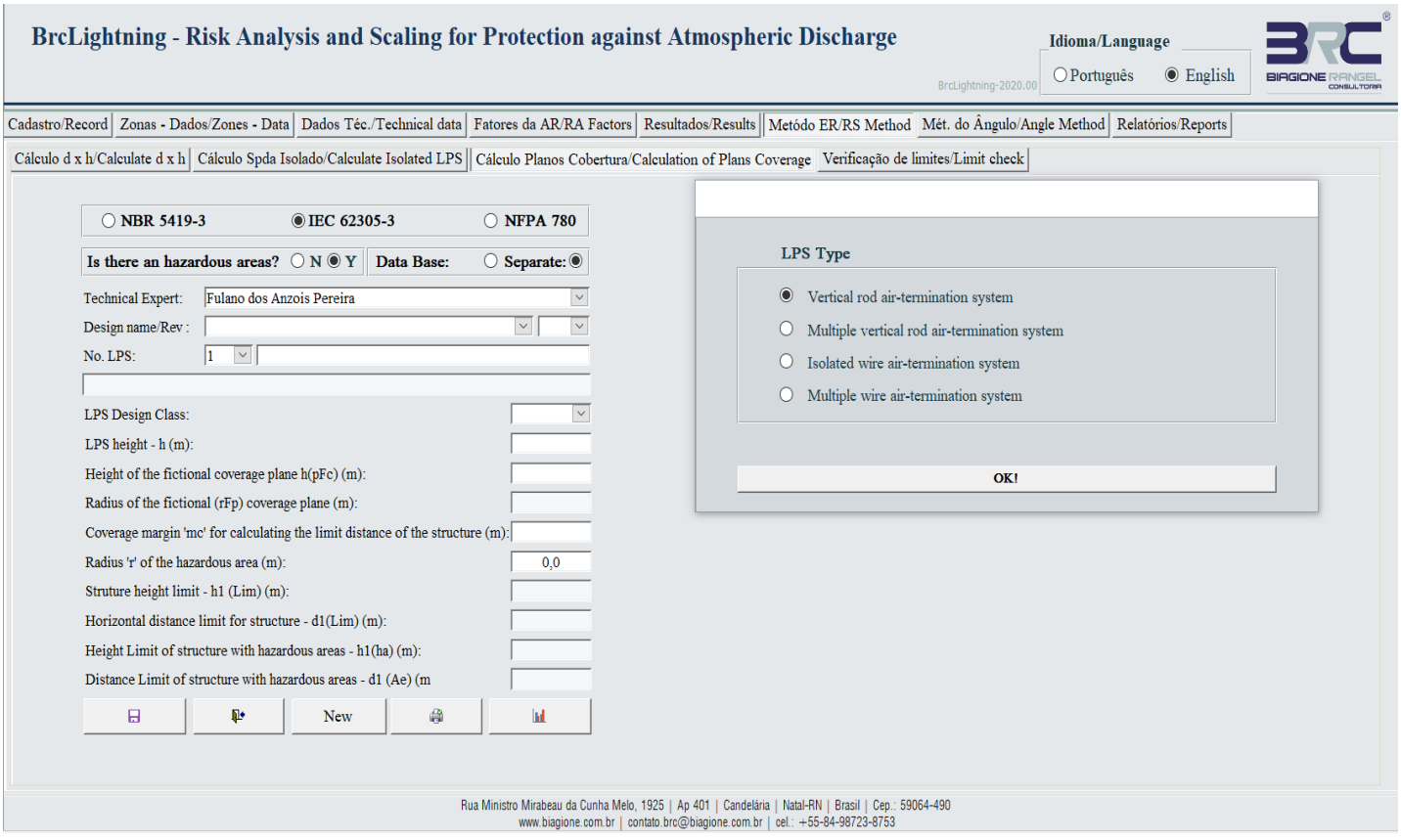


Figure 18: Screen with pop-up to define the LPS type

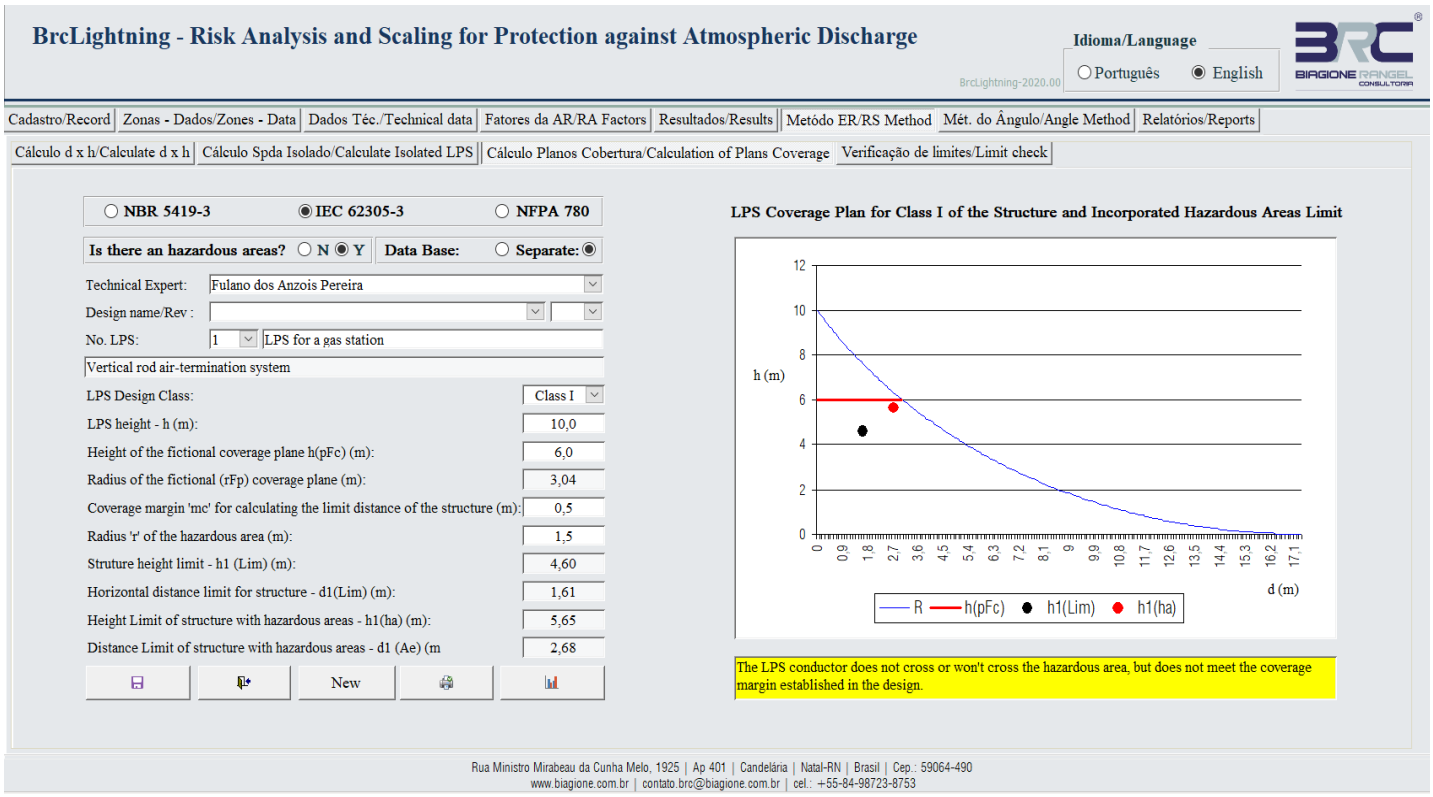


Figure 19: Calculation of the Fictional Coverage Plan with alert

5.2. Evaluation section of the limits of an installed LPS

This last section of the Rolling Sphere Module is intended to assess the protection that a specific LPS can provide, considering all the Classes of Protection established by NBR-5419-3 [4]; IEC 62305-3 [5] and NFPA 780 [6]. Improvements have been made to the definition of the LPS type and alert with a technical note, as shown in Figure 20.

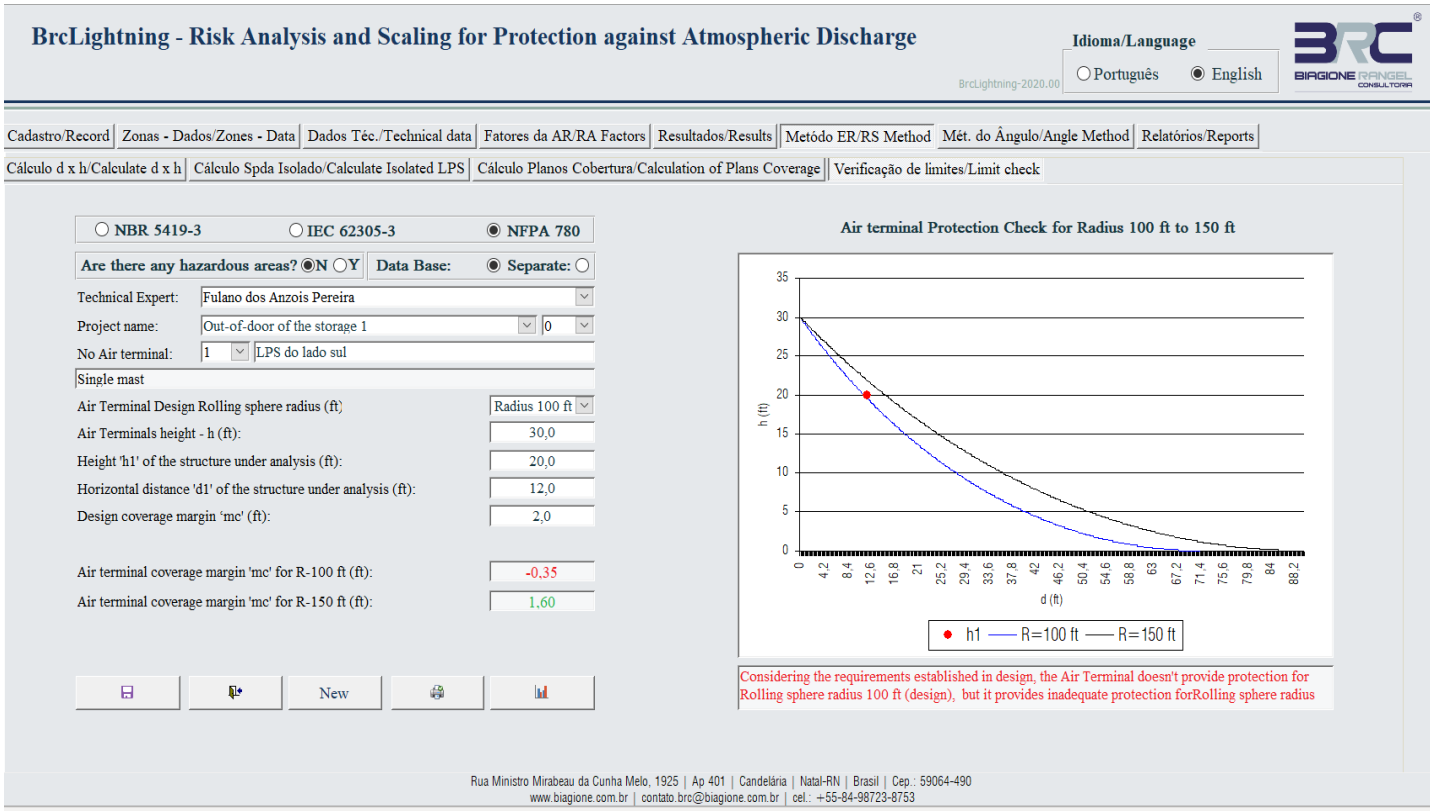


Figure 20: Limits evaluation module screen

6. Angle Method Module

This module maintains the three Sections of the original version presented in SIPDA XV [1] and in accordance with NBR 5419-3 [4]; IEC 62035-3 [5]. However, it was introduced in all of them the possibility to carry out the sizing separately and printing the report of this. Considering that this method is not widely used, only an improvement was introduced, in addition to those reported above, in Section 1, as detailed below.

6.1. Height or Distance Sizing (Angle Method)

This Section calculates the horizontal distance protected at ground level or from a reference surface when the height of the LPS is known. It also allows you to calculate the height required for an LPS when you know the horizontal distance that needs to be protected, as shown in Figure 21.

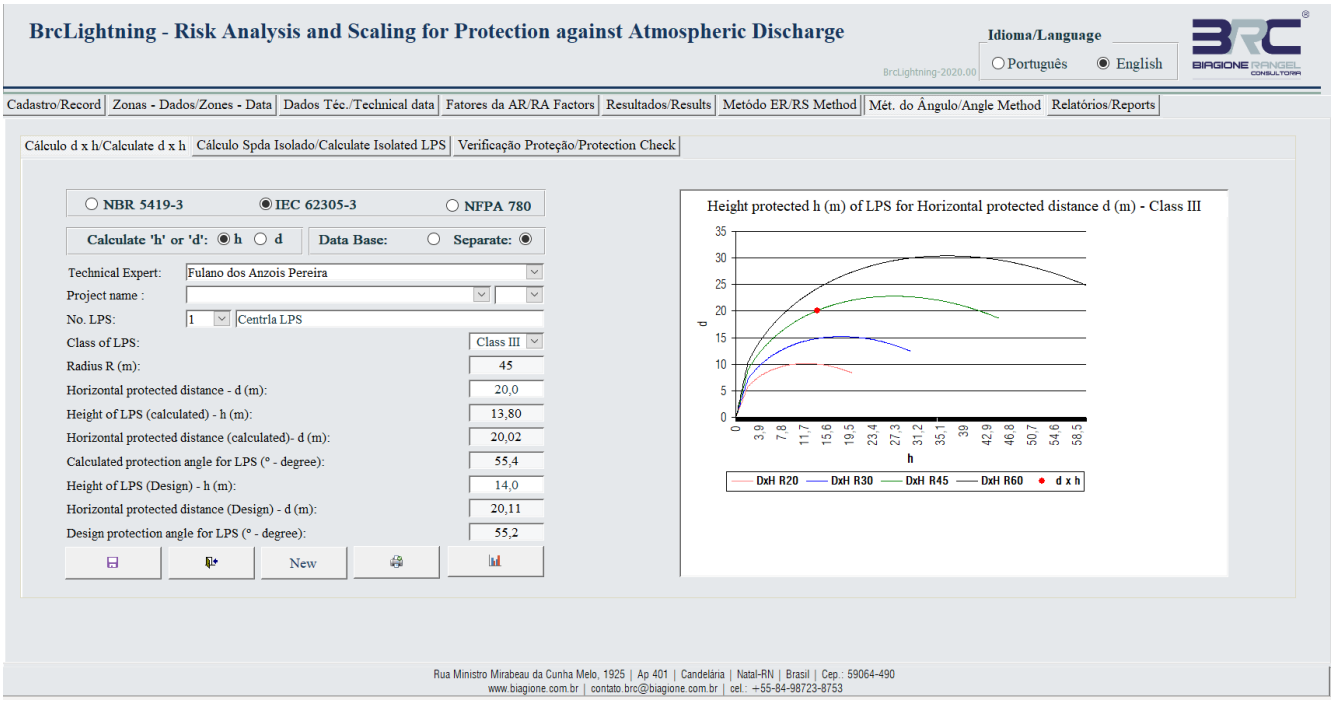


Figure 21: Calculation of the height x distance

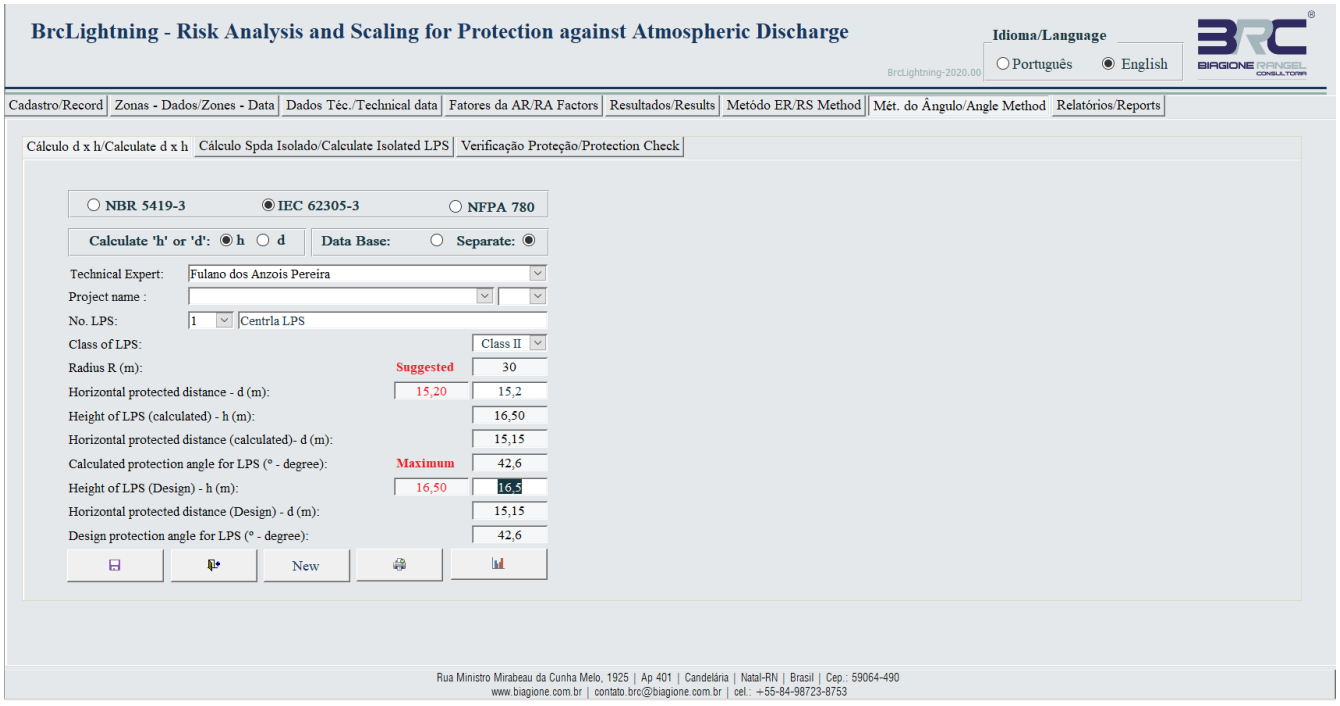


Figure 22: Calculation of the height x distance - Alert

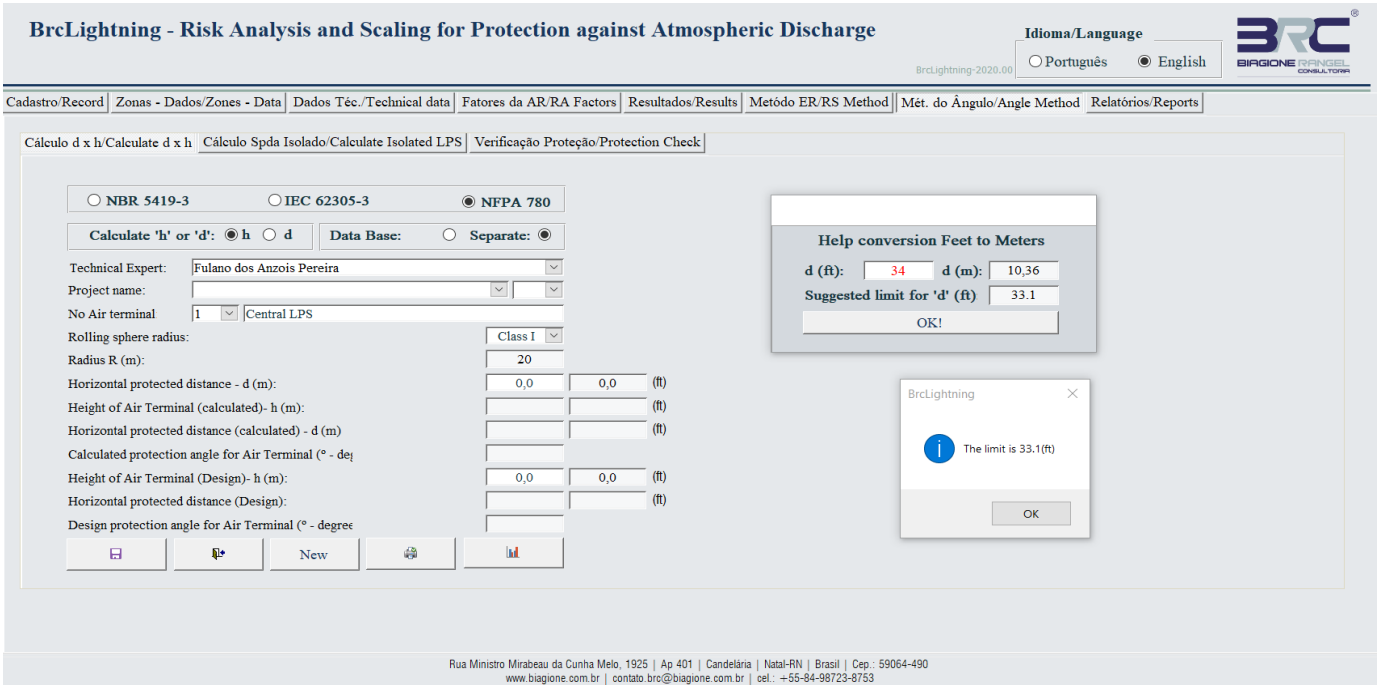


Figure 23: Calculation of the height x distance with alert

In order to comply with NFPA 780 [6], the improvement for converting units of measure in meters was introduced, as shown in Figure 23.

The system for this calculation only allows input in meters, thus, with this improvement, it becomes possible to insert the values in feet unit. To do so, two clicks are applied in the input field and one pop-up appears (as shown in Figure 23), providing an input field with the unit in feet. After this, when the pop-up is closed, the unit of measurement in meters. In case of exceeding the suggested limit, an alert will be issued.

7. Report Module

This Module includes options for the various preformat of reports that the system, currently, can provide: Risk Analysis Report; LPS project report by Rolling Sphere Method; Module Report using angle method, etc. The system screen where these options are available is shown in Figure 24 and several report templates are shown in Figure 25a Figure 25b. The improvements introduced in this module were composed by the increase in filter options to allow for more appropriate selectivity and the inclusion of the project revision number, as shown in Figure 24.

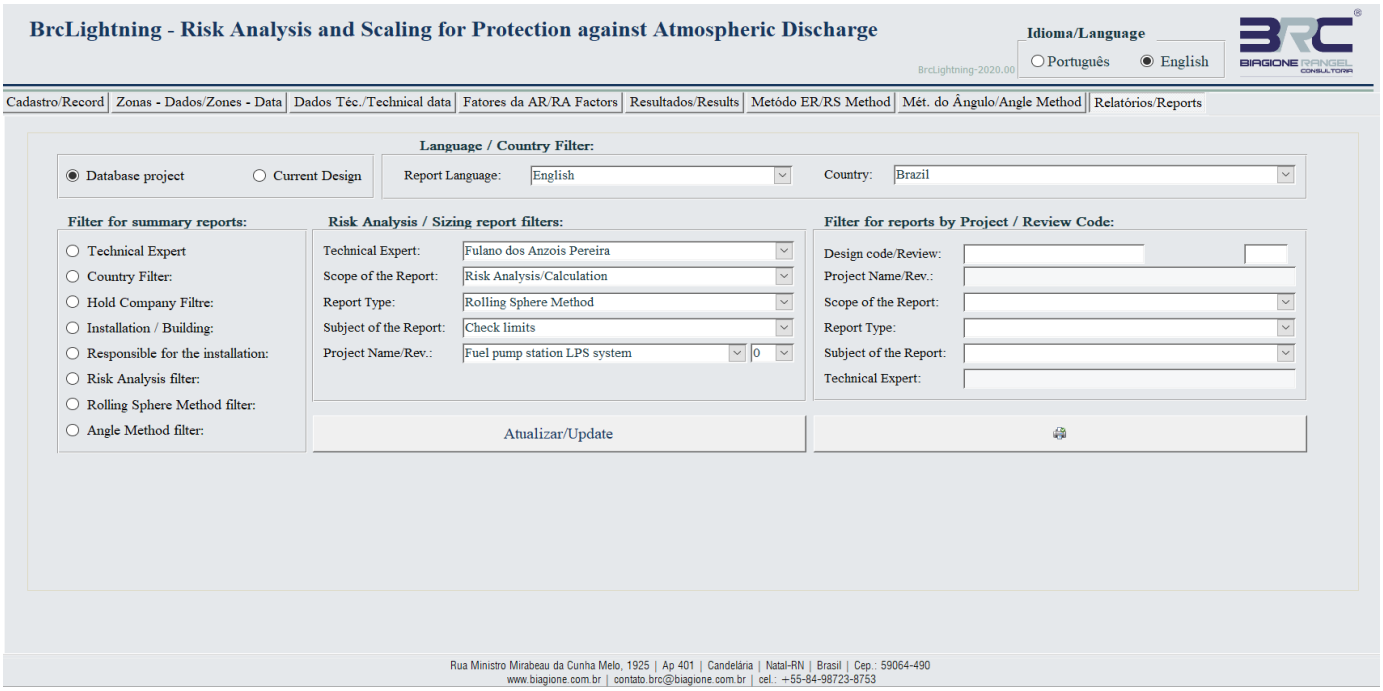
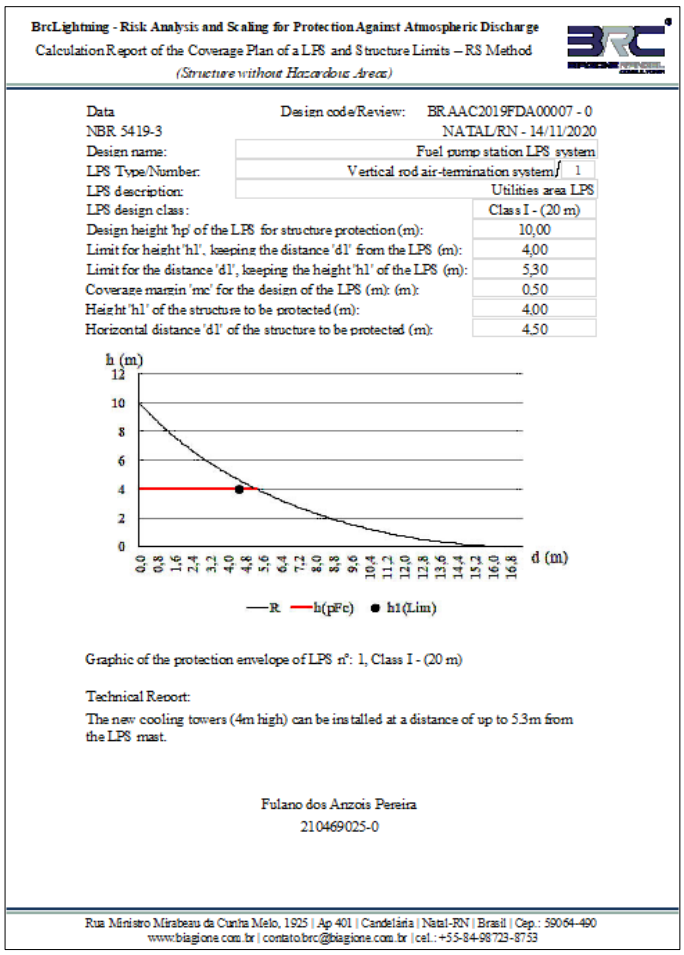
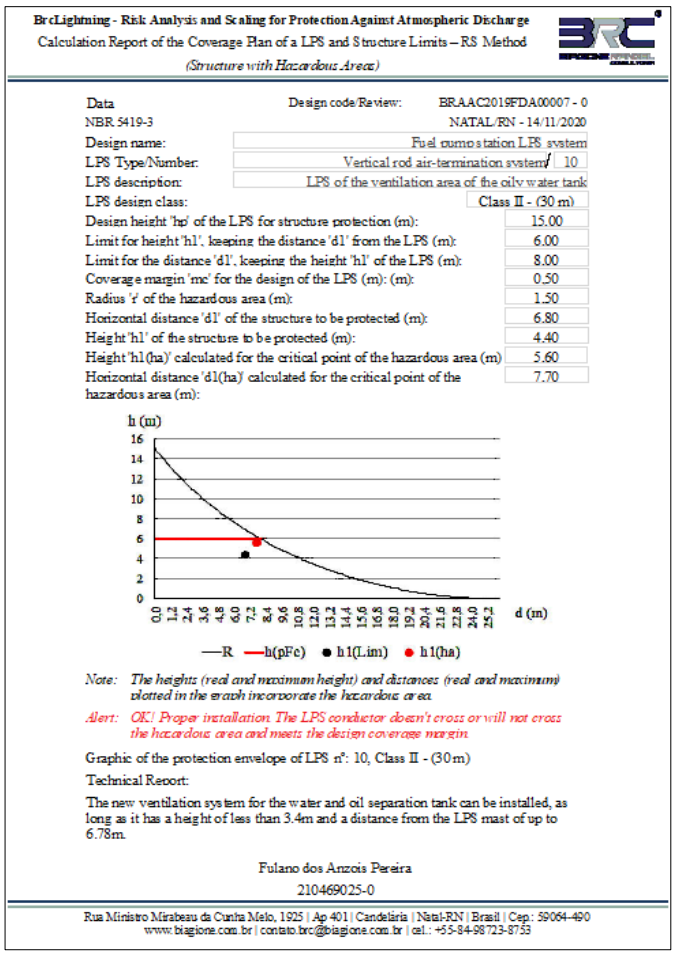


Figure 24: Screen of the reporting module



(a)



(b)

Figure 25 (a): System report templates formatted without hazardous areas (b): System report templates formatted with hazardous area

8. Improvements already identified for future insertion

8.1. Module for sizing the surge protective device

Considering that one of the most important mitigating measures identified in the Risk Analysis is the installation of Surge Protective Devices - DPS, it was understood that the inclusion of a module to meet this need is extremely important and, as the methodology and formulas already exist and are supported by NBR-5419-4 [7] and IEC 62305-4 [8], the insertion of a module with this function was included in the planning of future improvements.

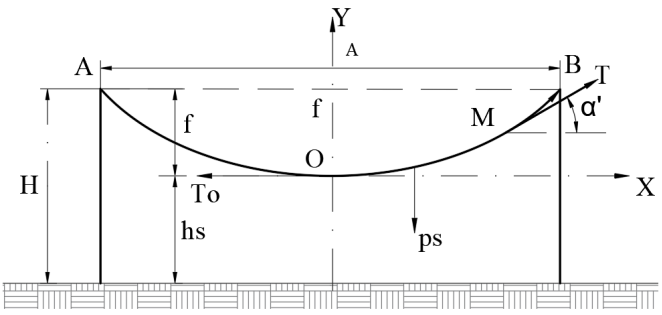


Figure 26: Efforts in wire air termination with supports of the same height

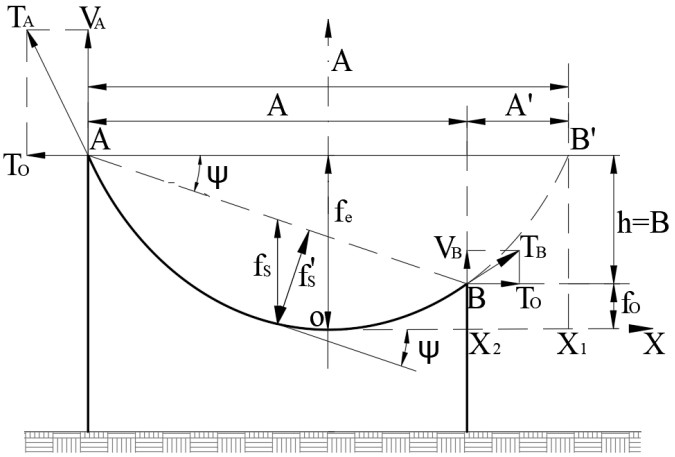


Figure 27: Efforts in wire air termination with supports of different heights

8.2. Module for sizing wire air-termination support structures

Considering that many of the SPDA systems are composed of wire air-termination, it was identified that it is convenient to have a tool to scale the efforts requested from supports of this kind of captor, especially when there is variation in room temperature. Considering, also, that the guarantee of the volume of protection derives from a projection of the height of a captor element and that the control of this lowest point is critical. So, equations were

developed to calculate these efforts from the catenary arrow that is required for the wire air-termination. Therefore, it was reversed the practice of calculating this type of sizing, where the effort is calculated and after that, the arrow of a catenary of the cable is calculated. The developed equations include both for supports for the same height and different heights, as shown in Figure 26 and Figure 27. However, these equations are not yet coded in a programming language.

The mathematical approach to dimension the efforts in the wire air-termination was developed in a spreadsheet and the results are in tables such as the Table 1. Other temperature ranges are possible. This is for information only.

8.3. Module to consolidate the list of materials of a project

The insertion of the LPS type in the sizing sections already allows to insert in the database the list of materials. Thus, it will be possible from commercially standardized systems to associate the spare parts list required for that dimensioning, when applicable. This module should have a restricted section to address the material list of grounding systems. There is no prediction to perform the calculation of grounding systems because several software are already dedicated to this purpose. Hence, the module includes only the integration of materials to the LPS design.

8.4. Module for sizing parallel wire air-termination system

We already have several mathematical models responsible for sizing this kind of LPS by parallel wire air-termination in several configurations for this type of systems, as shown in Figure 28 and Figure 29 (other details can be found in the XIII SIPDA [9]). However, we do not yet have the conversion of equations into programming codes, but this is not the biggest problem for insertion. The understanding is that the greatest difficulty is to develop the graphic formulation because it is more complex and I still do not have training in this area, and the addition of this qualification will require planning.

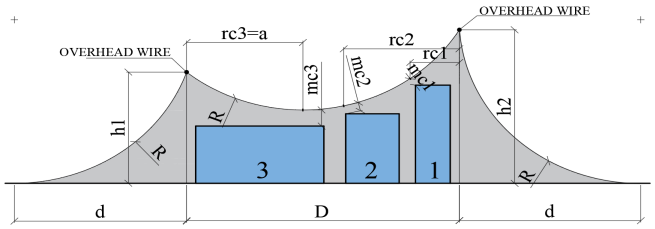


Figure 28: Representation for different levels of parallel wire air termination

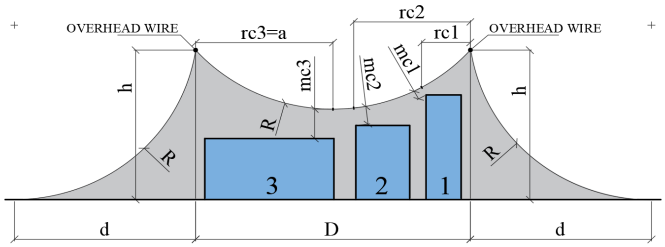


Figure 29: Representation for equal levels of parallel wire air termination

The mathematical approach to determine the limits of a parallel wire air-termination design, considering a possible unevenness of the terrain – at the time of the project implementation or spacing between the wire air-termination or both – produces a table as the one presented in Table 2. This is only an example.

Table 1: Calculated forces for the wire air-termination

Span: 20,30 (m)			Arrow of a catenary: 0,20 (m)		Length: 20,30 (m)		Reference temperature: 26ºC
Item	Temp.: (ºC)	Length: A (m)	f (m)	To (kgf)	T (kgf)	p_s (kgf)	Transverse force (kgf)
01	16	20,303	0,153	137,63	137,63	4,151	2,85
02	20	20,304	0,175	120,55	120,55	4,151	2,85
03	26	20,305	0,203	103,78	103,79	4,151	2,85
04	30	20,306	0,220	95,84	95,85	4,151	2,85
05	36	20,308	0,243	86,75	86,75	4,151	2,85
06	40	20,309	0,257	81,95	81,95	4,151	2,85
07	45	20,310	0,274	76,94	76,95	4,151	2,85
08	22	20,304	0,185	138,38	138,38	4,151	2,85
09	36/16	20,301	0,075	280,12	280,12	4,151	2,85
10	36/22	20,302	0,128	200,00	200,01	4,151	2,85

Note: Reference temperature: 26ºC

Table 2: Calculations of the LPS Limits for Figure 28

Designation		Distance between LPS)					
		30.0 (m)		30.5 (m)		31.0 (m)	
Reduce level	Height of the lower LPS	a	hpc	a	hpc	a	hpc
-	16.00	13.23	10.50	13.55	10.21	13.87	9.92

0.50	15.50	13.03	10.18	13.35	9.89	13.68	9.60
1.00	15.00	12.83	9.85	13.16	9.56	13.49	9.27
1.50	14.50	12.63	9.51				

8.5. Development of tools to integrate with software graph

Another important tool is the integration with design software, such as Autocad. Therefore, the future idea is to develop tools to migrate the sizing data of the various Modules and Sections to enable the plotting of drawings with the horizontal coverage planes for the various Covering projection for the various classes of protection of 15 meters high LPS over structures with a height of 3.5 m. These tools should incorporate the generation of 3D plants, which the resources necessary to rotate and promote cuts of the protected volume.

9. Software security requirements

This software is intended to be a tool used by professionals who design systems to protect against atmospheric discharges – either by direct or nearby impacts – and these might cause harmful effects with high probability of causing losses either by L_1 - losses of human life (including permanent injury); L_2 - losses of services to the public; L_3 - losses of cultural heritage and L_4 - losses of economic value. Thus, it is necessary for these professionals to carry out a very detailed and safe risk analysis. Therefore, safety requirements were established, with the goal of blocking errors or deviations from the requirements established by the standards on which this application was based. These requirements are as follows:

9.1. Controle na determinação das Zonas

The System alerts and blocks Zones from being repeated in the Risk Analysis, as shown in Figure 31.

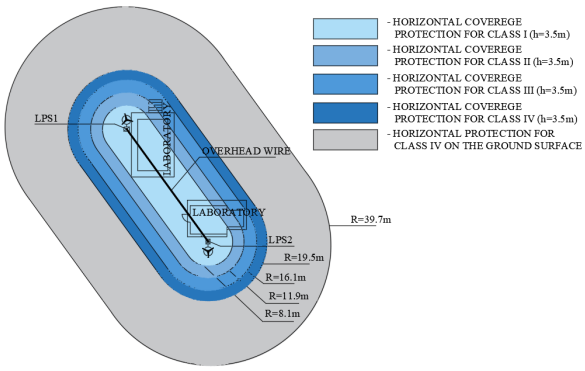


Figure 30: Horizontal projection screen of an LPS project (only for guidance)

9.2. Control of the total number of people in the Zones

When the number of people distributed in all Zones is inserted, the System verifies consistency. So, in case of divergence with the total number reported in the Registry (Figure 4), it issues an alert and blocks for further correction, either for more (Figure 32) or for less (Figure 33):

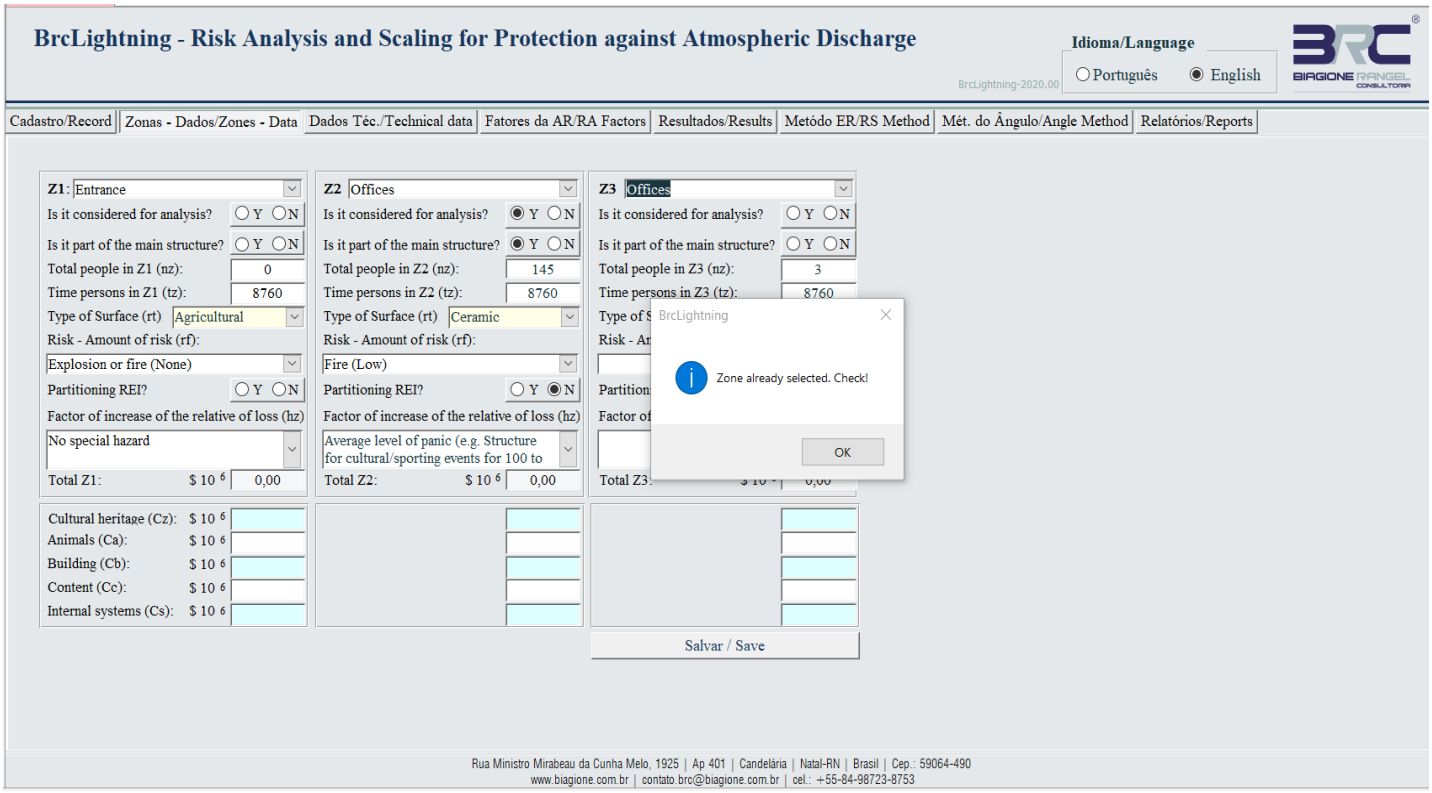


Figure 31: Zone alert already selected

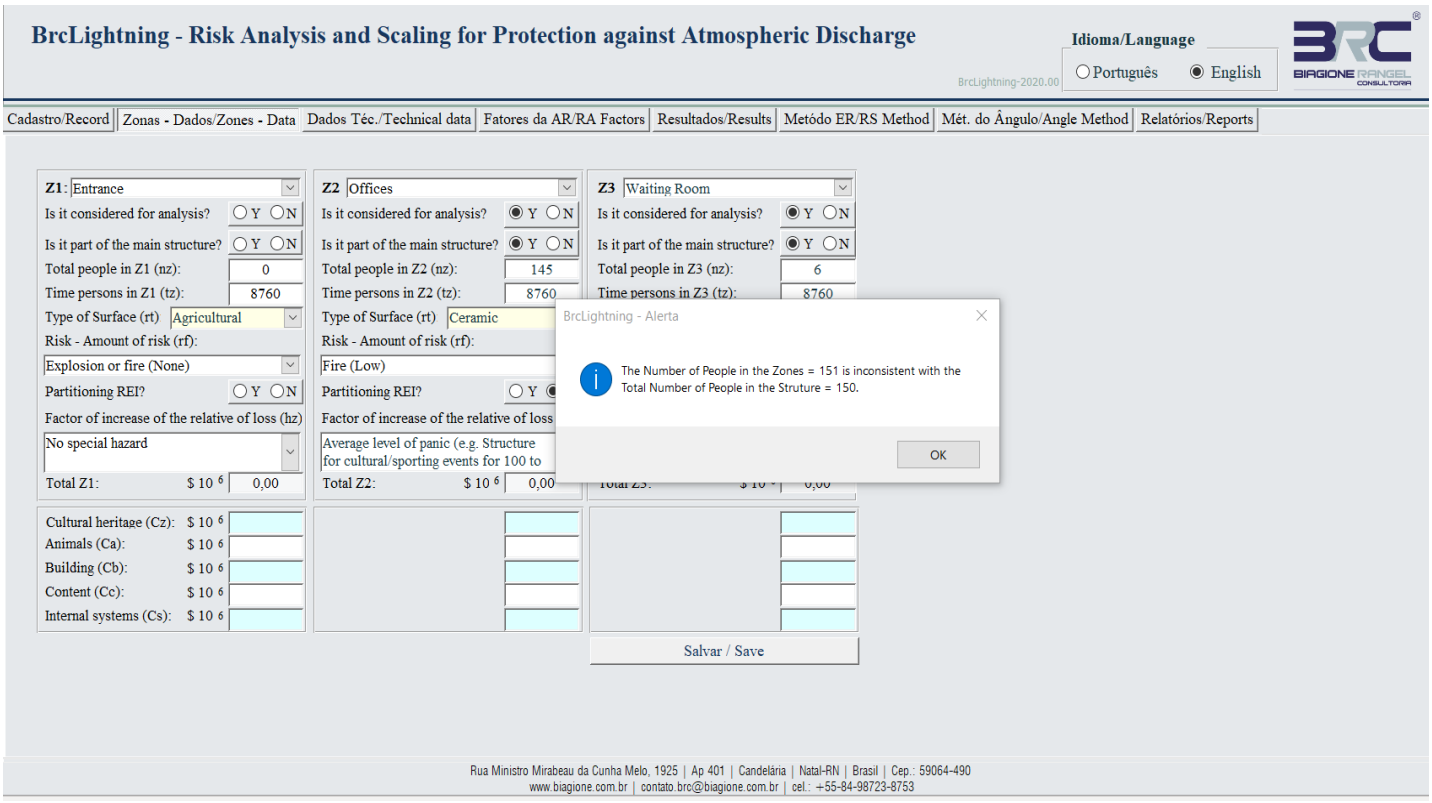


Figure 32: Alert of more people in the Zones than the registered total

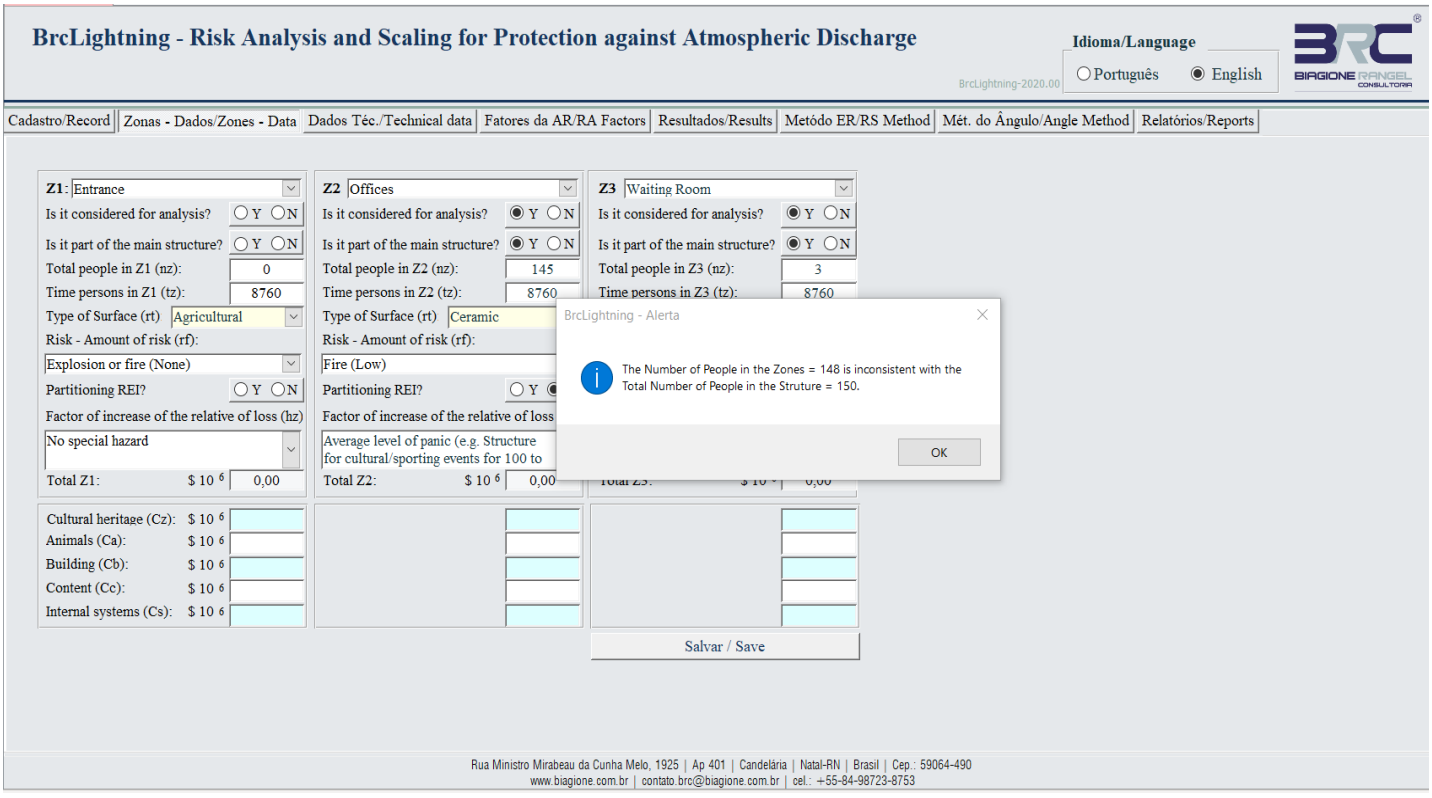


Figure 33: Alert of fewer people in the Zones than the total registered

9.3. Consistencies of line parameters

The System filters the options of values of the probability PLD depending on the resistance RS of the cable screen and the impulse withstand voltage UW of the equipment, based on the line type, routing, shielding, and bonding conditions, as can be seen in Figure 34. Power line is not bonded to the same bonding bar as equipment and telecom is bonded to the same bonding bar as equipment.

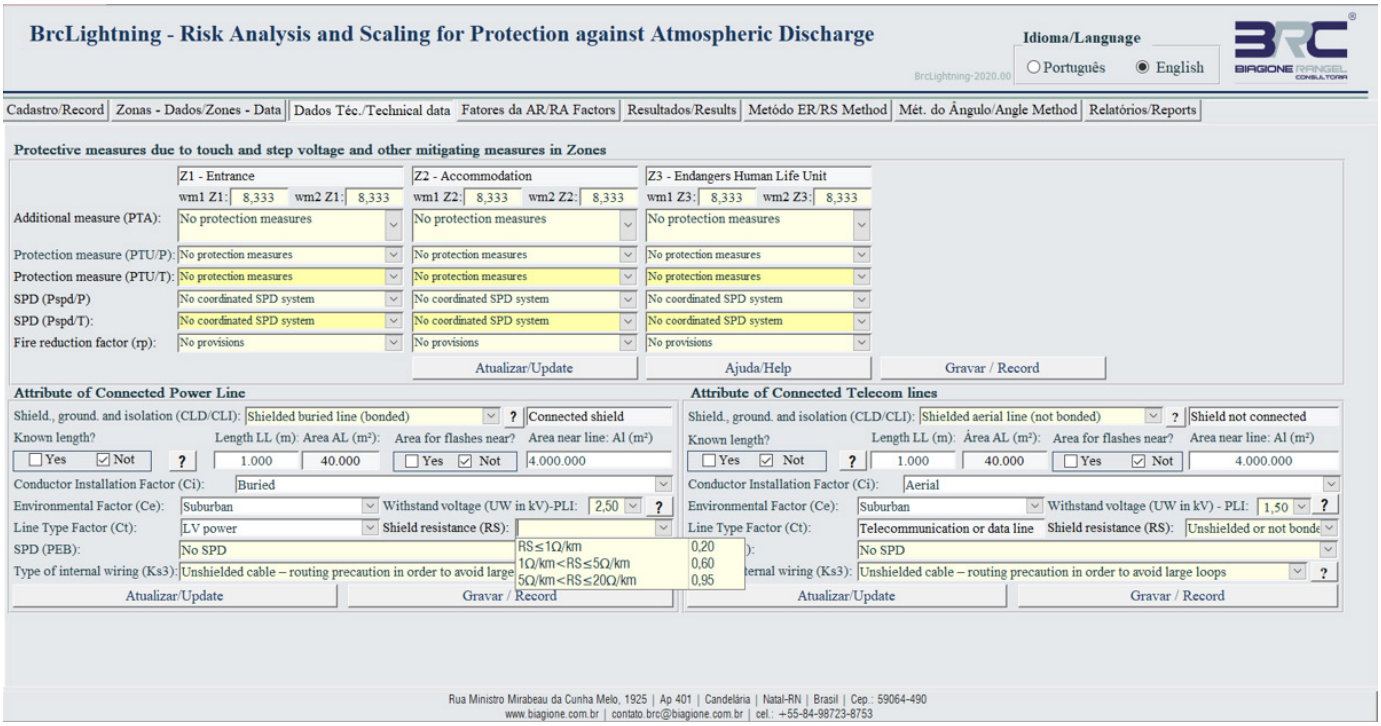


Figure 34: Filters for line parameters

9.4. Restrictions for changing typical loss values

When it comes to loss of human life (L1), the typical LT loss values (loss due to injuries from electric shock) cannot be changed (the system freezes). The same applies to the typical LF loss value (loss in a structure due to physical damage) when it is indicated that there is a risk of explosion in the Zone or when it is indicated that the place is a hospital, hotel, school or civic building.

In addition, changes in typical LO loss values (loss in a structure due to failure of internal systems) are blocked when it comes to areas that are part of the hospital or life support structure. When dealing with L2 and L3 losses, the values of LF and LO are also blocked for changes. For losses of type L4, permissions similar to L1 losses are applied. Blocking alerts are via a pop-up message, as shown in Figure 35.

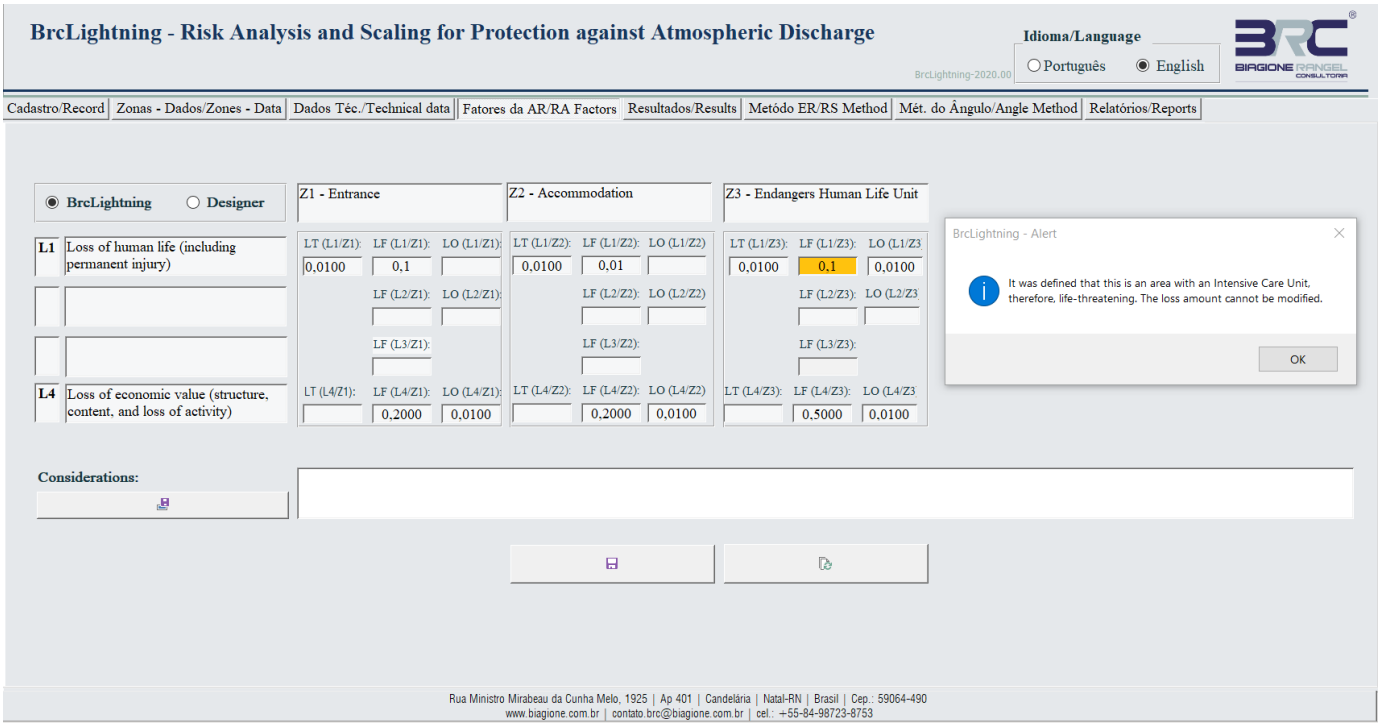


Figure 35: Blocking alert screen for changing typical loss value

9.5. Permission to modify data and customize

It is important to emphasize that to perform any customization and or modify parameters predefined in the System, there is a need for the professional to make, in advance, his or her registration in the system and, after that, he or she will have to define on the initial screen (Figure 4) the designer responsible for the project. This way, the system will save the modifications in an associated way with the person responsible for performing them.

10. Acknowledgment

This is a work that I have been developing alone, from the need I had to carry out the evaluation of projects of Protection Against Atmospheric Discrimination. Therefore, to be certain about my work, I developed the methodology and then the migration to the programming codes of this system, having paid every expense with my own resources, including all the development so far and the costs regarding my participation in all the previous symposia and the development of the papers.

11. Conclusion

This paper is in line with the conclusion of the work originally presented at the SIPDA XV symposium [1], as it incorporates new resources to provide LPS designers with more efficient and safe tools and means for project development, in accordance with Brazilian, International and American standards.

In this context, the BrcLightning application continues, also, with the objective of consolidating itself as an integrated software, having as a support, for some of its modules, the dimensioning of LPS based on the mathematical approach, which was developed and presented at the SIPDA XIII [9] and SIPDA XIV [10] symposia. Thereby, we hope to be contributing to a better understanding of the standards and to the development of safer and more effective projects for protection against lightning strikes.

Also, its modulated design and the format of a database allows the development of applications to be used on the internet, online, and by app systems used on cell phones or other mobile devices.

References

[1] Araujo, R. Biagione, " BrcLightning - Risk Analysis and Scaling for Protection against Atmospheric Discharge," International Symposium on Lightning Protection (XV SIPDA), **1**(0), 8-17, 2019.

[2] ABNT, "NBR-5419-2. Lightning protection Part 2: Risk management," ABNTCatálogo, p. 116, 2015.

[3] IEC, "IEC 62305-2. Protection against lightning - Part 2: Risk management," IEC, 2012.

[4] ABNT, "NBR-5419-3. Lightning protection – Part 3: Physical damage to structures and life hazard," ABNTCatálogo, 2015.

[5] IEC, "IEC 62305-3 - Protection against lightning – Part 3: Physical damage to structures and life hazard," IEC, 2012.

[6] NFPA, " NFPA 780. Standard for the Installation of Lightning Protection Systems. s.l. : IHS," NFPA, 2014.

[7] ABNT, "NBR-5419-4. Lightning protection – Part 4: Electrical and electronic systems within structures. n.l.," ABNTCatálogo, 2015.

[8] IEC, "IEC 62305-4 - Protection against lightning – Part 4: Electrical and electronic systems within structures. n.l. : IEC,," IEC, 2010.

[9] Araujo, Biagione R, Oliveira Jose T. , "Mathematical Approach Methodology to Analysis and Design of LPS," International Symposium on Lightning Protection, **13**(SIPDA), 2015.

[10] Araujo, Biagione R, "Mathematical Modeling for Analysis and Design of LPS Angle Method," International Symposium on Lightning Protection (SIPDA XIV), 2017.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 629-634 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Comparative Study of Control Algorithms Through Different Converters to Improve the Performance of a Solar Panel

Zouirech Salaheddine[*,1], El Ougli Abdelghani[2], Belkassem Tidhaf[3]

[1]Team SEERIA, ENSAO School, Electronics and Systems Laboratory, Faculty of Science, Mohammed First University, Oujda, Morocco

[2]Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, 30050, Morocco

[3]National School of Applied Sciences, Mohammed First University, Oujda, 60020, Morocco

ARTICLE INFO

Article history:
Received: 26 December, 2020
Accepted: 22 February, 2021
Online: 20 March, 2021

Keywords:
Fuzzy logic 'FL'
Perturb-observe 'P&O'
MPPT
Boost Buck SEPIC CUK

ABSTRACT

This article aims at comparing two controls to follow the maximum power point, making use of DC-DC converters for PV uses. All transformers operate continuously. To fulfil maximum power, we will exploit two MPPT controls: a traditional perturb – observe 'P&O' and a smart one – the fuzzy logic 'FL'. The goal of this article is two-fold: to scrutinize the efficiency of DC-DC transformers (Boost, Buck, Cuk and SEPIC), and to assess the outcomes of the simulation. For the construction of models and simulations, the Matlab / Simulink environment is employed.

1. Introduction

There has been recently a crucial advancement in renewable energies. With its permanent potential and without adverse effect on the external environment, new energies are an adequate and affordable technique for development. The main lines of research concerning the conversion of new energies are as follows: recovery of basic energy, mechanical transformation, electrical transformation, electricity production, conversion, and injection into the network. The main axis is to determine the quality of "green" energy generation [1, 2].

Transformers DC-DC are an essential constituent in power generation. These converters are chiefly used in connection with batteries, wind turbines, hybrid systems, solar panel. Transformers DC-DC are to match the tension between the input stage and the output stage of a system. In this work we will make a comparison between different converters: Boost, Buck, Sepic and Cuk for two commands perturb – observe 'P&O' and fuzzy logic 'FL', in order to obtain the maximum power of a solar panel source for meteorological variations conditions [3, 4].

2. Modeling of Photovoltaic System

The fundamental element of the solar panel is the semiconductor, consisting of many cells mounted in shunt or mounted in cascade. The photovoltaic cell is schematized by a circuit: a parallel resistance, a series resistance, and a single diode, represented in figure 1 [5].

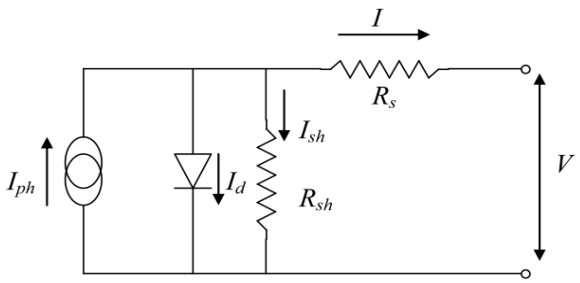


Figure 1: Circuit of solar panel cell.

The relationship of the electrical intensity current I_{ph}:

$$I_{ph} = [I_{sc} + K_I(T_c - T_{ref})]\frac{G}{G_{ref}} \quad (1)$$

whereby K_I is the temperature coefficient [%/K]; I_{sc} is the short intensity current [A]; T_c is the module temperature [K], G is the sunlight solar effect [W/m²]; T_{ref} = 298 K and G_{ref} = 1000 W/m². [6, 7]

*Corresponding Author: Zouirech Salaheddine, zouirechsalaheddine@gmail.com

https://dx.doi.org/10.25046/aj060272

The formula giving the electrical intensity I is:

$$I = I_{ph} - I_d - I_{sh} \quad (2)$$

whence

$$I = [I_{sc} + K_I(T_c - T_{ref})]\frac{G}{G_{ref}} - I_s\left[\exp(\frac{q.(V + R_s I)}{N.K.T} - 1\right] - \frac{(V + R_s I)}{R_{sh}} \quad (3)$$

where

I_{ph} the short electrical intensity,
I_s the electrical intensity saturation of the diode;
q the elementary charge of the electron ($1.60.10^{-19}$ C);
V the diode tension (V);
K the Boltzmann's universal constant ($1.38.10^{-23}$ J/K);
T the temperature,
N Ideality factor of the diode,
R_s the cascade resistance,
R_{sh} is the parallel resistance,

3. Prepare Transformers DC-DC analysis

The transformers DC-DC is a very important element in power generation of a solar panel. It is a block between the load and the photovoltaic panel that allows to extract the maximum power for the panel.

The transformers DC-DC consists of a capacitor, a diode, an inductor, commanded switch (controlled by PWM signal) and uncontrolled switch.

3.1. Buk Transformer

The buck transformer makes it possible to reduce the input tension [8].

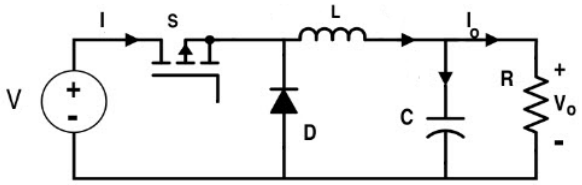


Figure 2: Buck transformer.

The relationship between input tension and output tension is:

$$V_0 = DV \quad (4)$$

where D represents the duty cycle.

3.2. Boost transformer

The boost transformer allows to increase the input tension [9, 10].

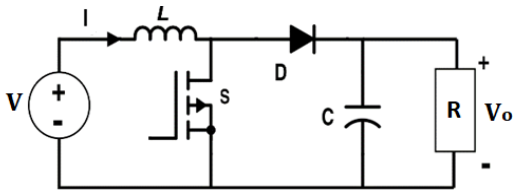


Figure 3: Boost transformer.

The relationship between input tension and output tension is:

$$V_0 = \frac{1}{1-D}V \quad (5)$$

3.3. Cuk transformer

Cuk transformer converts the input tension into a tension of the opposite sign. [11, 12].

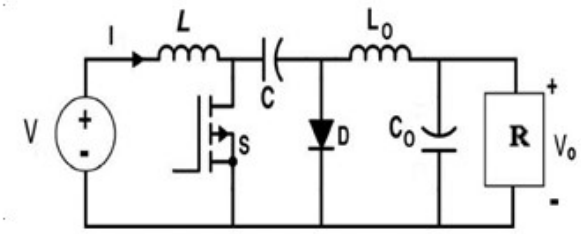


Figure 4: Cuk transformer.

The relationship between input tension and output tension is:

$$V_0 = -\frac{D}{1-D}V \quad (6)$$

3.4. SEPIC transformer

The SEPIC transformer is drawn from of the Cuk which gives a positive output tension [13].

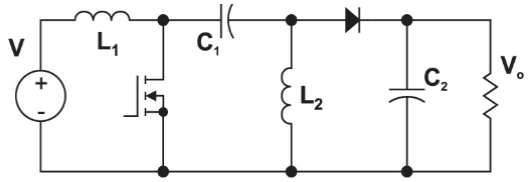


Figure 5: Sepic transformer

The relationship between input tension and output tension is:

$$V_0 = \frac{D}{1-D}V \quad (7)$$

4. MPPT techniques

The aim of the maximum power point tracking technique is to determine the maximum power operating point by changing the duty – cycle ratio of the DC-DC transformer, and then moving to the relevant point.

In this article two MPPT techniques are used: Perturb – Observe and fuzzy – logic 'FL'.

4.1. Technique P and O

The perturb and observe P&O command is ranked among the most used techniques. As its name suggests, this technique involves disturbing the V_{PV} voltage of a low amplitude around its initial value, and examining the impact on the power at the output of the Solar panel.

Figure 6 shows the structure of the Perturb – Observe command [14,15].

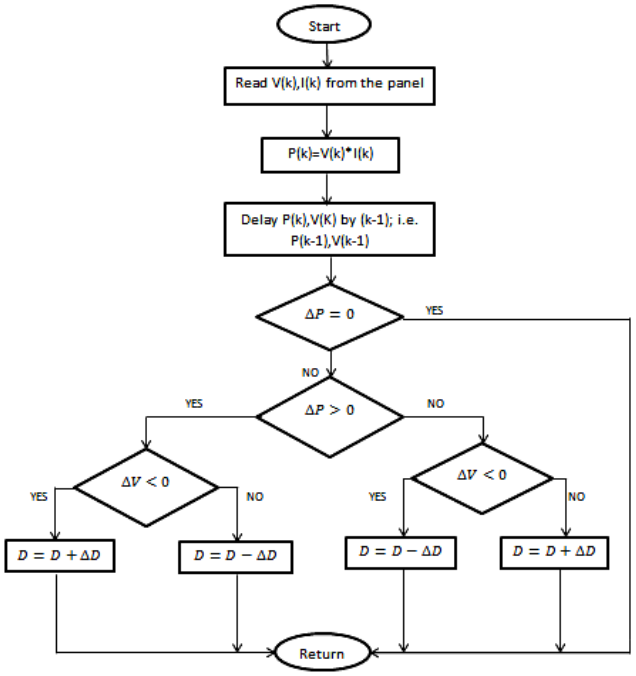


Figure 6: Organization chart P&O

4.2 Fuzzy – Logic 'F-L'

The 'F-L' technique is used in a point tracking system maximum MPPT power, this technique offers the importance of being a powerful technique and relatively simple to develop and it does not require exact specification of the model to be regulated. The 'F-L' technique is more suitable for non-linear systems than photovoltaic systems. The process of this technique is executed in three blocks: Fuzzification, defuzzification and inference. This figure supplies the clue 7 [16].

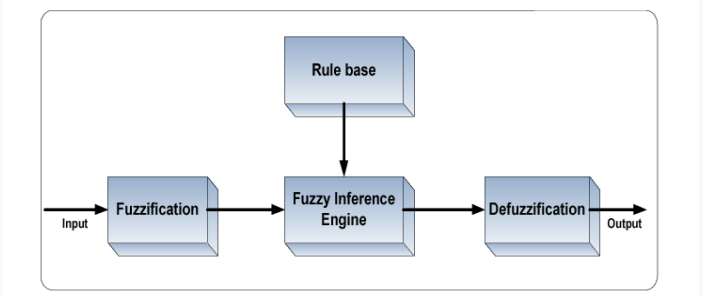


Figure 7: 'F-L' diagram

The 'F-L' diagram input parameters are: the variation of the error CE and the error E giving by the formula:

$$E = \frac{P(k) - P(k-1)}{V(k) - V(k-1)} \quad (8)$$

$$CE = E(k) - E(k-1) \quad (9)$$

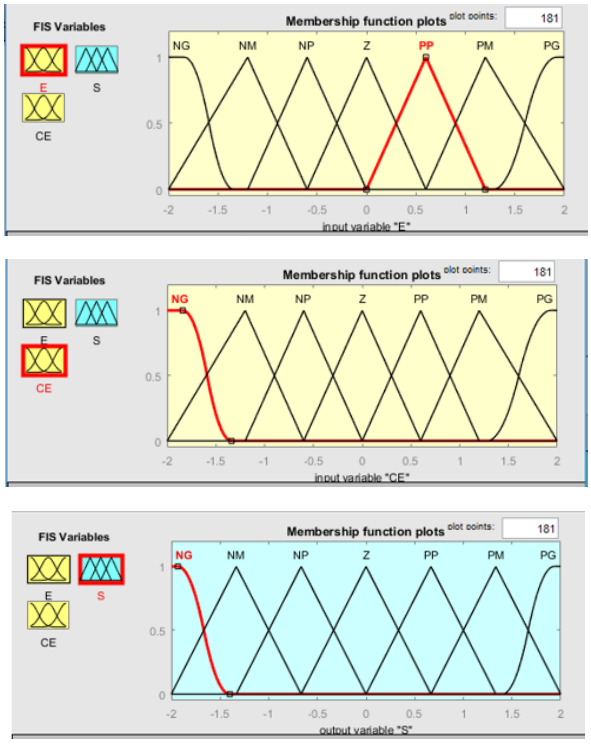


Figure 8: Output and inputs parameter

Table I below shows the different combination inference rules that link the linguistic parameters E (Error), CE (Variation of error) and output S (Output Variable).

Table 1: 'F- L' INFERENCE

E / CE	NG	NM	NP	Z	PP	PM	PG
NG	NG	NG	NG	NM	NM	NP	Z
NM	NG	NG	NM	NP	NP	Z	PP
NP	NG	NM	NM	NP	Z	PP	PM
Z	NM	Z	NP	Z	PP	PM	PM
PP	NM	NP	Z	PP	PM	PM	PG
PM	NP	Z	PP	PM	PM	PG	PG
PG	Z	PP	PM	PM	PG	PG	PG

5. Simulation and Results

Under Matlab Simulink environment the PV system is modeled. In this block we find mainly: the equivalent diagram of the converter, the command used and a PV system predefined in Matlab simulink.

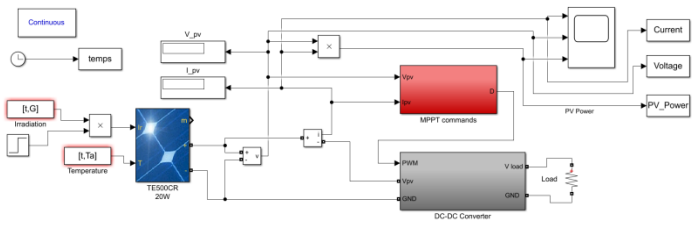


Figure 9: Diagram Block of the Perturb—Observe and fuzzy—Logic commands for the different transformers processed

Figure 9 presents an overview of all blocks: photovoltaic panel, DC-DC converter, MPPT commands, irradiation data et temperature data.

The Solar panel used in the simulations is P20-36-TDC-36-20W.

Table 2: P20-TDC-36-20W

No-Load tension V_{OC} (V)	21,20
Tension at maximum power V_{mp} (V)	17,20
Short-intensity current I_{sc} (A)	1,280
Current at maximum power I_{mp} (A)	1,170
Puissance maximum (W)	20,0
The Coefficient of temperature of V_{OC} (%/deg.C)	0.360990
Coefficient of temperature of I_{sc} (%/deg.C)	0.0650

Figures 10 and 11 depicts the change of the real solar radiation the and temperature data over a day.

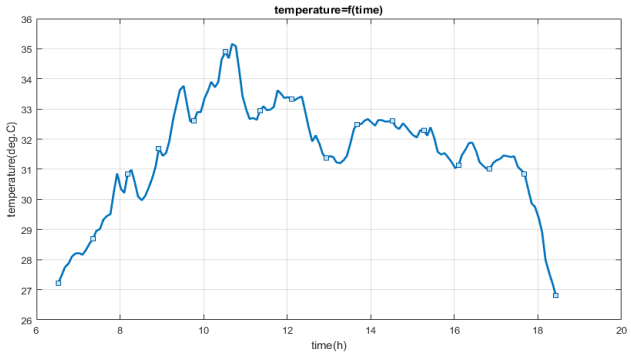


Figure 10: The change of temperature

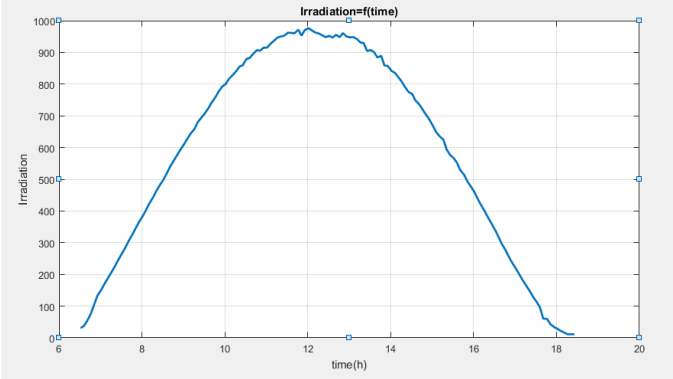


Figure 11: The change of solar radiation

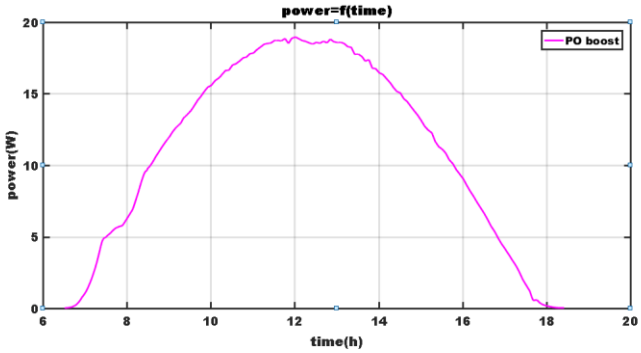


Figure 12: Power with Perturb-observe technique for the Boost transformer

Figures 12,13,14 and 15 enumerates the outcomes of the different transformers: Boost, Buck, CUK and SEPIC, using as a technique the Perturb and Observe.

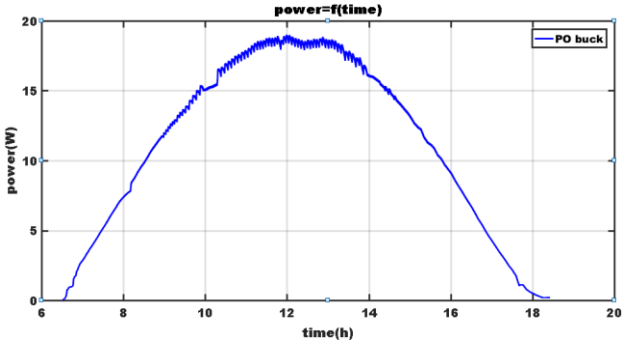


Figure 13: Power with Perturb-observe technique for the Buk transformer

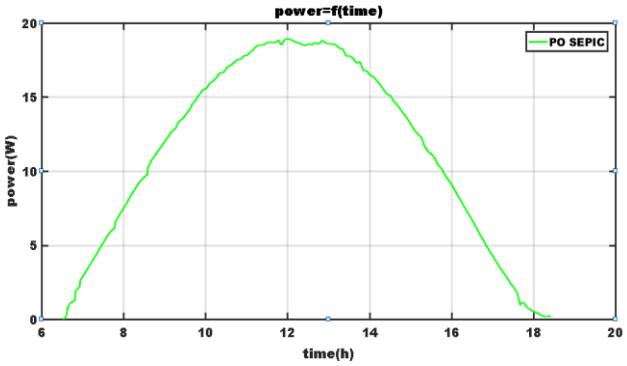


Figure 14: Power with Perturb-observe technique for the Sepic transformer

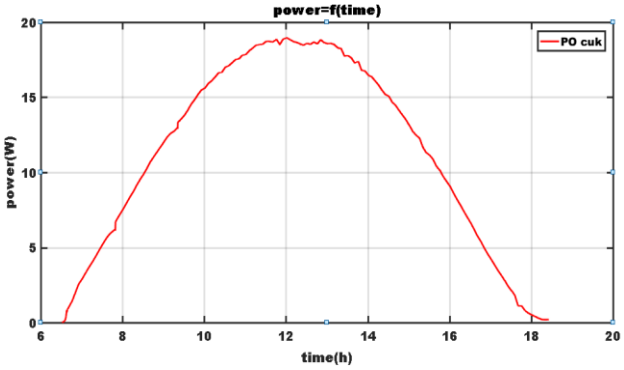


Figure 15: Power with Perturb-observe technique for the Cuk transformer

For the second simulation figures 16,17,18 and 19 the fuzzy logic command was used.

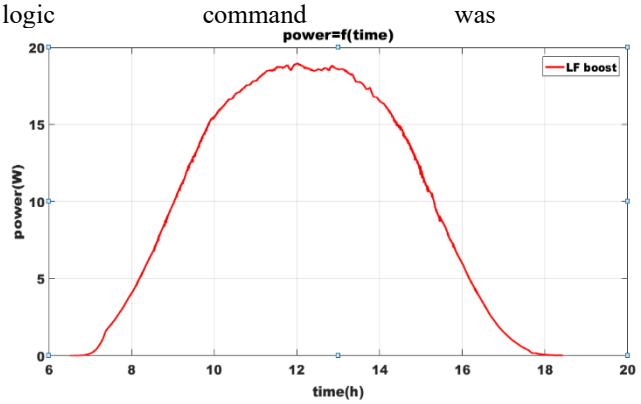


Figure 16: Power with 'F-L' technique for the Boost transformer

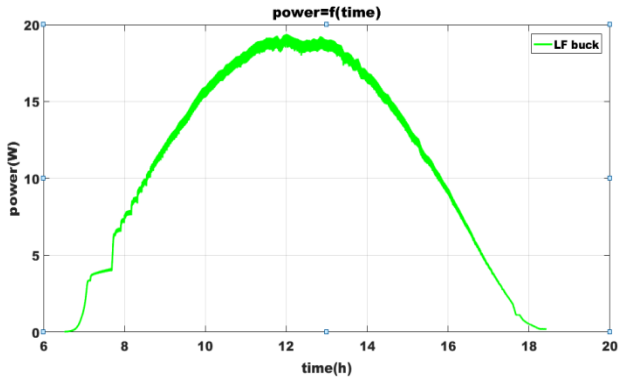


Figure 17: Power with 'F-L' technique for the Buck transformer

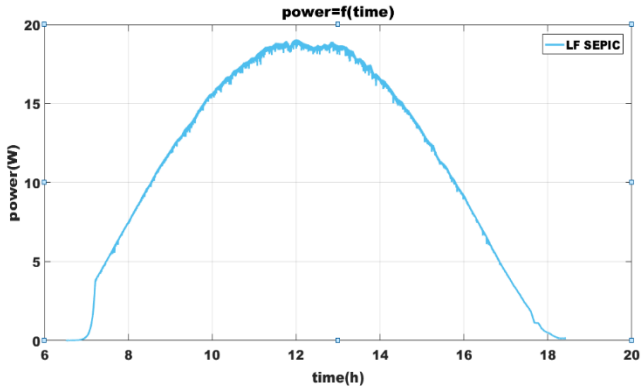


Figure 18: Power with 'F-L' technique for the Sepic transformer

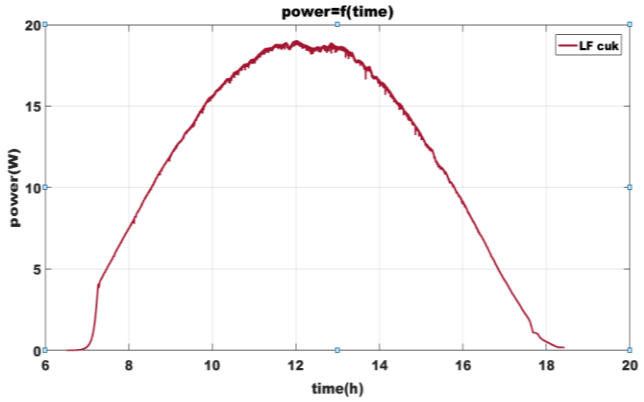


Figure 19: Power with 'F-L' technique for the for the Cuk transformer

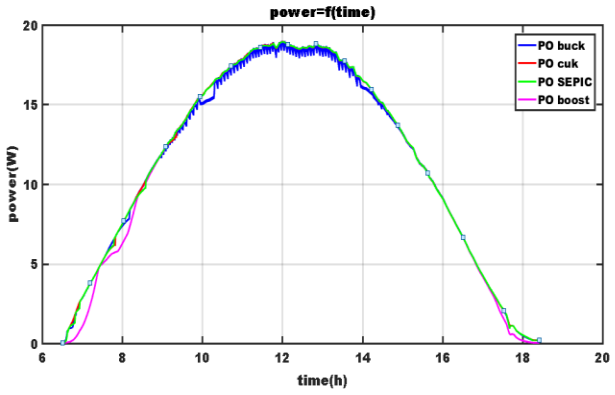


Figure 20: Evolution of power for the four transformer with Perturb-Observe

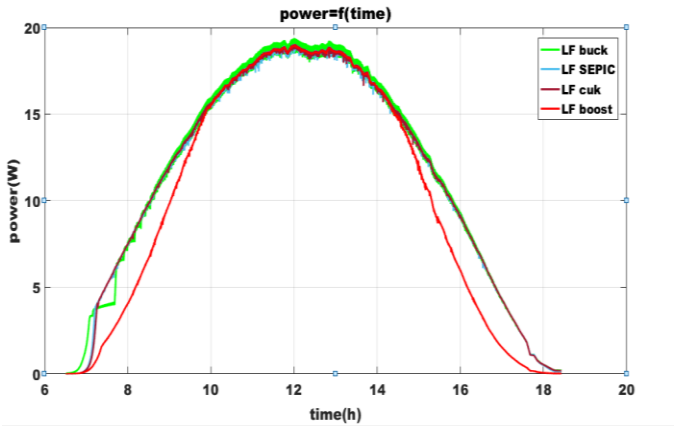


Figure 21: Evolution of power for the four transformer with 'F-L'

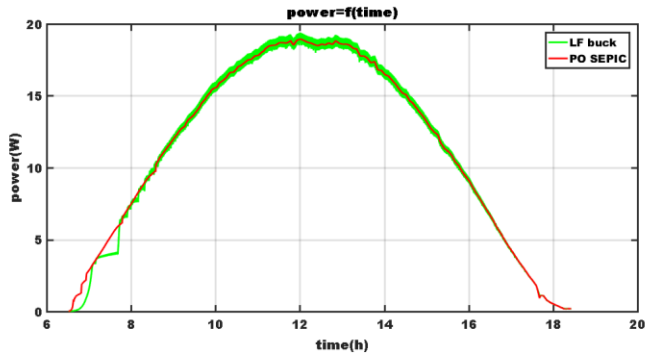


Figure 22: The best power produced with Perturb-Observe and fuzzy-logic

In the Perturb-Observe technique simulation context, the power of photovoltaic panel in the Sepic transformer – is maximal with regard to other transformers, but it is more stable in the boost transformer.

In the 'F-L' technique, the power photovoltaic panel in the Buck transformer generate more power compared to other transformers; it is also more stable, but in Boost transformer context.

6. Conclusion

This article studied two maximum power point tracking techniques – a traditional Perturb-Observe and a smart one: 'F-L' control to obtain the maximum power of the panel photovoltaic.

In this work, four DC-DC transformers are used to improve the quality of the MPP tracking by analyzing the power produced and the stability of each converter.

Simulation results gives an overview of the power evolution for a temperature and solar effect irradiation data for a day.

References

[1] S. Zouirech, M. Zerouali, A. El Ougli, B. Tidhaf, "The Impact of the Type of Converter and the Algorithm of the Control on the Production of Maximum Power by a Photovoltaic System, " 7th International Renewable and Sustainable Energy Conference (IRSEC). IEEE, 1-5, 2019, doi:10.1109/IRSEC48032.2019.9078204.

[2] M. Zerouali, S. Zouirech, A. El Ougli, B. Tidhaf, H. Zrouri, "Improvement of Conventional MPPT Techniques P&O and INC by Integration of Fuzzy Logic, " 7th International Renewable and Sustainable Energy Conference (IRSEC). IEEE, 1-6, 2019, doi:10.1109/IRSEC48032.2019.9078330.

[3] S. Saravanan, N.R. Babu, "Maximum power point tracking algorithms for photovoltaic system," Renewable and Sustainable Energy Reviews, **57**, 192-204, 2016, doi:10.1016/j.rser.2015.12.105.

[4] J.P. Ram, T.S. Babu, N. Rajasekar, "A comprehensive review on solar PV maximum power point tracking techniques,". Renewable and Sustainable Energy Reviews, 67, 826-847, 2017, doi:10.1016/j.rser.2016.09.076

[5] S. Miqoi, A. El Ougli, B. Tidhaf, A.Rabhi, "Application of fuzzy logic on a PV water pumpingsystem," 3rd International Symposium Environmental Friendly Energies and Applications (EFEA), St. Ouen, 1-6, 2014, doi: 10.1109/EFEA.2014.7059961.

[6] T. Arunkumari, V. Indragandhi, S. Sreejith, "Topologies of a DC–DC Converter for Micro-grid Application and Implementation of Parallel Quadratic Boost Converter,". Advances in Smart Grid and Renewable Energy, Springer, Singapore, 633-644, 2018, doi:10.1007/978-981-10-4286-7_63.

[7] S. Zouirech, M. Zerouali, H. Elaissaoui, A. El Ougli, B. Tidhaf, "Application of Various Classical and Intelligent MPPT Tracking Techniques for the Production of Energy through a Photovoltaic System, " 7th International Renewable and Sustainable Energy Conference (IRSEC). IEEE, 1-6, 2019, doi:10.1109/IRSEC48032.2019.9078154.

[8] A.P.K. Yadav, S. Thirumaliah, G.Haritha, "Comparison of mppt algorithms for dc-dc converters based pv systems," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, **1**(1), 18-23, 2012., ISSN: 2278– 8875.

[9] S. Nemsi, L. Barazane, S. Diaf, A. Malek, "Comparative study between two maximum power point tracking (MPPT) techniques for photovoltaic system," Journal of renewable energies, **16** (4), 773-782, 2013, ISSN: 1112-2242 . doi:10.17/1875-10-4286-2369.

[10] M. Boutouba, A. El Ougli, S.Miqoi, B.Tidhaf, "Design and Experimentation of a Control System Implemented on Raspberry Pi 3 Board for Photovoltaic Systems Using SEPIC Converter," Journal of Electrical Systems, **13**(4), 661-677, 2017, ISSN 1112-5209.

[11] J. Chauhan, P. Chauhan, T. Maniar, A. Joshi, "Comparison of MPPT algorithms for DC-DC converters based photovoltaic systems,". International Conference on Energy Efficient Technologies for Sustainability IEEE, 476-481, 2013, doi: 10.1109/ICEETS.2013.6533431.

[12] R. Patii, H. Anantwar, "Comparitive analysis of fuzzy based MPPT for buck and boost converter topologies for PV application," International Conference On Smart Technologies for Smart Nation (SmartTechCon), Bangalore, 1479-1484, 2017, doi: 10.1109/SmartTechCon.2017.8358610.

[13] L. Jotham Jeremy, C.A. Ooi, J. Teh, "Non-isolated conventional DC-DC converter comparison for a photovoltaic system: A review," Journal of Renewable and Sustainable Energy, **12**(1), 013502, 2020, doi:10.1063/1.5095811.

[14] M.M. Algazar, H. Al-Monier, H.A. El-Halim, M.E.E.K. Salem, "Maximum power point tracking using fuzzy logic control," International Journal of Electrical Power & Energy Systems, **39**(1), 21-28, 2012, doi:10.1016/j.ijepes.2011.12.006.

[15] D. Verma, S. Nema, A.M. Shandilya, S.K. Dash,"Maximum power point tracking (MPPT) techniques: Recapitulation in solar photovoltaic systems,". Renewable and Sustainable Energy Reviews, **54**, 1018-1034, 2016, doi:10.1016/j.rser.2015.10.068.

[16] S. Vinu, V. Anoop, P. Dixit, A. Arjaria; U. Chourasia; P. Bhambri; R. MR, R. Sundararaj, "CCGPA-MPPT: Cauchy preferential crossover based global pollination algorithm for MPPT in photovoltaic system," Progress in Photovoltaics: Research and Applications, **28**(11), 1128-1145, 2020, doi:10.1002/pip.3315.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 631-635 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Modelling Human-Computer Interactions based on Cognitive Styles within Collective Decision-Making

Nina Bakanova[1], Arsenii Bakanov[2], Tatiana Atanasova[3,*]

[1]*Keldysh Institute of Applied Mathematics – RAS, 125047, Russia*

[2]*Institute of Psychology of Russian Academy of Sciences, 129366, Russia*

[3]*Modelling and Optimization Department, Institute of Information and Communication Technologies – BAS, 1113, Bulgaria*

ARTICLE INFO

Article history:
Received: 01 September, 2020
Accepted: 13 January, 2021
Online: 05 February, 2021

Keywords:
Human–computer communication
Information flows
Cognitive styles
Collective decision-making

ABSTRACT

The article proposes an approach to evaluate human-computer interaction in the collective decision-making model. It is believed that all team members interact with each other through a distributed information system. The approach involves considering, when modelling, the personality characteristics of perception, each member of the team as a set of cognitive styles. Within the scope of the proposed technique, it is believed that information flows are interconnected with the processes of collective decision-making, which makes it possible to model the process of collective decision-making, monitor and analyse the effectiveness of the collective's activities. Experimental studies accomplished with statistical data processing were carried out and discussed.

1. Introduction

This paper is an extension of the work originally presented in Big Data, Knowledge and Control Systems Engineering (BdKCSE'2019) Conference [1]. The article further develops aspects of human-machine interaction in a distributed information environment. Here the emphasis is on collective decision-making taking into account the cognitive features of perception of team members.

The relevance of research in this area is due to a number of reasons, including the processes of globalization of the economic development of industrial sectors, the creation of a great number of large-scale, distributed corporations, concerns, holdings, and, accordingly, the decentralization of management processes controlled by parent organizations. Information interaction of such corporations is carried out through distributed information systems. The distributed information landscape is a genuine development of computerized, data and interaction technologies. The challenge of human communication with the distributed information ecosystem is gaining growing methodical and pragmatic importance.

A significant problem in the management of large-scale, distributed organizations are collective decision-making processes, which require the interaction of many people; this inevitably leads to the complication and delay of the decision-making process, as well as to the possible emergence of contradictions in the process of solving the problem.

The importance of taking into account the opinion of each member of the team, the importance of obtaining a coordinated, balanced solution to complex problems, the importance of ensuring transparency of the decision-making process in the team - all of the above emphasizes the relevance of developments in the field of creating automated tools to support collective decision-making. The development of mechanisms and techniques that provide modelling and visualization of collective decision-making processes will increase the efficiency and effectiveness of decision-making in a geographically distributed team through the implementation of modes of human-computer interaction.

When studying the processes of interaction in large-scale distributed information networks, we can talk about information flows connecting all users of the information system. As a result of the advance of information and computer technologies, now there are opportunities for the development of intelligent modules that allow to analyse, visualize and improve the efficiency of interaction processes in a team, including such as collective decision-making processes.

This article is devoted to the development of methods for modelling and presenting processes related to collective decision

*Corresponding Author: Tatiana Atanasova, atanasova@iit.bas.bg

making. First, the development of decision theory over time is discussed. It is then proposed to use a cognitive modelling methodology to rank team members in the collective decision-making process. The reflection of the cognitive styles of the participants in the human-computer interaction with distributed information system is emphasized in further consideration. Experimental studies accomplished with statistical data processing were carried out. The cognitive styles of the collective members were investigated through an experimental study to be included in the cognitive map modelling. Finally, the results of the experiments are discussed.

2. The Basis of Decision Theory

The fundamental basis of decision theory in mathematics, computer science and economics is the provision of rational human behaviour and the theory of utility.

The game theory was introduced in 1944 by John von Neumann and Oskar Morgenstern in their book, entitled "Game Theory and Economic Behavior". However, in a number of researches works, for example in [2], it is shown that human behaviour is not always rational, and often, on the contrary, is irrational (from the point of view of gaining a win). As an example, to demonstrate the features of human behaviour, we can cite the paradox proposed by [3]. As a result of presenting lotteries to various groups of people, it was shown that people prefer the lottery where the risk of losing is excluded or minimal. Whereas the computer calculates probabilities and acts rationally, i.e., maximizes utility (or gain).

A study of the evolutionary development of work in the expansion of decision support systems as a scientific direction shows that from the moment Atanasoff and Berry began to develop the first digital computing device in 1939 (at Iowa State College) and before the 1980s, the main goal was "to teach the computer to think like a person". Since the early 1980s up to the present time, the central paradigm has been the concept of "artificial intelligence" [4-10]. According to artificial intelligence professionals, the effectiveness of intelligent systems is determined both by formal mathematical schemes based on mathematical logic and by the knowledge of experts in a specific subject area, which can be explicated in the process of knowledge extraction. The analysis of scientific works in the field of knowledge extraction shows that, despite the significant success achieved in the advancement of intelligent, information systems, the key problem of creating any intelligent system is the process of extracting expert knowledge, as well as original heuristics used by specific experts to solve problems [11-15].

To date, technologies and methodologies for the design of knowledge-based intelligent systems have been developed, in which the sequential passage of the stages of acquiring expert knowledge is realized:

- the stages of extracting [16] and
- conceptual analysis of expert knowledge [17-21].

However, the actual problem of modelling and visualizing the collective decision-making process, as well as modelling and visualizing the information interaction between experts, has not received sufficient consideration, partly because this problem is interdisciplinary, as it is at the intersection of information technology, mathematics, sociology and psychology.

3. Cognitive Maps for Modelling

The use of models to study an object, phenomenon or process is a well-established and well-proven research method. By the term model, in this article, we mean a formal representation of an object, (process, phenomenon) in some form, intended for the study of this object. Formal representation of an object (process, phenomenon) can be:

- mathematical,
- simulation,
- sign-symbolic.

In the present study, cognitive maps were used as a basis for the model of making decision in information interaction within a distributed information system. Using of models based on cognitive maps made it possible to visualize information processes of remote (distributed) interaction in a team, including the processes of creating groups, monitoring remote activities of team members, etc. For visualization and ranking of team members, it is proposed to use the methodology of cognitive modelling [11, 22-27]. According to the methodology, it is supposed to build a fuzzy model, visualized as a fuzzy graph (cognitive map), in which the vertices are team members, and the weighted arcs are information flows, relationships and social significance (weight, authority, awareness, experience, etc.) of each member of the team, the weight of each arc of the graph reflects the strength of the influence of a particular individual on the process of making a collective decision.

Information flows between team members, in general, can be verbal and non-verbal, mediated by the electronic information environment and not mediated. In this work, only flows mediated by a distributed information system are investigated. The considered information flows are subdivided into explicit and implicit, stable and unstable, as well as formal and informal.

Formally, a cognitive map can be symbolized as a directed sign graph *(F, W),* where F is the set of vertices - team members, $W=|w_{ij}|$ - adjacency matrix. The dynamics of the decision-making process is presented as a sequential set of situations $X(t–1), X(t), X(t+1),..., X(t+n),$ which are the situation state vectors at successive discrete times: $t–1, t, t+1, ..., t+n$, where t is some number of such moment in time.

The interactions can be determined by summing all of them, according to the formula (1):

$$X_i(t+1) = X_i(t) + \sum_{i=1}^{n}(X_{i-1}(t+1) - X_{i-1}(t))W_{i,i-1} \quad (1)$$

The use of cognitive maps allows modelling and visualizing not only the information flows, but also the configuration of the team in the process of collective decision taking, that is, the use of cognitive maps allows displaying the dynamic processes occurring in the team in the process of deciding.

4. Cognitive Styles of the Collective Members in the Model

The need for specific information (information flow) for each team member is determined by several factors influencing the

request for information of a team member, namely: knowledge, skills, job responsibilities, style of activity, and personality characteristics of perception. By personal characteristics of perception, we mean the set of cognitive styles inherent in each member of the group. At the same time, we make it a condition that we consider in the model only personal characteristics of perception, since a person's personality traits are not limited to a set of cognitive styles.

4.1. Cognitive-Style Characteristics

According to several scientists, individual ways of processing information are largely determined by cognitive styles [28, 29]. Cognitive styles actively participate in the process of choice and making decisions in work, manage the emotional and behavioural characteristics of human activity [30]. Scientific works show the essential part of the cognitive and stylistic aspects of the subject in the execution of actions that involve self-government and answerability in during process of making decisions, particularly in highly uncertain conditions. In parallel, the assignment of psychological patterns as form-creating (combining, central) elements in the process of choosing an appropriate option is highlighted [31]. Of the entire set of cognitive styles given in the work [32], we will consider only three cognitive styles, namely: the "field dependence/field independence" style, the "narrow/wide range of equivalence" style, the "impulsivity/reflectivity". Detailed justification and appropriateness of the choice of the above styles are given in the work [33].

4.2. Identifying the Cognitive-Style Characteristics

As follows, we briefly describe the cognitive styles studied and present the methods used for the research.

The style "field dependence/field independence" is traditionally considered as a way of an individual to solve perceptual problems, and the presence of "the ability to overcome a complex context" (according to G. Witkin). To diagnose cognitive-style features, G. Witkin's "*Included Figures*" technique was used (evaluation of field dependence - field independence, individual variant) [34].

The style "narrow/wide range of equivalence" reflects the predominant orientation of the individual to the features of similarity or differences of classified objects, their obvious or hidden features. To diagnose these cognitive-style features, the method "*Free sorting of objects*" by R. Gardner and V. Kolga was used.

The style "impulsivity/reflectivity" characterizes individual differences in the speed and correctness of decisions made in situations of uncertainty and the presence of many alternatives. To diagnose these cognitive-style features, the method "*Comparison of similar drawings*" by J. Kagan (assessment of impulsivity - reflectivity as the cognitive pace of decision-making) was used [35].

The choice of these particular styles for research is due to the fact that the team members united in the work on the project must have a good idea of the general structure of the entire project, understand the specifics of the work and the distinct (individual) tasks of other team members within the framework of this project, and have a proportionate cognitive pace of decision-making, carry out synthesis and analysis of project tasks.

In the framework of this study, the relationship between the cognitive-style characteristics of team members and the effectiveness of decision-making when working on a common project was studied.

5. Experimental results

To examine the cognitive-style characteristics of team members in the model, an experimental study was conducted. The most important phases of human interaction with the information system were modelled as the subject was needed to:

- familiarize s/himself with the content of the texts specially developed for this study,
- analyse the content of the texts,
- classify the content of the texts,
- then answer the proposed questions.

Also, during the study, the above mentioned cognitive-style characteristics were identified.

For this experiment, several texts from 2 to 8 pages in volume (250 - 300 words per one page) were prepared specifically. Throughout the experiment, the subject of investigation has to read these specifically arranged texts. After reading, the subject of the investigation was asked to answer questions about the structure of the text and to give a quantitative calculation of the discovered options. The volume of the presented material was taken into account to model the interaction with the information system. Gaze movements were recorded using specialized SMI equipment (http://www.smivision.com), which tracks the trajectory of the subject's gaze.

Several quantitative criteria for assessing the activity of information interaction have been established [1]. Statistical methods for data processing were applied together with using Student's t-test and Mann-Whitney U-test. Spearman's correlation coefficient was used to assess the relationship between variables. Student's t-test was used to analyse two independent data samples to compare the mean of the two populations. Mann-Whitney U-test determines whether the area of overlapping values between the two series is small enough.

The research results are shown in the Table. Questions of type 1 were used for the quantitative assessment, and questions of type 2 were questions about the structure of the text.

Note: Indicators for the assessment of the success and effectiveness of the activity are time for reading of the presented text (document); the keyword usage rate, i.e., how often the subject rereads the reference words; time for answering questions; coefficient of correct answer; coefficient of confidence in decision making i.e., how long the subject analyses the alternatives.

As an outcome of the analysis of the results obtained, the following interdependencies were revealed: 1) the more work experience, the less time during which the subjects look at the text of the document presented on the computer screen, and the less time spent on making a decision regarding the choice of an alternative, 2) on the contrary, the "coefficient of correctness of answers" is directly proportional to the indicator "work experience". That is, subjects with work experience more often solve problems correctly.

Table 1: Correlation relationships between performance criteria and characteristics of socio-demographic status.

Indicators	Gender	Age	Education	Experience
Stage I. Reading text without pivot words				
Reading time	0,61	0,54	0,21	-0,72
Answering time of quantitative questions	0,17	0,27	-0,03	-0,41
Correctness of answer to quantitative questions	-0,04	-0,41	0,28	0,68
Decision confidence factor	0,43	-0,05	-0,01	0,06
Answering time of questions about text structure	0,35	-0,16	0,31	0,10
Correctness of answer to questions about structure	-0,41	0,15	-0,07	-0,37
Coefficient of confidence of answer about text structure	0,05	0,15	-0,37	-0,22
Stage II. Reading text with pivot words				
Reading time	0,26	0,37	-0,03	-0,52
Answering time of quantitative questions	0,17	0,22	0,07	-0,52
Correctness of answer to quantitative questions	-0,61	-0,44	-0,41	0,41
Decision confidence factor	0,48	0,26	0,04	-0,46
Answering time of questions about text structure	-0,43	-0,32	-0,08	0,72
Correctness of answer to questions about structure	-0,10	0,20	0,40	0,17
Coefficient of confidence of answer about text structure	0,78	0,68	0,48	-0,68

It should also be noted that statistical data processing not only made it possible to reveal the relationship between the "coefficient of correctness of answers" and the availability of work experience - this fact is quite expected and natural, but it made it possible to use the cognitive characteristics of the user in developing the model, making them dependent on the style of decision-making in the process of collective activity, with the efficiency and effectiveness of collective decision-making.

6. Conclusion

This article proposes an approach to developing a model of collective decision-making, which makes it possible to consider the cognitive features of a user in the process of interacting with a distributed information system. In the course of the research, it was revealed that respondents characterized by the breadth of the equivalence range (i. e. "synthetics") were the most successful in dealing with the intended tasks on reading, analysing and classifying the proposed texts, as compared with the respondents, style features of which are more inherent in the area of "narrow range of equivalence" (i.e. "analytics"). The results obtained also indicate that such a property as "field independence" allows individuals to cope with tasks more successfully than "field dependent" research participants.

In the process of modelling, analysis and monitoring of information flows were carried out to support the processes of remote group interaction based on a distributed information system. Information flows are interconnected with collective decision-making processes, which allow monitoring and analysing the effectiveness of the team's activities. For this purpose, it is possible to use the method of extracting specialized data samples from information arrays [36], where based on the analysis of the relationship of information flows; two classes-functions of collective information interaction were isolated:

- function-tasks - an interconnected sequence of information flows due to a sequence of actions team members aimed at solving a specific problem, as a result of which they have their own aspects and specifics in each case;
- functions-operations - an interconnected sequence of information flows due to the sequence of actions of team members, which are formal, standard and universal operations in the process of information interaction and are not aimed at solving a specific problem [37].

Based on the monitoring and analysis of the developed model, systematization, ranking and quantitative assessment of information flows were conducted.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] A. Bakanov, T. Atanasova, N. Bakanova, "Cognitive Approach to Modeling Human-Computer Interaction with a Distributed Intellectual Information Environment," in 2019 Big Data, Knowledge and Control Systems Engineering, BdKCSE 2019, 2019, doi:10.1109/BdKCSE48644.2019.9010597.

[2] A. Tversky, D. Kahneman, "Judgment under uncertainty: Heuristics and biases," Science, 1974, doi:10.1126/science.185.4157.1124.

[3] M. Allais, "Le Comportement de l'Homme Rationnel devant le Risque:

Critique des Postulats et Axiomes de l'Ecole Americaine," Econometrica, 1953, doi:10.2307/1907921.
[4] O.I. Larichev, Current Methodological Problems of Systems Analysis and Its Application, 1984, doi:10.1016/b978-0-08-030830-2.50013-1.
[5] A.B. Petrovsky, "Multiple criteria decision making: Discordant preferences and problem description," Journal of Systems Science and Systems Engineering, 2007, doi:10.1007/s11518-007-5032-z.
[6] A.B. Petrovsky, "Multi-method technology for group multi-attribute choice," in RPC 2018 - Proceedings of the 3rd Russian-Pacific Conference on Computer Technology and Applications, 2018, doi:10.1109/RPC.2018.8482143.
[7] A.K. Goel, J. Davies, Artificial intelligence, Book, 2019, doi:10.1017/9781108770422.026.
[8] F. Hayes-Roth, "Knowledge Based Expert Systems," Computer, 1984, doi:10.1109/MC.1984.1658976.
[9] C. Fyfe, "Advanced Information and Knowledge Processing," General Systems, 2010.
[10] B.R. Gaines, "Modeling practical reasoning," International Journal of Intelligent Systems, 1993, doi:10.1002/int.4550080105.
[11] R. Axelrod, Structure of decision: The cognitive maps of political elites, 2015, doi:10.2307/2616237.
[12] J.H. Boose, "A survey of knowledge acquisition techniques and tools," Knowledge Acquisition, 1989, doi:10.1016/S1042-8143(89)80003-2.
[13] J.H. Boose, B.R. Gaines, "Knowledge Acquisition for Knowledge-Based Systems: Notes on the State-of-the-Art," Machine Learning, 1989, doi:10.1023/A:1022662924615.
[14] S.J. Grossman, O.D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, 1986, doi:10.1086/261404.
[15] R. Kaluri, C.H. Pradeep Reddy, "An enhanced framework for sign gesture recognition using hidden markov model and adaptive histogram technique," International Journal of Intelligent Engineering and Systems, 2017, doi:10.22266/ijies2017.0630.02.
[16] R. Kaluri, C.H.P. Reddy, "Optimized feature extraction for precise sign gesture recognition using Self-Improved Genetic Algorithm," International Journal of Engineering and Technology Innovation, 2018.
[17] G. Lbov, V. Berikov, "Construction of an event tree on the basis of expert knowledge and time series," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2011, doi:10.1007/978-3-642-22140-8_21.
[18] V. Desnitsky, I. Kotenko, "Expert knowledge based design and verification of secure systems with embedded devices," in Lecture Notes in Computer Science, 2014, doi:10.1007/978-3-319-10975-6_15.
[19] T. Atanasova, "Integrated semantic-based processes in digital home," Informatyka Ekonomiczna, 2014, doi:10.15611/ie.2014.3.02.
[20] H. Suzuki, R. Hishiyama, "An analysis of expert knowledge transmission using machine translation services", in SoICT '16 - Proceedings of the Seventh Symposium on Information and Communication Technology, 2016, doi.org/10.1145/3011077.3011085
[21] V.N. Golovachyova, G.Z. Menlibekova, N.F. Abayeva, T.L. Ten, G.D. Kogaya, "Construction of expert knowledge monitoring and assessment system based on integral method of knowledge evaluation," International Journal of Environmental and Science Education, 2016, doi:10.12973/ijese.2016.705a.
[22] T. Temel, F. Karimov, "Information systems model for targeting policies: A graph-theoretic analysis of expert knowledge," Expert Systems with Applications, 2019, doi:10.1016/j.eswa.2018.11.014.
[23] B. Kosko, "Fuzzy knowledge combination," International Journal of Intelligent Systems, 1986, doi:10.1002/int.4550010405.
[24] L.A. Ginis, "The use of fuzzy cognitive maps for the analysis of structure of social and economic system for the purpose of its sustainable development," Mediterranean Journal of Social Sciences, 2015, doi:10.5901/mjss.2015.v6n3s5p113.
[25] J.P. Carvalho, "Rule based fuzzy cognitive maps in humanities, social sciences and economics," Studies in Fuzziness and Soft Computing, 2012, doi:10.1007/978-3-642-24672-2_16.
[26] M. León, C. Rodriguez, M.M. García, R. Bello, K. Vanhoof, "Fuzzy cognitive maps for modeling complex systems," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2010, doi:10.1007/978-3-642-16761-4_15.
[27] N. Abramova, Z. Avdeeva, S. Kovriga, D. Makarenko., Subject-formal Methods Based on Cognitive Maps and the Problem of Risk Due to the Human Factor, 2010, doi:10.5772/7118.
[28] N. Volkova, A. Gusev, "Cognitive styles: Controversial issues and research problems," National Psychological Journal, 2016, doi:10.11621/npj.2016.0203.
[29] M. Kozhevnikov, C. Evans, S.M. Kosslyn, "Cognitive Style as Environmentally Sensitive Individual Differences in Cognition," Psychological Science in the Public Interest, 2014, doi:10.1177/1529100614525555.
[30] J.W. Atkinson, Motivation for achievement, 2015, doi:10.4324/9781315720951.
[31] A.R. Masalimova, M.N. Mikhaylovsky, A. V. Grinenko, M.E. Smirnova, L.B. Andryushchenko, M.A. Kochkina, I.G. Kochetkov, "The interrelation between cognitive styles and copying strategies among student youth," Eurasia Journal of Mathematics, Science and Technology Education, 2019, doi:10.29333/ejmste/103565.
[32] E.Y. Korjova, O. V. Rudykhina, "The possibilities of typological approach in the study of epistemological styles," Vestnik of Saint Petersburg University. Psychology, 2016, doi:10.21638/11701/spbu16.2016.302.
[33] A. Bakanov, M. Zelenova Cognitive Styles as Determinants of Success in Professional Activity, Социальная Психология и Общество, 2015, doi:10.17759/sps.
[34] H.A. Witkin, D.R. Goodenough, "Cognitive styles: essence and origins. Field dependence and field independence.," Psychological Issues, 1981.
[35] J. Kagan, "Reflection-impulsivity: The generality and dynamics of conceptual tempo," Journal of Abnormal Psychology, 1966, doi:10.1037/h0022886.
[36] N.B. Bakanova, V.M. Vishnevskii, "Modeling document movement in corporate systems document circulation," Automation and Remote Control, 2008, doi:10.1134/S0005117908090142.
[37] N. B. Bakanova, T. V. Atanasova, "Method for Automated Analysis of Users' Requests to Service Centre of Information Networks in OIS," Problems of Engineering Cybernetics and Robotics, 2020, doi: 10.7546/PECR.74.20.04.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 636-650 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Comparison of Machine Learning Parametric and Non-Parametric Techniques for Determining Soil Moisture: Case Study at Las Palmas Andean Basin

Carlos López-Bermeo*, Mauricio González-Palacio, Lina Sepúlveda-Cano, Rubén Montoya-Ramírez, César Hidalgo-Montoya

Universidad de Medellín, Facultad de Ingenierías, Medellín, 050026, Colombia

ARTICLE INFO

Article history:
Received: 07 October, 2020
Accepted: 09 January, 2021
Online: 05 February, 2021

Keywords:
Soil moisture
Machine Learning
Regression

ABSTRACT

Soil moisture is one of the most important variables to monitor in agriculture. Its analysis gives insights about strategies to utilize better a particular area regarding its use, i.e., pasture for cows (or similar), production forests, or even to answer what crops should be planted. The vertical structure of the soil moisture plays an important role in several physical processes such as vegetation growth, infiltration process, soil – atmosphere interactions, among others. Despite a set of tools are currently being evaluated and used to monitor soil moisture, including satellite images and in-situ sensor, several drawbacks are still persisting. In situ data is expensive for high spatial monitoring and vertical measurements and satellite data have low spatial resolution and only retrieval information of soil moisture for the top few centimeters of the soil. The present work shows an experiment design for collecting soil moisture data in a specific Andean basin with in-situ sensors in different kinds of soils as a promising tool for reproducing soil moisture profiles in areas with scarce information, employing only surface soil moisture and simple soil characteristics. Collected data is used to train machine learning supervised parametric (Multiple Linear Regression - MLR) and non-parametric models (Artificial Neural Networks - ANNs and Support Vector Regression - SVR) for soil moisture estimation in different depths. Conclusions show that parametric methods do not meet goodness of fit assumptions; so, non-parametric methods must be considered, and SVR outperforms parametric methods regarding regression accuracy allowing to reproduce the soil moisture content profiles. The proposed SVR model represents a high potential tool to replicate the soil moisture profiles using only surface information from remote sensing or in-situ data.

1. Introduction and problem statement

This paper is an extension of work originally presented in *CISTI 2020* [1]. Soil moisture is one of the most critical variables to be monitored in soils [2]. Understanding the behavior of the soil moisture makes it possible to determine the soil use [3], i.e., if it is suitable for cropping, for animal's pastures [4], and even, as a critical variable to understand if a particular terrain may be considered to real estate projects [5]. Another important application of soil moisture analysis belongs to risk management, such as landslides prevention and prediction [6]. For hydrologic applications, the soil moisture represents one of the most important variables controlling the interactions between the atmosphere and land through the evapotranspiration and evaporation processes. Additionally, the soil moisture represents a key variable for the infiltration process and direct surface runoff production in hydrological models.

In that way, different knowledge disciplines, such as hydrology [7], environmental management [8], geology [9], topography [10], among others, deal with soil moisture to deeply study its effects in plenty of different applications. Nonetheless, getting information about soil moisture is commonly a difficult task: available sources are scarce and are associated with satellite images, which causes low resolution regarding spatial distribution [11], that is, soil moisture is measured in broad areas and cannot be explicitly determined in points. During the last decades satellite images have emerged as a powerful tool to retrieve information about soil moisture in the superficial soil surface, even though data in other depth horizons is mandatory to model the behavior in a lot of applications.

*Corresponding Author: Carlos López-Bermeo, celopez@udem.edu.co

Given the importance and drawbacks mentioned above, several techniques based on Artificial Intelligence (AI) theory are currently employed to estimate and predict the soil moisture for several engineering and scientific applications. Knowledge-based systems that apply AI are considered as easy and user-friendly tools [12, 13], as they have advantages in terms of neither requiring a pre-defined conceptual relationship between the input-target parameters nor requiring expensive experimental and field measurement apparatus [14].

In this sense, several studies have been proposed to determine the best Machine Learning (ML) or (AI) model to estimate the soil moisture, considering different variables like meteorological data and satellite images, among others. There are two main ML models used: Support Vector Regression (SVR) and Artificial Neural Networks (ANN) in different configurations. It is the case of [15], three approaches are proposed to generate the SVR regression model for soil moisture estimation: in the first approach, the authors use the soil moisture and meteorological data (air temperature, relative humidity, average solar radiation, and soil temperature at five and 10cm) at time step $t-1$ and t for predict soil moisture at $t+4$ and $t+7$, where t is in days. For the second approach, they use only the meteorological data, while in the third approach, they use only soil moisture at the same time step. The results of this study show that the SVR models performed better forecasting than a simple ANN model. In the same line, [16] proposes the use of SVR for soil moisture prediction using remote sensing data coming from 10 sites in the Lower Colorado River Basin (US). The data includes backscatter and incidence angle from Tropical Rainfall Measuring Mission (TRMM), and Normalized Difference Vegetation Index (NDVI) from Advanced Very High-Resolution Radiometer (AVHRR). The authors trained an SVR with five years of data (time series) and tested on three years of data. The results show a Root Mean Square Error (RMSE) less than 2%. Additionally, results are compared with an ANN and a Multivariate Linear Regression Model (MLR), showing that the SVR has a better performance than the other models.

The latest studies show a tendency for the use of deeper ANN like shows [17] in their review. The authors classify the use of ANN in three categories according to the types of training data. For the first category, the ANN is trained with model-generated data. For the second category case, the ANN is trained with in-situ measurements, with the restriction that the spatial scale could mismatch between point-scale measurements. Finally, the third category encloses the ANN trained with global Land Surface Model (LSM) simulations for soil moisture estimation at large scales.

Other useful applications using ANN are related to the infilling missing soil moisture records such as the research presented in [18], who were using five statistical methods, and nine ANN categorized into Feedforward, Dynamic, and Radial Basis Network methods estimate missing soil moisture records. The obtained values were validated against known values for 13 soil moisture monitoring stations for three different soil layer depths in the Yanco region in southeast Australia. The results show that the nonlinear autoregressive neural network performs in similar quality than other typical methods such as the rough sets method, and monthly replacement. Other studies employing ANN for hydrologic variables are [19–22].

Although SVR and ANN are the most used techniques, several works can be found related to soil moisture prediction that use other techniques of AI and ML. Some authors include multiple techniques, like, where Classification and Regression Trees (CART), Boosted Regression Trees (BRT), Random Forest (RF), Multivariate Adaptive Regression Splines (MARS), and Flexible Discriminant Analysis (FDA), are tested, all with promising results. In this sence, [14] using several input variables such as dielectric constant, soil bulk density, clay content, and organic matter for 1155 soil samples, proposed a hybrid Adaptative Neuro-Fuzzy Inference System (ANFIS) - Grey Wolf Optimization (GWO) intelligent model for simulating soil moisture content. The results based on several statistical parameters verified the feasibility of these kinds of models improving accuracy by around 50% when compared with other similar models as ANN, SVR, and standalone ANFIS models.

Despite the increasing usage of Artificial Intelligence (AI) theory tools related to soil moisture estimation, most of the studies have the same thing in common; none of them is focused on the estimate of the soil moisture at different depths. The data estimate corresponds to a depth between 5 and 10 cm approximately or the root zone. Few studies such as [23], employed surface soil moisture (0–5 cm) values and Hydrological Soil Groups (HSGs) information to perform a Statistical Soil Moisture Profile (SSMP) model to transfer the spatial variations of soil moisture profile with the change in soil hydraulic properties. The proposed model was based on correlation techniques with preceding time steps of soil moisture fields. In another study, [24] employed multispectral imagery from Unmanned Aerial Vehicle (UAVs) and a method based on the combination of an evolutionary algorithm and artificial intelligence called genetic programming (GP), to propose a methodology to simulate soil moisture at different levels. The results were compared to ANN and SVR methods.

With this idea in mind, the present work shows a methodology to fit some ML models with in-situ sensor data, which transmit real-time information by using IoT tools and methods to overcome limitations associated with the number of sensors to determine soil moisture at different depths using surface topsoil moisture estimates. The proposed methodology represents a high potential tool for reproducing the soil moisture profiles using only surface information from satellite or in-situ data for scarce information areas. Specifically, we use the most known statistics-based parametric method: Multiple Linear Regression (MLR) and perform the corresponding Analysis of Variance (ANOVA) [25] as well as the Goodness of Fit metrics (normality, independence, homoskedasticity). We also use two widespread ML non-parametric methods: SVR and ANN [26, 27]. All the models are primarily assessed by a case study, with collected data in the Las Palmas basin, at the Andean region, in Medellín, Colombia.

The rest of this paper is organized as follows: section 2 shows a short theoretical framework; section 3 stands the methodology and experiment design; section 4 shows the models fitting; in

section 5, the analysis over results is carried out. Finally, conclusions, future work, acknowledgments, and references are presented.

2. Theoretical framework

In this section, some key concepts will be defined to understand the work better.

2.1. Soil moisture

A given portion of soil is composed of solid particles and the rest of the voids. A part of the holes is occupied by water and the rest by air. The volume occupied by water is measured using the soil moisture content, which is defined as θ as in equation (1) :

$$\theta = \frac{Volume_{water}}{Volume_{total}} \quad (1)$$

The study and knowledge of soil moisture are essential due to its influence on hydrological processes and energy flows on the earth's surface. It is also critical because of its connection with precipitation, runoff generation, nutrient transport, and groundwater [28–30]. Recent research identifies that a surplus or lack of soil moisture can favor floods or droughts occurrence [31]. Likewise, the feedback of soil moisture on evapotranspiration is essential for temperature variation and the appearance and persistence of heatwaves, as well as for precipitation generation and location [32]. Besides, the role of soil moisture in photosynthesis, ecosystem dynamics, soil respiration, and the terrestrial carbon balance is undeniable.

Therefore, soil moisture variation is critical in hydrological, ecological, and environmental studies [33] and, in particular, as a support for agriculture and biomass production [34]. Combinations of the factors mentioned above cause variations in spatial and temporal soil moisture content, which makes it a significant limiting factor for crop growth. Indeed, low crop yields should be more often related to insufficient soil moisture than insufficient rainfall [35,36], which causes that soil moisture plays a vital role as a critical resource for vegetation growth that supports agricultural production. In this sense, it is identified that soil moisture contributes in a crucial way to understanding the global climate system. For this reason, it has been highlighted by the Global Climate Observing System (GCOS) as one of the "essential climate variables". Therefore, monitoring temporal and spatial soil moisture variability is essential to estimate water availability limits and to quantify its climatic variations sensitivity and human pressures [32].

2.1.1. Soil moisture sensing technologies

Currently, several techniques are used to measure soil moisture, allowing monitor of humidity on a large scale: in-situ sensors (to take humidity measurements at the site where the sensor is located), and remote sensing from towers, aircraft, and satellites, using radiometers in the microwave region, scatterometers, synthetic aperture radars, and radar combinations-radiometers. According to [28], the products obtained with passive and active sensors have different correlation values around the world, when compared with in-situ measurements.

Although there are many possible sources of information, in the case of the tropics and specifically in the study area, there could be limitations for the use of satellite information. Some derive from the possibility or not of having reliable and spatially coherent in situ soil measurements of soil moisture, which makes it challenging to carry out a cohesive evaluation of the accuracy and information content of remote sensing products. Besides, the use of inadequate techniques for downscaling can generate alterations in data quality. Another difficulty could arise due to the impossibility of obtaining quality data with fair spatial and temporal resolution in the study area, primarily due to its geographical location, which may limit some sensors use, especially optical ones. In this sense, in the visual spectrum, one of the significant limitations is the limited surface penetration due to high cloudiness. However, visible and infrared sensors and microwave sensors are not limited by cloud cover and night conditions [35]. Observations can be made at any time of day and are not dependent on sunlight [37].

Soil moisture in-situ sensing, In Colombia and many other tropical regions, there is very little information available on soil moisture taken from in-situ monitoring stations that can be used to understand the behavior of this variable for different applications. Although it is true that in-situ measurements require a costly investment and have low coverage of the territory, they allow the analysis at a local scale (as is the case of a basin). Additionally, they are necessary to validate the data from remote sensors that generally have a low spatial resolution.

The primary function of a soil moisture sensor is to report its current state, by using an electric variable (commonly voltage), to any acquisition system, i.e., datalogger, or IoT end node. The voltage reported is scaled to get the corresponding measurement, which is commonly given in cubic meters of water over cubic meters of soil material (m^3/m^3).

The most used operation principle in soil moisture sensors is the variation of capacitance as the moisture changes. Soil acts as a dielectric material between two plates, and it changes when this variable varies. However, conductivity sensors are also applied to sense moisture. Water content allows flowing electric currents between two electrodes, and it is proportional to soil moisture.

2.2. Parametric and non-parametric methods for data fitting

To fit some models for regression of moisture data and perform the comparison, we have chosen some of the most used methods in the literature. MLR, SVR, and ANN algorithms will be reviewed in this section.

2.2.1. Multiple Linear Regression (MLR)

MLR is one of the most used parametric strategies for data fitting. The general model is shown in equation (2):

$$\hat{Y} = \sum_{i=0}^{n} [\hat{\beta}_i X_i] + \varepsilon \quad (2)$$

Where $\hat{Y}$ is the regression value of the dependent variable, $\hat{\beta}_i$ is a set of n+1 predictors (coefficients) which are determined by using least-squares optimization, X_i are n independent (predicting) variables, and ε is the regression error. If a categorical variable is in the predictors, a set of dummy variables is included, depending

on the number of levels of the factor. The goal, then, is finding the values of each $\hat{\beta}_i$ such that the global error ε is minimal.

Once the predictors are determined, an Analysis of Variance (ANOVA) must be performed [25], to determine if the means of the prediction variables are equal, that is, if all the variables are representative for the analysis.

Finally, to check if the model is statistically valid, some goodness-of-fit tests must be performed: normality, by using the Kolmogorov-Smirnoff test [38]; autocorrelation, by using Durbin-Watson test [37]; and homoskedasticity, by using Breusch Pagan test [39]. Due to the model usage is limited to the goodness of fit, many datasets do not meet this strategy.

2.2.2. Support Vector Regression (SVR) [40, 41]

The goal of this supervised learning strategy is to find one (or more) hyperplanes that separate previously tagged classes. By finding the support vectors, it is possible to predict, in the case of regression, the value of a new continuous variable based on a set of inputs. Figure 1 shows the geometric concept for two input characteristics.

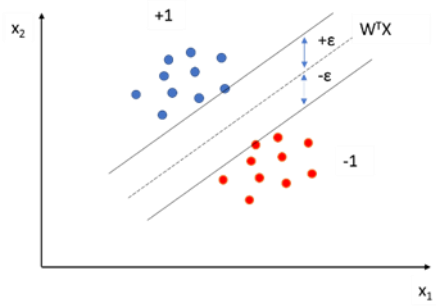


Figure 1: SVR geometric concept

The objective is to find the hyperplane that separates two classes with the maximum margin between them, therefore, finding the model is reduced to an optimization problem, where the coefficients w must be found, as shown in equation (3).

$$\begin{aligned} Min\ Z &= \frac{1}{2}\|W\|^2 \\ St: |y_i - w_i x_i| &\leq \varepsilon \end{aligned} \tag{3}$$

where y_i are the predicted variables, x_i are the independent variables (or features) and ε is the minimum tolerance between the separation hyperplanes and the features. For SVR, a slack variables set is added to equation (3), such that, for each value that falls outside ε, its deviation is recorded to make it minimal. The hyperparameter cost (C) is introduced to penalize said deviations, which is tuned through cross-validation. Thus, equation (3) is re-written, as shown in equation (4) (called the primal problem):

$$\begin{aligned} Min\ Z &= \frac{1}{2}\|W\|^2 \quad + C\sum_{i=0}^{n}|\xi_i| \\ St: |y_i - w_i x_i| &\leq \varepsilon + |\xi_i| \end{aligned} \tag{4}$$

However, in many cases, the classes are not linearly separable, for this reason, it is recommended to use kernel functions, K [41, 42], to perform a transformation that increases the number of dimensions, allowing separating such classes, as shown in equation (5):

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \tag{5}$$

where ϕ is a nonlinear mapping function and $\langle , \rangle$ is the inner product operator.

The dual problem to the primal problem is given by equation (6):

$$\begin{aligned} Min\ & \frac{1}{2}\alpha^T Q\alpha \ - e^T\alpha \\ St: \quad & y^T\alpha = 0 \\ & 0 \leq \ \alpha_i \leq C, i = 1, \dots, n \end{aligned} \tag{6}$$

where e is a vector of all ones, Q is a positive semidefinite matrix with $Q_{ij} = y_i y_j K(x_i, x_j)$ and α_i are the dual coefficients.

Using this model, the prediction function of the equation (7) is obtained:

$$f(x) = \sum_{i=1}^{n/2} y_i \alpha_i K(x_i, x) + b \tag{7}$$

where b is the bias term.

It is important to remark that the standard SVR model in equation (7) considers only real-valued functions [43]. Additionally, it is a discriminant method, i.e. produces a mapping from the data points to the class labels without computing probability distributions, allowing the method to be less computationally expensive, but more sensitive to noise than generative methods [44]. . The main advantage of SVR is that can deal with sparsity, non-linearity, and high dimensionality of the input data [45].

Regarding the SVR hyperparameter tuning (section 4.2), it is performed through cross-validation, and finally, the RMSE that exhibits the best results will be chosen. In the same way that MLR, if there are categorical variables in the set of X, *dummy variables* can be used.

2.2.3. Artificial Neural Networks (ANN). [40, 41]

This bio-inspired strategy is used in a lot of applications, where regression and classification are needed. The structure of a neural network emulates a meta-heuristic concept, based on interactions of neurons within the brain. The most known ANN is known as Multilayer Perceptron, and its architecture is shown in Figure 2. From this figure, a set of m input variables (x) in the Input layer are used to predict the values of n output variables (y) in the output layer. The prediction is achieved by choosing the correct number of neurons in different hidden layers h, which are fully connected with both previous and next layers. The effects of each connection are weighted by a set of constants W, which are determined by an optimization iterative method (commonly Gradient Descend). The main advantages of this kind of ANN (feedforward networks) are that they do not have a priori assumptions about the relationships between the independent and dependent variables [46], their

nonlinear modeling capability [47], and their minimal assumptions needed about the data [48]. In the same way that MLR, if there are categorical variables in the set of *X*, dummy variables can be used.

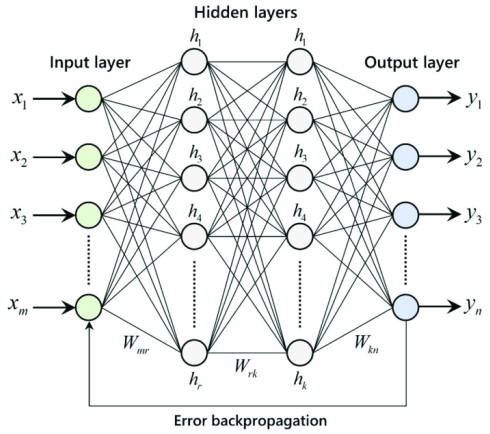


Figure 2: Multilayer Perceptron Architecture with backpropagation feedback [49]

Although using ANN can outperform other methods regarding its accuracy, its application must be extensively checked using cross-validation, since a common phenomenon known as *overfitting* makes that ANN cannot be a general model sometimes.

3. Methodology and experiment design

This section shows the steps to fit the different models. First, a measurement phase is proposed, where both instruments and locations are chosen. Then, with the set of data from in-situ sensors, some models are fitted. Finally, the performance evaluation is carried out, choosing the model which exhibits fewer regression errors.

3.1. Measurement phase

In this section, some considerations regarding the measurement are presented.

3.1.1. Location selection

Soil moisture is highly influenced by climatic characteristics of a region, specifically by rain amount and intensity, evapotranspiration [50], vegetation type, topography [51], soil properties (apparent density, porosity, organic matter content, texture, and structure), among other factors [52].

This is how changes in land use and land cover, in interaction with soil physical characteristics and climate, play a key role in soil moisture variations at different scales [53]. In the Colombian case, the soil moisture variability has been less studied and there is not enough information to understand its spatial and temporal dynamics, and much less has its contribution to water flows regulation been quantified.

This is the case in some areas of the Antioquia region, where precipitation is the only source of soil water and recharge of its surface layers. Therefore, soil moisture in deep layers cannot be replenished with contributions from rain and groundwater. For this reason, soil wet front movement study at different depths is key to understanding its variability and impact on agricultural production and ecosystems in tropical basins.

In different studies [54–56] it has been found that drought conditions affect variations in soil moisture content, while topography and vegetation type are the dominant factors that control the humidity in different soil layers. Specifically, soil water storage is mainly affected by topography at shallow depths, while in deep soil layers it is mainly controlled by vegetation type.

Considering the above, for the present study Las Palmas basin was selected, which has as characteristics of interest that it is in the tropical Andean zone of the municipality of Envigado, Antioquia, Colombia (coordinates: 6 ° 11'26.1 "N 75 ° 31'47.6" W), and which has undergone processes of change in land use and land cover, among which transition from crops and pastures to recreational and commercial uses stands out. There are also some small fragments of forest and secondary vegetation. This basin has an area of 31.31 km^2, with elevations from 2,500 to 2,600 meters above sea level and its predominant climate is cold humid to very humid and with an average annual rainfall of 2781 mm.

The basin also has different soil associations (Figure 3), which determine physical characteristics, as well as different depths in soil profiles [57]. The Tequendamita Association (TE) is found mainly in the Basin, formed by deep to moderately deep soils, well-drained, with medium textures, low to moderate fertility, mild to moderate erosion. A characteristic of these soils is that they are formed by metamorphic rocks (schists, gneisses) covered with volcanic ash. Soil samplings were made, and results were verified with previous studies such as the "General Study of Soils and Zoning of Lands of the Department of Antioquia" [58], and the" Semi-detailed study of soils in zone 13 of the municipality of Envigado for potential use purposes" [59].

Considering that the soil moisture content suitable for plant growth depends on the type of soil [60] and the physical explanation mentioned above related to the main variables affecting the soil moisture variability through the soil layers (vegetation, soil type, topography, among others), such parameters were considered to determine the sites where soil moisture monitoring stations were installed. Therefore, three types of vegetation cover representative of the basin (pastures, crops, and forest) were selected and combined with three different phases of the Tequendamita Association and one phase of the La Ceja Association (see Table 1). It is important to mention that the sensors were located having the characterization of soil profiles in the basin, which allowed determining the average depth of same, in each of the points where the monitoring stations were located. This means that an edaphological criterion was used for the location of each of the sensors. Additional criteria used to guarantee that the data collected were representative are the following: *i)* The points were located near access roads and agricultural areas located toward the higher part of the catchment where installation and security could be achieved with lower cost, *ii)* availability of connection to the telecommunications network to transmit data in real-time, *iii)* accessibility.

3.1.2. IoT end nodes

Soil moisture stations were designed and built using IoT tools and methods. Four stations are installed in the upper part of the basin, where a mobile network operator offers full 2G/3G coverage. At each point, three soil moisture sensors were put into

operation at different depths, depending on land cover (crops, pastures, and forest), and soil physical properties. The first layer, the depths per sensor, according to land cover, are shown in Table 1. Installed sensors are from the brand Meter Group, reference EC-5, measurement range from 0 up to 100%, with a resolution of 0.001 m^3/m^3 of Volumetric Water Content (VWC) and accuracy of ±0.02 m^3/m^3. Sensor output is voltage, from 0 up to 2 Vdc, and is collected by a microcontroller unit ESP32, with an Analog-to-Digital Converter of 12-bit resolution. To ensure that the measurements are stable to be transmitted, the next conditioning protocol is used: *i)* take *n* measures from the sensor, *ii)* arrange samples from the lowest to highest, *ii)* extract the samples in the positions *n/2* and *1+n/2,* and *iii)* take the mean. Engineering units are scaled from ADC units to VWC by using a two-point scheme, by comparing readings of moisture in both water and air and using a Meter Group EM-50 as a calibration pattern. In-situ end-nodes are depicted in Figure 4.

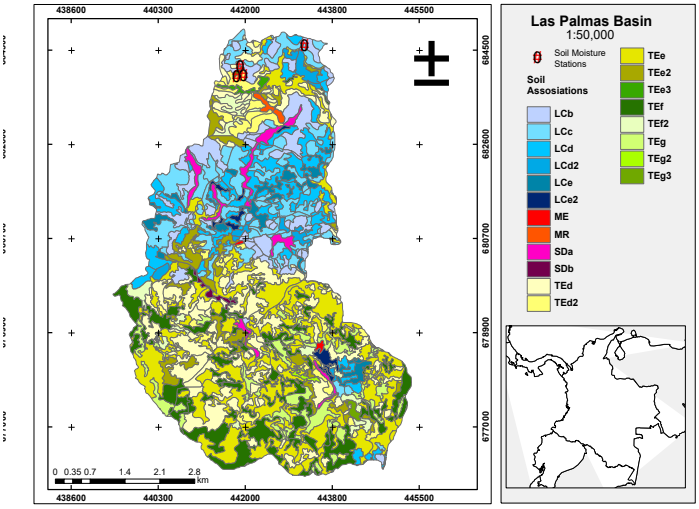
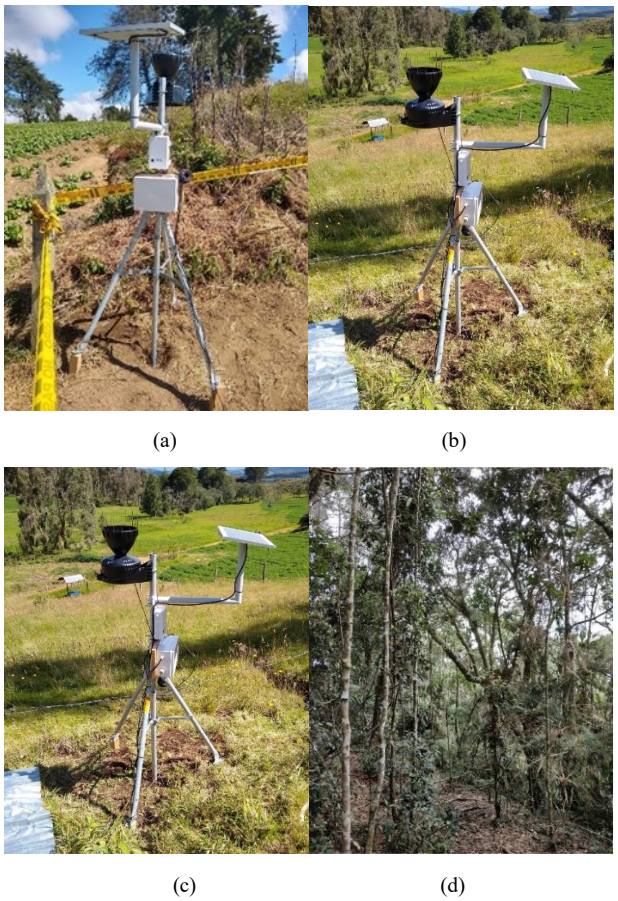


Figure 3: Study area location in Antioquia and Colombia and soil associations present in the Las Palmas basin

After collecting measurements, data is sent to an IoT backend by using the protocol Message Queue Telemetry Transport (MQTT). The chosen IoT platform is Ubidots [61], and it also helps to store information and to provide a frontend user interface to visualize real-time indicators, as shown in Figure 5.

Table 1: Soil moisture sensor depths according to soil use

Land Cover	Soil association	Sensor 1 depth range (cm)	Sensor 2 depth range (cm)	Sensor 3 depth range (cm)
Crop	La Ceja Consociation (LCb)	0-23	23-37	>37
Pasture 1	Tequendamita Association (Ted2)	0-22	22-33	>33
Pasture 2	Tequendamita Association (Ted3)	0-33	33-60	>60
Forest	Tequendamita Association (Ted)	0-16	16-35	>35



(a) (b)

(c) (d)

Figure 4: IoT soil moisture end nodes in (a) crops, (b) pasture type 1, (c) pasture type 2 and (d) forest

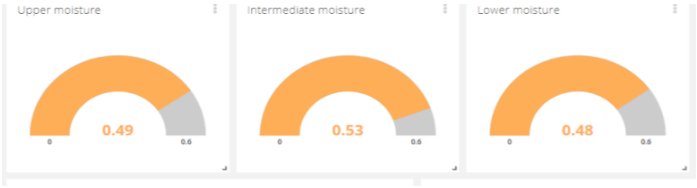


Figure 5: IoT backend / frontend platform [61]

3.2. *Model fitting phase*

First, all data is extracted from the Ubidots platform, using a temporal window from December of 2019 to June 2020, with an equi-temporal sampling rate of 15 minutes. All data is put together into a unique database table, including timestamp, moisture values per sensor, depths per sensor, and soil type. The models are fitted to output a moisture value at a particular depth, regarding the readings of the other two sensors, their corresponding depths, and the soil use.

To fit all the models, R Studio V 1.2.5019 was used. All the basic statistics measurements were inspected: quartiles, means, medians, and standard deviations (for continuous data), and counters (for categorical data). Regression expressions were adapted with *dummy variables* to allow that the categorical variable *soil type* could involve all the factors with different coefficients. Then, a preprocessing phase for continuous variables was carried out with two different strategies, according to the equations (8) and (9):

$$x_i' = \frac{x_i - \bar{x}}{s^2} \quad (8)$$

$$x_i' = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (9)$$

where x_i' is the i^{th} transformed sample, x_i is the i^{th} sample of the continuous variable x from the dataset, $\bar{x}$ is the mean of x, s^2 is the standard deviation of x, $min(x)$ is the minimum value of x, and $max(x)$ is the maximum value of x.

Finally, the model fitting for MLR, SVN, and ANN was performed according to the corresponding method. The database has 77848 observations and is split into two sets, with a pseudo-random strategy: training (70%, that is, 54493 examples) and test (30%, that is, 23355 examples). Moreover, some additional tests were run for MLR, to ensure goodness of fit: Kolmogorov-Smirnov [38], Durbin-Watson [37], and Breusch Pagan [39].

3.3. Results analysis phase

After the models were fitted, the regression was carried out with each model, by using the test set. An initial graphical inspection was performed, and three performance metrics are used: RMSE (Root Mean Square Error), Index of Agreement (Willmott), and R^2. These metrics and the graphical approximation served to determine the accuracy of the models regarding field measurements. The pertinent analysis was developed.

4. Model fitting

For the sake of clearness, we will adopt the following notation: M_i is the *ith* sensor, D_i is the depth of *ith* sensor, i is an integer index to identify the sensor position ($i = 1$ is the superficial sensor, $i = 2$ is the middle sensor, and $i = 3$ is the deepest sensor), and P_1, P_2, and F are dichotomic *dummy variables* for representing the factors *pasture type 1*, *pasture type 2* and *forest*, in the categorical variable *soil type*. All the processes are executed in a laptop Lenovo G40, with a processor Intel Core i7 4500 @ 1.8 GHz and RAM = 12 GB.

A previous graphical inspection is performed to qualitatively behold if there are relationships among different moistures, as shown in Figure 6. Some statistical metrics are presented in Table 2.

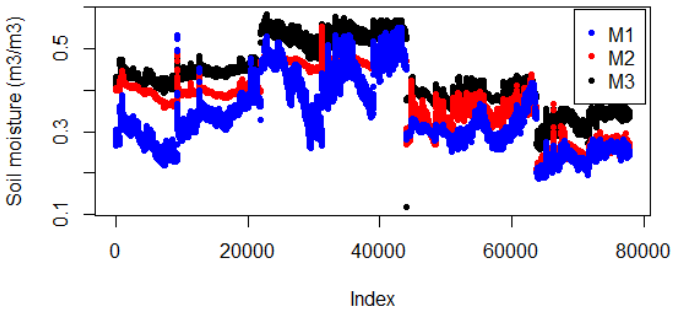


Figure 6. Trends of moisture data. M_1 is the superficial sensor, M_2 is the middle sensor, and M_3 is the deepest sensor. Index interval from 0 up to 21.000 corresponds to crops, from 21.000 to 44.000 corresponds to pasture 1, from 44.000 up to 64.000 corresponds up to pasture 2, and from 64.000 up to 79.000 corresponds to the forest.

From Figure 6, it can be noticed that effectively there is a relationship between the moistures at different depths since when an external variable, i.e., rainfall, causes a perturbation, each moisture varies. Moreover, according to Table 2, medians and means of moistures are arranged from M_1 to M_3 from the least to the greater, indicating that the deep the soil is, the high VWC is. It also can be noticed that the standard deviation of M_3 is the least, supporting that more superficial moisture changes in a wider range than the others, so external variables (rainfall or solar radiation) affect the most to the moistures closer to the soil surface. To validate the hypothesis of correlation among moistures, by using the Pearson correlation coefficient, the next results are obtained: $cor(M_1,M_2) = 0.85$; $cor(M_1,M_3) = 0.91$; $cor(M_2,M_3) = 0.97$.

Table 2: Statistical metrics for soil moistures

	M_1	D_1	M_2	D_2	M_3	D_3
Min	0.047	0.08	0.1918	0.255	0.117	0.33
Q1	0.27	0.11	0.332	0.275	0.382	0.33
Median	0.312	0.115	0.388	0.3	0.425	0.37
Mean	0.3308	0.1215	0.3772	0.3321	0.4322	0.4212
Q3	0.383	0.165	0.457	0.465	0.508	0.6
Max	0.552	0.165	0.553	0.465	0.582	0.6
SD	0.0786	0.029	0.0786	0.085	0.0735	0.113

4.1. MLR

Taking into account that there is a direct correlation among soil moistures, MLR can be fitted according to the process in section 2.2.1. So, the general model in equation (2) is re-written to the particular dataset, as shown in equation (10):

$$\begin{aligned} \widehat{M}_2 &= \hat{\beta}_{0,2} + \hat{\beta}_{1,2}M_1 + \hat{\beta}_{2,2}D_1 + \hat{\beta}_{3,2}D_2 + \hat{\beta}_{4,2}P_1 \\ &\quad + \hat{\beta}_{5,2}P_2 + \hat{\beta}_{6,2}F + \varepsilon_2 \\ \widehat{M}_3 &= \hat{\beta}_{0,3} + \hat{\beta}_{1,3}M_1 + \hat{\beta}_{2,3}D_1 + \hat{\beta}_{3,3}D_3 + \hat{\beta}_{4,3}P_1 \\ &\quad + \hat{\beta}_{5,3}P_2 + \hat{\beta}_{6,3}F + \varepsilon_3 \end{aligned} \quad (10)$$

where all the $\hat{\beta}$ are the model estimators, and ε is the error model. Note that, in the first instance, it is needed to fit two models for M_2 and M_3, depending on the variables associated with the superficial moisture and the *soil type*. However, due to the high correlation between M_2 and M_3, it is essential to find the optimal $\hat{\beta}$ values, the strategy of minimization of *least squares* is applied [62], searching for a hyperplane that minimizes the cumulative errors from measurements and the predicted variable. The R function *lm* is used, getting the results in Table 3, for training data without preprocessing, and with the preprocessing strategies in the equations (8) and (9). From this table, the dichotomic coefficients $\hat{\beta}_{4,2}, \hat{\beta}_{5,2}, \hat{\beta}_{4,3}$, and $\hat{\beta}_{5,3}$ are not determined, due to there are a multicollinearity effect with the other variables; that is, such variables explain the effects thoroughly without needing the inclusion of the variable *soil type* if it is a crop, pasture 1, or pasture 2. The multicollinearity effect is tested by using the *alias* method proposed in [63]. This method shows how a collinear variable is explained by the others, as shown in Table 4.

According to Table 4, for example, the coefficient $\hat{\beta}_{5,2}$ (which corresponds to the binary variable P_2, pasture type 2) is explained by the others, being D_1 (superficial moisture sensor depth), the variable that explains the most the behavior of P_2. A similar analysis can be performed for other variables.

Table 3: Regressor values for the model in equation (8). P-values less than 0.05 indicate that the regressor is significative

Raw data			
Coefficient	Estimate	Std. Error	p-value
$\hat{\beta}_{0,2}$	0.6537	0.0019	<2 x 10^{-16}
$\hat{\beta}_{1,2}$	0.2615	0.1395	<2 x 10^{-16}
$\hat{\beta}_{2,2}$	-17.6344	0.0691	<2 x 10^{-16}
$\hat{\beta}_{3,2}$	5.6271	0.0203	<2 x 10^{-16}
$\hat{\beta}_{4,2}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{5,2}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{6,2}$	-0.4906	0.0015	<2 x 10^{-16}
$\hat{\beta}_{0,3}$	0.5464	0.0014	<2 x 10^{-16}
$\hat{\beta}_{1,3}$	0.3519	0.0010	<2 x 10^{-16}
$\hat{\beta}_{2,3}$	-8.8458	0.0356	<2 x 10^{-16}
$\hat{\beta}_{3,3}$	2.1546	0.0074	<2 x 10^{-16}
$\hat{\beta}_{4,3}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{5,3}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{6,3}$	-0.3496	0.0011	<2 x 10^{-16}
Max/Min preprocessing			
Coefficient	Estimate	Std. Error	p-value
$\hat{\beta}_{0,2}$	0.2479	0.0008	<2 x 10^{-16}
$\hat{\beta}_{1,2}$	0.3660	0.0019	<2 x 10^{-16}
$\hat{\beta}_{2,2}$	-4.1683	0.0162	<2 x 10^{-16}
$\hat{\beta}_{3,2}$	3.2853	0.0118	<2 x 10^{-16}
$\hat{\beta}_{4,2}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{5,2}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{6,2}$	-13634	0.0044	<2 x 10^{-16}
$\hat{\beta}_{0,3}$	0.1363	4.6 x 10^{-4}	<2 x 10^{-16}
$\hat{\beta}_{1,3}$	0.3834	0.0011	<2 x 10^{-16}
$\hat{\beta}_{2,3}$	-1.6136	0.0064	<2 x 10^{-16}
$\hat{\beta}_{3,3}$	1.2485	0.0042	<2 x 10^{-16}
$\hat{\beta}_{4,3}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{5,3}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{6,3}$	-0.7496	0.0024	<2 x 10^{-16}
SD/Mean preprocessing			
Coefficient	Estimate	Std. Error	p-value
$\hat{\beta}_{0,2}$	1.1298	0.0037	<2 x 10^{-16}
$\hat{\beta}_{1,2}$	0.2608	0.0013	<2 x 10^{-16}
$\hat{\beta}_{2,2}$	-6.6737	0.0261	<2 x 10^{-16}
$\hat{\beta}_{3,2}$	6.0411	0.0217	<2 x 10^{-16}
$\hat{\beta}_{4,2}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{5,2}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{6,2}$	-6.2131	0.0201	<2 x 10^{-16}
$\hat{\beta}_{0,3}$	0.8646	0.0029	<2 x 10^{-16}
$\hat{\beta}_{1,3}$	0.3761	0.0010	<2 x 10^{-16}
$\hat{\beta}_{2,3}$	-3.6	0.0144	<2 x 10^{-16}
$\hat{\beta}_{3,3}$	3.3201	0.0113	<2 x 10^{-16}
$\hat{\beta}_{4,3}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{5,3}$	*N/A*	*N/A*	*N/A*
$\hat{\beta}_{6,3}$	-4.7565	0.0157	<2 x 10^{-16}

Table 4: Alias test for checking multicollinearity of non-determined coefficients in MLR

	$\hat{\beta}_{0,2}$	$\hat{\beta}_{1,2}$	$\hat{\beta}_{2,2}$	$\hat{\beta}_{3,2}$	$\hat{\beta}_{6,2}$
$\hat{\beta}_{4,2}$	-3.2352	0	58.8235	-11.7647	1.5294
$\hat{\beta}_{5,2}$	-9.3529	0	388.2353	-117.6471	8.2941
	$\hat{\beta}_{0,3}$	$\hat{\beta}_{1,3}$	$\hat{\beta}_{2,3}$	$\hat{\beta}_{3,3}$	$\hat{\beta}_{6,3}$
$\hat{\beta}_{4,3}$	-3.2352	0	47.0588	-5.8823	1.5294
$\hat{\beta}_{5,3}$	-9.3529	0	270.5882	-58.8235	8.2941

Once MLR models are fitted, an Analysis of Variance must be performed for the raw data and the two preprocessing strategies in the equations (8) and (9), to check if all the variables are significant. This analysis is carried out by using the R function *ANOVA,* obtaining the results in Table 5, showing that hypothesis tests performed to validate that all the model variables are significant, excepting the *soil type* for crops, pasture type 1, and pasture type 2.

Table 5: ANOVA for MLR in (a) estimation for M_2 and (b) estimation for M_3

(a)

Raw data					
Variable	Df	Sum squares	Mean square	F value	p-value
M_1	1	244.84	244.8495	125163	<2 x10^{-16}
D_1	1	54.990	54.9909	281105	<2 x10^{-16}
D_2	1	9.8928	9.8928	50570	<2 x10^{-16}
soil type	1	18.6491	18.6491	95331	<2 x10^{-16}
Residuals	54488	10.6591	0.000195		
Max-Min preprocessing					
Variable	Df	Sum squares	Mean square	F value	p-value
M_1	1	1881.34	1881.34	125387	<2 x10^{-16}
D_1	1	422.75	422.75	281755	<2 x10^{-16}
D_2	1	76.29	76.29	50829	<2 x10^{-16}
soil type	1	144.07	144.07	96022	<2 x10^{-16}
Residuals	54488	81.75	0		
SD / Mean preprocessing					
Variable	Df	Sum squares	Mean square	F value	p-value
M_1	1	39213	39213	124894	<2 x10^{-16}
D_1	1	8779	8779	279632	<2 x10^{-16}
D_2	1	1594	1594	50777	<2 x10^{-16}
soil type	1	3005	3005	95714	<2 x10^{-16}
Residuals	54488	1711	0		

(b)

Raw data					
Variable	Df	Sum squares	Mean square	F value	p-value

M_1	1	240.583	240.5830	230668	<2 x10^{-16}
D_1	1	37.9458	37.9458	363820	<2 x10^{-16}
D_2	1	0.3932	0.3932	70.236	<2 x10^{-16}
soil type	1	9.4532	9.4532	90636.	<2 x10^{-16}
Residuals	54488	5.6830	0.0001043		

Max Min preprocessing

Variable	Df	Sum squares	Mean square	F value	p-value
M_1	1	1115.72	1115.72	234484	<2 x10^{-16}
D_1	1	174.04	174.04	365759	<2 x10^{-16}
D_2	1	1.78	1.78	3750	<2 x10^{-16}
soil type	1	43.53	43.53	91487	<2 x10^{-16}
Residuals	54488	25.93	0		

SD / Mean preprocessing

Variable	Df	Sum squares	Mean square	F value	p-value
M_1	1	39213	39213	124894	<2 x10^{-16}
D_1	1	8779	8779	239632	<2 x10^{-16}
D_2	1	1594	1594	50777	<2 x10^{-16}
soil type	1	3005	3005	95714	<2 x10^{-16}
Residuals	54488	1711	0		

To assess the best preprocessing strategy, we calculate the R^2, the RMSE and the Index of Agreement for each model with the training data, for both M_2 and M_3 (Table 6). Because the preprocessing strategies change the scale of the measurements, it is evident that the R^2 and the Index of Agreement are similar for all the subsets, however, the RMSE is less for the raw data, since all the moistures are in the range 0.2 up to 0.6 m^3/m^3, and because the standard deviations are from 0.029 to 0.113 (Table 2), the span of the transformed sets is higher than the raw data, according to equations (8) and (9).

Table 6: Performance metrics for the sets of raw and preprocessed data for (a) M_2 and (b) M_3

(a)

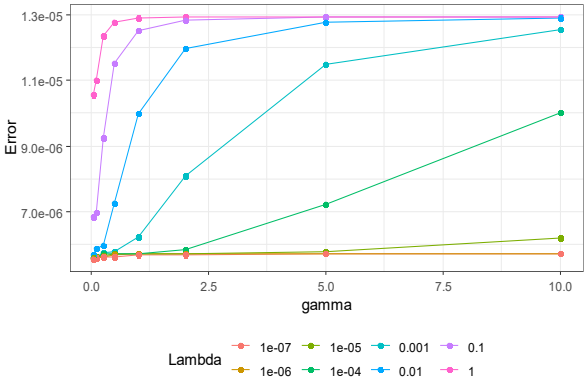
	R^2	RMSE	Index of Agreement
Raw data	0.9685	0.14	0.9217
Max / Min	0.9685	0.0387	0.9217
SD / Mean	0.9683	0.177	0.9217

(b)

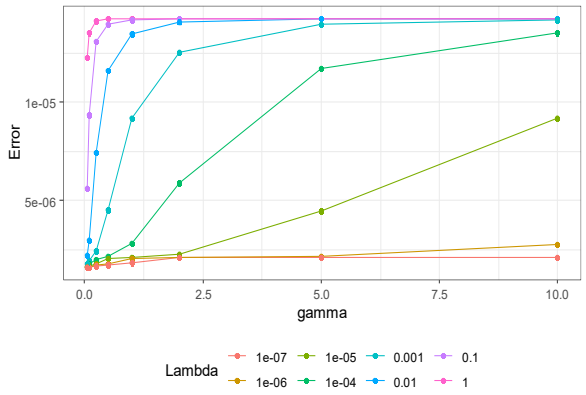
	R^2	RMSE	Index of Agreement
Raw data	0.9808	0.0101	0.9401
Max / Min	0.9808	0.0218	0.9402
SD / Mean	0.9807	0.1387	0.94

4.2. SVR

An SVR is trained to perform the regression of M_2 and M_3, based on M_1, D_1, D_2 or D_3, F, P_1, and P_2. Three types of kernels are tested: linear, polynomial, and Radial Basis Function (RBF) [64]. A cross-validation process is carried out to determine which of the kernels exhibits better performance, by using the function *tune* in R Studio, considering the inverse of the regularization parameter cost, known as lambda (λ) and kernel width (γ) as hyperparameters. Since predictions with linear and polynomial kernels did not exhibit acceptable errors, we selected the RBF kernel to show the regressor tuning process. The steps developed to tune the SVR are as follows (for each preprocessing scheme): *i)* split raw data and preprocessed data into training and test sets, *ii)* configure a 2D grid with the values of the hyperparameters λ = *(0.0001, 0.001, 0.01, 0.1, 1, 10, 100)* and γ = *(0.001, 0.01, 0.1, 1, 10, 100), iii)* determine an SVR model for each subset of hyperparameters, *iv)* evaluate the regression errors of each model, and *v)* choose the model with the least regression error. After tuning all the SVRs, we found the best hyperparameters configuration. Regarding the training sets, Figure 7(a) shows the regression errors for the raw data of M_2 (which exhibited the least regression error = 4.36 x 10^{-6}), considering the hyperparameters λ =1 x 10^{-7} and γ = 0.001, and Figure 7(b) shows the regression errors for the raw data of M_3 (which exhibited the least regression error = 1.59 x 10^{-6}), considering the hyperparameters λ =1 x 10^{-5} and γ = 0.1. Besides, the regression errors for each preprocessing configuration (for the training sets) are shown in Table 7. Finally, for M_2, we obtained R^2= 0.985, RMSE = 0.0115, and Index of Agreement = 0.9422. Similarly, for M_3, we obtained R^2= 0.9885, RMSE = 0.0078, and Index of Agreement = 0.9542.



(a) Raw data for M_2



(b) Raw data for M_3

Figure 7: Hyperparameters selection for M_2 and M_3 for the training subset

Table 7: Regressor errors for each preprocessing scheme for the training subset

M_2	
Preprocessing	Regressor error
Raw	1.43 x 10^{-5}
Max/min	4.36 x 10^{-6}
SD/Mean	5.0387 x 10^{-6}
M_3	
Preprocessing	Regressor error
Raw	1.59 x 10^{-6}
Max/min	7.47 x 10^{-6}
SD/Mean	0.00032046

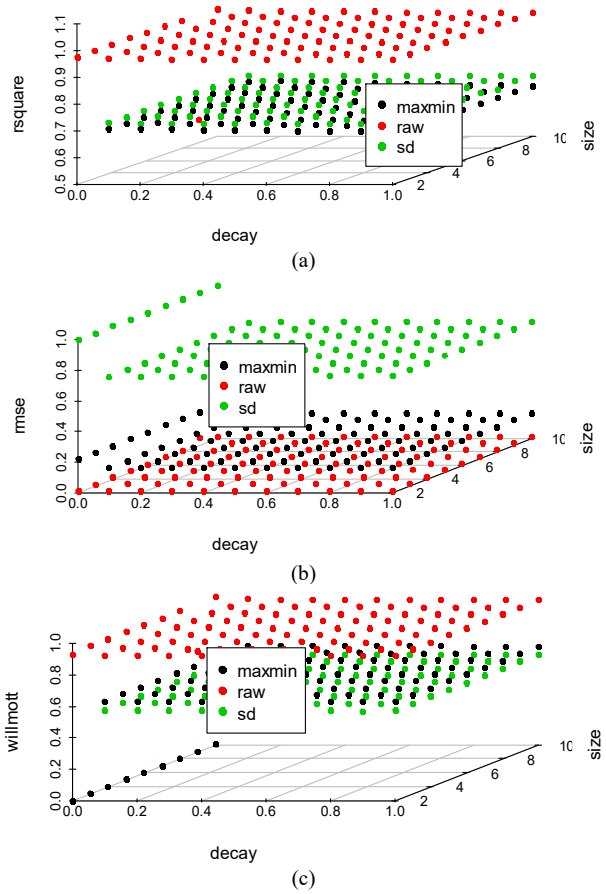


Figure 8: ANN regression performance for each preprocessing scheme for (a) R^2, (b) RMSE, and (c) Index of Agreement (Willmott) for M_2

4.3. ANN

An ANN is proposed to predict soil moisture. Two multilayer perceptrons are trained to perform the regressions of M_2 and M_3, based on M_1, D_1, D_2 or D_3, F, P_1, and P_2. A cross-validation process was carried out aiming to tune the hyper parameters of the ANN, i.e., *i)* best number of hidden layers, taking into account a tradeoff between low estimation error and simplicity, *ii)* weights values for each connection and *iii)* activation function the neurons (binary step, linear, sigmoid, tanh and Rectified Linear Unit - RELU- were considered). We have followed the next procedure to choose the best configuration: *i)* take raw data, or preprocessed data with the equations (8) or (9) and divide them into training (70%) and test (30%) sets, *ii)* train a perceptron with a hidden layer and a number of neurons from 2 up to 10, and a decay from 0 up to 1 with steps of 0.1, *iii)* for each pair of hyperparameters, evaluate R^2, RMSE and the Index of Agreement (Willmott) for the training set, *iv)* inspect and select the best preprocessing scheme, *v)* plot errors for the selected preprocessing scheme according to the set of hyperparameters chosen, and *vi)* select the configuration of the neural network. Figure 8 and Figure 9 show different regression errors for the different preprocessing schemes for M_2 and M_3.

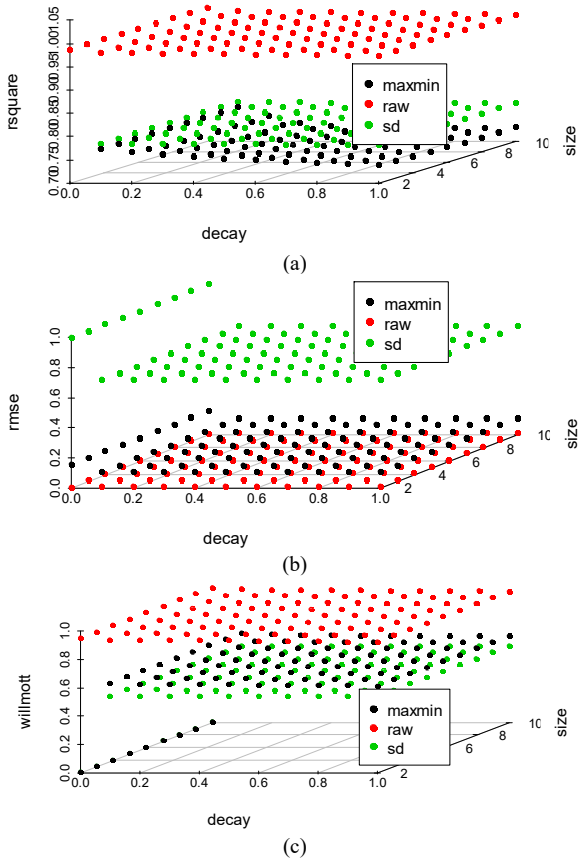
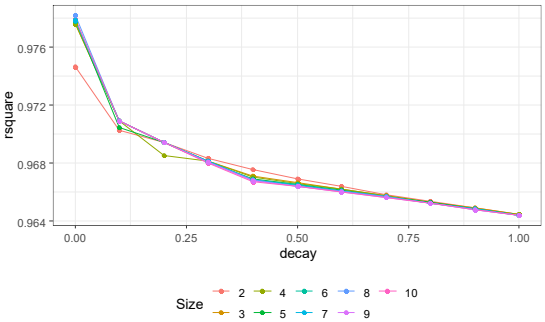
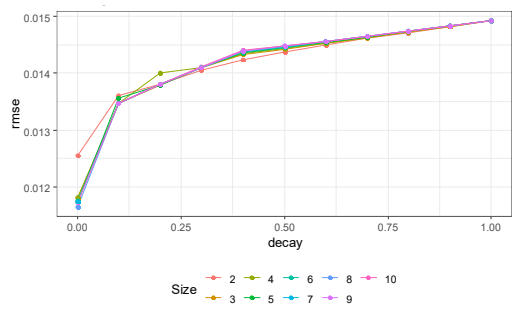


Figure 9: ANN regression performance for each preprocessing scheme for (a) R2, (b) RMSE, and (c) Index of Agreement (Willmott) for M_3

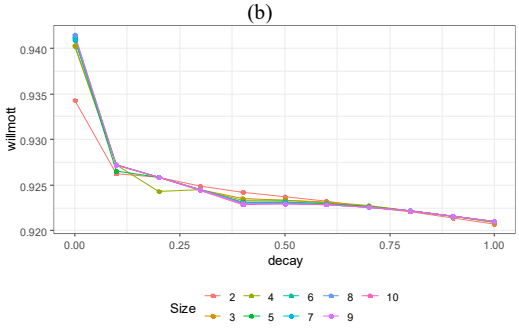
From Figure 8 and Figure 9, it can be seen that all the configurations of ANNs have better performance with raw data, since R^2 and Index of Agreement are close to 1, and RMSE is close to 0. Thus, we selected the raw data to get the best ANN configuration, as shown in Figure 10 and Figure 11.



(a)

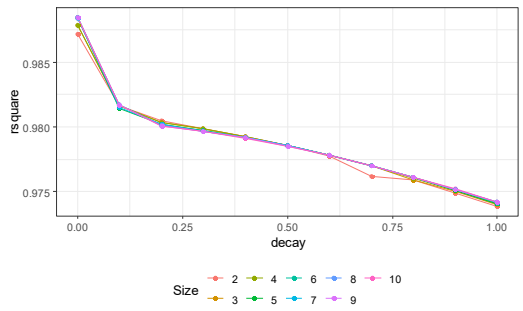


(b)

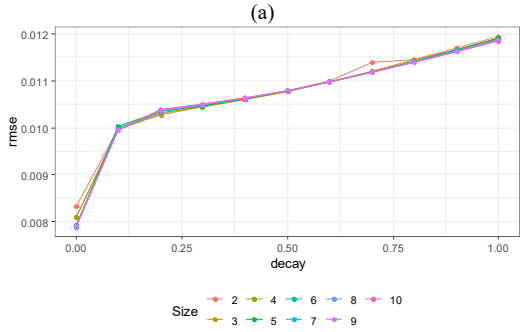


(c)

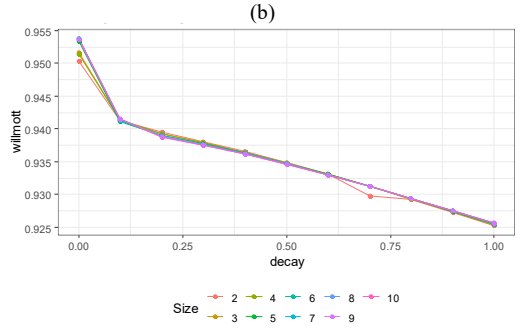
Figure 10: ANN regression performance for raw data for (a) R^2, (b) RMSE, and (c) Index of Agreement (Willmott) for M_2



(a)



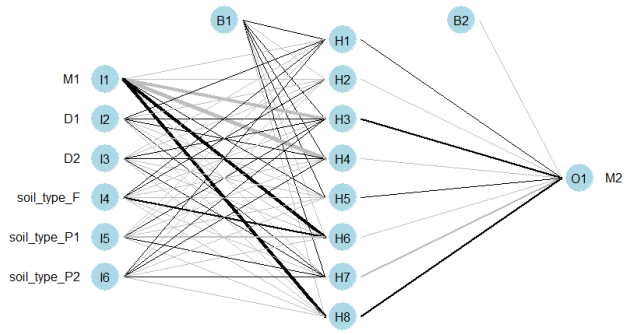
(b)



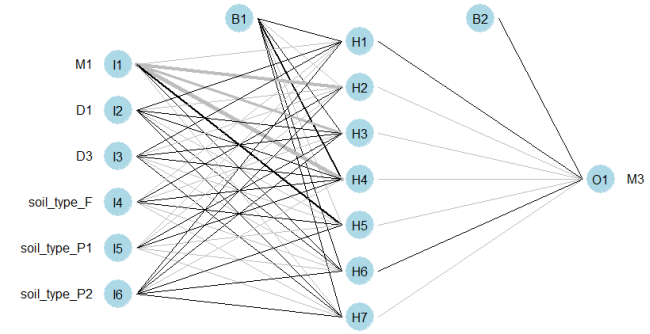
(c)

Figure 11: ANN regression performance for raw data for (a) R^2, (b) RMSE, and (c) Index of Agreement (Willmott) for M_3

From Figure 10, we obtained the best ANN configuration for M_2 as follows: an input layer with six neurons, a hidden layer with eight neurons, an output layer with one layer, and a decay = 0 (R^2=0.9781, RMSE=0.0116, and Index of Agreement=0.9414). Similarly, for M_3, we found the best configuration with a hidden layer of 7 neurons and a decay = 0 (R^2=0.9884, RMSE=0.0079, and Index of Agreement=0.9534). The ANNs' configurations are depicted in Figure 12.



(a)



(b)

Figure 12: Multilayer perceptron ANN for (a)M_2 estimation and (b) M_3 estimation

In Figure 12, the darker and broader the line between two neurons is, the greater the weight is, which represents the relative importance (positive) of such connection. On the other hand, the grayer and broader the line between two neurons is, the lesser the weight is, which represents the relative importance (negative) of such connection. The tuned weights are shown in Table 8.

Table 8: Tuned weights for neurons connections of the ANN. Table (a) shows the weights from the input layer to the hidden layer for M_2 estimation, table (b) shows the weights from the hidden layer to the output for M_2 estimation, table (c) shows the weights from the input layer to the hidden layer for M_3 estimation, table (d) shows the weights from the hidden layer to the output for M_3 estimation.

(a)

	B1	I1	I2	I3	I4	I5	I6
H1	0.54	-2.61	0.15	-0.90	4.58	-8.25	1.71
H2	-5.51	-5.17	-0.63	-2.37	4.15	-6.48	-0.29
H3	11.76	-49.2	2.19	5.04	-9.21	8.75	0.46
H4	13.42	-54.7	1.68	2.96	-7.80	-11.8	0.25
H5	2.61	10.10	-0.60	-0.37	-2.94	-9.21	-5.50
H6	-13.4	52.54	-2.25	-6.89	14.50	-16.1	-0.90
H7	1.84	-2.64	0.86	1.68	-5.57	5.91	3.73
H8	-12.6	49.26	-2.26	-4.89	-5.09	-8.00	-0.66

(b)

	B2	H1	H2	H3	H4	H5	H6	H7	H8
O1	-2.9	7.86	-2.07	17.10	-7.78	2.27	-7.57	-14.44	16.95

(c)

	B1	I1	I2	I3	I4	I5	I6
H1	4.40	-6.37	0.81	1.36	-2.21	-0.84	0.66
H2	-1.44	-36.24	-0.70	-0.80	5.94	-0.48	2.11
H3	5.53	-24.60	0.78	4.76	-1.36	9.97	0.05
H4	15.28	-45.70	1.70	6.19	4.33	3.73	-4.13
H5	-4.00	17.68	-0.98	-2.84	2.12	-5.66	0.38
H6	0.14	11.19	0.51	-2.04	-8.67	-7.35	4.88
H7	2.07	-1.97	0.64	0.63	-9.89	-3.19	0.51

(d)

	B2	H1	H2	H3	H4	H5	H6	H7
O1	1.47	2.42	-11.36	-1.99	-0.28	-3.38	5.10	-6.77

Table 9: Performance metrics for regression models (test set) ANN, SVR and MLR

	R^2	RMSE	Index of Agreement
Test set ANN M_2	0.978	0.0116	0.9411
Test set MLR M_2	0.9686	0.014	0.9217
Test set SVR M_2	0.9786	0.0115	0.9423
Test set ANN M_3	0.988	0.008	0.9551
Test set MLR M_3	0.9885	0.01	0.9411
Test set SVR M_3	0.9808	0.007	0.9542

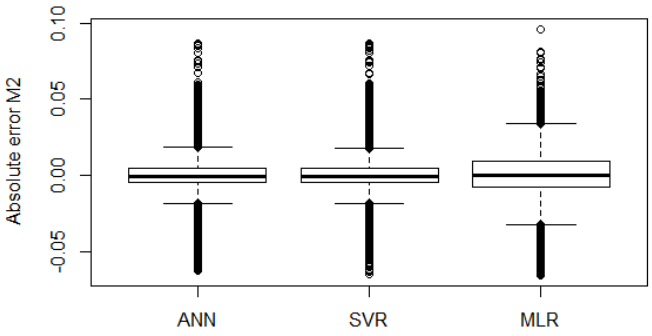


Figure 13: Absolute errors for the test set of M_2

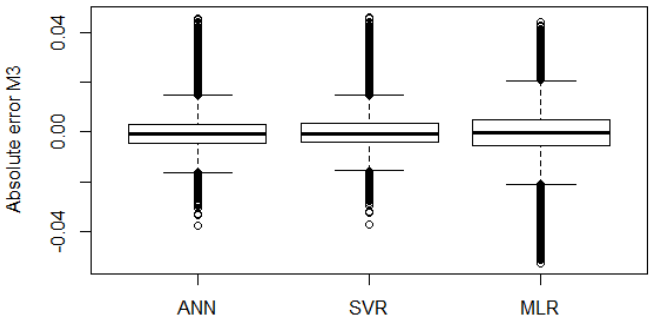


Figure 14: Absolute errors for the test set of M_3

5. Results

After fitting the parametric and non-parametric methods, some performance metrics will be applied to test what method could be used with better results to represent the soil moisture variability toward the soil profile for different type of soil and land cover configurations. It will indicate what model could be applied to extrapolate the soil moisture profiles toward areas with scarce information. To evaluate the models' performances, R^2 , RMSE, and the Index of Agreement are proposed, as shown in Table 9. It can be noticed that, at first glance, the general performance of the methods is almost the same with slightly worse results for the MLR method for the sites presented. A deeper analysis can be carried out from the error dispersions of each method, as depicted in Figure 13 and Figure 14, where it can be noticed that error dispersion is higher for MLR, and the absolute error performances of ANNs and SVR are similar. Furthermore, when plotting the time-series (Figure 15 and Figure 16), it can be noticed that a deeper analysis should be carried out regarding the soil types.

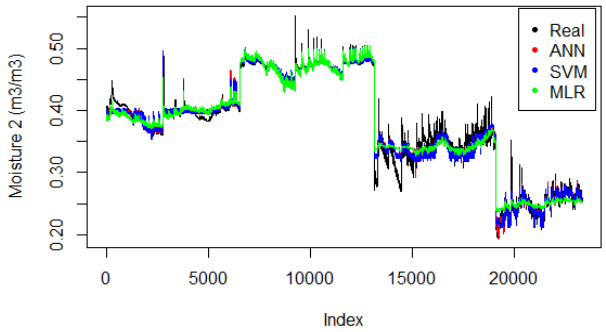


Figure 15: Regression results for the test set of M_2. Index interval from 0 up to 6.500 corresponds to crops, from 6.501 to 13.100 corresponds to pasture 1, from 13.101 up to 19.100 corresponds up to pasture 2, and from 19.101 up to 23.355 corresponds to the forest

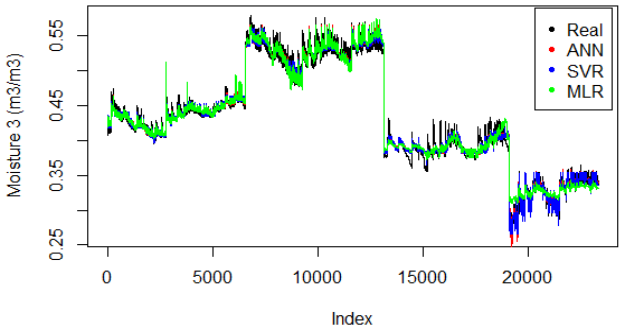


Figure 16: Regression results for the test set of M_3 . Index interval from 0 up to 6.500 corresponds to crops, from 6.501 to 13.100 corresponds to pasture 1, from 13.101 up to 19.100 corresponds up to pasture 2, and from 19.101 up to 23.355 corresponds to the forest.

It can be noticed that the methods (ANN, MLR, and SVR) exhibit similar results for the crop, pasture 1, and pasture 2 in terms of average variability of the soil moisture (Figure 15 and Figure 16). Nonetheless, SVR outperforms the other two methods in terms of high variability when comparing versus the measured data. In the case of the forest, it is more evident that SVR can better predict the real values of both M_2 and M_3. On the other hand, none of the methods are accurate enough to predict behaviours in pasture 2. Due to this visual inspection, RMSE, R^2, and Index of Agreement are recomputed but segmented by each *soil type,* as shown in Table 10.

Table 10: Performance metrics for regression models ANN, MLR, and SVR differenced by *soil type*

Soil type – Moisture	Performance metric	R^2	RMSE	Index of Agreement
Crop M_2	Test ANN	0.8095	0.0425	0.7922
	Test MLR	0.8129	0.0424	0.8027
	Test SVR	0.6542	0.1168	0.6031

Crop M_3	Test ANN	0.9307	0.0044	0.8725
	Test MLR	0.9325	0.0043	0.8782
	Test SVR	0.9344	0.0044	0.8993
Pasture 1 M_2	Test ANN	0.8176	0.0583	0.7722
	Test MLR	0.8174	0.0588	0.7695
	Test SVR	0.8246	0.0049	0.7715
Pasture 1 M_3	Test ANN	0.8439	0.0077	0.7824
	Test MLR	0.8444	0.0081	0.7902
	Test SVR	0.8585	0.0074	0.8136
Pasture 2 M_2	Test ANN	0.2836	0.0551	0.2526
	Test MLR	0.2770	0.0551	0.2468
	Test SVR	0.3386	0.0196	0.3121
Pasture 2 M_3	Test ANN	0.3749	0.0125	0.3185
	Test MLR	0.3560	0.0128	0.3054
	Test SVR	0.4419	0.0119	0.4112
Forest M_2	Test ANN	0.8247	0.0785	0.8063
	Test MLR	0.8245	0.0795	0.7985
	Test SVR	0.8264	0.0079	0.8111
Forest M_3	Test ANN	0.8247	0.0185	0.7954
	Test MLR	0.3245	0.0796	0.2212
	Test SVR	0.9207	0.0059	0.8924

It can be noted, according to Table 10, that SVR outputs better results in most of the cases, excepting for M_2 in the crop, where MLR fits better data with the test set. In the rest of the cases, ANN outperforms MLR. Since the R^2 value depends on the individual standard deviations, it is possible that the segmentations conducted for analyzing the individual performance regarding the soil type present different fitting values than the case of the whole dataset.

For the sake of statistical validity, some goodness-of-fit tests must be carried out for the parametric method (MLR): normality, autocorrelation, and homoskedasticity. All the tests are calculated using residual errors. For normality, the Kolmogorov Smirnov test [38] is applied, obtaining a p-value = 2×10^{-16}, so residuals are not normal. For autocorrelation, the Durbin test is applied, obtaining a DW = 0.02. According to [37], if the value is too far from 2, there is a autocorrelation. Finally, for homoskedasticity, the Breusch Pagan test [39] is carried out, obtaining a p-value=2×10^{-16}, which means that residuals are heteroskedastic. Because there is no normality, correlation, nor homoskedasticity, the linear regression model could not be applied.

6. Conclusions

In this paper, a comparison between parametric and non-parametric methods for fitting soil moistures in Las Palmas Andean Basin. This research was motivated due to scarce information from satellite images and in-situ measurements toward the tropical areas where the soil moisture plays an important role in climate variability, so more in-depth soil moisture research is not possible.

The proposed methodology based on parametric and non-parametric methods employing a database collected from three typical land covers: crop, pasture, and forest, and typical soil types demonstrated that there are high correlations between superficial and intermediate soil moistures, as well as superficial and deepest soil moistures, so it is possible to propose prediction models to avoid the need to install many instruments per point or even the high ability of the proposed method to use surface information from satellite data to obtain soil moisture profiles.

A model fitting using MLR, SVR, and ANN was carried out, aiming to predict the soil moisture, by using as predictor variables the superficial soil moisture, its depth, the depth of the soil moisture that wants to be predicted, and the *soil type*. All models offer similar performance metrics, except for forest, where the SVR outperforms the other methods. However, goodness-of-fit tests are not met by residual errors of the MLR model, so it cannot be used with statistical validity. In that way, SVR is the best model to fit moistures based on superficial soil moisture.

Finally, it is important to state that the traditional hydrological models had been calibrated employing few information of the full complex and non-linear hydrological rainfall runoff transformation process (normally flow in specific points), avoiding the understanding of the spatial variability along the full catchment and imposing calibration parameters representing just a small part of the physical phenomenon. The inclusion of distributed spatial parameters in the calibration process such as infiltration parameters or soil moisture profiles can increase strongly the quality of the results obtained along the full catchment or even increase the understanding of the physics involved. In this direction the proposed methodology represents an important contribution to several engineering and hydrological applications.

7. Future work

We are planning to install two additional soil moisture stations in another forest, and residential soil uses, to better tune the models.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Special thanks to the University the Medellín for financing the research project “Interacciones entre la cobertura vegetal, la variabilidad climática y el contenido de humedad en el suelo, en cuencas agrícolas neotropicales”, Project code 1070.

References

[1] M. Gonzalez-Palacio, L. Sepulveda-Cano, J.D. Valencia-Calvo, J. Quiza-Montealegre, “System dynamics baseline model for determining a multivariable objetive function in Wireless Sensor Networks,” in CISTI 2020, 2020.

[2] V.R. Pauwels, R. Hoeben, N.E. Verhoest, F.P. De Troch, “The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation,” Journal of Hydrology, **251**, 88–102, 2001.

[3] H. Sharma, M.K. Shukla, P.W. Bosland, R. Steiner, “Soil moisture sensor

calibration, actual evapotranspiration, and crop coefficients for drip irrigated greenhouse chile peppers," Agricultural Water Management, **179**, 81–91, 2017, doi:10.1016/j.agwat.2016.07.001.

[4] S. Walther, G. Duveiller, M. Jung, L. Guanter, A. Cescatti, G. Camps-Valls, "Satellite Observations of the Contrasting Response of Trees and Grasses to Variations in Water Availability," Geophysical Research Letters, **46**(3), 1429–1440, 2019, doi:10.1029/2018GL080535.

[5] H. Janssen, G.A. Scheffler, R. Plagge, "Experimental study of dynamic effects in moisture transfer in building materials," International Journal of Heat and Mass Transfer, **98**, 141–149, 2016, doi:10.1016/j.ijheatmasstransfer.2016.03.031.

[6] L. Zhuo, Q. Dai, D. Han, N. Chen, B. Zhao, M. Berti, "Evaluation of Remotely Sensed Soil Moisture for Landslide Hazard Assessment," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **12**(1), 162–173, 2019, doi:10.1109/JSTARS.2018.2883361.

[7] L. Brocca, L. Ciabatta, C. Massari, S. Camici, A. Tarpanelli, "Soil moisture for hydrological applications: Open questions and new opportunities," Water (Switzerland), **9**(2), 2017, doi:10.3390/w9020140.

[8] X. Huang, Z.H. Shi, H.D. Zhu, H.Y. Zhang, L. Ai, W. Yin, "Soil moisture dynamics within soil profiles and associated environmental controls," Catena, **136**, 189–196, 2016, doi:10.1016/j.catena.2015.01.014.

[9] I. V. Florinsky, Digital Terrain Analysis in Soil Science and Geology: Second Edition, Elsevier Inc., 2016.

[10] J. Liu, B.A. Engel, Y. Wang, Y. Wu, Z. Zhang, M. Zhang, "Runoff Response to Soil Moisture and Micro-topographic Structure on the Plot Scale," Scientific Reports, **9**(1), 2019, doi:10.1038/s41598-019-39409-6.

[11] M. Pan, E.F. Wood, "Impact of Accuracy, Spatial Availability, and Revisit Time of Satellite-Derived Surface Soil Moisture in a Multiscale Ensemble Data Assimilation System," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **3**(1), 49–56, 2010, doi:10.1109/JSTARS.2010.2040585.

[12] B. Kuang, Y. Tekin, A.M. Mouazen, "Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content," Soil and Tillage Research, **146**(PB), 243–252, 2015, doi:10.1016/j.still.2014.11.002.

[13] K. Were, D.T. Bui, Ø.B. Dick, B.R. Singh, "A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape," Ecological Indicators, **52**, 394–403, 2015, doi:10.1016/j.ecolind.2014.12.028.

[14] S. Maroufpoor, E. Maroufpoor, O. Bozorg-Haddad, J. Shiri, Z. Mundher Yaseen, "Soil moisture simulation using hybrid artificial intelligent model: Hybridization of adaptive neuro fuzzy inference system with grey wolf optimizer algorithm," Journal of Hydrology, **575**, 544–556, 2019, doi:10.1016/j.jhydrol.2019.05.045.

[15] M.K. Gill, T. Asefa, M.W. Kemblowski, M. McKee, "Soil moisture prediction using support vector machines," Journal of the American Water Resources Association, **42**(4), 1033–1046, 2006, doi:10.1111/j.1752-1688.2006.tb04512.x.

[16] S. Ahmad, A. Kalra, H. Stephen, "Estimating soil moisture using remote sensing data: A machine learning approach," Advances in Water Resources, **33**(1), 69–80, 2010, doi:10.1016/j.advwatres.2009.10.008.

[17] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," Remote Sensing of Environment, **241**, 2020, doi:10.1016/j.rse.2020.111716.

[18] G. Dumedah, J.P. Walker, L. Chik, "Assessing artificial neural networks and statistical methods for infilling missing soil moisture records," Journal of Hydrology, **515**, 330–344, 2014, doi:10.1016/j.jhydrol.2014.04.068.

[19] M. Khalil, U.S. Panu, W.C. Lennox, "Groups and neural networks based streamflow data infilling procedures," Journal of Hydrology, **241**(3–4), 153–176, 2001, doi:10.1016/S0022-1694(00)00332-2.

[20] F.D. Mwale, A.J. Adeloye, R. Rustum, "Infilling of missing rainfall and streamflow data in the Shire River basin, Malawi - A self organizing map approach," Physics and Chemistry of the Earth, **50**–**52**, 34–43, 2012, doi:10.1016/j.pce.2012.09.006.

[21] T.R. Nkuna, J.O. Odiyo, "Filling of missing rainfall data in Luvuvhu River Catchment using artificial neural networks," Physics and Chemistry of the Earth, **36**(14–15), 830–835, 2011, doi:10.1016/j.pce.2011.07.041.

[22] P. Coulibaly, N.D. Evora, "Comparison of neural network methods for infilling missing daily weather records," Journal of Hydrology, **341**(1–2), 27–41, 2007, doi:10.1016/j.jhydrol.2007.04.020.

[23] M. Pal, R. Maity, "Development of a spatially-varying Statistical Soil Moisture Profile model by coupling memory and forcing using hydrologic soil groups," Journal of Hydrology, **570**, 141–155, 2019, doi:10.1016/j.jhydrol.2018.12.042.

[24] M. Aboutalebi, N. Allen, A.F. Torres-Rua, M. McKee, C. Coopmans, "Estimation of soil moisture at different soil levels using machine learning techniques and unmanned aerial vehicle (UAV) multispectral imagery," SPIE-Intl Soc Optical Eng: 26, 2019, doi:10.1117/12.2519743.

[25] R. Girden, "ANOVA: Repeated measures," Computer Science, 1991.

[26] N. Rodríguez-Fernández, P. de Rosnay, C. Albergel, P. Richaume, F. Aires, C. Prigent, Y. Kerr, "SMOS neural network soil moisture data assimilation in a land surface model and atmospheric impact," Remote Sensing, **11**(11), 2019, doi:10.3390/rs11111334.

[27] X. Dai, Z. Huo, H. Wang, "Simulation for response of crop yield to soil moisture and salinity with artificial neural network," Field Crops Research, **121**(3), 441–449, 2011, doi:10.1016/j.fcr.2011.01.016.

[28] W.E.H. Blum, Functions of soil for society and the environment, Reviews in Environmental Science and Biotechnology, **4**(3), 75–79, 2005, doi:10.1007/s11157-005-2236-x.

[29] K. Liao, X. Lai, Z. Zhou, Q. Zhu, "Applying fractal analysis to detect spatio-temporal variability of soil moisture content on two contrasting land use hillslopes," Catena, **157**, 163–172, 2017, doi:10.1016/j.catena.2017.05.022.

[30] J. Geris, D. Tetzlaff, J.J. McDonnell, C. Soulsby, "Spatial and temporal patterns of soil water storage and vegetation water use in humid northern catchments," Science of the Total Environment, **595**, 486–493, 2017, doi:10.1016/j.scitotenv.2017.03.275.

[31] L. Brocca, T. Moramarco, F. Melone, W. Wagner, "A new method for rainfall estimation through soil moisture observations," Geophysical Research Letters, **40**(5), 853–858, 2013, doi:10.1002/grl.50173.

[32] W. Dorigo, W. Wagner, C. Albergel, F. Albrecht, G. Balsamo, L. Brocca, D. Chung, M. Ertl, M. Forkel, A. Gruber, E. Haas, P.D. Hamer, M. Hirschi, J. Ikonen, R. de Jeu, R. Kidd, W. Lahoz, Y.Y. Liu, D. Miralles, T. Mistelbauer, N. Nicolai-Shaw, R. Parinussa, C. Pratola, C. Reimer, R. van der Schalie, S.I. Seneviratne, T. Smolander, P. Lecomte, "ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions," Remote Sensing of Environment, **203**, 185–215, 2017, doi:10.1016/j.rse.2017.07.001.

[33] H. Lin, H.J. Vogel, J. Phillips, B.D. Fath, Complexity of soils and hydrology in ecosystems, Ecological Modelling, **298**, 1–3, 2015, doi:10.1016/j.ecolmodel.2014.11.016.

[34] X. Jia, M. Shao, Y. Zhu, Y. Luo, "Soil moisture decline due to afforestation across the Loess Plateau, China," Journal of Hydrology, **546**, 113–122, 2017, doi:10.1016/j.jhydrol.2017.01.011.

[35] S. Zhang, W. Fan, Y. Li, Y. Yi, "The influence of changes in land use and landscape patterns on soil erosion in a watershed تاثیر تغییرات کاربری اراضی و الگوی شکل شناسی زمین برروی فرسایش خاک در حوزه," Science of The Total Environment, **574**, 34–45, 2017.

[36] L. Gao, Y. Lv, D. Wang, T. Muhammad, A. Biswas, X. Peng, "Soil water storage prediction at high space-time resolution along an agricultural hillslope," Agricultural Water Management, **165**, 122–130, 2016, doi:10.1016/j.agwat.2015.11.012.

[37] K.J. White, "The Durbin-Watson Test for Autocorrelation in Nonlinear Models," The Review of Economics and Statistics, **74**(2), 370, 1992, doi:10.2307/2109675.

[38] M. F.J.Jr., "The Kolmogorov-Smirnov test for goodness of fit," Journal of the American Statistical Association, **56**(1951), 68–78, 1951.

[39] D.M. Waldman, "A note on algebraic equivalence of White's test and a variation of the Godfrey/Breusch-Pagan test for heteroscedasticity," Economics Letters, **13**(2–3), 197–200, 1983, doi:10.1016/0165-1765(83)90085-X.

[40] W.S. Noble, What is a support vector machine?, Nature Biotechnology, **24**(12), 1565–1567, 2006, doi:10.1038/nbt1206-1565.

[41] N. Deng, Y. Tian, C. Zhang, Support vector machines: Optimization based theory, algorithms, and extensions, CRC Press, 2012, doi:10.1201/b14297.

[42] T. Kavzoglu, I. Colkesen, "A kernel functions analysis for support vector machines for land cover classification," International Journal of Applied Earth Observation and Geoinformation, **11**(5), 352–359, 2009, doi:10.1016/j.jag.2009.06.002.

[43] W. Zhou, L. Zhang, L. Jiao, J. Pan, "Support vector regression based on unconstrained convex quadratic programming," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag: 167–174, 2006, doi:10.1007/11881070_27.

[44] M. Awad, R. Khanna, Support vector regression, Apress, Berkeley: 67–80, 2015, doi:doi.org/10.1007/978-1-4302-5990-9_4.

[45] L. Tian, X. ZHANG, A Convergent Nonlinear Smooth Support Vector Regression Model, 205–207, 2015, doi:10.2991/978-94-6239-102-4_43.

[46] Y.O. Ouma, C.O. Okuku, E.N. Njau, "Use of Artificial Neural Networks and Multiple Linear Regression Model for the Prediction of Dissolved Oxygen in Rivers: Case Study of Hydrographic Basin of River Nyando, Kenya," Complexity, **2020**, 2020, doi:10.1155/2020/9570789.
[47] A. Landi, P. Piaggi, M. Laurino, D. Menicucci, "Artificial neural networks for nonlinear regression and classification," in Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications, ISDA'10, 115–120, 2010, doi:10.1109/ISDA.2010.5687280.
[48] A. Biglarian, E. Bakhshi, A.R. Baghestani, M.R. Gohari, M. Rahgozar, M. Karimloo, "Nonlinear survival regression using artificial neural network," Journal of Probability and Statistics, 2013, doi:10.1155/2013/753930.
[49] P.L. Fernández-Cabán, F.J. Masters, B.M. Phillips, "Predicting roof pressures on a low-rise structure from freestream turbulence using artificial neural networks," Frontiers in Built Environment, **4**, 2018, doi:10.3389/fbuil.2018.00068.
[50] B. Liu, M. Shao, "Modeling soil-water dynamics and soil-water carrying capacity for vegetation on the Loess Plateau, China," Agricultural Water Management, **159**, 176–184, 2015, doi:10.1016/j.agwat.2015.06.019.
[51] H. YiLong, C. LiDing, F. BoJie, H. ZhiLin, G. Jie, L. XiXi, "Effect of land use and topography on spatial variability of soil moisture in a gully catchment of the Loess Plateau, China.," Ecohydrology, **5**(6), 826–833, 2012.
[52] X. Fang, W. Zhao, L. Wang, Q. Feng, J. Ding, Y. Liu, X. Zhang, "Variations of deep soil moisture under different vegetation types and influencing factors in a watershed of the Loess Plateau, China," Hydrology and Earth System Sciences, **20**(8), 3309–3323, 2016, doi:10.5194/hess-20-3309-2016.
[53] C. Zhu, Y. Li, "Long-Term Hydrological Impacts of Land Use/Land Cover Change From 1984 to 2010 in the Little River Watershed, Tennessee," International Soil and Water Conservation Research, **2**(2), 11–21, 2014, doi:10.1016/S2095-6339(15)30002-2.
[54] L. Gao, M. Shao, "Temporal stability of soil water storage in diverse soil layers," Catena, **95**, 24–32, 2012, doi:10.1016/j.catena.2012.02.020.
[55] X. Mei, Q. Zhu, L. Ma, D. Zhang, H. Liu, M. Xue, "The spatial variability of soil water storage and its controlling factors during dry and wet periods on loess hillslopes," Catena, **162**, 333–344, 2018, doi:10.1016/j.catena.2017.10.029.
[56] B. Yang, X. Wen, X. Sun, "Seasonal variations in depth of water uptake for a subtropical coniferous plantation subjected to drought in an East Asian monsoon region," Agricultural and Forest Meteorology, **201**, 218–228, 2015, doi:10.1016/j.agrformet.2014.11.020.
[57] Soil Survey Staff, Keys to soil taxonomy, 2014.
[58] IGAC, General Study of Soils and Land Zoning: Department of Antioquia (In Spanish), Instituto Geografico Agustin Codazzi, Bogota, Colombia, 2007.
[59] IDEA, Semi-detailed Study of Soil in Zone 13 of the Municipality of Envigado for Potential Use Purposes (In Spanish), 2014.
[60] Y. Zhang, L. Qiao, C. Chen, L. Tian, X. Zheng, "Effects of organic ground covers on soil moisture content of urban green spaces in semi-humid areas of China," Alexandria Engineering Journal, 2020, doi:10.1016/j.aej.2020.08.001.
[61] Unidots, Ubidots IoT Platform, 2020.
[62] R.L. Burden, F. J.D., Numerical Analysis, Brooks/Cole, Cengage Learning, 2011.
[63] J.M. Chambers, A.E. Freeny, R.M. Heiberger, Analysis of variance; designed experiments, CRC Press: 145–193, 2017, doi:10.1201/9780203738535.
[64] S. Fine, K. Scheinberg, "Efficient svm training using low-rank kernel representations," Journal of Machine Learning Research, **2**(Dec), 243–264, 2002.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 637-644 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Overcome Discrimination: A Logistic Regression with 10-year Longitudinal Investigation of Emo Kids' Facebook Posts

Proud Arunrangsiwed[1,*], Yothin Sawangdee[2]

[1]Faculty of Management Science, Suan Sunandha Rajabhat University, Bangkok, 10300, Thailand

[2]Institute for Population and Social Research, Mahidol University, Nakhon Pathom, 73170, Thailand

ARTICLE INFO

Article history:
Received: 29 July, 2020
Accepted: 25 September, 2020
Online: 05 October, 2020

Keywords:
Emo subculture
Facebook
Logistic regression

ABSTRACT

This study primarily aims to identify the factors that helped emo kids in 2010 move through the emo-identity discrimination and be able to obtain a certain level of achievement. Facebook is the social network that allows users to track friends' posts back over 10 years. Content analysis was conducted by using two coders to rate 1,432 Facebook posts published during 2010 until 2019, based on the variables: (1) emo-related content, (2) emotion expression, (3) being a part of fandom, (4) the group of fandom, (5) friend(s) and (6) family appearance. The results from logistic regression analysis reveal that past emo kids' Facebook posts are increasing in the contents about friends, family, and fan objects over time. However, the number of posts with emo music or emo fan object was reduced. Having social support and belonging to social group, like fandom, might help past emo kids overcome the hard time that they had got prejudiced. Future studies should develop a model or existing theory to explain the complexity of individuals' overlapping identities blended in social networking profiles.

1. Introduction

Emo is a subculture and a music genre being popular for teenagers in the first decade of 2000s [1]. It was linked to negative deviant behaviors, such as suicide and self-harm [2], [3]. The members of subculture or "emo kids" in those days currently become young adults [4]. However, these people initially experienced music-taste-based discrimination, and they are negatively stereotyped by people in mainstream culture [5], [6]. This study aims to track their social network posts to see the transformation of their identity and the evolution of relationship with family and peers. These would help identify the factors that save these subculture members from unfair prejudice, which result in their well-being in society.

Emo is considered as a subculture because it has its own values, norm, belief, sets of behaviors, and communication patterns [1], [7]. This subculture was spread via online media, such as Facebook, Myspace, and YouTube. The members shared common interest toward emo, punk music, and artists, such as Fall Out Boy, We the Kings, The Academy Is…, Dashboard Confessional, My Chemical Romance, Paramore, Mayday Parade, Taking Back Sunday, Panic! At the Disco, The All-American Rejects, Boys Like Girls, etc. They would not just perform as fans of these artists, but they also need to be able to identify with their music [1], [8]. These fans believe that their favorite songs can help them with difficult life event regarding school and parents. Their common apparels were black T-shirt and skinny jeans, and they usually applied black eyeliner with dyed black hair covering one eye [9], [10]. Emotional expression is acceptable for both female and male ones, e.g., subculture members would comfort a crying emo boy, rather than blaming him regarding the lack of masculinity. As the result of these characteristics, emo boys were often stereotyped by outsiders as gay men, which could cause serious cyberbullying.

Other than the mentioned practices, suicide and self-injury are perceived as important norms of emo subculture [11], [12]. After the suicide of two Australian female students, Jodie Gater and Stephanie Gestier, in 2017, emo kids were not viewed only as subculture members or fans of modern punk music, but folk devils [13], [14]. They used social network and allowed emo texts to alter their thoughts to associate with self-injury and suicidal idea [8], [15]. This contradiction against social norm is the cause that public marked them as deviants rather than a subculture [16], [17]. Mass media have played an important role to create moral panic regarding the link between music taste and anti-social behaviors

*Corresponding Author: Proud Arunrangsiwed, Suan Sunandha Rajabhat University, email: proud.ar@ssru.ac.th, parunran@nyit.edu

[18] as they connected the case of Columbine school shooting to metal music [19], [20], and the suicide of two female teenagers to emo music fandom [1], [21]. Although metal fans were initially judged by public that their music taste was related to aggression [19], [22], this group of fans discriminate against emo kids based on similar stereotypes [19], [23]. Metal fans also differentiate themselves from emo teens that they were stronger and more masculine than emo kids [24]. Whether emo boys were straight or had an LGBT identity, they were often prejudiced as gay [23]-[26]. On the other hand, as a part of punk music genre [27], punk musicians, fans, and subculture members have displayed their feminine aesthetic by applying cosmetics and female clothing for long time before the birth of emo music [28], [29]. In other words, how people viewed emo kids as gay or problematic is a kind of the discrimination toward a group.

2. Hypotheses Construction

In aforementioned section, emo is both music genre and subculture, some members or fans of which engage self-mutilation and suicide, while public stereotyped them -as a whole- to be folk devils with aggression against self [30], [31]. Problem music, whether punk, emo, or metal, has been targeted as the cause of negative behaviors [32]. Later, it was found that low self-esteem, not problem music, could directly bring about self-harm [33], and there is no relationship found between problem music use and low self-esteem [34], [35]. This implies that emo music may not be the cause of teens' suicide and self-harm, but it might be a part of coping process that pushes them through their actual initial life problem, that had diminished their self-esteem or had led them to self-mutilation [36].

While people outside subculture understand that teenagers are reinforced with suicidal thought and self-injury by emo music, emo fan scholars reported that these fans stopped cutting themselves after being coped by such the music [21], [37]. To identify with sad music is a part of coping skills that help people accept and make sense with negative event in their life [38], [39]. Sad emo songs were easily used as coping tools by their fans, since the lyrics are about problems with friends, family, heartbreaking, and loneliness [27], [40], [41]. Moreover, a part of subculture norm is to express emotion and identify with emotional emo songs [1], [8]. This helps construct the hypotheses that prior emo kids may continuously listen to emo music or being emo after the age of emo boom (2010), and they have been likely to express their emotion in their social network posts. Hence, the following hypotheses were formulated.

H1: Emo kids have posted emo-related texts on their social network (Facebook) profile since 2010 (roughly the last year that emo subculture was popular) until 2019.

H2: Emo kids have expressed their emotion in their social network posts since 2010 until 2019.

Although emo music helps teenagers cope with personal troubles, they may not avoid negative stereotype constructed by public. Therefore, it is important to identify the factors that help them walk through this discrimination. Racial discrimination could result in lower self-esteem and higher depression [42], [43]. Parents should not only help their children prepare for being discriminated at school [44], but should also teach them coping skills after ad hominem attack [45]. Other family members and people with similar identity as the victims could also help reduce stress from discrimination and find the way to solve such the problem for the victims [46] – [48]. The above literature could indicate that family and peers might be an important part that saved emo kids from the discrimination. This means emo kids may have lived with their family and communicated with their friends during 2010 to 2019. The particular evidence eventually entails the following hypotheses.

H3: Emo kids have posted texts (alphabetic text, image, or video) about their friends since 2010 until 2019.

H4: Emo kids have posted texts about their family since 2010 until 2019.

However, Strauss [49] stated that self-harm in emo kids was rooted from their lack of social and family support, and this is opposite from H3 and H4. If these adolescents were negatively stereotyped, low in self-esteem, and have lack of social support; being a part of fandom or fanship might be the factor that heightens their self-esteem [50], [51], and allows them to healthily grow up as well as people in mainstream culture [52]. Joining fan community could provide fans with sense of belonging and that brings about self-esteem [53]-[55]. While fans listen to their favorite songs, the music could stimulate their self-esteem [35], especially for metal music fans who extremely believe in ones' own authenticity [56]. Authenticity refers to having realness value, or freedom of expression regardless others' attitude or hegemony [57]. Eminem's authenticity, for instance, is scorn and crudeness [58]. Hall [29] found that emo musicians and also fans accepted their difference apart from outsiders in mainstream culture. This is close to the term of self-concept clarity toward group identity, which is associated with having self-esteem [59], [60]. Recently, it was found that emo fans could earn fan-related self-esteem during attending emo-punk concert [61]. These could derive a hypothesis that emo kids have been part of fandom or fanship whether it is directly about emo music or other objects of interest.

H5: Emo kids have posted texts related to any kind of fandom or fanship from 2010 until 2019.

Because identity can be changed over time [62], [63], to track Facebook posts of past emo kids could help understand their identity transformation related to their fan identification, emotion expression as subculture value, and how they obtain family and peer support. This may help explain how they overcome the discrimination and can live normally as well as general adults.

3. Method

The current study built on the previous study, which reported that emo kids who engaged in emo online community in past 10 years could handle the emo-identity discrimination and grow up and work in proper careers [64]. The current study would track the social network posts of the same group of past emo kids in mentioned study to seek for the factors that may allow them to move pass through the discrimination, and consequently, they could reach their life achievement.

Based on the reviewed articles [37], [46], [48], [51], fan identity, family and peers, and emo music consumption could be the important factors that help emo kids overcome the

discrimination and other life problems. Hence, the major variables that would be explored are (1) emo music or emo-subculture-related content in the Facebook post, (2) the level of emotion expression in the post, (3) being fan or being a part of fandom, (4) the group of fandom that the samples belong to, (5) friend(s) and (6) family appearance in Facebook post.

3.1. Samples

Because emo subculture was risen and grown in the Internet [14], data collection of the current study would be done online, as well as earlier studies of Overell [1], and Zdanow and Wright [8]. The cases of these researchers were prepared in the time that emo subculture and music were popular, or around 2006-2010, but for the current study, the data collection was done in 2019. The samples needed to be active social network users since 2010 until 2019. Because people could change their identities over time, to use the continuous-active users could help confirm their original identification as a part of emo subculture.

Therefore, the first researcher of the present study used one's own Facebook friend list containing emo kids in 2010 as the site for data collection. The majority of samples were white and yellow. Few were from Africa, Middle East, and India. This recruitment method might be the only way to ensure that these users had had emo identity during 2010. Other strategies to recruit the cases might be impossible, because it is hard to know the exact date that users like an emo-related Facebook page or join in a group. For Myspace, although it was the major social network considered as social capital for emo kids in those days, most users have now been inactive [1], [14], [24], [41].

The mentioned friend list had contained more than 300 users, but in 2019, the number of users left in this list is 273. The users who did not post any new post in the latest two months were excluded from the study. Only 62 users were active and selected as the samples. After 62 active users were added into the codesheet, we prepared the codesheet by listing the randomly selected timeline (wall) posts. Twenty-nine posts were chosen from each user. Three posts were selected each year, from 2010 until 2019. The only year with 2 selected posts is 2019, because the data collection was done in July, 2019. The process of the randomization was that each selected post in each year had equally gap of time, which was about 4 months. For instance, if the first selected post was published in early January, the next selected posts would be published in early May and early September, respectively. Therefore, the possible highest number of cases is 1,798. Because some users did not post anything in some month(s), the number of valid cases left for the analysis is 1,432. The 366 cases, which were blank, got removed from the final codesheet.

3.2. Data Collection

Data collection of the present study was done in Facebook. The raters needed to gather the information from "About" section of profile, and the selected posts in main timeline or wall. These posts could be in many forms, such as text, image, video, link from other websites, and mixing between two or more mentioned types. After codesheet preparation, two raters coded 7 variables as follows:

- Gender. Raters could look for this information in "about" part of Facebook profile. Female was coded as 0, and male was coded as 1.
- Emo-music-or-subculture-related content in the post. Raters could consider adding the value into this variable based on text, image, or video in each post. If the media contain emo music, music video, emo apparels, emo facial makeup, and hairstyle, the particular post would be marked as emo-related. Some phases or signs in text were also associated with emo-style discourse, e.g., xoxo, rawr, etc. This variable is for checking if users still have the same identity as they had had 10 years ago. If the post applied to the given condition, the raters would code 1, and 0 for not applied.
- The level of emotion expression in the post. Unlike the previous items, this variable is not dichotomous. If the user did not display any emotion at all, both verbally and nonverbally (use photos, sticker, Facebook mood, or emoticon), the raters would code the post as 0. The small, moderate, and extreme amount of emotion expression were coded as 1, 2, and 3, respectively.
- Friend(s) appearance in Facebook post. If friend(s)' name(s) was mentioned or tagged in the post, or friend(s)' photo(s) was published, the raters would code that post as 1. On the contrary, if the post was not related to any friend, it would be coded as 0.
- Family appearance in Facebook post. The coding rule was similar with the previous item. If there are one or more family members in the post, it would be coded as 1, and if not, it would be coded as 0.
- Being fan or being a part of fandom. Whether the user displayed oneself as emo fan or fan of other media object, the post would be coded as 1, and if the post has nothing to do with any fandom or idol, it would be coded as 0. It might be important to note that fandom-related posts could be a shared official post by their favorite artist, photos of artists, album cover, shared news, and also fan art or fan Photoshop manipulation.
- The group of fandom that the samples (each of their posts) belong to. If the previous item, "being fan or being a part of fandom," were coded as 0, the raters would leave this part of codesheet blank. If the raters found that the post was related to emo fandom, it would be coded as 0. If metal, death metal, or melodic death metal music was found in the post, it would be coded as 1. If the post was about pop or R&B music, it would be coded as 2. If the post was about sport, movie, Japanese anime/manga, games, and politic or government, it would be coded as 3, 4, 5, 6, and 7 respectively. Posts about politics and government could be considered as fandom-related posts, because it consists of the acts of fan and anti-fan, to support or to protest against some policy. Moreover, Bronstein [65] discovered that many people treated politicians the same way as they treated celebrities, and vice versa. Politicians could post non-politic-related photos, such as vacation with their family, and their supporters also commented and shared the posts.

Another variable is the number of month(s) counted from the beginning of 2010. This variable is the major predictor in all regression models. If the post was published in March, 2010, for instance, the value of this variable would be 3. If the post was published in February, 2011, the value will be 14. Hence, the smallest possible value is 1 month, and the highest is 115 (9 years and 7 months).

Case ID	year	month	Gender	Friend	Family	Emotional Expression Level	Fandom-related	Emo-subculture-related	Type of fandom	Month Pass Counted from Jan 2010
115	2018	Sep	0	0	1	2	0	0		105
116	2019	Feb	0	1	0	2	0	0		110
117	2019	July	0	0	1	2	0	0		115
118	2010	Sep	0	1	0	2	1	0	3	9
119	2011	Feb	0	0	0	3	1	0	3	14

Figure 1: Codesheet preparation

3.3. Analysis

Most parts of the data were presented using descriptive statistics. Logistic regression analysis was used to understand the probability of the variables, emo-music-or-subculture-related contents, being a part of fandom or fanship, friend(s), and family appeared in posts over time during 2010 to 2019. This implies that in the analysis, independent variable is the number of month counted from the beginning of 2010, and the dependent variables are all 4 mentioned dichotomous variables. At the same time, linear regression analysis was used to understand the effect of passing time on the level of emotion expression in social network post.

Because logistic regression analysis was employed for data analysis, the rated values of dichotomous variables cannot be averaged. There were 24 cases with disagreement, so the values of one rater were chosen to paste in final codesheet.

4. Findings

To answer if these past emo kids still listen to emo music or having emo identity (H1), the findings show that there are 341 posts related to emo music and emo subculture from 2010 until 2019. These are 23.8 percent of all post (n=1432), as presented in Table 1. However, lately, the number of this kind of post had been slightly decreased (Odds Ratio=.995; p=.006), especially in last 4 years (2016-2019).

According to Table 1. it is noticeable that some groups of variables are overlapping, and the most overlapping cases are between the groups, fandom and emo-related posts (n=150), and following by the overlapping of fandom, emo-related posts, and friend appearance in posts (n=33).

As shown in Figure 2., both male and female tracked emo kids have similar decreasing tendency in posts related to emo music and subculture, and women generally have higher numbers of this kind of posts than men. It was interesting to look into the demographic data regarding gender. After checking carefully in these 62 selected users, there are only two genders appearing which are female and male. Based on these 1432 cases, 570 cases were posts by female users, and 862 cases were posts by male users. This offers an opportunity for future researchers to understand more about emo masculinity. Men, who are emo or were previously emo, may not have an LGBT identity, but with the subculture value, it allows men to display their feminine characteristics.

Table 1: Frequency of, and friend-and-family-related posts, fandom-related post, and emo-related posts

Friend Appearance	Family Appearance	Fandom-related	Emo-related	n	percent
✓	-	-	-	154	10.75%
-	✓	-	-	83	5.80%
✓	✓	-	-	13	0.91%
-	-	✓	-	343	23.95%
✓	-	✓	-	27	1.89%
-	✓	✓	-	8	0.56%
✓	✓	✓	-	1	0.07%
-	-	-	✓	122	8.52%
✓	-	-	✓	22	1.54%
-	✓	-	✓	4	0.28%
-	-	✓	✓	150	10.47%
✓	-	✓	✓	33	2.30%
-	✓	✓	✓	5	0.35%
✓	✓	✓	✓	5	0.35%
-	-	-	-	462	32.26%
255	119	572	341	1432	100%
17.81%	8.31%	39.94%	23.81%	100%	

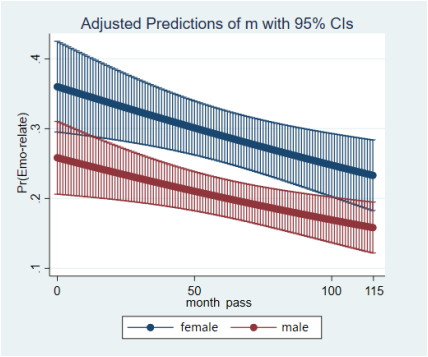


Figure 2: Margins plots for emo-related posts

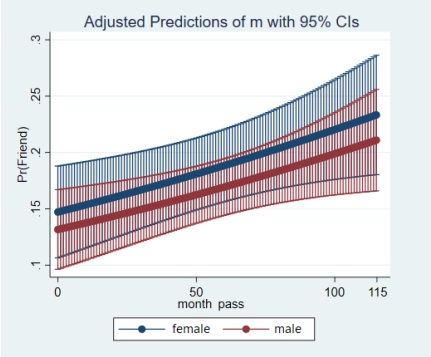


Figure 3: Margins plots for friend-related posts

Although emo media consumption and identification toward subculture were somewhat reduced, they still have emo core value, which is vulnerability (H2). Many posts were high in emotion expression until 2019 (Mean=1.661; S.D.=1.129). The linear

regression reveals no difference in the level of emotion expression in the post over time (F=3.595; R2=.003; Beta=-.050), but with small decreasing lately.

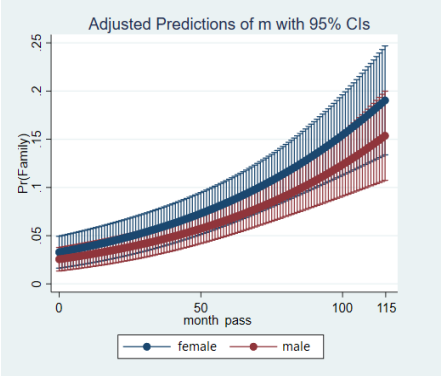


Figure 4: Margins plots for family-related posts

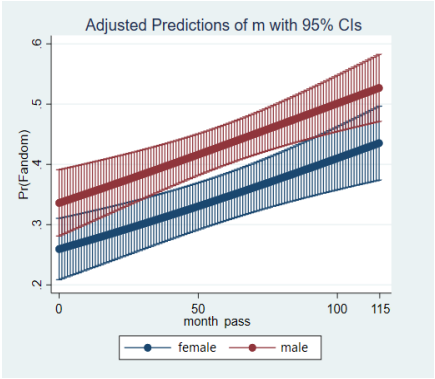


Figure 5: Margins plots for fandom-related posts

Unlike the expectation (H3, H4), these tracked emo kids rarely posted text and photos about their family and friends. Table 1. shows that there are only 17.8 percent of posts containing text, image, or video related to friend(s), and only 8.3 percent related to family. Based on the results of logistic regression in Table 2., only small increasing in friend-related posts was found (OR=1.005; p=.022), while there is a strong significant increasing of family-related posts (OR=1.017; p=.000). As seen in Figure 3. and 4., there is an actual increasing of friend and family contents in their Facebook posts. Women are higher in these types of posts them men. This could be because women have more willingness to disclose personal information in social network than men [66]. This might be related to gender norm and social acceptance that women can freely display their friendship, intimacy, and relationship with their family and peers.

Table 2: Univariate logistic regression
with the number of month counted from the beginning of 2010 as the predictor.

Dependent Variables	Emo-Related	Friend(s)	Family	Fandom
Overall OR (95%CI)	.995** (.991-.998)	1.005* (1.001-1.009)	1.017*** (1.011-1.023)	1.007*** (1.004-1.010)
Female OR (95%CI)	.996 (.990-1.001)	1.006 (.999-1.012)	1.014** (1.005-1.024)	1.011*** (1.006-1.017)
Male OR (95%CI)	.994* (.989-.999)	1.004 (.999-1.010)	1.019*** (1.010-1.028)	1.004* (1.000-1.008)
Overall OR (95%CI) Controlled by Gender	1.005** (1.002-1.009)	.995* (.991-.999)	.983*** (.977-.989)	1.445*** (1.159-1.800)

*p<.05; **p<.01; ***p<.001, n=1,432, n_{female}=570, n_{male}=862

The examples of family members found during the coding are wife, husband, grandmother, and uncle. The later marriage might be the reason that the number of family-related posts are lately greater. After reviewing the actual data, it could also be noticed that some who were married often changed their music taste and group of fandom to match their life partner.

For the fifth hypothesis (H5), these past emo kids belong to groups of fandom, whether it is emo music or other kind of media object fandom. This means that these people have had their fan identity for entire 10-year period tracked in the current study. The frequency of posts related to each group of fandom was presented in Table 3.

Table 3: Frequency of the posts
shown the groups of fandom which the cases belong to

Belonging to fandom	Number of Posts	Percent
Emo music	266	46.50
Metal/Death Metal	20	3.50
Pop or R&B	140	24.48
Sport	66	11.54
Movie	62	10.84
Japanese anime/manga	48	8.39
games	63	11.01
politic or government	37	6.47
Total (with overlapping cases)	702	112.73[1]
Total (all fandom-related post)	572	100.00

[1] The percentage is over 100, because some posts contained more than one type of fandom-related content.

Around 2013-2019, they posted texts and images about their fan objects more frequent than the first 3 years, as shown in Figure 5. Most of them still prefer emo music and bands until 2019, but they have also belonged to more groups of fandom. Like, after the first year, 2010, many of them started to listen to softer music, like pop and R&B. At the same time, some of them were interested in games, and Japanese comics and animation. Some showed that they have started to watch sport matches since 2015, such as soccer and American football. Similarly, some began to be interested in famous movies and Hollywood actors in 2015. In late 4 years, 2016-2019, some turned themselves to be the fans of politic. These past emo kids might increase their online civic engagement because they got older.

5. Discussion and Suggestion

As shifting away from the findings and suggestions of various studies that family plays an important role to solve the problem regarding discrimination in teenagers and students [45], [46], [48], the current study indicates that family and friends may not be a prominent factor that helps emo kids overcome the discrimination. This is because there had been lack of family-and-friend-related Facebook posts, especially at the beginning of the tracked period, 2010-2019. It is necessary to note that emo-identity discrimination had been peak at the beginning of the time period. One observation that can be explained this issue is that at the time that they got discriminated or cyberbullied, they may get lack of family support. Therefore, this part of findings suggests that it was other factor(s) that helps them overcome the discrimination.

Since the current study was conducted with content analysis and the cases were the public posts on Facebook, it might yield some imprecise data. Many emo kids, back in those days, were teenagers, and most of them might live with their parents and

communicate with them, but they might not need to show the photos of their parents in online space where other members of emo subculture would see. Like, middle school students do not want their parents to show up at school, because they may want to look independent and confident [67]. Therefore, these samples, with high-school age (around 2010), may not want to show their family on Facebook, too. To choose only one right conclusion among the finding (H3, H4) and this assumption that emo kids might live with parents in early 2010s, the in-depth interview should be employed in future studies.

Supported by numerous fan studies [50], [51], [56] and psychological studies [52], [68], these tracked emo kids may be able to move pass though the obstruction regarding discrimination, because they have been a part of fandom(s) (H5). Because fan activities [69], fans' media consumption [50], and sense of belonging in fan community [55] have a significant impact on fans' self-esteem, these emo kids may obtain the self-esteem from these fan-related factors. To have high self-esteem, people will be low in depression, and high in many positive life outcomes [52], [70].

Listening to emo music and identifying with it is also a kind of coping skills [21], [37]. Emo kids who faced negative life events were mentally healed by listening to emo music [21], [37]. The result of this study reveals that the tracked emo kids have continuously consumed emo music and expressed their emo identity (H1). Emo music might be the factor that helps them overcome the discrimination. It might be unusual to conclude that emo fans were prejudiced because of their music taste, and they could overcome the discrimination because of listening to emo music. An alternative suggestion is that fans used emo music as a part of coping process according to serious event in their real life. To be discriminated online might become small issue compared to the main problem they had had. This could be supported by the finding of Hall [29] that emo artists understood and showed the difference between emo subculture members and outsiders. Moreover, listening to ones' favorite music could help heighten self-esteem [35], so if emo fans were high in self-esteem strongly toward music and their subculture, the external factor might not be able to harm them.

Because emo kids have not only communicated with only people in the same subculture, it is still important to understand how they communicated with other people. Emo music and its fandom have been great parts of emo subculture. Fan communication may be able to help explain the practice in emo subculture. Fans normally communicated to 4 groups of people regarding their fan object and identity, which are (1) in-group members with the same identity, (2) hegemony, – or in this case is emo artists and emo bands, – (3) general out-group members, e.g. the friends and parents, and (4) out-group members who are their enemy, like people who discriminated against them [71].

To explain emo subcultural identity and their communication pattern might be somewhat complex. In Facebook or other social networking system (SNS) blends together many contexts people live in, in their real life. A Facebook friend might be an in-group member of emo subculture, but one might be an out-group member of workplace context. Future researchers may try to use social identities theory to explain all these complexities of the way many identities of an individual are mixed and blended in single SNS profile.

In the current study, the findings allow us to see the transformation of the identity from one to the others, from emo fans to other types of fandom, from young adolescents to adults with family and larger group of friends, and especially from vulnerable teenagers to stronger mature people.

Engaging in fandom might be the most important factor that helps emo kids move forward the discrimination, while the overall result could suggest that the way they move on and the transformation of their identity may be developed alongside with one another. Social network does not only help scholars to understand a subculture, but it actually helps and individual to see their developing identity, help them to understand themselves better, and that might be a part of how they move forward from discrimination, too.

An unexpected finding beside the major part of hypothesis testing is that many emo kids in 2010 had left their Facebook accounts, which causes a lot of missing cases in the present data sheet. Burke [72] found that many graduate students left their Facebook profile and signed up for other SNS such as LinkedIn because they need to look more professional, and some needed to switch to use other SNS to match their colleagues' choice.

The identity as an emo subculture member cannot be considered as a professional. Instead of editing the existing Facebook profile, some may choose to create a new profile. With the same group of samples but earlier study [64], many of them have worked in professional occupation such as, teachers, lecturers, engineers, designers, and some are entrepreneur of real estate business. In other words, while some people decided to leave their Facebook profile, some chose to continue using the same profile, but they turned it from an emo profile to be an expert one.

Even though the data of this study may consist of selection bias, we could learn that social network is an important part to spread, support, and maintain many subcultures, and one of these is emo subculture. We could see emo kids through their social network profiles that they have grown up to be young adults, used social capital in order to overcome the discrimination, coped with possible depression, and finally adjusted their life as subculture members to fit into the world complex society. Future study should continue looking at these emo kids' identity development and communication pattern through their adulthood and older adult ages. This way could help understand the strategies to achieve the long-term well-being of people who had experienced taste-and-identity-based discrimination.

6. Limitations

The major limitation of the present study is the small number of samples, which is 62 users. This may not be comparable to the actual population or emo kids before 2010. Moreover, these 62 users are the researcher's online friend, which might not represent general emo kids. Additionally, the data collection was content analysis done in online space, Facebook, where people may construct another identity such as ideal-self.

To solve this issue, researchers of future studies should hire computer experts to recruit more prior emo kids in 2010 who are

still active on social networking today. This would help increase sample size and strengthen the statistical conclusions of the current study. It is also important to note that many genders are available for Facebook users to choose. Therefore, samples with LGBT identity and asexual should be also included in future studies. As a result of this, the particular research would achieve a better generalization.

References

[1] R. Overell, "Emo online: networks of sociality/networks of exclusion," Perfect Beat, **11**(2), 141–162, 2010, doi: 10.1558/prbt.v11i2.141.

[2] M. Definis-Gojanović, D. Gugić, D. Sutlović, "Suicide and Emo youth subculture–a case analysis," Collegium antropologicum, **33**(2), 173–175, 2009.

[3] R. Young, N. Sproeber, R.C. Groschwitz, M. Preiss, P.L. Plener, "Why alternative teenagers self-harm: exploring the link between non-suicidal self-injury, attempted suicide and adolescent identity," BMC psychiatry, **14**(1), 137, 2014, doi: https://doi.org/10.1186/1471-244X-14-137.

[4] K. Moulton, Dashboard Confessional's Chris Carrabba Refuses to Age Out of Emo. Westword, (2018). Retrieved from https://www.westword.com/music/dashboard-confessionals-chris-carrabba-wont-age-out-of-emo-10161529

[5] S.J. Minton, Alterophobic bullying and violence. in J.L. Ireland, P. Birch, C.A. Irland, (eds.), The Routledge International Handbook of Human Aggression: Current Issues and Perspectives, Routledge, 2018, ISBN: 978-1-315-61877-7.

[6] D. Hernández-Rosete, "Bullying of emos in Mexico City: an ethnographic analysis," Cadernos de Saude Publica, **33**(12), 2017, doi: 10.1590/0102-311x00080116.

[7] T.D. Baker, S. Smith-Adcock, V.R. Glynn, "The Ghost of Emo: Searching for Mental Health Themes in a Popular Music Format," Journal of School Counseling, **11**(8), 1–35, 2013.

[8] C. Zdanow, B. Wright, "The representation of self-injury and suicide on emo social networking groups," African Sociological Review/Revue Africaine de Sociologie, **16**(2), 81–101, 2012.

[9] M. Heřmanský, There and Back Again: Linking Online and Offline Spaces in/of Czech Emo Subculture. in Hebdige and Subculture in the Twenty-First Century (pp. 169-205), Palgrave Macmillan, Cham, 2020.

[10] C.U. Gielis, It's not a fashion statement. An exploration of masculinity and femininity in contemporary emo music, Master thesis, Radboud University, 2018.

[11] P. Arunrangsiwed, R. Arunrangsiwed, Stereotyped Emo Kids: A literature review. in The 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), IEEE, 2019, doi: 10.1109/TIMES-iCON47539.2019.9024586.

[12] C. Partridge, Suicide, Angst, and Popular Music, Coward-Gibbs, M. (Ed.) Death, Culture & Leisure: Playing Dead (Emerald Studies in Death and Culture), Emerald Publishing Limited, 189-208, 2020, https://doi.org/10.1108/978-1-83909-037-020201020.

[13] M. Phillipov, "Just emotional people? Emo culture and the anxieties of disclosure," M/C Journal: A Journal of Media and Culture, **12**(5), 2009.

[14] M. Phillipov, "Generic misery music? Emo and the problem of contemporary youth culture," Media International Australia, **136**(1), 60–70, 2010, doi: doi.org/10.1177/1329878X1013600109.

[15] R. Trnka, M. Kuška, K. Balcar, P. Tavel, "Understanding death, suicide and self-injury among adherents of the emo youth subculture: A qualitative study," Death studies, **42**(6), 337–345, 2018, doi: 10.1080/07481187.2017.1340066.

[16] M.D. Daschuk, It's not a fashion statement, it's a death wish: subcultural power dynamics, niche-media knowledge construction, and the emo kid folk-devil, Master's thesis, University of Saskatchewan, 2009.

[17] S.F. Williams, "A walking open wound: Emo rock and the "crisis" of masculinity in America," in F. Jarman-Ivens (Ed.), Oh Boy! Masculinities and Popular Music (pp. 153–168). Taylor & Francis Group, 2013.

[18] R. Burns, C. Crawford, "School shootings, the media, and public fear: Ingredientsfor a moral panic," Crime, law and social change, **32**(2), 147–168, 1999.

[19] M.A. Miernik, "The evolution of emo and its theoretical implications," Polish Journal for American Studies, **7**, 175–188, 2013.

[20] G.W. Muschert, "Frame-changing in the media coverage of a school shooting: The rise of Columbine as a national concern," The Social Science Journal, **46**(1), 164–170, 2009, doi: 10.1016/j.soscij.2008.12.014.

[21] R. Haydn, My Chemical Romance: This Band Will Save Your Life, Plexus Publishing, 2015.

[22] C. Baker, B. Brown, "Suicide, self-harm and survival strategies in contemporary heavy metal music: a cultural and literary analysis," Journal of medical humanities, **37**(1), 1–17, 2016, doi: 10.1007/s10912-014-9274-8.

[23] R.L. Hill, "Is Emo Metal? Gendered Boundaries and New Horizons in the Metal Community," Journal for cultural research, **15**(3), 297–313, 2011, doi: 10.1080/14797585.2011.594586.

[24] K. Schmitt, Exploring dress and behavior of the emo subculture. Master's thesis, Kennesaw State University, 2011.

[25] H.T. Sternudd, A. Johansson, The girl in the corner: Aesthetics of suffering in a digitalized space, in L.G. Guerra, and J.A. Nicdao (Eds.), Narrative of Suffering: Meaning and Experience in a Transcultural Approach (pp. 105–115). Inter-Disciplinary Press, 2014.

[26] B.M. Peters, "Emo gay boys and subculture: Postpunk queer youth and (re) thinking images of masculinity," Journal of LGBT Youth, **7**(2), 129–146, 2010, doi: 10.1080/19361651003799817.

[27] B. Bailey, Emo Music and Youth Culture. in S. Steinberg, P. Parmar, B. Richard (Eds.), Encyclopedia of Contemporary Youth Culture, Greenwood Press, 2005.

[28] C.A. Aberle, Serious is Bad: A Queer Reading of Punk, Midwest Emo, and Connecticut DIY, Honors thesis, Wesleyan University, 2019.

[29] N. Hall, Representations of the Outsider in David Bowie's Glam Period and its Continuation through Punk, Goth, and Emo: Thematic, Aesthetic, and Subcultural Considerations, Master's thesis, University of Ottawa, 2015.

[30] M. Chęć, A. Potemkowski, M. Wąsik, A. Samochowiec, "Parental attitudes and aggression in the Emo subculture," Psychiatria polska, **50**(1), 19–28, 2016, doi: 10.12740/PP/36316.

[31] M.A. Hughes, S.F. Knowles, K. Dhingra, H.L. Nicholson, P.J. Taylor, "This corrosion: A systematic review of the association between alternative subcultures and the risk of self-harm and suicide," British journal of clinical psychology, **57**(4), 491–513, 2018, doi: 10.1111/bjc.12179.

[32] J. Lozon, M. Bensimon, "Music misuse: a review of the personal and collective roles of problem music," Aggression and Violent Behavior, **19**(3), 207–218, 2014, doi: 10.1016/j.avb.2014.04.003.

[33] A.C. North, D.J. Hargreaves, "Problem music and self-harming," Suicide and life-threatening behavior, **36**(5), 582–590, 2006, doi: 10.1521/suli.2006.36.5.582

[34] E. Bodner, M. Bensimon, "Problem music and its different shades over its fans," Psychology of Music, **43**(5), 641–660, 2015, doi: 10.1177/0305735614532000

[35] J. Kneer, D. Rieger, "The memory remains: How heavy metal fans buffer against the fear of death," Psychology of Popular Media Culture, **5**(3), 258, 2016, doi: 10.1037/ppm0000072.

[36] L. Meono, Using music-based interventions with adolescents coping with family conflict or parental divorce: A resource manual, Ph.D. Thesis, Pepperdine University, 2015.

[37] R. Hill, Emo saved my life: challenging the mainstream discourse of mental illness around My Chemical Romance, in Can I Play With Madness? Metal, Dissonance, Madness and Alienation (143–153), Inter-Disciplinary Press, Oxford, 2011.

[38] A.J. van den Tol, "The appeal of sad music: A brief overview of current directions in research on motivations for listening to sad music," The Arts in Psychotherapy, **49**, 44–49, 2016, doi: 10.1016/j.aip.2016.05.008.

[39] A.J. van den Tol, J. Edwards, N.A. Heflick, "Sad music as a means for acceptance-based coping," Musicae Scientiae, **20**(1), 68–83, 2016, doi: 10.1177/1029864915627844.

[40] T. Kelley, L. Simon, Everybody hurts: An essential guide to Emo culture. Harper Collins, 2007.

[41] R.A. Tanjung, The Commodified Emo Subculture In America Through The Used's Music Videos, Ph.D. thesis, Sebelas Maret University, 2011.

[42] E.H. Mereish, H.S. N'cho, C.E. Green, M.M. Jernigan, J.E. Helms, "Discrimination and depressive symptoms among Black American men: Moderated-mediation effects of ethnicity and self-esteem," Behavioral Medicine, **42**(3), 190–196, 2016, doi: 10.1080/08964289.2016.1150804.

[43] T. Yip, "The effects of ethnic/racial discrimination and sleep quality on depressive symptoms and self-esteem trajectories among diverse adolescents," Journal of youth and adolescence, **44**(2), 419–430, 2015, doi: 10.1007/s10964-014-0123-x.

[44] C.D. Presseau, Racial Discrimination and Mental Health for Transracially Adopted Adults: The Roles of Racial Identity and Racial Socialization, Ph.D. thesis, Lehigh University, 2016.

[45] S.M. Schires, N.T. Buchanan, R.M. Lee, M. McGue, W.G. Iacono, S.A. Burt, "Discrimination and ethnic-racial socialization among youth adopted from South Korea into white American families," Child development, **91**(1), e42–

e58, 2020, doi: 10.1111/cdev.13167.
[46] M. Banerjee, J.S. Eccles, Perceived racial discrimination as a context for parenting in African American and European American youth, in H.E. Fitzgerald, D.J. Johnson, D.B. Qin, F.A. Villarruel, and J. Norder (Eds.), Handbook of Children and Prejudice (233–247), Springer, 2019.
[47] L.A. Leslie, J.R. Smith, K.M. Hrapczynski, D. Riley, "Racial socialization in transracial adoptive families: Does it help adolescents deal with discrimination stress?," Family Relations, **62**(1), 72–81, 2013.
[48] M. Svetaz, "Tips to help teen patients deal with discrimination," Family practice management, **23**(4), 44, 2016.
[49] E. Strauss, Early adolescent boys' perceptions of the Emo youth subculture, Ph.D. Thesis, North-West University, 2012.
[50] J. Phua, "Consumption of sports team-related media: Its influence on sports fan identity salience and self-esteem," in Annual Meeting of the International Communication Association **21**, 2008.
[51] Y.C. Rhee, J. Wong, Y. Kim, "Becoming sport fans: Relative deprivation and social identity," International Journal of Business Administration, **8**(1), 118–134, 2017, doi: 10.5430/ijba.v8n1p118.
[52] U. Orth, R.W. Robins, K.F. Widaman, "Life-span development of self-esteem and its effects on important life outcomes," Journal of personality and social psychology, **102**(6), 1271–1288, 2012, doi: 10.1037/a0025558.
[53] P. Arunrangsiwed, C.S. Beck, "Positive and Negative Lights of Fan Culture," Suan Sunandha Rajabhat University Journal of Management Science, **3**(2), 40–58, 2016.
[54] H. Jenkins, Textual poachers: Television fans and participatory culture, Routledge, 2012, ISBN: 9780203114339.
[55] D.L. Wann, J. Hackathorn, M.R. Sherman, "Testing the team identification–social psychological health model: Mediational relationships among team identification, sport fandom, sense of belonging, and meaning in life," Group Dynamics: Theory, Research, and Practice, **21**(2), 94–107, 2017, doi: 10.1037/gdn0000066.
[56] R.E.S. Breen, We Mean It Maaan: Deconstructing Authenticity in the Punk and Metal Discourse, Master's thesis, Wesleyan University, 2017.
[57] M. Corciolani, "How do authenticity dramas develop? An analysis of Afterhours fans' responses to the band's participation in the Sanremo music festival," Marketing Theory, **14**(2), 2014, doi: 10.1177/1470593114521454.
[58] E.G. Armstrong, "Eminem's construction of authenticity," Popular Music and Society, **27**(3), 335–355, 2004, doi: 10.1080/0300776041000173317O.
[59] J.D. Campbell, P.D. Trapnell, S.J. Heine, I.M. Katz, L.F. Lavallee, D.R. Lehman, "Self-concept clarity: Measurement, personality correlates, and cultural boundaries," Journal of personality and social psychology, **70**(1), 141–156, 1996.
[60] J. Wu, D. Watkins, J. Hattie, "Self-concept clarity: A longitudinal study of Hong Kong adolescents," Personality and Individual Differences, **48**(3), 277–282, 2010, doi: 10.1016/j.paid.2009.10.011.
[61] P. Arunrangsiwed, P. Lekyan, "Being Unique does not make an Identity Conflict: A Narcissistic Fan of Mayday Parade," in The 6th International Conference on Management Science, Innovation and Technology, 2019.
[62] B.R. Johnson, G. Duwe, M. Hallett, J. Hays, S.J. Jang, K.R. Kerley, Faith and Service: Pathways to Identity Transformation and Correctional Reform. in K. R. Kerley (Ed.), Finding Freedom in Confinement: The Role of Religion in Prison Life (pp. 3–23), ABC-CLIO, 2018.
[63] N.L. Young, "Identity trans/formations," Annals of the International Communication Association, **31**(1), 250–299, 2007, doi: 10.1080/23808985.2007.11679068.
[64] P. Arunrangsiwed, K. Utapao, P. Bunyapukkna, K. Cheachainart, N. Ounpipat, "Emo Myth: 10-Year Follow-Up Stereotype Test of Emo Teens in 2000s," in International Conference on Innovation, Smart Culture and Well-Beings (ICISW2018), Bangkok, Thailand, 2018.
[65] J. Bronstein, "Like me! Analyzing the 2012 presidential candidates' Facebook pages," Online Information Review, **37**(2), 173-192, 2013, doi: 10.1108/OIR-01-2013-0002.
[66] N.R. Raselekoane, T.J. Mudau, P.P. Tsorai, "Gender differences in cyber-bullying among first-year University of Venda students," Gender and Behaviour, **17**(3), 13848-13857, 2019, ISSN: 1596-9231.
[67] L. Bailey, "Parental Involvement and Student Achievement: A Review of the Literature," Culminating Experience Action Research Projects, **6**, 19-44, 2004, doi: 10.1007/978-3-319-28064-6_2.
[68] J.F. Sowislo, U. Orth, "Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies," Psychological bulletin, **139**(1), 213–240, 2013, doi: 10.1037/a0028931.
[69] K. Bahoric, E. Swaggerty, "Fanfiction: Exploring in-and out-of-school literacy practices," Colorado Reading Journal, **26**, 25–31, 2015.
[70] A.A. Gardner, C.A. Lambert, "Examining the interplay of self-esteem, trait-emotional intelligence, and age with depression across adolescence," Journal of adolescence, **71**, 162–166, 2019, doi: 10.1016/j.adolescence.2019.01.008.
[71] P. Arunrangsiwed, R. Komolsevin, C.S. Beck, "Fan Activity as Tool to Improve Learning Motivation," Suan Sunandha Rajabhat University Journal of Management Science, **4**(2), 16–32, 2017.
[72] V.C. Burke, Email is Alive: How to Communicate with Graduate College Students, Doctoral thesis, UNLV, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 642-651 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Performance Analysis and Enhancement of Spline Adaptive Filtering based on Adaptive Step-size Variable Leaky Least Mean Square Algorithm

Sethakarn Prongnuch[*,1], Suchada Sitjongsataporn[2]

[1]*Department of Computer Engineering, Faculty of Industrial Technology, Suan Sunandha Rajabhat University Bangkok, 10300, Thailand*

[2]*Department of Electronic Engineering, Mahanakorn Institute of Innovation (MII), Faculty of Engineering, Mahanakorn University of Technology, Bangkok, 10530, Thailand*

ARTICLE INFO

Article history:
Received: 28 September, 2020
Accepted: 11 November, 2020
Online: 24 November, 2020

Keywords:
Spline adaptive filter
Variable leaky
Nonlinear filter

ABSTRACT

This paper presents an adaptive step-size and variable leaky least mean square algorithm based on nonlinear adaptive filter with the adaptive lookup table using spline interpolation. An adaptive step-size approach is proposed with the energy of squared previous and present errors to boost up the convergence rate. A modified variable leaky mechanism is proposed with the optimal leaky parameter by using the recursion form. The proposed algorithm merges an adaptive step-size and a modified variable leaky method with least mean square algorithm for linear and nonlinear network part of spline adaptive filtering in term of fast convergence enhancement. Experimental results demonstrate that proposed algorithm can notably achieve a competitive performance on the convergence rate in comparison with the conventional least mean square algorithm for spline adaptive filter. Simulation results suggest that mean square error performance of proposed algorithm can be partially assessed using adaptive step-size with the variable leaky parameters indicating better than the conventional least mean square algorithm by 16.76%.

1 Introduction

This paper is an extension of work originally presented in the International Electrical Engineering Congress [1], which has been proposed the variable leaky mechanism based on the least mean square (LMS) algorithm for spline adaptive filter. The combination of a linear finite impulse response (FIR) filter and a nonlinear adaptive lookup table (LUT) based on the spline interpolation is called the spline adaptive filtering (SAF) with the adaptation process [2]. It is a class of nonlinear adaptive filtering with a spline function [1]-[12].

According to the practical models, the linear adaptive filter should be inadequate [3]. The models of nonlinear dynamic systems have demonstrated with more robust performance. Linear adaptive filtering should be insufficient in real-world models. Moreover, many dynamic systems using model of nonlinear structure have been extended to the operating model.

Based on the SAF, many research works have achieved efficiently in the practical system identification. In [2], the authors have presented the sign approach with the normalized version of least mean square (NLMS) based on Wiener spline adaptive filter in order to fight against the impulsive noise environment. In [4], the authors have analyzed the convergence and stability analysis of SAF based on LMS algorithm. A steady-state performance of SAF has been examined in [5].

In order to model the nonlinear system identification, the SAF is more attention for the practical use [6],[7]. The linear time-invariant model with the cubic spline function [8] can obtain the good performance working with the adaptive lookup table (LUT) [9] and the spline basis matrix on the adaptive control points coefficient vector.

Applications of SAF architecture have been applied in the infinite impulse response [10] and the system environment with impulsive noise [11]. A set-membership mechanism with the Wiener spline adaptive filtering and normalized least M-estimate algorithm have been obtained significantly the fast convergence in the environment of impulsive noise [11]. In [12], the authors have proposed the Hammerstein function with SAF based on LMS algorithm for improving the convergence rate. Further, a SAF based on the maximum correntropy criterion [13], [14] has been discussed that correntropy is robustness to non-Gaussian noise.

Least mean square (LMS) algorithm which is simple and low computation has been widely used [15]. Most researchers have modified accurately on LMS and NLMS algorithms [16] and [17], the variable step-size LMS [18], the sign mechanism with NLMS [2] and so on.

[*]Corresponding Author: Sethakarn Prongnuch, Suan Sunandha Rajabhat University, Bangkok, Thailand, sethakarn.pr@ssru.ac.th

As stated in a steady-state of convergence analysis, a variable leaky mechanism has been depicted in use of tracking the systems. In [19], a variable leaky LMS algorithm based on the greedy heuristic algorithm has been leveraged to enhance with the high eigenvalue input. In [20], the authors have implemented the variable leaky LMS in the field programmable gate array implementation. The results have been shown that the algorithm can explore the various digital filter architecture.

Based on the variable leaky LMS algorithm which has been orchestrated against and attenuating the drifting [21]. In the field of adaptive signal processing, a variable leaky based on the orthogonal gradient adaptive algorithm with the minimized cost function has been applied for the wireless communications [22].

The novelty of this paper is to merge an adaptive step-size and a modified variable leaky method with least mean square algorithm for linear and nonlinear network part of spline adaptive filtering in term of fast convergence enhancement. The analysis of the convergence performance of the proposed algorithm is derived in forms of the properties and mean square error performance.

In this paper, we organize this paper as follows. Section 2 explains shortly in the structure of SAF. Next, Section 3 introduces the modified adaptive step-size mechanism and the modified variable leaky criterion on the minimization cost function based on the LMS algorithm. Section 4 describes the performance analysis of the proposed adaptive step-size and variable leaky LMS algorithm in terms of the properties and the mean square error performance. Further, the experiment results and discussion show in Section 5 and Section 6, respectively. Finally, Section 7 summarizes the proposed algorithm.

2 Spline Adaptive Filter

Spline adaptive filter (SAF) is namely a combination of linear and nonlinear structures shown in Fig. 1, where $\mathbf{x}_k$ is the input of SAF structure and y_k is the output of system.

The objective is that adaptive lookup table in the nonlinear structure generates an output of SAF y_k nearly to a desired sequence d_k as

$$d_k = y_k + e_k \,, \tag{1}$$

where the error of system e_k should be small. Thus, the adaptive FIR filter brings an output ϕ_k in the linear structure, while the input signal is a sequence of x_k at the linear structure.

An adaptive FIR filter output ϕ_k is given by [3]

$$\phi_k = \mathbf{w}_k^T \, \mathbf{x}_k \,, \tag{2}$$

where $\mathbf{w}_k$ is the update FIR coefficient vector.

Following [3], the SAF output y_k can be examined by following

$$y_k = \mathbf{u}_k^T \, \mathbf{C}_b \, \mathbf{g}_{m,k} \,, \tag{3}$$

$$\mathbf{v}_k = [v_k^3, \; v_k^2, \; v_k, \; 1]^T \,, \tag{4}$$

where $\mathbf{g}_{m_k}$ is the coefficient of control point vector and $\mathbf{C}_b$ is a spline basis matrix.

The local parameter v_k and index m are usually indicated as [5]

$$v_k = \frac{\phi_k}{\Delta x} - \left\lfloor \frac{\phi_k}{\Delta x} \right\rfloor \,, \tag{5}$$

$$m \; = \left\lfloor \frac{\phi_k}{\Delta x} \right\rfloor + \frac{P-1}{2} \,, \tag{6}$$

where Δx is the uniform space between two adjacent-coefficients of control points [6], P is the size of control point coefficient, and the floor operator $\lfloor \cdot \rfloor$ is applied.

By using the minimized objective function based on least mean square (LMS) [5], it becomes

$$J_{\mathbf{w},\mathbf{g}}^{\mathrm{LMS}}(k) = \min_{\mathbf{w},\mathbf{g}} \left\{ \frac{1}{2} \mid e_k \mid^2 \right\} , \tag{7}$$

where *a priori* error e_k is defined as

$$e_k = d_k - y_k = d_k - \mathbf{v}_k^T \, \mathbf{C}_b \, \mathbf{g}_{m,k} \,. \tag{8}$$

The update FIR coefficient vector $\mathbf{w}_k$ can compute by the gradient vector of (7) with respect to the coefficients $\mathbf{w}_k$ in terms of the recursion form as [3]

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_w \, \nabla J_{\mathbf{w}}^{\mathrm{LMS}} \,, \tag{9}$$

where μ_w is a step-size and $\nabla J_{\mathbf{w}}^{\mathrm{LMS}}(k)$ is the gradient vector for $\mathbf{w}_k$.

Similarly, the update control points vector $\mathbf{g}_{m,k}$ can be expressed by the gradient vector of (7) with respect to the coefficients $\mathbf{g}_{m,k}$ [3]

$$\mathbf{g}_{m,k+1} = \mathbf{g}_{m,k} - \mu_g \, \nabla J_{\mathbf{g}}^{\mathrm{LMS}}(k) \,, \tag{10}$$

where μ_g is a step-size and $\nabla J_{\mathbf{g}}^{\mathrm{LMS}}(k)$ is the gradient vector for $\mathbf{g}_{m,k}$. Thus, the FIR weight vector $\mathbf{w}_k$ and the control points update coefficient vector $\mathbf{g}_{m,k}$ are the particularly simple forms as [5]

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mu_w \, \mathbf{v}_k' \, \mathbf{C}_b \, \mathbf{g}_{m,k}^T \, \mathbf{x}_k \, e_k \,, \tag{11}$$

$$\mathbf{g}_{m,k+1} = \mathbf{g}_{m,k} + \mu_g \, \mathbf{C}_b^T \, \mathbf{v}_k \, e_k \,, \tag{12}$$

where $\mathbf{v}_k'$ is given by

$$\mathbf{v}_k' = [3 \, v_k^2, \; 2 \, v_k, \; 1, \; 0]^T \,. \tag{13}$$

and the local parameter $\mathbf{v}_k$ is given in (5).

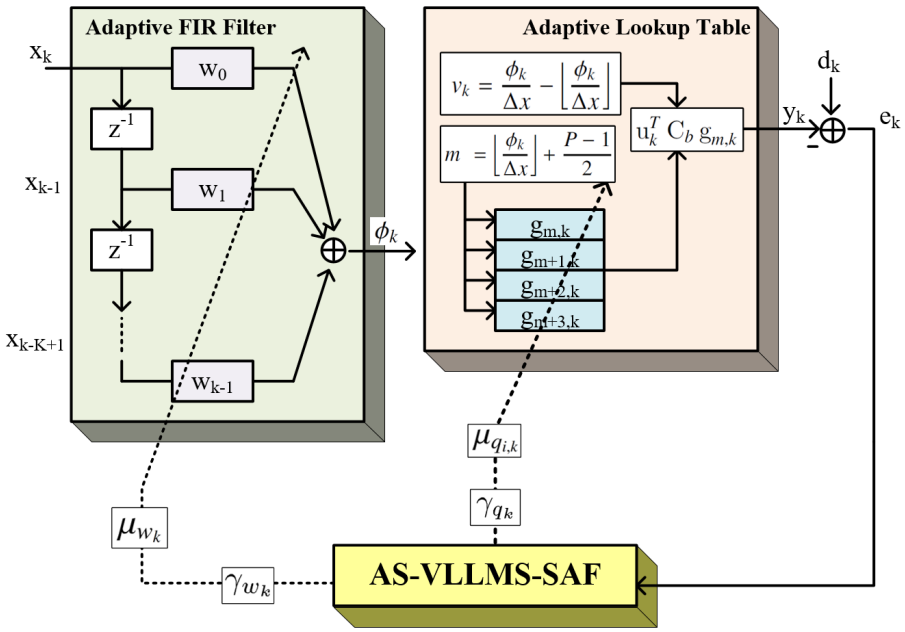


Figure 1: Linear-Nonlinear structure of proposed adaptive step-size variable leaky LMS algorithm for spline adaptive filter (AS-VLLMS-SAF).

3 Proposed Adaptive Step-size and Variable Leaky Least Mean Square based on Spline Adaptive Filter

Based on the least mean square (LMS) algorithm, the advantage of the leaky LMS algorithm is that it avoids the drift of weights [23]. Meanwhile, the step-size parameter is an efficient approach to improve the convergence rate [16].

Following [1] and [11], the cost function using adaptive step-size and variable leaky criterion for SAF based on least mean square (AS-VLLMS-SAF) algorithm can be minimized as

$$J(\mathbf{w}_k, \mathbf{g}_{m,k}) = \min_{\mathbf{w}_k, \mathbf{g}_{m,k}} \left\{ \frac{1}{2} \mid e_k \mid^2 + \gamma_w \|\mathbf{w}_k\|^2 + \gamma_g \|\mathbf{g}_k\|^2 \right\}, \tag{14}$$

where γ_w and γ_g are the leaky paramters for the linear FIR coefficient vector $\mathbf{w}_k$ and the control points weight vector $\mathbf{g}_{m,k}$, respectively. The *a priori* error of system e_k is given as

$$e_k = d_k - \mathbf{v}_k^T \, \mathbf{C}_b \, \mathbf{g}_{m,k} \, . \tag{15}$$

Considering the chain rule on the cost function in (14) with respect to (w.r.t) $\mathbf{w}_k$, we get

$$\begin{aligned} \frac{\partial J(\mathbf{w}_k, \mathbf{g}_{m,k})}{\partial \mathbf{w}_k} &= \left\{ - e_k \frac{\partial y_k}{\partial \mathbf{v}_k} \frac{\partial \mathbf{v}_k}{\partial \phi_k} \frac{\partial \phi_k}{\partial \mathbf{w}_k} + \gamma_w \mathbf{w}_k \right\} \\ &= \left\{ - \frac{e_k}{\Delta x} \mathbf{v}'_k \, \mathbf{C} \, \mathbf{g}_{m,k} \, \mathbf{x}_k + \gamma_w \mathbf{w}_k \right\}, \end{aligned} \tag{16}$$

where $\mathbf{v}'_k$ is defined as

$$\mathbf{v}'_k = [3v_k^2 \;\; 2v_k \;\; 1 \;\; 0] \, . \tag{17}$$

The gradient of cost function in (14) w.r.t $\mathbf{g}_{m,k}$ can be expressed by using the chain rule in the form of vector below

$$\begin{aligned} \frac{\partial J(\mathbf{w}_k, \mathbf{g}_{m,k})}{\partial \mathbf{g}_{m,k}} &= \left\{ - e_k \frac{\partial y_k}{\partial \mathbf{v}_k} \frac{\partial \mathbf{v}_k}{\partial \phi_k} \frac{\partial \phi_k}{\partial \mathbf{g}_{m,k}} + \gamma_g \mathbf{g}_{m,k} \right\} \\ &= \{ - e_k \, \mathbf{C}^T \, \mathbf{v}_k + \gamma_g \mathbf{g}_{m,k} \} \, . \end{aligned} \tag{18}$$

Hence, the proposed update linear FIR coefficient vector $\mathbf{w}_k$ of AS-VLLMS-SAF algorithm is the stochastic adaptation formula as

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_{w_k} \frac{\partial J(\mathbf{w}_k, \mathbf{g}_{m,k})}{\partial \mathbf{w}_k} \, , \tag{19}$$

where μ_{w_k} is a adaptive step-size parameter at symbol k.

By substituting (16) into (19), the proposed update tap-weight vector $\mathbf{w}_k$ can be expressed as

$$\therefore \mathbf{w}_{k+1} = (\mathbf{1} - \mu_{w_k} \gamma_{w_k}) \, \mathbf{w}_k + \frac{\mu_{w_k}}{\Delta x} \, \mathbf{v}'_k \, \mathbf{C}_b \, \mathbf{g}_{m,k} \, \mathbf{x}_k \, e_k \, , \tag{20}$$

where γ_{w_k} is a variable leaky parameter for $\mathbf{w}_k$. It is noticed that $(1 - \mu_{w_k} \gamma_{w_k})$ is also defined as the leakage factor [23] for updated weight $\mathbf{w}_k$ in which its value is generally close to 1.

The adaptive control points coefficient vector $\mathbf{g}_{m,k}$ of AS-VLLMS-SAF becomes

$$\mathbf{g}_{m,k+1} = \mathbf{g}_{m,k} - \mu_{\mathbf{g}_k} \frac{\partial J(\mathbf{w}_k, \mathbf{g}_{m,k})}{\partial \mathbf{g}_{m,k}} \, , \tag{21}$$

By substituting (18) into (21), the proposed control points vector $\mathbf{g}_{m,k}$ is given as

$$\therefore \mathbf{g}_{m,k+1} = (\mathbf{1} - \mu_g \gamma_{g_k})\, \mathbf{g}_{m,k} + \mu_q\, \mathbf{C}_b^T\, \mathbf{v}_k\, e_k\ , \tag{22}$$

where μ_g is a step-size parameter and γ_{g_k} is a variable leaky parameter for $\mathbf{g}_{m,k+1}$. It is noted that $(1 - \mu_g \gamma_{g_k})$ is also defined as the leakage factor [23] for updated weight $\mathbf{g}_{m,k+1}$ in which its value is generally close to 1.

3.1 Modified Variable Leaky mechanism

Following [24], we introduce the modified variable leaky algorithm with the optimal leaky parameter for weights $\mathbf{w}_k$ and $\mathbf{g}_{m,k}$ in the recursion form as

$$\gamma_{w_k} = \gamma_{w_{k-1}} + \rho_w\, \gamma_w^{\text{opt}}\ , \tag{23}$$

$$\gamma_{g_k} = \gamma_{g_{k-1}} + \rho_g\, \gamma_g^{\text{opt}}\ , \tag{24}$$

where ρ_w, ρ_g and γ_w^{opt}, γ_g^{opt} are the adaptation rate and the optimal leaky parameters of weights $\mathbf{w}_k$ and $\mathbf{g}_{m,k}$.

We rewrite (20) as

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_{w_k} \gamma_{w_k} \mathbf{w}_k + \frac{\mu_{w_k}}{\Delta x}\, \mathbf{v}_k'\, \mathbf{C}_b\, \mathbf{g}_{m,k}\, \mathbf{x}_k\, e_k\ , \tag{25}$$

Assumption 1: *We consider the steady-state value of* $E\{\boldsymbol{w}_{k+1}\}$ *for* $k \to \infty$ *by*

$$E\{\mathbf{w}_{k+1}\} \simeq E\{\mathbf{w}_k\}.$$

We determine the optimal leaky parameter for $\mathbf{w}_k$ by using this assumption above in (25), we arrive at

$$\gamma_w^{\text{opt}} \simeq \frac{\mathbf{v}_k'\, \mathbf{C}_b^T\, \mathbf{g}_{m,k}\, \mathbf{w}_k^{-1}\, \mathbf{x}_k\, e_k}{\Delta x}\ . \tag{26}$$

Therefore, the modified variable leaky algorithm in the recursion form can be expressed as

$$\gamma_{w_k} = \gamma_{w_{k-1}} - \rho_w\, \frac{\mathbf{v}_k'\, \mathbf{C}_b^T\, \mathbf{g}_{m,k}\, \mathbf{w}_k^{-1}\, \mathbf{x}_k\, e_k}{\Delta x}\ , \tag{27}$$

where ρ_w is an adaptation rate for $\mathbf{w}_k$.

Consequently, the *a posterior* AS-VLLMS error [1] is rewritten as

$$e_{p_k}^{\text{VL-LMS}} = (1 - \mu_{g_k}\, \mathbf{v}_k\, \mathbf{C}_b^T\, \mathbf{v}_k\, \mathbf{C}_b)\, e_k + \mu_{g_k}\, \gamma_{g_k}\, \mathbf{v}_k\, \mathbf{C}_b^T\, \mathbf{g}_{m,k}\ . \tag{28}$$

Assumption 2: *We consider the steady-state value of proposed algorithm is stable for* $k \to \infty$ *by*

$$E\{e_{p_k}^{\text{VL-LMS}}\} \simeq E\{e_k\}.$$

By using Assumption 2 in (28), the optimal leaky parameter can be defined as

$$\gamma_g^{\text{opt}} \simeq \frac{\Omega_k^T\, \Omega_k\, e_k}{\Omega_k^T\, \mathbf{g}_{m,k}}\ , \tag{29}$$

where $\Omega_k = \mathbf{v}_k^T\, \mathbf{C}$.

Table 1: Proposed Variable Leaky Least Mean Square based on Spline Adaptive Filtering with the adaptive step-size (AS-VLLMS-SAF) algorithm.

Initialize : $\mathbf{w}(0) = \delta\, \mathbf{I}$ and $\mathbf{I}$ is the identity matrix.

$$\mathbf{w}_k = [\ w_0\ w_1\ \ldots\ w_{K-1}\ ]$$
$$\mathbf{x}_k = [\ x_k\ x_{k-1}\ \ldots\ x_{k-K+1}\ ]$$
$$\mathbf{g}_{m,k} = [\ g_{m,k}\ \ g_{m+1,k}\ \ g_{m+2,k}\ \ g_{i+3,k}\ ]$$

for $k = 0, 1, 2, \ldots, K-1$.

$$\phi_k = \mathbf{w}_k^T\, \mathbf{x}_k$$
$$v_k = \frac{\phi_k}{\Delta x} - \left\lfloor \frac{\phi_k}{\Delta x} \right\rfloor$$
$$m\ = \left\lfloor \frac{\phi_k}{\Delta x} \right\rfloor + \frac{P-1}{2}$$
$$\mathbf{v}_k = [v_k^3,\ v_k^2,\ v_k,\ 1]^T$$
$$\mathbf{v}_k' = [3v_k^2,\ 2v_k,\ 1,\ 0]^T$$
$$e_k = d_k - \mathbf{v}_k^T\, \mathbf{C}_b\, \mathbf{g}_{m,k}$$
$$\gamma_{w_k} = \gamma_{w_{k-1}} + \rho_w\, \frac{\mathbf{v}_k'\, \mathbf{C}_b^T\, \mathbf{g}_{m,k}\, \mathbf{w}_k^{-1}\, \mathbf{x}_k\, e_k}{\Delta x}$$
$$\gamma_{q_k} = \gamma_{q_{k-1}} + \rho_q\, \frac{\Omega_k^T\, \Omega_k\, e_k}{\Omega_k^T\, \mathbf{g}_{m,k} + \epsilon}$$
$$\Omega_k = \mathbf{v}_k^T\, \mathbf{C}_b$$
$$\mu_{w_k} = \alpha_w\, \mu_{w_{k-1}} + \beta_w\, \xi_k^2$$
$$\xi_k = \lambda\, \xi_{k-1} + (1-\lambda)\, e_{k-1}^*\, e_k$$
$$\mu_{q_k} = \alpha_q\, \mu_{q_{k-1}} + \beta_q\, e_{k-1}^*\, e_k$$
$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_{w_k} \gamma_{w_k} \mathbf{w}_k + \frac{\mu_{w_k}}{\Delta x}\, \mathbf{v}_k'\, \mathbf{C}_b^T\, \mathbf{g}_{m,k}\, \mathbf{x}_k\, e_k$$
$$\mathbf{q}_{i,k+1} = \mathbf{g}_{m,k} - \mu_{q_k} \gamma_{q_k} \mathbf{g}_{m,k} + \mu_{q_{m,k}}\, \mathbf{C}_b^T\, \mathbf{v}_k\, e_k$$

end

Therefore, the modified variable leaky algorithm can be expressed in the recursion method as

$$\gamma_{g_k} = \gamma_{g_{k-1}} - \rho_q\, \frac{\Omega_k^T\, \Omega_k\, e_k}{\Omega_k^T\, \mathbf{g}_{m,k} + \epsilon}\ , \tag{30}$$

where ρ_g is an adaptation rate for $\mathbf{g}_{m,k}$ and ϵ is a regularization parameter with a small constant.

3.2 Modified Adaptive Step-size Approach

According to the better convergence of proposed AS-VLLMS-SAF algorithm, the step-size parameter should be adapted recursively.

The behavior of convergence rate is which the algorithm starts converging, the value of step-size parameter should be large in order to boost up the learning rate of convergence. At closely to the steady-state, the step-size parameter might be decreased adequately to get the lower adjustment. This leads to adjust accordingly the step-size parameter.

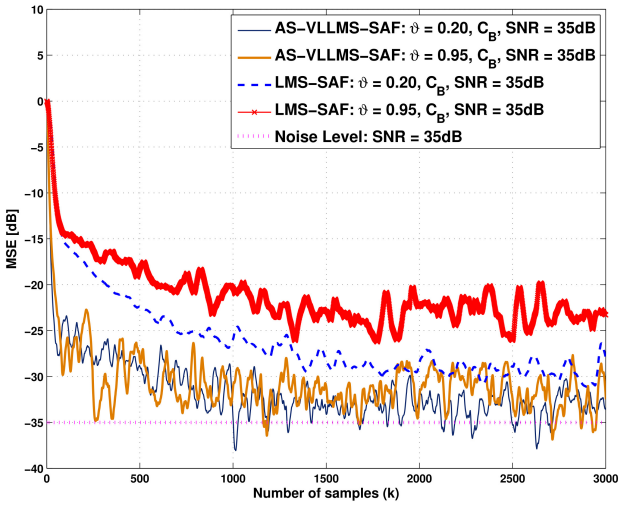


Figure 2: MSE of proposed AS-VLLMS-SAF and LMS-SAF [3] on the *B-spline* matrix $\mathbf{C}_B$ with the different $\vartheta = 0.20, 0.95$ and SNR = 35dB.

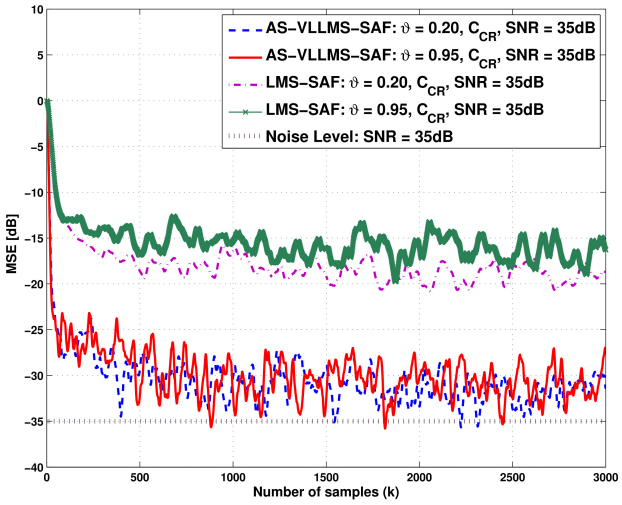


Figure 3: MSE of proposed AS-VLLMS-SAF and LMS-SAF [3] on the *Catmul-Rom* matrix $\mathbf{C}_{CR}$ with $\vartheta = 0.20, 0.95$ and SNR=35dB.

We assume that the step-size parameters μ_{w_k} and μ_{g_k} are regulated by squared estimated error. Hence, the proposed adaptive step-size μ_{w_k} for weight $\mathbf{w}_k$ based on AS-VLLMS algorithm is introduced for spline adaptive filtering as [4]

$$\mu_{w_k} = \alpha_w \mu_{w_{k-1}} + \beta_w \xi_k^2 , \tag{31}$$

$$\xi_k = \lambda \xi_{k-1} + (1-\lambda) e_{k-1}^* e_k , \tag{32}$$

and the proposed adaptive step-size μ_{g_k} for weight $\mathbf{g}_{m,k}$ can be expressed as

$$\mu_{g_k} = \alpha_g \mu_{g_{k-1}} + \beta_g e_{k-1}^* e_k , \tag{33}$$

A summary of proposed variable leaky least mean square for spline adaptive filtering based on adaptive step-size algorithm (AS-VLLMS) is depicted in Table 1.

4 Performance Analysis

In this section, we analyze the convergence performance of the proposed variable leaky LMS based on SAF. Compared to the LMS-SAF algorithm, the proposed AS-VLLMS algorithm is presented by introducing an additional concept.

Variable leaky algorithm is used in the iterative of weights $\mathbf{w}_k$ and $\mathbf{g}_{m,k}$, respectively. So, the computational complexity of AS-VLLMS-SAF is more complex than that of the LMS-SAF algorithm.

4.1 Properties of AS-VLLMS Algorithm

The properties of adaptive step-size and variable leaky LMS algorithm may be derived by examining the behaviour of $E\{\mathbf{w}_k\}$ and $E\{\mathbf{g}_{m,k}\}$ of SAF.

Taking expectation value and using the independence assumption [25], we then have

$$E\{\mathbf{w}_{k+1}\} = (1 - \mu_{w_k}\gamma_{w_k})\, E\{\mathbf{w}_k\} + \frac{\mu_w}{\Delta x} E\{\mathbf{v}_k' \mathbf{C}_b \mathbf{g}_{m,k} \mathbf{x}_k e_k\} . \tag{34}$$

Specially, we note that if $\mathbf{w}_k$ converges to the steady-state solution, we get

$$\lim_{k\to\infty} E\{\mathbf{w}_k\} = \frac{\mathbf{v}_k' \mathbf{C}_b \mathbf{g}_{m,k}^T \mathbf{x}_k e_k}{\gamma_{g_k} \Delta x} . \tag{35}$$

In a similar fashion, we substitute *a priori* error e_k in (15) into (22), then we have

$$\mathbf{g}_{m,k+1} = \left[\mathbf{I} - \mu_{g_k} \{\gamma_{g_k} \mathbf{I} + \mho_k^T \mho_k\}\right] \mathbf{g}_{m,k} + \mu_{g_k} \mho_k d_k , \tag{36}$$

where $\mho_k$ is given by

$$\mho_k = \mathbf{C}_b^T \mathbf{v}_k . \tag{37}$$

Thus, we take the expectation value and using the independence assumption [25], it becomes

$$E\{\mathbf{g}_{m,k+1}\} = \left[\mathbf{I} - \mu_{g_k}\left\{\mathbb{R}_\mho + \gamma_{g_k} \mathbf{I}\right\}\right] E\{\mathbf{g}_{m,k}\} + \mu_{g_k} \mathbf{r}_{d\mho} , \tag{38}$$

where $\mathbb{R}_\mho$ and $\mathbf{r}_{d\mho}$ are defined by

$$\mathbb{R}_\mho = \mho_k^T \mho_k , \tag{39}$$

$$\mathbf{r}_{d\mho} = d_k \mho_k . \tag{40}$$

Clearly, we note that if $\mathbf{g}_{m,k}$ converges to the steady-state solution, we have

$$\lim_{k\to\infty} E\{\mathbf{g}_{m,k}\} = \left(\mathbb{R}_\mho + \gamma_{g_k} \mathbf{I}\right)^{-1} \mathbf{r}_{d\mho} . \tag{41}$$

It is noticed that the leakage coefficient introduces a bias into the steady-state solution.

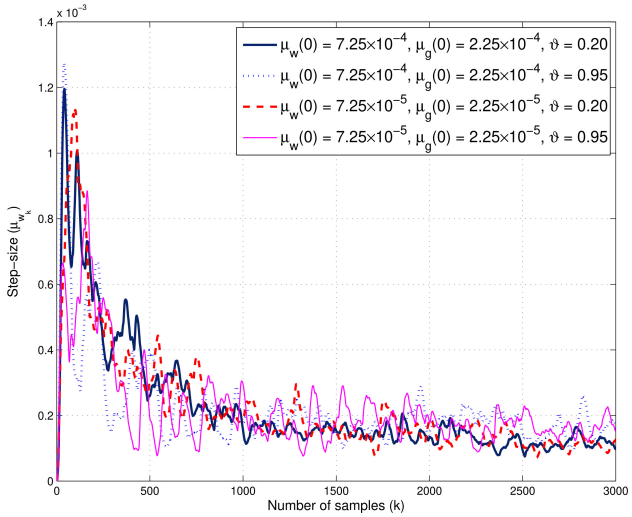


Figure 4: Learning curves of adaptive step-size μ_{w_k} of coefficient vector $\mathbf{w}_k$ of proposed AS-VLLMS-SAF on the *B-spline* matrix $\mathbf{C}_B$ with $\alpha = 0.20, 0.95$ and SNR = 35dB.

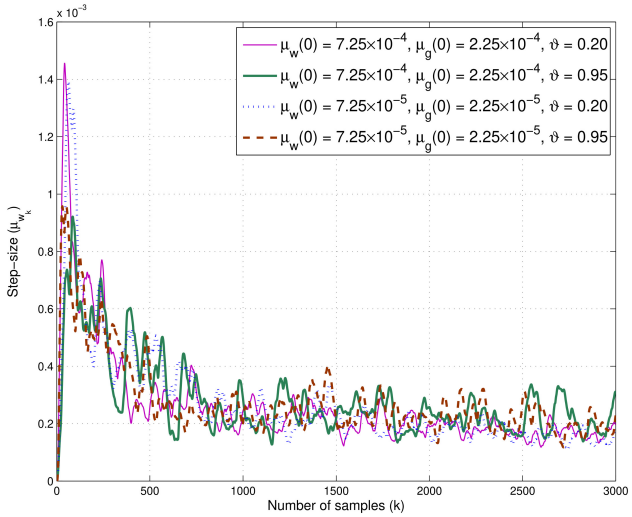


Figure 6: Learning curves of adaptive step-size μ_{w_k} of coefficient vector $\mathbf{w}_k$ of proposed AS-VLLMS-SAF on the *Catmul-Rom* matrix $\mathbf{C}_{CR}$ with $\alpha = 0.20, 0.95$ and SNR = 35dB.

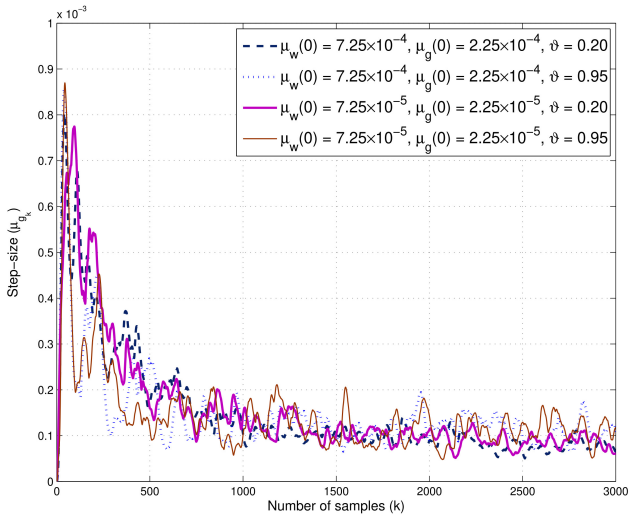


Figure 5: Learning curves of adaptive step-size μ_{g_k} of coefficient vector $\mathbf{g}_{m,k}$ of proposed AS-VLLMS-SAF on the *B-spline* matrix $\mathbf{C}_B$ with $\alpha = 0.20, 0.95$ and SNR = 35dB.

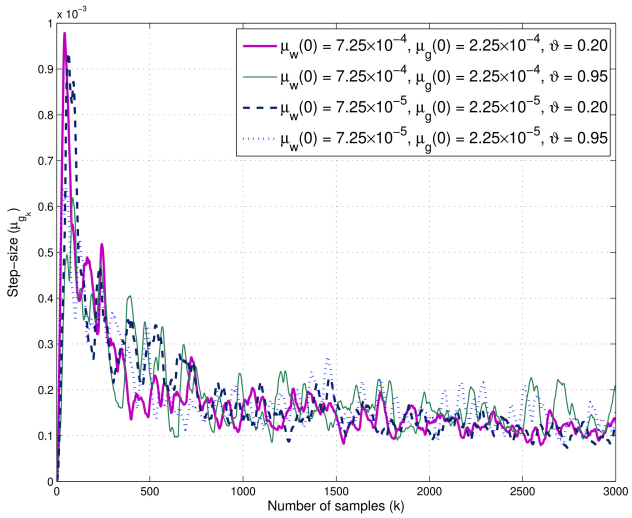


Figure 7: Learning curves of adaptive step-size μ_{g_k} of coefficient vector $\mathbf{g}_{m,k}$ of proposed AS-VLLMS-SAF on the *Catmul-Rom* spline matrix $\mathbf{C}_{CR}$ with $\alpha = 0.20, 0.95$ and SNR = 35dB.

4.2 *The Leaky Adjustment*

In this section, the leaky adjustment is investigated in form of the *a posterior* error compared with *a priori* error in the spline adaptive filtering. We determine the *a posteriori* LMS error as [1]

$$\tilde{e}_k^{LMS} \triangleq d_k - \mathbf{v}_k^T \, \mathbf{C}_b \, \mathbf{g}_{m,k+1} = e_k \left(1 - \mu_g \, \mho_k^T \, \mho_k \right), \qquad (42)$$

where $\mho_k$ is described in (37).

Correspondingly, we examine the *a posteriori* variable leaky least mean square (VL-LMS) error as

$$\begin{aligned} \tilde{e}_k^{VLLMS} &\triangleq d_k - \mathbf{v}_k^T \, \mathbf{C}_b \, \mathbf{g}_{m,k+1} \\ &= e_k(1 - \mu_{g_k} \mho_k^T \, \mho_k) + \mu_{g_k} \gamma_{g_k} \mho_k^T \mathbf{g}_{m,k} \, . \end{aligned}$$

It is noticed that the leaky adjustment will be achieve based on a greedy heuristic algorithm with each iteration, which can get the appropriate leaky value, if $|\tilde{e}_k^{VLLMS}| < |\tilde{e}_k^{LMS}|$. That means the proposed VL-LMS algorithm would allow to get the efficient LMS algorithm. Otherwise, the leak parameter should be diminished.

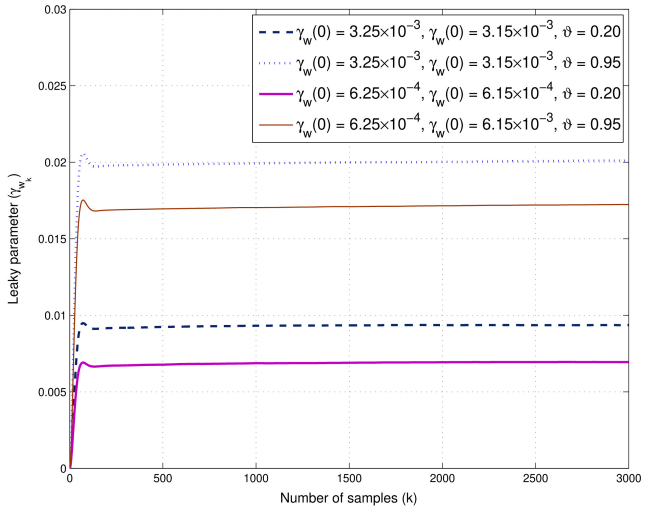


Figure 8: Learning curves of variable leaky γ_{w_k} of coefficient vector $\mathbf{w}_k$ of proposed AS-VLLMS-SAF on $\mathbf{C}_B$ with $\vartheta = 0.20, 0.95$ and SNR = 35dB.

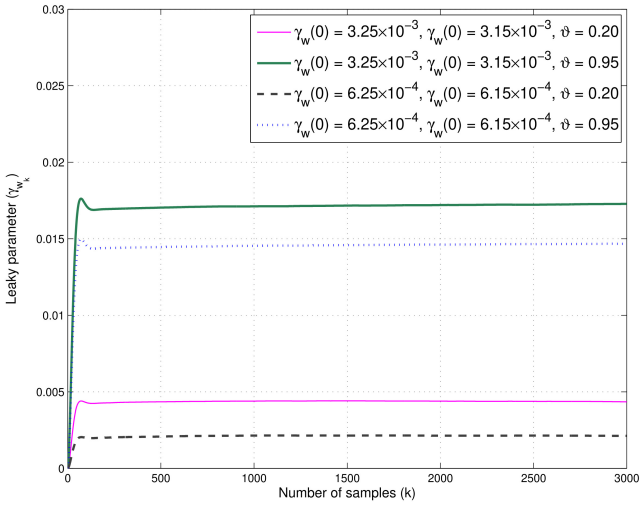


Figure 10: Learning curves of variable leaky γ_{w_k} of coefficient vector $\mathbf{w}_k$ of proposed AS-VLLMS-SAF on $\mathbf{C}_{CR}$ with $\vartheta = 0.20, 0.95$ and SNR = 35dB.

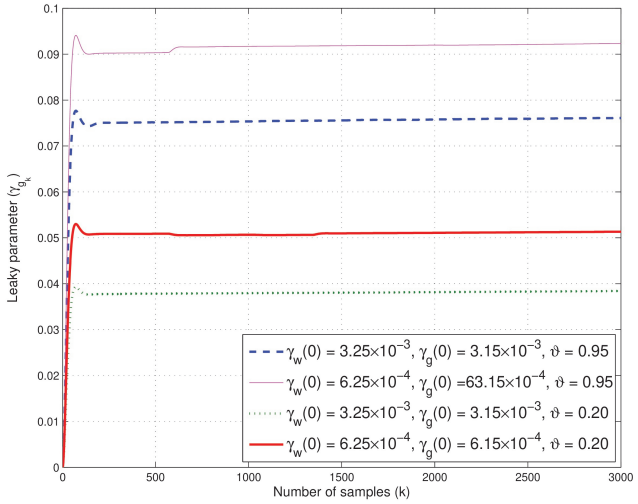


Figure 9: Learning curves of variable leaky γ_{g_k} of coefficient vector $\mathbf{g}_{m,k}$ of proposed AS-VLLMS-SAF on $\mathbf{C}_B$ with $\alpha = 0.20, 0.95$ and SNR = 35dB.

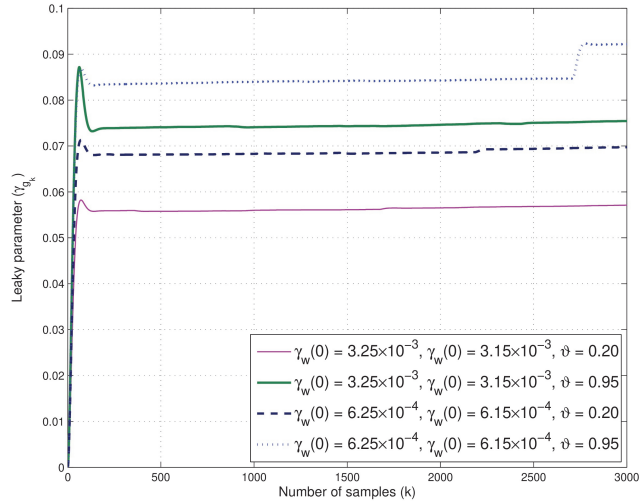


Figure 11: Learning curves of variable leaky γ_{w_k} of coefficient vector $\mathbf{w}_k$ of proposed AS-VLLMS-SAF on $\mathbf{C}_{CR}$ with $\alpha = 0.20, 0.95$ and SNR = 35dB.

4.3 Mean Square Error Performance

We investigate the mean square error (MSE) performance of the proposed AS-VLLMS-SAF algorithm at the steady-state.

Assumption 3: *We assume that the a priori and a posteriori optimal errors are identical, we have*

$$E\{\epsilon_g^{\text{opt}}\} \approx e_{\text{opt},g} \, .$$

Assumption 4: *We consider the convergence condition for $k \to \infty$, that is of*

$$E\{\epsilon_g^{\text{opt}}\} \to 0 \, , \text{ as } \; k \to \infty$$
$$E\{\mathbf{g}_{m,k}\} \to \mathbf{g}_{m,k}^{\text{opt}} \, , \text{ as } \; k \to \infty$$

Following [12], we decompose the MSE under these assumptions above as follows.

$$\mathbb{J}_{MSE}^{P} = \mathbb{J}_{MMSE}^{P} + \mathbb{J}_{EX}^{P} = E\{\|\epsilon_g\|^2\} \, , \tag{43}$$

where $\mathbb{J}_{MMSE}^{P}$ is the minimum mean square error (MMSE) as

$$\mathbb{J}_{MMSE}^{P} = E\{\|\epsilon_g^{\text{opt}}\|^2\} \, , \tag{44}$$
$$\epsilon_g^{\text{opt}} = d_k - \mathbf{v}_k^T \, \mathbf{C}_b \, \mathbf{g}_{\text{opt}} \, , \tag{45}$$

where ϵ_g^{opt} is the *a posteriori* optimal error of $\mathbf{g}_{\text{opt}}$.

Certainly, $\mathbb{J}_{EX}^{P}$ is the *a posteriori* excess mean square error (EMSE) given by

$$\mathbb{J}_{EX}^{P} = \mathbb{J}_{MSE}^{P} - \mathbb{J}_{MMSE}^{P} = E\{\|\epsilon_g\|^2\} - E\{\|\epsilon_g^{\text{opt}}\|^2\} \, . \tag{46}$$

4.4 Experimental conditions on parameters setting

In this section, the theoretical experiments are simulated in system identification over the random process and under the Gaussian noise scenario. We evaluate the performance of proposed adaptive step-size variable leaky least mean square algorithm based on spline adaptive filter (AS-VLLMS-SAF) as compared to the conventional least mean square algorithm based on spline adaptive filter (LMS-SAF) [3].

The coloured input signal generates for all experiments over 100 Monte Carlo trials and 5,000 samples used that is generated by [2]

$$x_k = \vartheta \cdot x_{k-1} + \sqrt{1-\vartheta^2}\zeta_k \,, \tag{47}$$

where ζ_k is a unitary variance type of zero mean white Gaussian noise, $\vartheta = [0, 0.95]$ and an interval sampling Δx is used at 0.2 [11].

An unknown Wiener system is composed by a linear component as [3]

$$\mathbf{w}_0 = [0.6,\ -0.4,\ 0.25,\ -0.15,\ 0.1]^T \,,$$

and a 23-point length lookup table (LUT) $\mathbf{g}_0$ is a nonlinear memoryless target function applied by

$$\mathbf{g}_0 = [\ -2.2,\ \ldots\ , -0.8,\ -0.91,\ -0.4,\ -0.2,\ 0.05,\ 0,\ -0.4,\ 1.0,\ 1.0,\ 1.2,\ \ldots\ , 2.2\ ] \,.$$

The linear filter $\mathbf{w}_0$ is initialzed as $\mathbf{w}_0 = [1,\ 0, \ldots,\ 0]$ and $\delta = 0.001$. Other parameters are fixed at the length of coefficient vector $T = 5$, a signal to noise ratio $SNR = 35$dB . The spline basis matrices are used as the *B-spline* matrix $\mathbf{C}_B$ and the *Catmul-Rom spline* matrix $\mathbf{C}_{CR}$ in [8] which are selected for simulation experiments as follows:

$$\mathbf{C}_B = \frac{1}{6}\begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -3 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix}, \tag{48}$$

$$\mathbf{C}_{CR} = \frac{1}{2}\begin{bmatrix} -1 & 3 & -3 & 1 \\ 2 & -5 & 4 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}. \tag{49}$$

The fixed parameters of proposed AS-VLLMS-SAF algorithm are as follows: $\alpha = 0.20, 0.95$, $\alpha_w = 7.55 \times 10^{-3}$, $\beta_w = 2.75 \times 10^{-3}$, $\alpha_g = 6.55 \times 10^{-3}$, $\beta_g = 1.85 \times 10^{-3}$, $\lambda = 0.97$, $\epsilon = 1 \times 10^{-6}$ and $\rho_w = 1.5 \times 10^{-6}$, $\rho_g = 1.125 \times 10^{-6}$. And the initial parameters for proposed AS-VLLMS-SAF are $\gamma_w(0) = 3.25 \times 10^{-2}$, $\gamma_g(0) = 3.25 \times 10^{-2}$, $\mu_w = 7.75 \times 10^{-4}$, $\mu_g = 2.25 \times 10^{-4}$. Other fixed parameter of SAF-LMS [3] are as: $\mu_w = 0.05$ and $\mu_g = 0.05$.

Table 2: Summary of MSE of proposed AS-VLLMS-SAF with the initial parameters: $\mu_w(0) = 7.25 \times 10^{-5}$, $\mu_g(0) = 2.25 \times 10^{-5}$, $\gamma_w(0) = 6.25 \times 10^{-4}$, $\gamma_g(0) = 6.15 \times 10^{-4}$, of VL-LMS in [1] with $\gamma_w(0) = 3.15 \times 10^{-4}$, $\gamma_g(0) = 3.15 \times 10^{-4}$ and of LMS-SAF in [3] with $\mu_w = \mu_g = 0.05$ over 100 Monte Carlo trials and 5,000 samples used at SNR = 35dB

Algorithm	ϑ in (47)	Spline matrix	MSE at steady-state condition	
			MSE	dB
AS-VLLMS	$\vartheta = 0.20$	$\mathbf{C}_B$	1.268×10^{-3}	-28.969
		$\mathbf{C}_{CR}$	1.448×10^{-3}	-28.392
	$\vartheta = 0.95$	$\mathbf{C}_B$	1.768×10^{-3}	-29.293
		$\mathbf{C}_{CR}$	2.342×10^{-3}	-26.304
VL-LMS in[1]	$\vartheta = 0.20$	$\mathbf{C}_B$	1.815×10^{-3}	-27.409
		$\mathbf{C}_{CR}$	1.931×10^{-3}	-27.144
	$\vartheta = 0.95$	$\mathbf{C}_B$	2.175×10^{-3}	-26.625
		$\mathbf{C}_{CR}$	2.592×10^{-3}	-25.862
LMS in[3]	$\vartheta = 0.20$	$\mathbf{C}_B$	3.303×10^{-3}	-24.811
		$\mathbf{C}_{CR}$	1.352×10^{-2}	-18.688
	$\vartheta = 0.95$	$\mathbf{C}_B$	8.345×10^{-3}	-20.786
		$\mathbf{C}_{CR}$	2.858×10^{-2}	-15.439

5 Simulation Results

For the experiment results, the mean square error (MSE) is simulated at $\vartheta = 0.20, 0.95$. Fig. 2 and Fig. 3 show the MSE convergence curves of proposed AS-VLLMS-SAF compared with the original LMS-SAF [5] based on the *B-spline* matrix ($\mathbf{C}_B$) matrix and the *Catmul-Rom* spline matrix ($\mathbf{C}_{CR}$) with the parameters $\vartheta = 0.2, 0.95$ in (47) shown in dB and SNR=35dB, respectively. We notice that the curves of MSE of proposed AS-VLLMS-SAF based on both spline basis matrices outperform when comparable to that of the LMS-SAF algorithm.

Furthermore, Fig. 4 and Fig. 5 based on the *B-spline* matrix and Fig. 6 and Fig. 7 based on the *Catmul-Rom* spline matrix demonstrate the curves of learning rate of μ_{w_k} of coefficient vector $\mathbf{w}_k$ and μ_{g_k} of control points $\mathbf{g}_{m,k}$ for the proposed AS-VLLMS-SAF at $\vartheta = 0.20, 0.95$ and SNR=35dB, respectively. Their learning curves are depicted to converge to their equilibrium at the steady-state, even the initial of μ_{w_k} and μ_{g_k} are varied.

Finally, Fig. 8 and Fig. 9 based on the *B-spline* matrix and Fig. 10 and Fig. 11 based on the *Catmul-Rom* spline matrix present the curves of learning parameters of γ_{w_k} of coefficient vector $\mathbf{w}_k$ and γ_{g_k} of adaptive control points vector $\mathbf{g}_{m,k}$ for the proposed AS-VLLMS-SAF at $\vartheta = 0.20, 0.95$ and SNR=35dB, respectively. It can seen clearly that their learning rates are shown to converge for the tracking ability at the steady-state with the different initial parameters.

Summary of MSE of proposed AS-VLLMS-SAF with the initial parameters as $\mu_w(0) = 7.25 \times 10^{-5}$, $\mu_g(0) = 2.25 \times 10^{-5}$, $\gamma_w(0) = 6.25 \times 10^{-4}$, $\gamma_g(0) = 6.15 \times 10^{-4}$ and of LMS-SAF in [3] with the fixed parameter as $\mu_w = \mu_g = 0.05$ over 100 Monte Carlo trials and 5,000 samples used at SNR = 35dB and $\vartheta = 0.20, 0.95$ is presented in Table. 2. Simulation results suggest that mean square error performance of proposed algorithm can be partially assessed using adaptive step-size with the variable leaky parameters indicating better than the conventional least mean square algorithm by 16.76% with the *B-spline* matrix.

6 Discussion

The comparison of MSE averaged over 100 trials are shown the robustness and superiority of proposed AS-VLLMS algorithm over the conventional LMS algorithm. We have plotted the learning curves of adaptive step-size μ_{w_k} of coefficient vector $\mathbf{w}_k$, μ_{g_k} of control points $\mathbf{g}_{m,k}$ for the proposed AS-VLLMS algorithm and of variable leaky γ_{w_k} of coefficient vector $\mathbf{w}_k$ and γ_{g_k} of adaptive control points vector $\mathbf{g}_{m,k}$ after 5,000 iterations, which are seen that their learning curves converges to their equilibrium, even the initial values are assigned to be varied.

7 Conclusion

We have orchestrated an adaptive step-size and a variable leaky approach based on least mean square algorithm for spline adaptive filtering. The proposed AS-VLLMS-SAF algorithm has been explained how to derive using the variable leaky mechanism. We have designed an adaptive step-size algorithm with the methods of an energy of squared previous and present estimated error. We have designed a modified leaky algorithm with the methods of an optimal leaky parameter. Simulation experiments have shown that the proposed AS-VLLMS-SAF algorithm can perform well with the corresponding LMS-SAF algorithm.

In general, spline adaptive filtering structure is already fascinatingly applied in many applications such as signal processing for communications in nonlinear channel estimation and acoustic processing in bio-acoustic signal.

Conflict of Interest The authors declare no conflict of interest.

References

[1] A. Saengmuang, S. Sitjongsataporn, "Spline Adaptive Filtering based on Variable Leaky Least Mean Square Algorithm," in Proc. IEEE International Electrical Engineering Congress (iEECON), Chiang Mai, Thailand, 2020.

[2] C. Liu, Z. Zhang, X. Tang, "Sign Normalised Spline Adaptive Filtering Algorithms Against Impulsive Noise", Signal Processing, **148**, 234-240, 2018, doi:10.1016/j.sigpro.2018.02.022.

[3] M. Scarpiniti, D. Comminiello, R. Parisi, A. Uncini, "Nonlinear spline adaptive filtering," Signal Processing, **93**(4), 772-783, 2013, doi:10.1016/j.sigpro.2012.09.021.

[4] A. Saengmuang, S. Sitjongsataporn, "Convergence and Stability Analysis of Spline Adaptive Filtering based on Adaptive Averaging Step-size Normalized Least Mean Square Algorithm ", International Journal of Intelligent Engineering and Systems, **13**(2), 267-277, 2020, doi:10.22266/ijies2020.0430.26.

[5] M. Scarpiniti, D. Comminiello, G. Scarano, R. Parisi, A. Uncini, "Steady-State Performance of Spline Adaptive Filters", IEEE Transactions on Signal Processing, **64**(4), 816-828, 2015, doi:10.1109/TSP.2015.2493986.

[6] M. Scarpiniti, D. Comminiello, R. Parisi, A. Uncini, "Novel cascade spline architectures for the identification of nonlinear systems," IEEE Transactions on Circuits and Systems I: Regular Papers, **62**(7), 1825-1835, 2015, doi: 10.1109/TCSI.2015.2423791.

[7] M. Scarpiniti, D. Comminiello, R. Parisi, A. Uncini, "Hammerstein uniform cubic spline adaptive filtering : learning and convergence properties," Signal Processing, **100**, 112-123, 2014, doi:10.1016/j.sigpro.2014.01.019.

[8] S. Guan, Z. Li, "Normalised Spline Adaptive Filtering Algorithm for Nonlinear System Identification," Neural Processing Letter, **46**(2), 595-607, 2017, doi:10.1007/s11063-017-9606-6.

[9] S. Prongnuch, R.E. Valmoria, "Applied of Co-design in Reconfigurable System for Remote Image Noise Filtering via Ethernet Technology", in Proc. International Science, Social Science, Engineering and Energy Conference, 92-98, 2013.

[10] M. Scarpiniti, D. Commiiniello, R. Parisi, A. Uncini, "Nonlinear system identification using IIR spline adatpive filers," Signal Processing, **108**, 30-35, 2015, doi:10.1016/j.sigpro.2014.08.045.

[11] C. Liu, Z. Zhang, "Set-membership normalised least M-estimate spline adaptive filtering algorithm in impulsive noise," Electronics Letters, **54**(6), 393-395, 2018, doi:10.1049/el.2017.4434.

[12] S. Prongnuch, S. Sitjongsataporn, "Stability and Steady-State Performance of Hammerstein Spline Adaptive Filter Based on Stochastic Gradient Algorithm ", International Journal of Intelligent Engineering and Systems, **13**(3), 112-123, 2020, doi:10.22266/ijies2020.0630.11.

[13] W. Wang, H. Zhao, X. Zeng, K. Dogancay, "Steady-State Performance Analysis of Nonlinear Spline Adaptive Filter Under Maximum Correntropy Criterion", IEEE Transactions on Circuits and Systems II: Express Briefs, vol. **67**(6), 1154-1158, 2020, doi: 10.1109/TCSII.2019.2929536.

[14] W. Shi, Y. Li, Y. Wang, "Noise-Free Maximum Correntropy Criterion Algorithm in Non-Gaussian Environment", IEEE Transactions on Circuits and Systems II: Express Briefs, **67**(10), 2224-2228, 2020, doi:10.1109/TCSII.2019.2914511.

[15] D. Bismor, M. Pawelczyk, "Stability Conditions for the Leaky LMS Algorithm based on Control Theory Analysis", Archives of Acoustics, **41**(4), 731–739, 2016, doi:10.1515/aoa-2016-0070.

[16] S. Sitjongsataporn, W. Chimpat, "Adaptive Step-size Normalised Least Mean Square Algorithm for Spline Adaptive Filtering", in 2019 IEEE International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 544-547, 2019, doi:10.1109/ITC-CSCC.2019.8793383.

[17] L. Deng, Y. M. Huang, Q. Chen, Y. He, X. Sui, "Collaborative Blind Equalization for Time-Varying OFDM", IEEE Access, **8**, 103073-103087, 2020, doi: 0.1109/ACCESS.2020.2999387.

[18] K. Xiong, S. Wang, B. Chen, "Robust Normalized Least Mean Absolute Third Algorithm", IEEE Access, **7**, 10318-10330, 2019, doi:10.1109/ACCESS.2019.2891549.

[19] M.Kamenetsky, B.Widrow, "A Variable Leaky LMS Adaptive Algorithm," in *2004 IEEE Asilomar Conference on Signals, Systems and Compute (ASILOMAR)*, 125-128, 2004, doi:10.1109/ACSSC.2004.1399103.

[20] P. Kalkar, J.S.R.Alex, "Field Programmable Gate Array implementation of a Variable Leaky Least Mean Square Adaptive Algorithm," Asian Journal of Pharmaceutical and Clinical Research, 2017, doi:10.22159/ajpcr.2017.v10s1.19566.

[21] T.R.Gwadabe, M.S.Salman, H.Abuhilal, "A Modified Leaky-LMS Algorithm," International Journal of Computer and Electrical Engineering, **6**(3), 222-225, 2014, doi:10.1109/4234.1001664.

[22] S. Sitjongsataporn, P. Nurarak, "Low Complexity Variable Leaky Normalised Orthogonal Gradient Adaptive Algorithm", in 2017 IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA), 1-4, 2017, doi:10.1109/WAINA.2017.110.

[23] H. Kaur, K. Kaur, R. Mehra, "Adaptive Filter For Audio Signal Enhancement Using Leaky and non leaky algorithms," International Journal of Electrical and Electronic Engineering and Telecommunications, **3**(2), 114-122, 2014.

[24] S. Sitjongsataporn, "Modified adaptive forgetting-factor recursive least squares-based frequency-domain equalisation," in 2013 IEEE International Electrical Engineering Congress (iEECON), 1-4, 2013. doi:10.1109/iEECON.2013.110.

[25] M. H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley&Sons, 1996.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 651-657 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Novel Blockchain-Based Authentication and Access Control Model for Smart Environment

Nakhoon Choi, Heeyoul Kim*

Department of Computer Science and Engineering, Kyonggi University, Suwon 16227, Republic of Korea

ARTICLE INFO

Article history:
Received: 28 November, 2020
Accepted: 18 January, 2021
Online: 05 February, 2021

Keywords:
Blockchain
Access Control
Smart Environment

ABSTRACT

With the increase of smart factories and smart cities following the recent 4th industrial revolution, internal user authentication and authorization have become an important issue. The user authentication model using the server-client structure has a problem of forgery of the access history caused by the log manipulation of the administrator and unclearness of the responsibility. In addition, users must independently manage the authentication method for each service authentication. In this paper, to solve the above problem, the researchers propose an integrated ID model based on a hybrid blockchain. The proposed model is implemented as two layers of Ethereum and Hyperledger Fabric: the former layer is responsible for integrated authentication, and the latter layer is responsible for access control. The physical pass or application for user authentication and authorization are integrated to one ID through the proposed model. In addition, the decentralized blockchain ensures the integrity and transparency of the stored access history, and it also provides non-repudiation of authority and access history.

1. Introduction

This paper is an extension of work originally presented in ICACT 2020 (International Conference on Advanced Communication Technology) [1].

With the increase of smart factories and smart cities in accordance with the 4th industrial revolution, user authentication and access control are emerging as important issues. A smart factory automates factory facilities to reduce process failure rates and is operated by a small number of employees [2]. However, there is a problem of responsibility in case of an accident. In addition, users of smart city may need different authentication means for each city service, and there is a risk of manipulation of access history.

The administrator of the centralized access control model of the existing server-client structure can modify and delete the access history. With the above problem, it is possible to avoid responsibility through internal corruption in case of an accident. In addition, the user proves the authority through different authentication methods including physical means and mobile applications for each service. This makes it inconvenient for users to issue and manage various authentication methods, and it is troublesome for users to prove their authority history and career.

To solve the above problem, the authors in [3] proposed integrated authentication using smart cards for integrated access control of various services. However, there is a problem that the administrator can modify the authority and access details arbitrarily, such as the access control list or the capability list.

In order to solve the above problem through the blockchain, the model in [4] stores the information in DHT(Distributed Hash Tables) and performs access control for information based on the Bitcoin system. However, it is not possible to store access details, and there is a problem of processing speed with low TPS(Transaction per Second) of Bitcoin. Also, it is necessary to verify practicality and suitability. These studies are inadequate for authentication and access control in smart environments.

This paper proposes an integrated authentication and access control model in smart environments that solves the problems above with the help of blockchain technology. The proposed model processes user authentication and access request through a hybrid blockchain system using private and public blockchain together. The proposed model enables users to acquire and prove various access rights through one authentication means, thus it integrates the authentication of various smart factory services and smart city services. In addition, the proposed model prevents problems such as access history manipulation that occurs in the existing central model based on the integrity and transparency of the blockchain. Moreover, the user's history recorded in the blockchain can be easily used as the proof of the user's career.

*Corresponding Author: Heeyoul Kim, Department of Computer Science and Engineering, Kyonggi University, Suwon, 16227, Korea, Email: heeyoul.kim@kgu.ac.kr

The rest of this paper presents the following. Section 2 introduces blockchain platforms utilized in the proposed model. Section describes the architecture and process of the model. Section 4 displays the results of implementing the model. Section 5 provides application scenario using the model and Section 6 contains conclusions.

2. Background

The flexible connection of vertical or horizontal organizations became important in the wake of the Fourth Industrial Revolution [5]. This paper uses blockchain platform to build reconfigurable network of existing organizations. The integrity of information is secured in the relationship of organizations through the decentralized characteristics of blockchain. Users can also benefit from all the benefits without noticing that the blockchain has been used on the surface.

2.1. Ethereum

Ethereum [6] is a representative public blockchain using PoW(Proof of Work) [7] consensus algorithm like Bitcoin. PoW consensus algorithm is the process of selecting block creators among unreliable participants on blockchain networks. Participants compete to find a hash value that matches the target condition to elect a miner. The public blockchain introduces the economic concept of cryptocurrency, which leads to the correct consensus by compensating the cryptocurrency to the miner elected. The block generated through the consensus is distributed to the participants of the blockchain network and is connected to the previous block through the 'previous block hash' included in the block header to ensure the integrity of the previous block. All blocks and transactions are open to the public with ensuring transparency, and the decentralized network evolves continuously without administrators.

Ethereum supports the creation and distribution of smart contract [8] for the development of DApp(decentralized Application). A smart contract is written in Solidity, a Turing completeness language, and recorded in the distributed ledger with ensuring the integrity of the results according to automated execution and input [9]. Figure 1 represents the operation of the smart contract in Ethereum. The smart contract made with Solidity is compiled and stored in the form of the EVM(Ethereum Virtual Machine)-bytecode in the ledger. Calling the code through Ethereum clients such as Geth(Go-Ethereum) is executed in the EVM environment.

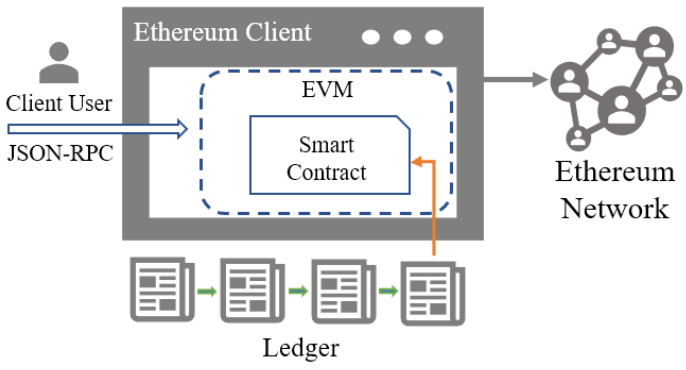


Figure 1: Operation of Smart Contract in Ethereum Client

2.2. Hyperledger Fabric

Hyperledger Fabric [10] (hereinafter referred to as "Fabric"), as part of the Linux Foundation's Hyperledger project, is a representative private blockchain platform commonly used for enterprise environments. Unlike public blockchains, it uses certificates and PKI(Public Key Infrastructure) to authenticate participants with a restricted network, and is managed through a MSP(Membership Service Provider), an authentication management system that defines the role and access rights of nodes in the network.

A peer is a node of organizations that compose the Fabric network and manages ledgers and chaincodes, and the peers are divided into endorser, committee, anchor, and leader according to their roles. Transactions for the execution of chaincodes in the network are performed through three steps: execution, ordering and validation. The chaincode plays the same role as Ethereum's smart contract and updates or queries data to the ledger. Currently, development of chaincodes is possible through Golang, Node.js, etc. Fabric does not require much resources in the consensus, unlike PoW, by ordering transactions through Ordering service to create blocks. Therefore, it has much better performance than the public blockchain platforms.

2.3. Metamask

The blockchain wallet stores the user's personal key as well as managing the balance of the blockchain account, and creates a transaction on behalf of the user and submits it to the network. In this sense, wallet is the easiest approach to using blockchain for normal users. The proposed model uses Metamask [11] for convenient use in PC or mobile environment.

Metamask is the most representative web-based Ethereum wallet extension program executed in browser, and is currently supported by Chrome, Firefox, Opera, etc., and beta test is being conducted in mobile environment. It supports Ethereum account management, digital signature generation, and transaction creation. users can easily access the smart contract of the blockchain through Metamask.

In order for users to access the Ethereum network, they must participate in the network as a blockchain node through the Ethereum client. This method lowers user accessibility, so Metamask uses Infura API [12], a cloud service for network connection to wallet users. Through the Metamask, users can access the Ethereum network in the web environment without operating the node directly, and operate the smart contract and receive the results.

3. A Blockchain-based Authentication and Access Control Model

3.1. Model Architecture

The proposed model uses a hierarchical hybrid blockchain where a public Ethereum network is connected to multiple private Fabric networks. Figure 2 shows the layered architecture of the proposed model. The public layer based on Ethereum is an integrated authentication management layer. The private layer is composed of various subsystems for smart factories and smart

cities (hereinafter referred to as "Organization"), and each organization operates Fabric network separately for authorization and access control.

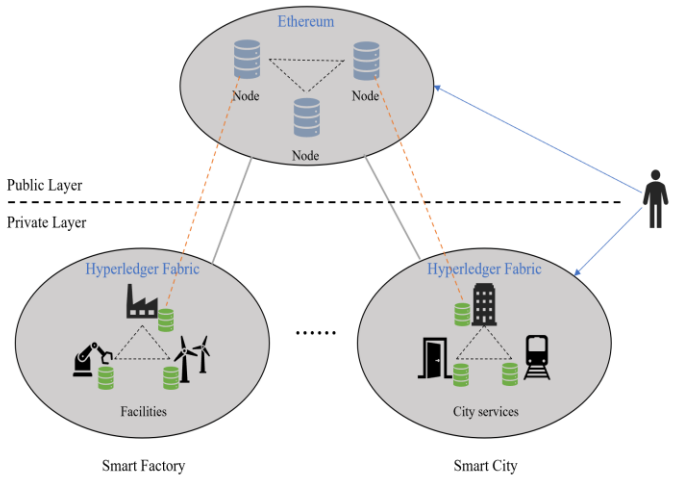


Figure 2: Layered Architecture of the Proposed Model

The proposed model uses a hierarchical hybrid blockchain where a public Ethereum network is connected to multiple private Fabric networks. Figure 2 shows the layered architecture of the proposed model. The public layer based on Ethereum is an integrated authentication management layer. The private layer is composed of various subsystems for smart factories and smart cities (hereinafter referred to as "Organization"), and each organization operates Fabric network separately for authorization and access control.

3.1.1. Public Layer - Integrated Authentication Management

Through this public layer, users create their own IDs based on their Ethereum accounts. The ID of a user is composed of the user's Ethereum account, a secret computed by hashing the user's personal information, a set of tokens used for user authentication by organizations, and a contact method (e.g., E-Mail) as shown in Figure 3. The proposed model provides a smart contract named IAM (Integrated Authentication Manager) utilized by the users to create and control their IDs. The functions in the IAM is described in Table 1, and these functions except for query functions have an access control mechanism on the IDs based on the function modifier in Solidity. In other words, it is guaranteed that only the legitimate owner of the ID can modify the secret and register tokens via these functions.

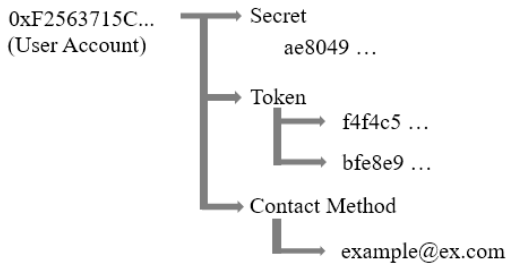


Figure 3: The structure of a user's ID

Table 1: Description of Functions in the IAM Smart Contract

Function	Purpose
createId(secret)	To create an ID and register the secret based on the user's Ethereum account
modifySecret(secret)	To modify the secret registered
regToken (token, tkName)	To register tokens and their aliases for the authentication by organizations
queryUser(secret)	To find the registered user by given secret and to retrieve user information
queryByToken(token)	To verify whether the token was successfully registered by the user

3.2.2. Private Layer – Access Control Management

In the private layer, there may be several Organizations managing smart environments respectively. Figure 4 represents the architecture of an Organization. Several nodes compose a Hyperledger Fabric network for the Organization. Among them, the admin node provides the ordering service for consensus and is handled by the administrator. The Ethereum sync node is connected to the Ethereum in the public layer and responsible for delivering the user's ID information for authentication.

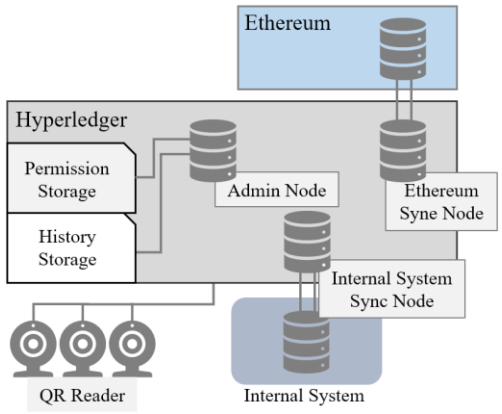


Figure 4: Architecture of an Organization in the Private Layer

Our model provides role-based access control (RBAC), and the role and access rights of each user are recorded in the Fabric blockchain via the chaincode named PM (Permission Manager). Table 2 shows the functions provided by the PM, and only the administrator can execute these functions for creation and modification of the authority. The access history is written by another chaincode called by the client application. The information stored in the Fabric blockchain is transparent and open to the legitimate participants in the Organization.

Table 2: Description of Functions in the PM Chaincode

Function	Purpose
userReg(secret)	To register a new user identified by the secret
permSet(account, role)	To set the role and authorization of the user having the given account
queryPerm(account)	To check and verify the user's role and access rights

3.2. Process Description

3.2.1. Creating an Integrated ID

A new user joining in the proposed model first performs the process of creating an integrated ID. The user first generates a secret by hashing his personal information as follows:

$$secret = Keccak\text{-}256(Birth, Name, Phone) \quad (1)$$

Formula 1 shows the creation of the user's secret, and Keccak-256 is a hash function adopted in the SHA-3 competition of the NIST (National Institute of Standards and Technology) [13].

Then, with assuming that the user already has an Ethereum account, he makes a transaction via Metamask to request the smart contract IAM for the creation of his integrated ID. Since the transaction is signed by his Ethereum private key, the IAM can identify and authenticate the requester. The secret can be seen by anyone after recorded in the blockchain. However, his personal information is not exposed due to the one-wayness of the hash function.

3.2.2. Registering User's Access Rights in Organization

When a user who has completed creating his ID newly joins in a specific Organization, his access right is registered by the process shown in Figure 5. The detailed process is as follows.

1) A new user Alice provides his personal data used for computing secret to the Organization.
2) The internal system of the Organization computes secret of the user by Equation (1), and then send it to the Ethereum sync node in the Fabric network.
3) The Ethereum sync node approaches the smart contract IAM to obtain the user's ID with the received secret.
4) The admin node issues a user-specific validation token T_{Alice}. The purpose of this token is to verify the ownership of the ID obtained in step 3). This token is the hash value of a seed generated per each user, and it is signed by the private key of the admin node (Equation (2)). Then, the token is sent to the user.

$$T_{Alice} = Sign(Keccak\text{-}256(seed_{Alice}), Org\ admin) \quad (2)$$

5) Alice registers the received token in his integrated ID through the smart contract IAM. This step only can be done by the creator of the ID because the IAM compares the requester of this register transaction with the owner of the target ID.
6) If the registration of T_{Alice} is successfully accepted by the IAM contract, the Ethereum sync node can recognize it and is convinced of Alice's ID ownership.
7) The administrator registers both the authority information of Alice and his Ethereum account in the permission storage of the organization.

3.2.3. Access Control

Our model provides a role-based access control with registered users' authorization information. Figure 6 shows the process when Alice requests access to the manufacturing control system in a smart factory. A QR reader attached to the control system is used to interact with the user. Our model determines whether the user's role has appropriate access right or not as follows.

1) Alice generates a QR code which includes both a timestamp for preventing replay attack and a digital signature generated by his Ethereum private key. Our model uses ECDSA [14] based on elliptic curve 'secp256k1' as well as Ethereum. Let G be the base point of the curve E where n is the order of G. Alice's private key is d, and his public key is $Q=dG$. The signature (r, s, v) is computed as follows.

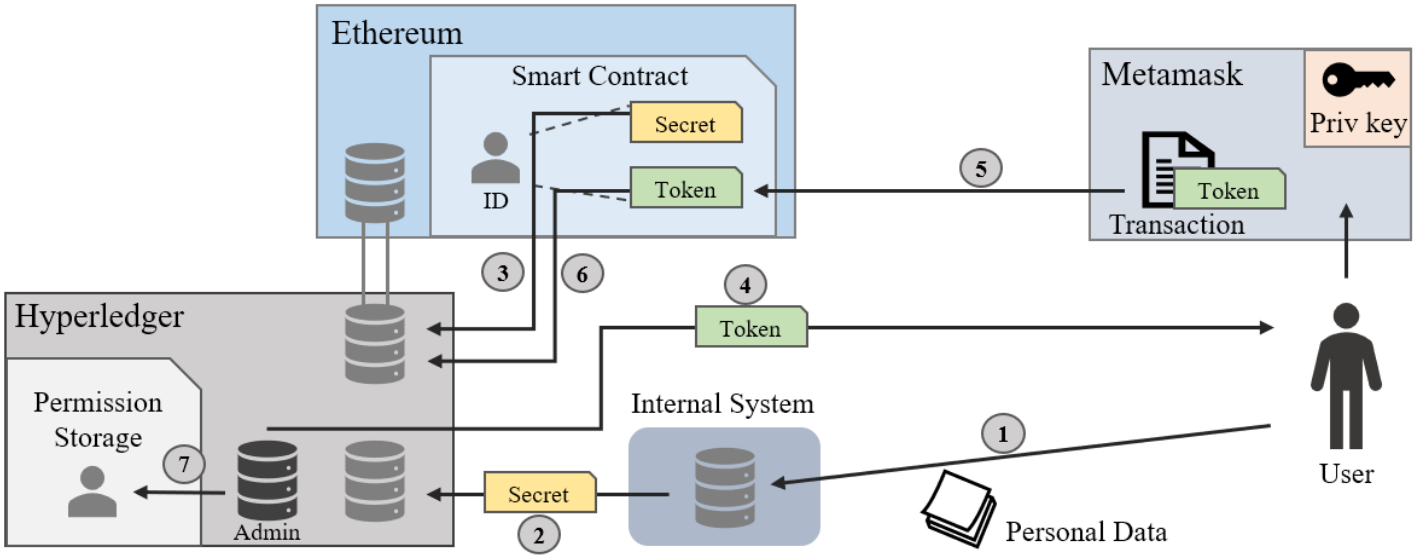


Figure 5: The Process of Registering User's Access Rights in an Organization

① Calculate the hash value $h = hash(m)$ of the message m to be signed, and let z be the l leftmost bits of h where l is the bit length of n.
② Generates a random number k within $[0 \sim n-1]$.
③ Compute $R = kG$ and obtain the x-coordinate r of point R.
④ Compute $s = k^{-1}(z + rd) \bmod n$.
⑤ Compute $v = 27 + (y \% 2)$ (y: y-coordinate of Q).
⑥ (r, s, v) is the digital signature value.

2) The QR reader scans the QR code generated by Alice, and as a Fabric client it delivers the timestamp, signature, and the *object_serial* (ID of the QR reader) to the Fabric node.
3) The Fabric node verifies received signature (r, s, v) with the help of Ethereum's *ecRecover()* method. If it is valid, as a result of this method call, the node discovers Alice's Ethereum account.
4) The authority information of the discovered Ethereum account is retrieved from the permission storage.
5) If Alice's role has been approved for access to the subject to which the *object_serial* is assigned, this request is granted.
6) Alice's access attempt is recorded in the history storage.

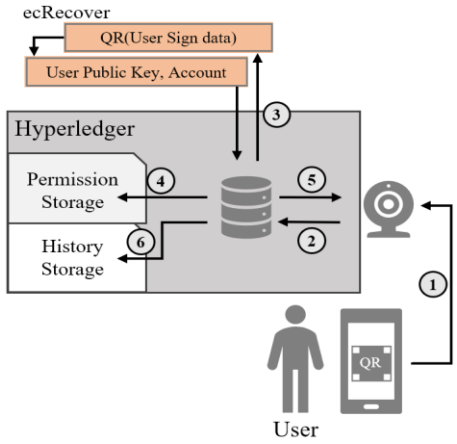


Figure 6: The Process of Access Control in an Organization

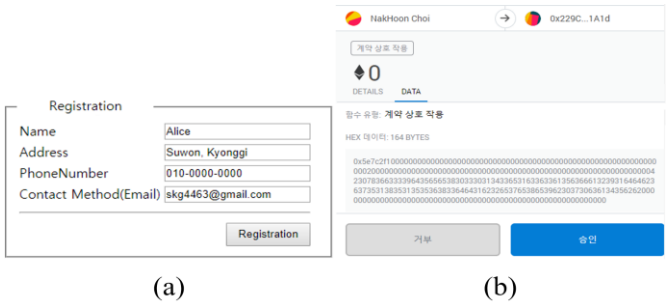
4. Implementation and Experiment

The proposed model was implemented, and its soundness was confirmed through an experiment. Table 3 describes the implementation environment of each component. The web dApp runs on the Chrome browser with Metamask installed, and it registers the user's secret for creating ID and the Organization's token for registering the user's access rights. The smart contract IAM and the chaincode PM provides the functions described in Section 3.1. QR generator is an application where the user's Ethereum private key is stored, and it generates a QR code containing the user's signature to obtain access right. QR reader is a Fabric client attached to the protected resources.

Table 3: Development Environment

Web dApp	Environment	Chrome, Metamask
	Language	JavaScript
Smart Contract IAM	Environment	Ethereum Ropsten Testnet
	Language	Solidity ^0.5.0
Chaincode PM	Environment	Hyperledger Fabric 2.1
	Language	Node.js
QR Generator	Environment	Android Emulator
	Language	Java
QR Reader	Environment	Raspberry Pi
	Language	Node.js

Figure 7 shows the screenshots when a user creates his integrated ID via the web dApp. As can be seen in Figure 8(a), a user inputs his personal data to compute secret. Then, the installed Metamask generates and transmits an Ethereum transaction requesting creation of the user's integrated ID.



(a) (b)

Figure 7: Screenshots when Creating Integrated ID. (a) Input of User Data for Computing Secret, (b) Generation of Ethereum transaction with Metamask

Figure 8 shows the logs when an organization registers a user Alice's role after verifying authentication. With given Alice's personal data, the organization computes Alice's secret (Figure 8(a)). Then, the organization checks whether Alice's integrated ID exists or not in the smart contract IAM by calling the *queryUser(secret)* function. If the ID exists, a validation token is issued for Alice (Figure 8(b)). After Alice successfully registers his token in his ID, the organization's Ethereum sync node can detect this event by implementing an event listener for IAM. As can be seen in Figure 8(c), the organization discovers Alice's Ethereum account from this event, and then assigns Alice's role and access right. In this experiment, the role 'level_2' is assigned to Alice and he is granted to access 'object 001'.

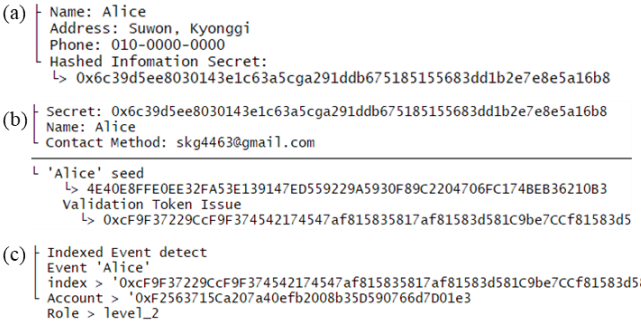


Figure 8: Logs when Registering Alice's Access Right in an Organization. (a) Computation of Alice's Secret by the Organization, (b) Issuance of a Validation Token for Alice, (c) Detection of Token Registration and Assignment of Alice's Role

Figure 9 shows an example QR code generated when Alice wants to access a protected resource in the organization. The QR generator installed in Alice's smartphone generates this QR code

by using of the stored private key (Figure 9(a)). As can be seen in Figure 9(b), the QR code contains two values (q0 and q1), where q0 is Alice's ECDSA signature and q1 is the Unix timestamp. Then, Alice provides this QR code to the QR reader attached to the resource.



(a)

"QR": {
"q0" : "0x9c304714ae532199b8cb0dd67eaf3248fe4d77b15672633855c8c579329eedff5145c2d9f502e5d80c1a4f8422dc426f7174161c701f1ea6b711adddb5f5ff621c",
"q1" : "1597140222"
}

(b)

Figure 9: An Example QR Code Generated for Access. (a) Screenshot of QR Generator, (b) The Information stored in the QR Code

The QR reader transmits both its *object_serial* ('0001' in Figure 10) and the signature obtained by scanning the QR code to the Fabric node. The node recovers Alice's Ethereum account by using of the *ecRecover* function. Also, it verifies the validity of both ECDSA signature and timestamp. Then, it can get Alice's registered role from the permission storage with the account, as shown in Figure 10. Finally, the node decides whether to grant or deny Alice's access request based on his role, and then it informs the QR reader of the decision result.

```
├ Access Detect > Object > '0001'
│ QR Signature: '0x9c304714ae532199b8cb0dd67eaf3248fe4d77b15672633855c8c579329e
ea6b711adddb5f5ff621c'
│ QR timestamp: '1606462209'
└ system timestamp: '1606464069'
  └ ecRecover
    └ recovering Account: '0xF2563715Ca207a40efb2008b35D590766d7D01e3', 'Alice'
    └ 'Alice' Role > 'level_2'

> User access request granted at: 1606464069
```

Figure 10: Logs when Alice's Access Request is Granted

5. Application Scenario using the Proposed Model

To help understanding the proposed model and claim its usefulness, we provide an application scenario using the model from a user point of view in this section. Let us assume that a new user Alice having an Ethereum account Eth_A decided to use the proposed model. He visits the web dApp and creates his own integrated ID (ID_A) by registering both his email address and secret based on his personal data.

Since then, Alice has moved into the smart city K-city (Organization #1) to seek employment. The K-city already has adopted the model extensively, and it participates as an organization. He visits the government office of the city to register himself. The office checks the existence of ID_A and then issues a token T_{Alice}^{Org1}. Alice registers this token in his ID_A by using of his Ethereum private key. Then, the office assigns the role 'citizen' to Alice and registers his access right to use various public facilities in the city.

Now Alice can prove his identity and use various facilities with his smartphone where the QR generator application is installed. Suppose Alice visits a public library to get hiring information. He just generates a QR code and shows it to the QR reader attached to the gate of the library. Then, his access is allowed, and he can enter the library.

Before long, Alice is hired by a smart factory F-factory (Organization #2). Similarly, in the enrollment process, he can easily register himself with a newly issued token T_{Alice}^{Org2}, without creating a new identity to be used in the factory. Then, the role 'engineer' is assigned to Alice and his access right for the control panel in the factory is registered. As a result, he can control the panel by generation a valid QR code whenever he needs to access it while at work.

Suppose someday Alice commits a mistake in operating some equipment in F-factory, which causes enormous damage to the factory. He gets into a panic, and he may try to avoid suspicion by deleting or manipulation evidences such as CCTV videos. However, since the access history of the users is stored in the immutable blockchain ledger of the F-factory nodes, he cannot hide the fact that he accessed the equipment at that time. Thus, the F-factory can identify Alice's behavior and call him to account.

In another case, suppose an attacker Trudy pretending as Alice tries to access the control panel. He cannot generate a valid QR code because he cannot discover Alice's private key. The attacker may eavesdrop Alice's QR code and try to use it later. However, the timestamp in the QR code prevents this kind of replay attack. Moreover, even if a security incident occurs, the factory can perform a security audit with the immutable access history.

6. Conclusion

Currently many companies and organizations use a centralized model for authentication and access control, so there is internal corruption caused by manipulating access history and it is possible to avoid responsibility in case of an accident. This paper proposes an access control model for smart factories and smart cities using blockchain, a decentralization platform, to solve the problems of existing centralized model and to ensure the integrity and immutability of access history.

The proposed model allows a user to prove their access rights through one integrated ID based on his Ethereum account without creating a new ID separately for each organization. In the aspect of an organization, it is needless to establish separate authentication mechanism, and the organization can easily authorize the users belong to it to access its protected resources. In addition, the proposed model records the access history in the internal blockchain of the organization not only to prevent arbitrary modification and deletion but also to provide non-repudiation.

The proposed model provides a clue to how various smart environments cooperate closely. Both the concept of integrated ID and the hybrid blockchain combining public blockchain and private blockchain can help secure, pragmatic, and user-friendly smart systems. In addition, flexible scalability in a smart environment is secured through a blockchain, and reliability of services and information provided to users is guaranteed.

Acknowledgement

This research was supported by Ministry of Science ICT Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education(2018R1C1B6002903). This work was supported by Kyonggi University's Graduate Research Assistantship 2020.

References

[1] N. Choi, H. Kim, "Hybrid Blockchain-based Unification ID in Smart Environment," International Conference on Advanced Communication Technology, ICACT, **2020**, 166–170, 2020, doi:10.23919/ICACT48636.2020.9061430.

[2] G. Büchi, M. Cugno, R. Castagnoli, "Smart factory performance and Industry 4.0," Technological Forecasting and Social Change, **150**(October 2019), 119790, 2020, doi:10.1016/j.techfore.2019.119790.

[3] G. Moukhliss, R. Filali Hilali, H. Belhadaoui, "A smart card digital identity check model for university services access," ACM International Conference Proceeding Series, **Part F1481**, 3–6, 2019, doi:10.1145/3320326.3320401.

[4] G. Zyskind, O. Nathan, A.S. Pentland, "Decentralizing privacy: Using blockchain to protect personal data," Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015, 180–184, 2015, doi:10.1109/SPW.2015.27.

[5] A. Mushtaq, I.U. Haq, "Implications of blockchain in industry 4.O," 2019 International Conference on Engineering and Emerging Technologies, ICEET 2019, 1–5, 2019, doi:10.1109/CEET1.2019.8711819.

[6] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," Ethereum Project Yellow Paper, **151**(2014), 1–32, 2014.

[7] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, Manubot, 2019.

[8] N. Szabo, "Smart contracts: building blocks for digital markets," EXTROPY: The Journal of Transhumanist Thought,(16), **18**(2), 1996.

[9] K. Christidis, M. Devetsikiotis, "Blockchains and Smart Contracts for the Internet of Things," IEEE Access, **4**, 2292–2303, 2016, doi:10.1109/ACCESS.2016.2566339.

[10] C. Cachin, "Architecture of the hyperledger blockchain fabric," in Workshop on distributed cryptocurrencies and consensus ledgers, 2016.

[11] Metamask Inc, Metamask, Website Https://Metamask.Io/,.

[12] The Infura Inc, INFURA, Website Https://Infura.Io/,.

[13] S. Chang, R. Perlner, W.E. Burr, J.M. Kelsey, L.E. Bassham, "Third-Round Report of the SHA-3 Cryptographic Hash Algorithm Competition NISTIR 7896 Third-Round Report of the SHA-3 Cryptographic Hash Algorithm Competition."

[14] D.R.L. Brown, "SEC 1: Elliptic curve cryptography," Certicom Research, v2, 2009.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 658-671 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Multiple Machine Learning Algorithms Comparison for Modulation Type Classification Based on Instantaneous Values of the Time Domain Signal and Time Series Statistics Derived from Wavelet Transform

Inna Valieva[1,*], Iurii Voitenko[2], Mats Björkman[1], Johan Åkerberg[1], Mikael Ekström[1]

[1]*School of Innovation, Design, and Engineering, Mälardalen University, Västerås, 721 23, Sweden*

[2]*Electronics Development, Wireless P2P Technologies, 79140, Sweden*

ARTICLE INFO

Article history:
Received: 29 November, 2020
Accepted: 21 January, 2021
Online: 05 February, 2021

Keywords:
Cognitive radio
Machine learning
Modulation classification

ABSTRACT

Modulation type classification is a part of waveform estimation required to employ spectrum sharing scenarios like dynamic spectrum access that allow more efficient spectrum utilization. In this work multiple classification features, feature extraction, and classification algorithms for modulation type classification have been studied and compared in terms of classification speed and accuracy to suggest the optimal algorithm for deployment on our target application hardware. The training and validation of the machine learning classifiers have been performed using artificial data. The possibility to use instantaneous values of the time domain signal has shown acceptable performance for the binary classification between BPSK and 2FSK: Both ensemble boosted trees with 30 decision trees learners trained using AdaBoost sampling and fine decision trees have shown optimal performance in terms of both an average classification accuracy (86.3 % and 86.0 %) and classification speed (120 0000 objects per second) for additive white gaussian noise (AWGN) channel with signal-to-noise ratio (SNR) ranging between 1 and 30 dB. However, for the classification between five modulation classes demonstrated average classification accuracy has reached only 78.1 % in validation. Statistical features: Mean, Standard Deviation, Kurtosis, Skewness, Median Absolute Deviation, Root-Mean-Square level, Zero Crossing Rate, Interquartile Range and 75th Percentile derived from the wavelet transform of the received signal observed during 100 and 500 microseconds were studied using fractional factorial design to determine the features with the highest effect on the response variables: classification accuracy and speed for the additive white gaussian noise and Rician line of sight multipath channel. The highest classification speed of 170 000 objects/second and 100 % classification accuracy has been demonstrated by fine decision trees using as an input Kurtosis derived from the wavelet coefficients derived from signal observed during 100 microseconds for AWGN channel. For the line of sight fading Rician channel with AWGN demonstrated classification speed is slower 130 000 objects/s.

1. Introduction

This paper is an extension of the work "Multiple Machine Learning Algorithms Comparison for Modulation Type Classification for Efficient Cognitive Radio" originally presented in Milcom 2019 [1].

The global market of mobile devices and services actively using the electromagnetic spectrum is continuously growing. Traditionally the electromagnetic spectrum utilization has been performed using a robust static approach developed almost a century ago: it rations access to the spectrum in exchange for the guarantee of interference-free communication spectrum is divided into the rigid, exclusively licensed bands, allocated over large, geographically defined regions. In conditions when some of these license bands are being nearly unused, while the others are overwhelmed, the problem of spectrum scarcity arises [2]. To cope with the increasingly populated spectrum, the electromagnetic spectrum utilization policies have been reformed

*Corresponding Author: Inna Valieva, inna.valieva@mdh.se

in recent years with the objective to allow the unlicensed secondary users to access licensed bands without causing interference to the licensed primary users [3].

Blind waveform estimation techniques can be used with an intelligent transceiver, yielding an increase in the transmission efficiency by reducing the overhead. Waveform information is critical to implement spectrum sharing scenarios like frequency hopping spread spectrum (FHSS) and dynamic spectrum access (DSA). Modulation type classification is a part of the waveform estimation together with the central frequency and symbol rate estimation. Multiple artificial intelligence algorithms have been successfully applied to solve the modulation classification problems. The availability of abundant data, breakthroughs of algorithms, and the advancements in hardware development during recent years have been driving forward vibrant development of deep learning [4]. Modulation classification using convolutional NN has been presented in [5] and [6]. Mathworks deployed a hands-on practical implementation of deep learning-based modulation classification in [7]. In [8], the authors have demonstrated successful launch of the optimized AlexNet on ZYNQ7045 FPGA: the inference execution times of CNNs in low density FPGAs has been improved using fixed-point arithmetic, zero-skipping and weight pruning. However, in this work the choice has been made to use less computationally demanding supervised ML algorithms. Conventional artificial intelligence (AI) algorithms like machine learning (ML) require preprocessing of the input signal and feature extraction. Received signal features used for the modulation classification could be classified into spectral-based and cyclostationary features. The spectral-based features exploit the unique spectral characters of different signal modulations in three physical aspects of the signal: the amplitude, phase, and frequency. The authors have summarized some of the well-recognized spectral features designed for modulation classification including the following and suggested a decision tree classification [9]. The authors have applied instantaneous amplitude, instantaneous phase, and spectrum symmetry together with the set of new features from both spectral and time domain including linear predictive coefficients, adaptive component weighting, zero-crossing ratio, linear frequency bank spectral coefficient for the classification of commonly used digital modulations including ASK, FSK, MSK, BPSK, QPSK, PSK, FSK4 and QAM-16 [10]. Gardner pioneered the area of cyclostationary signal analysis [11]. Gardner and Spooner first implemented cyclostationary analysis for modulation classification problems in [12].

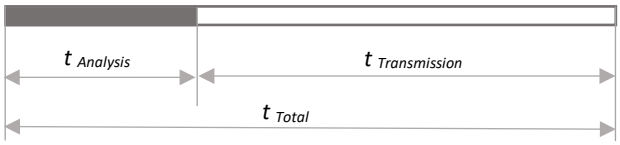


Figure 1: Optimizing sensing and transmission times.

The aim of this work is to determine the optimal modulation classification algorithm and suggest the subset of the strongest and robust features derived from the received signal that could be used to blindly classify the modulation type in our hardware application: a software-defined radio-based network consisting of multiple digital cognitive radio nodes. SDR-based nodes are operating in the frequency band from 70 MHz to 6 GHz with up to 56 MHz of instantaneous bandwidth are used as the hardware platform for the cognitive functionality deployment with FHSS capabilities. Available computational resources are Dual-core ARM Cortex A9 CPU, 2x512 MB of DDR3L RAM, and 512 MB of QSPI flash memory. The operation system is Embedded Linux. The radio part is based on Analog Devices AD9364 radio transceiver and Xilinx' Zynq-7020 FPGA. It is supporting multiple digital modulations including both linear: QPSK, BPSK, 8PSK, 16PSK, and non-linear: 2FSK. Symbol rate could be also adjusted between 10 KSymbols/s and 1 MSymbol/s to generate the cognitive waveform. Our target application predefines most of the boundary conditions and operational requirements such as required decision-making speed and computational resources available. In this work time required for radio scene analysis $t_{Analysis}$ is defined as a sum of time required for radio scene observation $t_{Observation}$ and processing $t_{Processing}$ for the received signal on the receiver front end. Time allocated for the active data transmission is data transmission time, $t_{Transmission}$ and total time is the sum of observation and transmission time as illustrated by Figure 8. Allocating the sensing time and transmission time at the MAC layer is involving a tradeoff between ensuring the PUs user's QoS requirements as opposed to maximizing the data throughput. To meet real-time operation requirements on the target hardware the following real-time performance characteristics must be met by the proposed algorithms. The maximum time available for radio-scene environment observation is $t_{observation}=500\times10^{-6}$ seconds. Our application is a time-slotted communication system, where 500 microseconds corresponds to one-time slot, also maximum processing time has been selected likewise $t_{processing}=500\times10^{-6}$ seconds, which requires the minimum classification speed of 2000 objects per second. Modulation classification is required to be performed for SNR values above the demodulation threshold of 12 dB, which corresponds to bit-error-rate BER=10^{-8} and BER=3.4×10^{-5} for BPSK and 2FSK, respectively.

The frequency of the radio scene environment sensing (how often sensing should be performed by the cognitive radio) and sensing time (the duration the sensing is performed) are key parameters affecting the throughput. While higher sensing times ensure more accurate radio scene environment sensing, this may result in a comparatively smaller duration for actual data transmission during the total time for which the spectrum may be used, thereby lowering the throughput [13]. To achieve the optimum between sensing time and throughput, for example, in IEEE 802.22 two-stage sensing (TSS) mechanism is implemented that includes: fast sensing, done at the rate of 1 ms/channel, and fine sensing performed on-demand. In this work, we have evaluated the possibility to perform the classification based on instantaneous values to shorten the spectrum sensing time and reduce computational costs. To improve classification accuracy, classification based on spectral-based statistical features derived from the wavelet transform of the received signal observed during the certain observation time frame has been proposed. In this study observation times of 100 and 500 microseconds have been used. The lowest observation time has been selected 100 microseconds to accommodate the lowest symbol rate of 10 KSymbols/second supported in our target application for cognitive waveform generation.

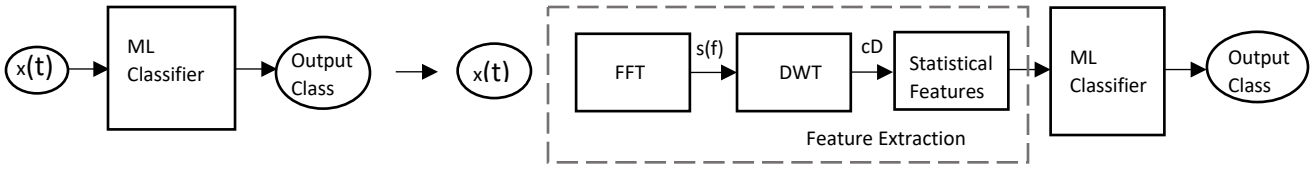


Figure 2: Modulation classification using a) instantaneous values of the time domain signal input; b) time series statistics input.

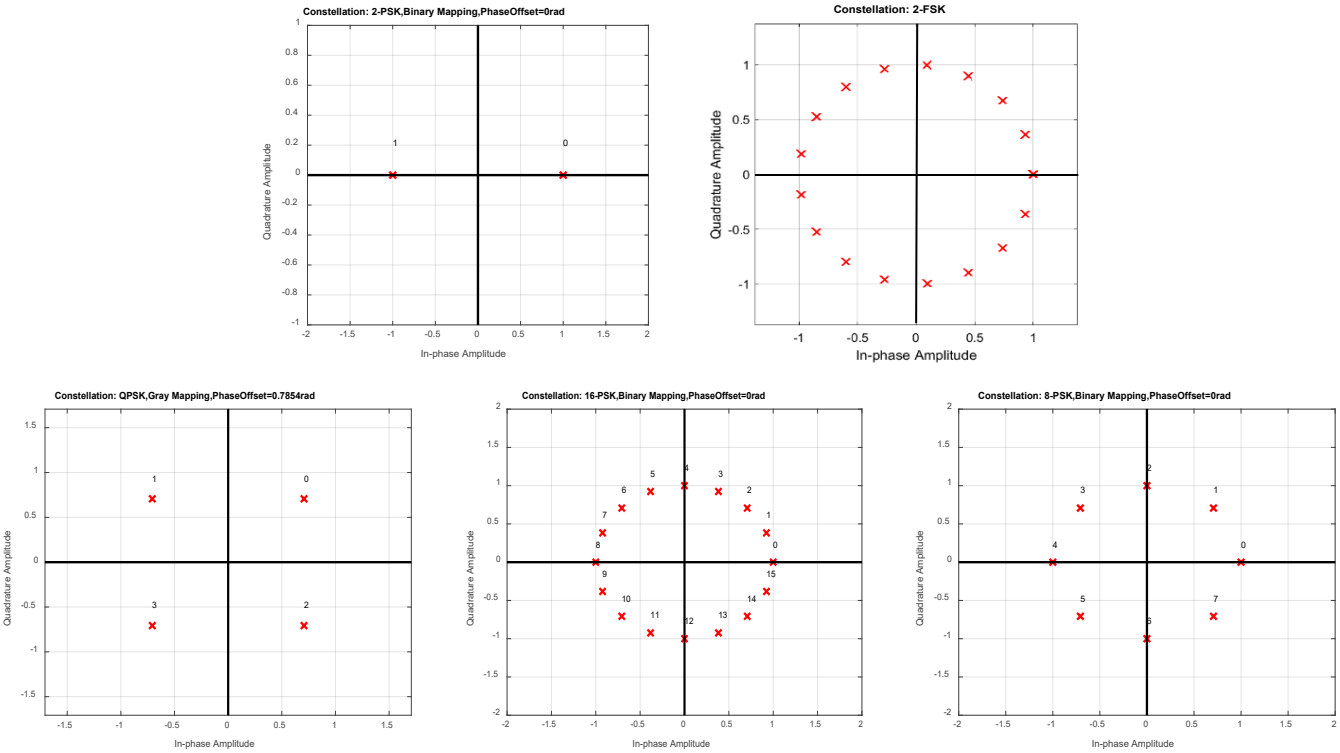


Figure 3: Constellation plots for studies modulation classes BPSK, 2FSK, QPSK, 8PSK, 16PSK.

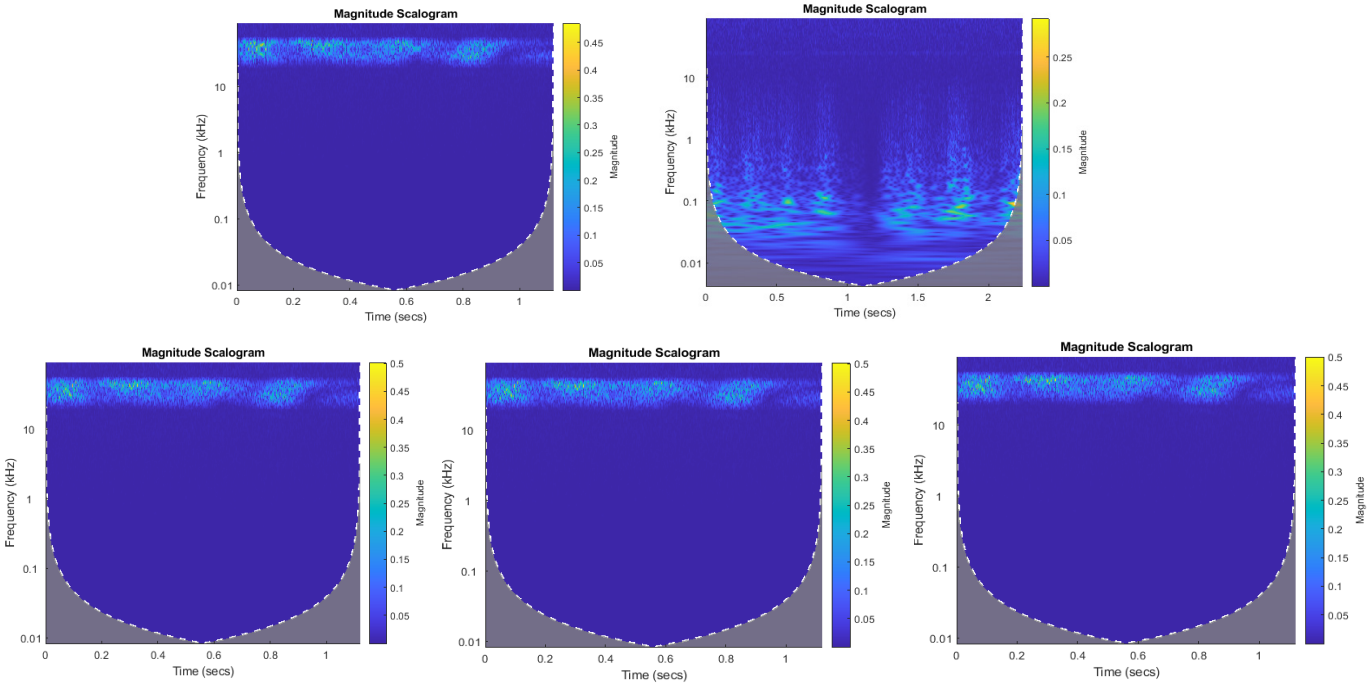


Figure 4: Scalograms obtained from Haar transform. wavelet coefficients for BPSK, 2FSK, QPSK, 8PSK, 16PSK. Observation time 500 microseconds signal.

The highest: 500 microseconds that correspond to one time slot in our time-slotted target application. Wavelet transform has been used in this work for the feature extraction, since obtained wavelet coefficients could be reused also to detect the vacant frequency channels and estimate the symbol rate, what could potentially save time and computational resources spent on radio scene analysis. Cyclostationary based algorithms have also been named in the literature as a powerful algorithm for modulation type and symbol rate estimation and vacant frequency channels detection. However, they are prone to cyclostationary noise and require longer observation time and is relatively computationally complex [14], and therefore this work has been focused on wavelet-based algorithms. Multiple supervised machine learning classifiers have been tested to perform classification based on instantaneous values and time-series statistics. The performance of tested classifiers has been evaluated in terms of classification accuracy and speed.

2. Methodology

Both feature extraction and classification algorithms applied in the literature to solve the modulation classification problem has been studied and evaluated. Primary, modulation type classification has been performed based on the input values of the instantaneous values of the in-phase and quadrature components of the time domain digital signal. Classification results have been satisfactory for the case of the binary classification between 2FSK and BPSK modulations. However, once the classification task has been extended to the higher-order modulations, the suggested classification approach has shown unacceptably low classification accuracy of 78.1 %. The highest classification accuracy of 84.9 % has been observed in the validation phase for classifying five modulations into two classes: linear and non-linear modulations using a fine SVM classifier.

Therefore, to improve the classification accuracy for higher-order modulations we have looked at multiple statistical features extracted from the time series containing the received signal observed during observation times of 100 and 500 microseconds. The following criteria have been considered when selecting classification features and feature extraction algorithm:

1. Robustness to noise; 2. Computational complexity; 3. Possibility to reuse preprocessed data for solving other radio scene environments observation tasks such as vacant bands detection and symbol rate estimation. Nine spectral-based statistical features derived from Haar wavelet transform coefficients calculated from the frequency domain signal observed during the selected observation time. The Haar transform has been selected as the simplest and less computationally demanding of all wavelet families. From the vector of wavelet coefficients, nine statistical features have been derived. To determine the strongest features to be used as inputs to ML classifier fractional factorial design analysis has been performed.

In statistics, a full factorial experiment is an experiment whose design consists of two or more factors, each with discrete possible values or "levels", and whose experimental units take on all possible combinations of these levels across all studied factors. Fractional factorial designs are experimental designs consisting of a carefully chosen subset or fraction of the experimental runs of a full factorial design. Significance is quantitively measured and referred to as the main effect [15]. A fractional factorial design approach has been used to analyze the main effects of every of nine statistical features and observation time on two studied response parameters: classification speed and accuracy in conditions of noise represented by two levels: AWGN and AWGN and multipath Rician channel model. Two levels have been applied and studied: first-level "-1" corresponds to not including the classification feature into classification and second level "+1" corresponds to including the studied feature into classification input. Observation time has also been varied according to two levels: first "-1"corresponds to 100 microseconds and the second level "+1" corresponds to 500 microseconds observation time. MATLAB and Simulink environment has been used for training and validation data sets generation including the modulator and propagation channel models, MathWorks classification learner ML Classification Learner tool has been used to train and validate classifiers. Twenty-four supervised ML classification algorithms have been studied and evaluated in terms of classification accuracy and classification speed for two studied channel models: AWGN and Rician multipath channels with AWGN.

3. Input data and feature extraction

To suggest optimal feature extraction and classification algorithms for our target radio application seven artificial data sets, consisting of signal samples modulated into 2FSK, BPSK, QPSK, 8PSK, 16PSK have been generated. Data sets are described in detail in Section 4. Figure 3 presents the constellation plots of five studied modulation types.

3.1. Classification using instantaneous values

The classification performed using instantaneous values of the time domain signal and SNR as classification inputs does not require any feature extraction: the raw values of the in-phase and quadrature component and SNR (recorded by transceiver as RSSI) have been used directly as an input to the classifier.

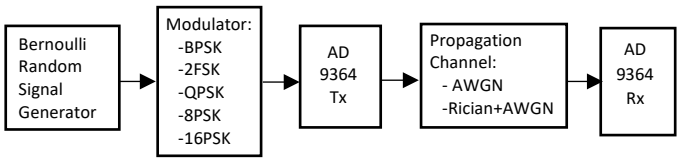


Figure 5: Generalized model used for data set generation.

Table 1: Spectral-based features derived from time series used for modulation type classification

N	Feature	Factor	Description
1	Mean	A	Mean value of the wavelet coefficients obtained from single-level discrete wavelet transform of the frequency domain signal.
2	Standard Deviation	B	Standard deviation of the wavelet coefficients obtained from single-level discrete wavelet transform of the frequency domain signal.

3	Kurtosis	C	Kurtosis is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution calculated from the wavelet coefficients obtained from single-level discrete wavelet transform of the frequency domain signal [16].
4	Skewness	D	Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. Calculated from the wavelet coefficients obtained from single-level discrete wavelet transform of the frequency domain signal [16].
5	Median absolute deviation	E	It is a measure of the statistical dispersion: a robust measure of the variability of a univariate sample of quantitative data. It is more resilient to outliers in a data set than the standard deviation. It is calculated from the wavelet coefficients obtained from single-level discrete wavelet transform of the frequency domain signal [17].
6	Root-mean-square level	F	The RMS value of a set of values (or a continuous-time waveform) is the square root of the arithmetic mean of the squares of the values, or the square of the function that defines the continuous waveform [18].
7	Zero crossing rate	G	It is the rate of sign-changes along a signal: the rate at which the signal changes from positive to zero to negative or from negative to zero to positive [19].
8	Interquartile range	H	The interquartile range (IQR) is a measure of variability: calculated based on dividing a data set into quartiles [20].
9	75^{th} Percentile	I	The percentile rank of a score is the percentage of scores in its frequency distribution that are equal to or lower than it [21].
10	SNR	J	Signal-to-Noise ratio corresponds to RSSI value for the received signal measured by transceiver.

Table 2: Classifiers performance reached in validation. Classification between FSK and BPSK modulations. AWGN channel (SNR ranging from 1 to30 dB). Data set 1

Classifier	With PCA (1 feature of 3)		No PCA (3 features of 3)		No PCA (2 features of 3, No SNR)	
	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s
Decision Trees: Fine	52.1	1600 000	86.0	1 200 000	84.2	2 000 000
Decision Trees: Medium	51.3	1200 000	85.2	1 000 000	83.3	2 100 000
Decision Trees: Coarse	50.7	1300 000	82.2	2 300 000	82.1	2 100 000
KNN: Fine	55.5	770 000	82.3	310 000	78.9	300 000
KNN: Medium	57.9	220 000	86.1	110 000	84.6	110 000
KNN: Coarse	61.5	55 000	86.8	28 000	85.5	27 000
KNN: Cosine	50.0	300	83.7	290	78.1	320
KNN: Cubic	57.9	150 000	86.1	29 000	84.6	46 000
KNN: Weighted	56.0	220 000	84.2	130 000	82.0	120 000
SVM: linear	50.1	47 000	55.3	3000	63.8	120 000
SVM: Quadratic	50.0	70 000	54.7	2300	42.8	92 000
SVM: Cubic	50.0	160 000	58.8	34 000	62.8	40 000
SVM: Fine Gaussian	49.5	230	86.9	790	85.4	970
SVM: Medium Gaussian	49.8	190	86.5	610	84.3	750
SVM: Coarse Gaussian	49.8	240	49.8	240	81.4	600
Ensemble Boosted Trees	51.3	110 000	86.3	120 000	85.0	120 000
Ensemble Bagged Trees	55.5	19 000	85.4	44 000	83.1	33 000
Ensemble Subspace Discriminant	49.8	100 000	50.3	86 000	50.5	120 000
Ensemble Subspace KNN	55.5	34 000	80.5	6100	69.5	17 000
Ensemble RUS Boosted Trees	51.3	120 000	85.1	150 000	83.3	130 000
Logistic Regression	49.8	1600 000	50.9	1 200 000	50.5	3 200 000
Linear Discriminant	49.8	1200 000	50.3	1 900 000	50.5	1 300000

3.2. Classification using time-series statistics.

Nine spectral-based statistical features have been derived from the received signal observed during the observation time. Primary fast Fourier transform has been applied to switch to the frequency domain. Then discrete wavelet transform has been applied to the frequency domain signal using Haar wavelet. This transform cross-multiplies a function against the Haar wavelet with various shifts and stretches. Figure 4 presents scalograms plots for studied modulations obtained by Haar wavelet transform. Then from the derived wavelet coefficients, nine spectral-based statistical features summarized in Table 1 have been calculated.

4. Data set

Seven data sets used for training and validation have been generated and are described briefly below. Three labeled artificial data sets consisting of three features: instantaneous values of in-phase and quadrature components of the time domain signal and SNR and four data sets consisting of nine statistical features extracted from time-series recorded during 100 and 500

microseconds and SNR for two channel models: AWGN and Rician channel with AWGN. Figure 5 describes the generalized model used to generate data sets.

Data Set 1: *Binary classification between BPSK and 2FSK using instantaneous values.* Data set for training and validation has been generated using a virtual model presented in Figure 5. The random input signal has been generated by Bernoulli binary random signal generator and modulated into BPSK or FSK. The modulated signal is transmitted by transmitter AD9364 TX. AWGN channel model has been selected to emulate the propagation environment, where SNR ranges from 30 to 1 dB. Data set consisting of 408000 rows and 4 columns, where two first columns correspond to the instantaneous values of the signal in the time domain: in-phase and quadrature components, the third column corresponds to SNR value and the fourth column represents the data label. 80 % of the data set has been used for training the classifiers and 20 % has been used for validation.

Data Set 2: *Classification of BPSK, 2FSK, QPSK, 8PSK, 16PSK into two classes linear and nonlinear using instantaneous values.* Data set of total size 910000 rows by 4 columns consisting of BPSK, 2FSK, QPSK, 8PSK, 16PSK has been generated in a similar way as Data Set 1: two first columns are corresponding to the instantaneous values of the signal in the time domain: in-phase and quadrature components, the third column corresponding to SNR value and the fourth column represents the data label corresponding to two classes: linear or non-linear modulation. Also, 80 % of the generated data set has been used for training the classifiers and 20 % of the generated data set has been used for validation of the classifiers.

Data Set 3: *Classification of BPSK, 2FSK, QPSK, 8PSK, 16PSK into five classes using instantaneous values.* Data set 3 consisting of five modulations described above has been used to train classifiers to classify all five modulation types. In this data set only the fourth column corresponding to the label has been changed to accommodate five output classes. Also, 80 % of the data set has been used for training of the classifiers and 20% for validation.

Table 3: Classifiers performance reached in validation. Classification between linear and non-linear modulations. Non-Linear: FSK, linear: BPSK, QPS, 8PSK, 16PSK modulations. AWGN channel (SNR ranging from 1 to 30dB). Data set 2.

Classifier	Average Accuracy, (%)	Prediction speed, (Objects/s)
Decision Trees: Fine	80.5	1 200 000
Optimized Trees	83.0	3 400 000
KNN: Fine	77.1	310 000
KNN: Medium	82.6	110 000
SVM: Fine Gaussian	84.9	790
Ensemble Boosted Trees	80.9	120 000
Ensemble Bagged Trees	81.8	44 000
Ensemble Subspace KNN	80.5	6100
Ensemble RUS Boosted Trees	78.1	150 000

Data Set 4: *Classification of BPSK, 2FSK, QPSK, 8PSK, 16PSK into five classes using statistical features derived from time series observed during 500 microseconds for AWGN channel.*

Data set 4 consists of 1500 samples (300 signal samples for every modulation type) resulting in a matrix of 1500 rows by 11 columns where the first nine columns correspond to statistical features column 10 corresponds to SNR and the last column corresponds to data class label. Statistical features for every signal sample have been derived from the received signal observed during 500 microseconds in conditions of the AWGN propagation channel (SNR ranging from 1 dB to 30 dB). Also, 80 % of the data set has been used for training, and 20 % for validation.

Data Set 5: *Classification of BPSK, 2FSK, QPSK, 8PSK, 16PSK into five classes using statistical features derived from time series observed during 100 microseconds. AWGN channel.* Data set 5 consisting of 1500 samples (300 signal samples for every modulation type) resulting in a matrix of 1500 rows by 11 columns where the first nine columns correspond to statistical features, column 10 corresponds to SNR and the last column corresponds to data class label. Statistical features for every signal sample have been derived from the received signal observed during 100 microseconds in conditions of AWGN propagation channel with

SNR ranging from 1 to 30 dB. Also, 80% of the data set has been used for training, and 20% for validation.

Data Set 6: *Classification of BPSK, 2FSK, QPSK, 8PSK, 16PSK into five classes using statistical features derived from the received signal observed during 500 microseconds for Rician multipath channel.* Data set 6 consisting of 1500 samples (300 signal samples for every modulation type) resulting in a matrix 1500 rows by 11 columns where the first nine columns correspond to statistical features, column 10 corresponds to SNR and the last column corresponds to data class label. Statistical features have been derived from time-domain signal observed during 500 microseconds in conditions of Rician multipath line-of-sight fading channel model with AWGN (SNR = 30). Three fading paths have been chosen with path delays selected for outdoor environments 0; 9×10^{-5} and 1.7×10^{-5}. Path delay ranging 10^{-5} is suggested for the mountains area. Average path gains have been chosen [0 -2 -10]. The Rician K-factor has been selected K-factor = 4, it specifies the ratio of specular-to-diffuse power for a direct line-of-sight path, it is usually in the range [1, 10] and 0 corresponds to Rayleigh fading. Maximum Doppler shift has been chosen 4 dB, which corresponds to a signal from a moving pedestrian [22]. Also, 80% of the data set has been used for training the classifiers and 20 % for validation.

Data Set 7: *Classification of BPSK, 2FSK, QPSK, 8PSK, 16PSK into five classes using statistical features derived from time series observed during 100 microseconds. Rician multipath channel.* Data set 7 consisting of 1500 samples (300 signal samples for every modulation type) resulting in a matrix 1500 rows by 11 columns where the first nine columns correspond to statistical features column 10 corresponds to SNR and the last column corresponds to data class label. Statistical features for every signal sample have been derived according to the steps summarized in feature extraction from time-domain signal observed during 100 microseconds in conditions of Rician multipath channel model with AWGN with SNR = 30 dB. The properties of the Rician channel have been selected the same as in data set 6. 80% of the data set has been used to train classifiers, 20% for validation.

Table 4: Classifiers performance reached in validation. Classification between FSK, BPSK, QPSK, 8PSK and 16PSK modulations using statistical features derived from the wavelet transform of the time series recorded during 500 microseconds. AWGN (SNR ranging from 1 to 30 dB) channel. Data set 4.

Classifier	With PCA (1 feature of 10)		No PCA (10 features of 10)		No PCA (5 features of 10)		No PCA Mean (1 features of 10)	
	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s
Decision Trees: Fine	100	15000	100	21000	100	20000	100	12000
Decision Trees: Medium	89.3	13000	99.7	23000	90.6	21000	79.9	18000
Decision Trees: Coarse	82.9	15000	88.1	19000	68.9	23000	67.5	19000
KNN: Fine	100	22000	100	42000	100	48000	100	51000
KNN: Medium	98.1	22000	96.6	26000	98.2	38000	97.2	56000
KNN: Coarse	80.6	18000	61.1	13000	61.7	21000	69.1	36000
KNN: Cosine	38.3	18000	96.9	22000	97.2	33000	21.3	38000
KNN: Cubic	98.1	17000	96.6	7800	98.5	20000	97.2	52000
KNN: Weighted	100	18000	100	36000	100	52000	100	52000
SVM: linear	57.0	15000	93.8	18000	75.1	21000	63.8	28000
SVM: Quadratic	51.2	20000	100	17000	98.0	17000	51.8	22000
SVM: Cubic	59.5	16000	100	16000	100	17000	47.7	27000
SVM: Fine Gaussian	79.9	13000	100	12000	99.2	16000	75.9	18000
SVM: Medium Gaussian	81.4	9400	97.9	16000	85.0	18000	69.4	20000
SVM: Coarse Gaussian	52.1	8200	88.9	14000	67.8	14000	65.9	16000
Ensemble Boosted Trees	93.0	9200	52.8	27000	99.9	10000	82.1	11000
Ensemble Bagged Trees	100	8200	100	11000	100	11000	100	12000
Ensemble Subspace Discriminant	56.3	6100	69.9	5900	59.3	6200	70.6	9500
Ensemble Subspace KNN	100	4400	100	4500	100	5000	100	6700
Ensemble RUS Boosted Trees	90.5	9800	83.6	61000	95.3	13000	80.4	11000
Linear Discriminant	56.3	13000	77.4	17000	62.5	17000	70.6	17000
Naïve Bayes	48.8	28000	60.9	60000	86.0	4800	73.8	18000

Table 5: Classifiers performance reached in validation. Classification between FSK, BPSK, QPSK, 8PSK and 16PSK modulations using statistical features derived from the wavelet transform of the time series recorded during 500 microseconds for Richian multipath with AWGN (SNR =30dB) channel model. Data set 5.

Classifier	With PCA (1 feature of 10)		No PCA (10 features of 10)		No PCA (5 features of 10)		No PCA Mean (1 feature of 10)	
	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s
Decision Trees: Fine	100	21000	100	27000	100	150000	93.0	160000
Decision Trees: Medium	49.5	98000	60.9	140000	55.6	150000	40.8	140000
Decision Trees: Coarse	41.6	11000	45.2	110000	41.5	120000	31.1	150000
KNN: Fine	100	22000	100	40000	100	79000	100	72000
KNN: Medium	95.7	23000	96	25000	95.7	45000	95.4	71000
KNN: Coarse	34.5	18000	36	16000	42.5	30000	37.7	36000
KNN: Cosine	37.7	17000	95.7	36000	95.7	30000	21.3	34000
KNN: Cubic	95.7	24000	95.6	2400	95.3	17000	95.4	54000
KNN: Weighted	100	29000	100	35000	100	38000	100	81000
SVM: linear	36.1	21000	36.6	26000	36.1	26000	21.8	40000
SVM: Quadratic	35.7	17000	66.6	21000	45.2	26000	23.5	29000
SVM: Cubic	36.5	19000	100	21000	61.4	30000	23.2	33000
SVM: Fine Gaussian	37.1	7800	91.8	11000	81.6	14000	38.0	11000
SVM: Medium Gaussian	36.2	8700	46.0	10000	44.2	11000	29.6	12000
SVM: Coarse Gaussian	35.0	8500	35.1	9700	34.9	11000	24.3	9700
Ensemble Boosted Trees	56.8	9400	86.4	12000	85.4	9700	49.9	12000

Ensemble Bagged Trees	100	9300	100	14000	100	8700	100	10000
Ensemble Subspace Discriminant	35.8	7800	39	8100	35.0	6400	22.5	8400
Ensemble Subspace KNN	100	5700	100	5300	100	4100	100	6100
Ensemble RUS Boosted Trees	61.7	11000	80.5	14000	86.0	14000	47.5	13000
Linear Discriminant	35.8	8600	37.9	18000	34.7	110000	22.5	100000
Naïve Bayes	39.9	9700	37.6	74000	37.9	110000	24.9	140000

5. ML classifiers training and validation results

Selected twenty-three classifiers have been trained and validated multiple times using every data set described above to investigate the effect of the extracted features and observation time on classification accuracy and speed. Among the trained classifiers we have studied decision trees, KNN with the various kernels, support vector machines (SVM), ensembles with bagging and boosting sampling, and discriminant methods. Also, principal component analysis PCA has been applied to keep enough components to explain 95 % of data variance for the dimension reduction. To protect against overfitting five-fold cross-validation method was used.

5.1. Classification using instantaneous values

Modulation type classification has been performed based on the input values of the instantaneous values of the in-phase and quadrature components of the time domain digital signal. Primary Classifiers have been trained three times using data set 1: 1. using PCA, 2. Using all three features: in-phase, quadrature components and SNR 3. Using two features including in-phase, quadrature components. Table 2 summarizes the validation results for the AWGN channel with SNR ranging from 1 to 30 dB. The effect of input features available in the data set 1 on classification accuracy, and speed has been studied. The highest classification accuracy has been achieved using all three available features: in-phase, quadrature components, and SNR. However, removing SNR from classification resulted in a reduction in the average classification accuracy by only 3 %. The highest average classification accuracy of 86.9 % was observed for the Fine Gaussian SVM, which on the other hand has been observed to be the slowest. Ensemble boosted trees with 30 decision trees learners trained using AdaBoost sampling and 20 splits and fine decision trees have shown optimal performance in terms of both classification accuracy and speed with an average classification accuracy of 86.3 % and 86.0 %, classification speed of 1200000 objects per second, which is faster than required 2000 objects per second.

Data set 3 containing also instantaneous values of the time domain signal and SNR consisting of five modulations has been used to train classifiers to classify all five modulation types. However, the average classification accuracy has not met the requirement of 85 %. The highest classification accuracy of 78.1 % has been reached by customized decision trees with the number of splits set to 2689, Gini's diversity index has been used as a split criterion.

Among the other tested classifiers, there were ensemble bagged trees, ensemble boosted trees, RUS boosted trees which have demonstrated even worse classification accuracy than decision trees. To capture more of the fine differences between the received signal modulated into different linear modulations it is suggested to use the spectral features derived from the signal observed during the selected observation time.

5.2. Classification using time-series statistics.

Training and validation for both propagation channels AWGN (SNR from 1 dB to 30 dB) and AWGN (SNR = 30 dB) with Rician fading have been performed using data sets 4-7. Primary, the training of the studied classifiers has been performed using time series recorded during observation time corresponding to 500 microseconds using data sets 4 and 5 for AWGN and Rician channel with AWGN, respectively. Classifiers have been trained four times: 1. using PCA, 2. Using all ten spectral-based statistical features 3. Using five features including mean, standard deviation, root-mean-square level, zero crossing rate, 75th percentile of a data set, and 4. Using only one input feature: the mean value of the wavelet coefficients. Tables 4-5 summarize the validation results for the AWGN channel with SNR ranging from 1 to 30 dB and Rician multipath with AWGN (SNR = 30 dB) channel. The best performing classifiers that have reached 100% classification accuracy have been retrained on the features derived from the time series observed during 100 microseconds. Tables 6 and 7 present the validation results for AWGN and multipath Rician channel trained and validated four times as for the time series recorded during 500 microseconds. Tables 6 and 7 show that four out of the five selected classifiers including fine decision trees, fine KNN, ensemble bagged trees, and ensemble subspace KNN that demonstrated classification accuracy of 100 % on the time series recorded during 500 microseconds have demonstrated classification accuracy close to 100 % when trained using PCA, using all ten features and using five features for both AWGN and multipath channel also when have been trained on time series recorded during 100 microseconds. However, SVM has shown worse performance when trained with a reduced number of features.

Applying PCA has resulted in a drastic decrease in the classification accuracy for ensemble subspace KNN for both AWGN and multipath Rician channels. For most classifiers, like fine decision trees, ensemble bagged trees applying PCA has resulted in the decreased classification speed. Fine KNN has shown 100% classification accuracy for classification using only one input to the classifier: mean value and the highest classification speed of 56000 and 91000 objects/s for AWGN and multipath fading channel, respectively. However, KNN is still slow even using only one feature than fine decision trees using five features classifying 110000 and 120000 for AWGN and multipath channel.

6. Classification feature analysis using fractional factorial design

The fractional factorial design has been applied to determine the most significant classification features referred to as the design.

Table 6: Classifiers performance reached in validation. Classification between FSK, BPSK, QPSK, 8PSK and 16PSK modulations using statistical features derived from the wavelet transform of the time series recorded during 100 microseconds for AWGN (SNR ranging from 1 to 30 dB) channel. Data set 6

Classifier	With PCA (1 feature of 10)		No PCA (10 features of 10)		No PCA (5 features of 10)		No PCA Mean (1 feature of 10)	
	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s
Decision Trees: Fine	100	9500	100	18000	100	110000	100	120000
KNN: Fine	100	21000	100	37000	98.6	19000	100	56000
KNN: Weighted	100	22000	100	34000	100	47000	100	58000
SVM: Cubic	95.7	23000	100	19000	100	24000	38.7	31000
Ensemble Bagged Trees	100	7600	100	8700	100	6800	100	11000
Ensemble Subspace KNN	36.4	5500	100	4700	100	4800	100	5400

Table 7: Classifiers performance reached in validation. Classification between FSK, BPSK, QPSK, 8PSK and 16PSK modulations using statistical features derived from the wavelet transform of the time series recorded during 100 microseconds for Rician multipath with AWGN (SNR = 30dB) channel model. Data set 7

Classifier	With PCA (1 feature of 10)		No PCA (10 features of 10)		No PCA (5 features of 10)		No PCA Mean (1 feature of 10)	
	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s	Average Accuracy, %	Prediction speed, Objects/s
Decision Trees: Fine	100	15000	100	23000	100	120000	94.2	160000
KNN: Fine	100	27000	100	24000	100	55000	100	91000
KNN: Weighted	100	25000	100	27000	100	48000	100	58000
SVM: Cubic	100	7900	100	16000	69.9	28000	69.9	28000
Ensemble Bagged Trees	100	5700	100	9900	100	11000	100	11000
Ensemble Subspace KNN	36.4	6900	100	5200	100	5000	100	6500

Table 8: Experimental factors and levels

N	Factor	Unit	Sym-bol	Coded level	
				-1	+1
1	Mean	-	A	on	off
2	Standard Deviation	-	B	on	off
3	Kurtosis	-	C	on	off
4	Skewness	-	D	on	off
5	Median absolute deviation	-	E	on	off
6	Root-mean-square level	-	F	on	off
7	Zero crossing rate	-	G	on	off
8	Interquartile range	-	H	on	off
9	75^{th} percentile of data set	-	I	on	off
10	SNR	dB	J	on	off
11	Observation time	μs	K	100	500

factors that have the highest main effect on the response parameters: classification accuracy and speed We have been looking at both the main effects of the independent variables and interactions between the input parameters (A-J) and observation time (K) for the best performing classifier in terms of accuracy and speed. We have 10 design factors corresponding to the features referred as (A-J) and observation time referred to as (K) with 2 levels for each design factor corresponding to: high (+1) the feature is used for classification or low (-1) the feature is not used for classification. Observation time has been also varied at two levels (+1) high corresponding to 500 microseconds and (-1) low corresponding to 100 microseconds. We have considered one noise factor corresponding to the propagation channel model with two levels corresponding to the AWGN (-1) channel with SNR ranging from 1dB to 30 dB and Rician multipath propagation and AWGN with SNR=30 dB (+1). The full factorial design will result in 2^{10} number of experiments, to reduce the number of experiments this study has been limited to fractional factorial design with 13 experiment runs performed, where the first ten runs with only one feature used for classification independently from the others for 100 microseconds observation time, since we are mostly interested to study the performance for the shortest observation time. The eleventh run has been selected to investigate the effect of observation time; the twelfth and thirteenth runs have been included to study the effect of observation time independently from the classification features with all nine spectral-based statistical features used as classification input.

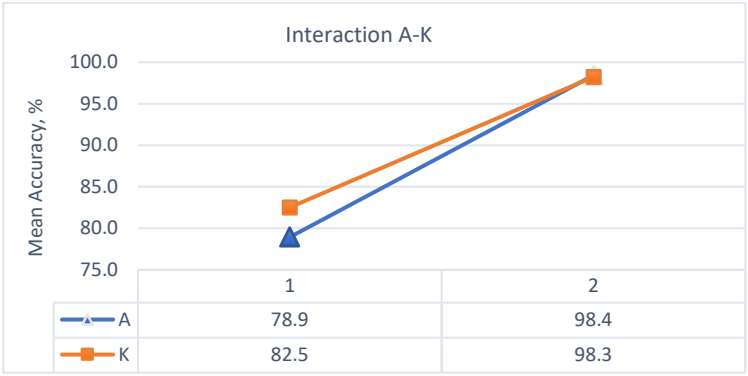
The main effect is defined as the overall effect of an independent variable in the complex design. The definition of an effect is the difference in the means between the high (+1) and the low level (-1) of a factor. From this notation, A is the difference between the mean values of the observations at the high level of A minus the average of the observations at the low level of A. Interaction is defined here as the effect of one independent variable

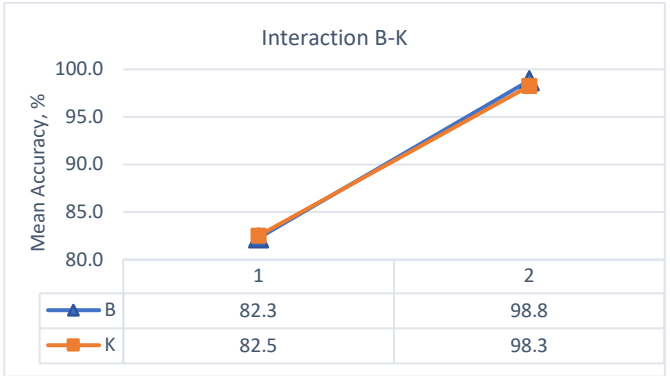
depending on the other independent variable, i.e it describes the combined effect of the independent variables considered simultaneously [15]. Tables 9-13 summarize the experimental design and results: main effect of every factor on classification accuracy and speed for five best performing classifiers: fine decision trees, fine KNN, weighted KNN, ensemble bagged trees with 59886 splits, and 30 learners and ensemble subspace KNN with 9 subspace dimensions and 30 learners.

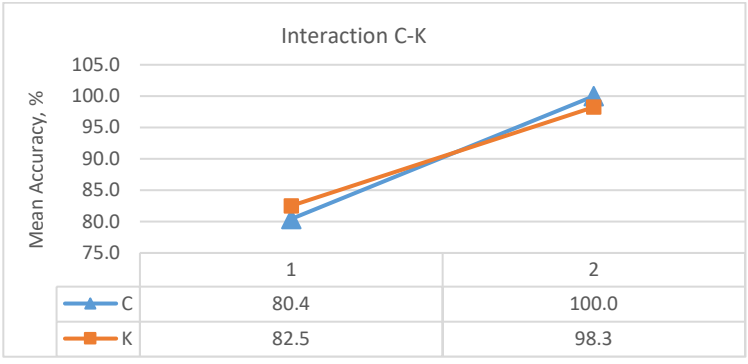
The overall highest main effects on classification accuracy and speed have been observed for the fine decision trees classifier. In tables 9-13, we have obtained a relatively high main effect for using/not using SNR for classification. In tables 9-13, summarizing the performance of the classifier, it is possible to see that removing the SNR value as an input to the classifier does not affect the response parameters significantly, on the other hand using SNR value only for classification as in the experiment run N10 (what in reality does not make any sense) result in the very low classification accuracy. Since the main effect is a difference between the mean values of the high and low, we observe here the high main effect of SNR as a classification input parameter. For example, the main effect of using SNR (J) as classification input on classification accuracy for fine decision trees is –20.0, while for Kurtosis (C) it is 19.6. Classification accuracy if we use only (C) for classification experiment run N3 is 100% for both Rician and AWGN channels and if we use only (J) for classification is 8.6% for both Rician and AWGN channels. This makes the highest main effect of SNR input a questionable result for practical application. Also, the value of the main effect should be interpreted only in combination with the value of the response parameter: classification accuracy and speed.

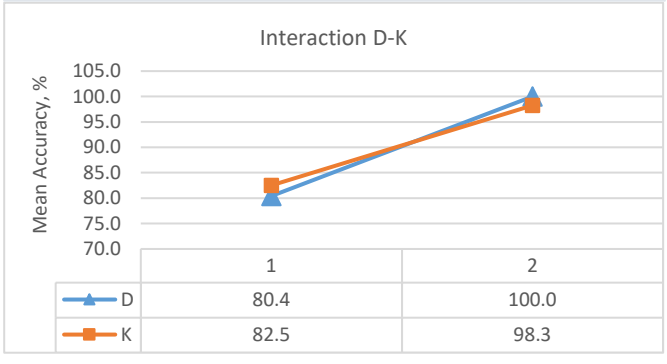
The highest classification speed of 170 000 objects per second for 100% classification accuracy has been demonstrated by fine decision trees using only one classification input the skewness of the wavelet coefficients derived from signal observed during 100 microseconds for AWGN channel model. For Rician and AWGN, channel classification speed has been slower 130 000 objects/s. Both skewness and kurtosis has shown the highest main effect on classification accuracy for fine trees. Using mean value as the only input parameter to fine trees classifier, however, has shown less robust results in terms of the classification accuracy: 94.5% has been demonstrated for Rician and AWGN channel, while for AWGN channel 100%. The mean value shows the second-highest main effect on the classification accuracy 19.5 for the fine decision trees and the highest main effect of 10.9, 11.1, 11.5, and 10.9 for the other four classifiers including fine KNN, weighted KNN, ensemble bagged trees and ensemble subspace KNN, respectively. It also shows the highest main effect on the classification speed for all five selected classifiers.

Slightly slower classification speed has been demonstrated by KNN classifiers: 110000 objects per second have been observed for fine KNN for AWGN channel and 89000 objects per second for AWGN channel with Rician multipath. The slowest from all studied are ensembles, for example, ensemble subspace KNN has demonstrated a classification speed of 4500-6800 objects per second. However, classification speed for ensembles does not decrease as much as for the fine trees with the increasing number of features. For ensemble bagged trees for AWGN channel the same classification speed of 11000 has been demonstrated for classification using only mean A and using all calculated features (A-J).









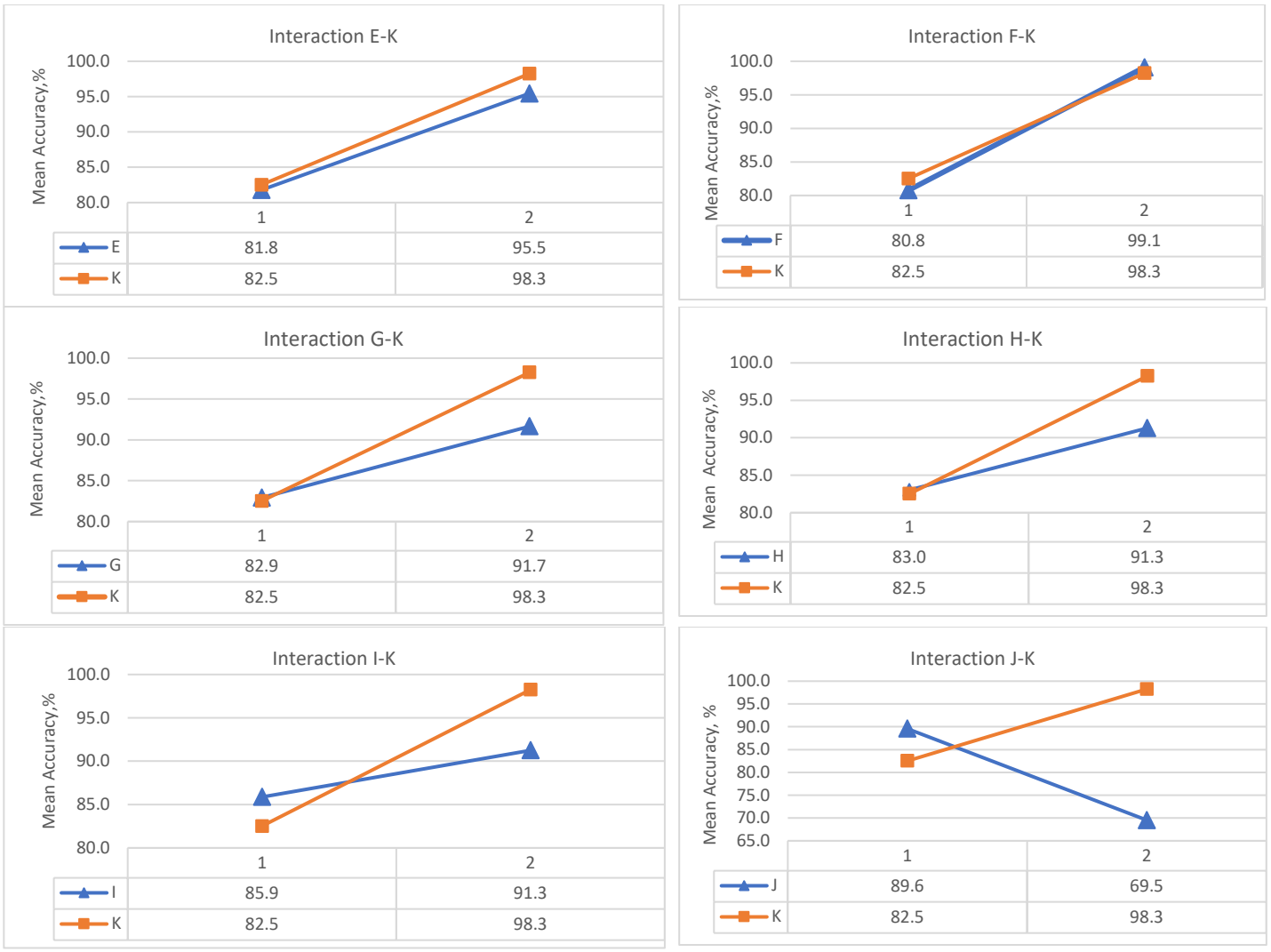


Figure 6: Main Effects and interactions between factors (A-J) and observation time (K) for fine decision trees

Table 9: Experimental design and results of fractional factorial design limited to 13 experimental runs for fine decision trees classifier

Factors												AWGN		AWGN + Rician		Mean Accuracy	STD Accuracy	Mean Speed	STD Speed
Run	A	B	C	D	E	F	G	H	I	J	K	Speed	Accuracy	Speed	Accuracy				
1	+	-	-	-	-	-	-	-	-	-	-	100000	100	160000	94.5	97.3	3.9	130000	113068
2	-	+	-	-	-	-	-	-	-	-	-	140000	100	140000	92.9	96.5	5.0	140000	98927
3	-	-	+	-	-	-	-	-	-	-	-	110000	100	150000	100	100.0	0.0	130000	105995
4	-	-	-	+	-	-	-	-	-	-	-	170000	100	130000	100	100.0	0.0	150000	91853
5	-	-	-	-	+	-	-	-	-	-	-	130000	93.8	98000	78.9	86.4	10.5	114000	69235
6	-	-	-	-	-	+	-	-	-	-	-	110000	100	140000	92.9	96.5	5.0	125000	98927
7	-	-	-	-	-	-	+	-	-	-	-	170000	75.4	130000	74.5	75.0	0.6	150000	91871
8	-	-	-	-	-	-	-	+	-	-	-	130000	72.9	160000	74.9	73.9	1.4	145000	113085
9	-	-	-	-	-	-	-	-	+	-	-	120000	73.4	99000	74.1	73.8	0.5	109500	69951
10	-	-	-	-	-	-	-	-	-	+	-	110000	8.6	180000	8.6	8.6	0.0	145000	127273
11	+	-	-	-	-	-	-	-	-	-	+	12000	100	160000	93	96.5	4.9	86000	113069
12	+	+	+	+	+	+	+	+	+	+	-	18000	100	23000	100	100.0	0.0	20500	16193
13	+	+	+	+	+	+	+	+	+	+	+	21000	100	27000	100	100.0	0.0	24000	19021
Main Effects Accuracy	19.5	16.5	19.6	19.6	13.7	18.3	8.7	8.3	5.4	-20.0	15.7								
Main Effects Speed	-69.2	-29.6	-29.5	-27.5	-35.6	-31.1	-35.8	-36.7	-40.3	-58.4	-24.7								

Table 10: Experimental design and results of fractional factorial design limited to 13 experimental runs for fine KNN classifier

Factors												AWGN		AWGN + Richian		Mean Accuracy	STD Accuracy	Mean Speed	STD Speed
Run	A	B	C	D	E	F	G	H	I	J	K	Speed	Accuracy	Speed	Accuracy				
1	+	-	-	-	-	-	-	-	-	-	-	73000	100	91000	100	100.0	0.0	36500	64276
2	-	+	-	-	-	-	-	-	-	-	-	69000	100	41000	100	100.0	0.0	55000	28921
3	-	-	+	-	-	-	-	-	-	-	-	83000	100	65000	100	100.0	0.0	74000	45891
4	-	-	-	+	-	-	-	-	-	-	-	54000	100	77000	100	100.0	0.0	65500	54377
5	-	-	-	-	+	-	-	-	-	-	-	91000	100	84000	100	100.0	0.0	87500	59326
6	-	-	-	-	-	+	-	-	-	-	-	110000	100	86000	100	100.0	0.0	98000	60740
7	-	-	-	-	-	-	+	-	-	-	-	72000	79.9	69000	80.6	80.3	0.5	70500	48734
8	-	-	-	-	-	-	-	+	-	-	-	80000	100	110000	100	100.0	0.0	95000	77711
9	-	-	-	-	-	-	-	-	+	-	-	83000	100	110000	100	100.0	0.0	96500	77711
10	-	-	-	-	-	-	-	-	-	+	-	93000	21.3	83000	21.3	21.3	0.0	88000	58675
11	+	-	-	-	-	-	-	-	-	-	+	51000	100	72000	100	100.0	0.0	61500	50841
12	+	+	+	+	+	+	+	+	+	+	-	37000	100	30000	100	100.0	0.0	33500	21142
13	+	+	+	+	+	+	+	+	+	+	+	42000	100	40000	100	100.0	0.0	41000	28214
Main Effects Accuracy	10.9	9.8	9.8	9.8	9.0	9.8	1.3	9.8	9.8	-24.3	9.0								
Main Effects Speed	-38.0	-34.1	-25.9	-29.6	-20.1	-15.5	-27.4	-16.8	-16.2	-19.8	0.0								

Table 11: Experimental design and results of fractional factorial design limited to 13 experimental runs for weighted KNN classifier.

Factors												AWGN		AWGN + Rician		Mean Accuracy	STD Accuracy	Mean Speed	STD Speed
Run	A	B	C	D	E	F	G	H	I	J	K	Speed	Accuracy	Speed	Accuracy				
1	+	-	-	-	-	-	-	-	-	-	-	62000	100	58000	100	100.0	0.0	60000	40941
2	-	+	-	-	-	-	-	-	-	-	-	63000	100	67000	100	100.0	0.0	65000	47305
3	-	-	+	-	-	-	-	-	-	-	-	90000	100	63000	100	100.0	0.0	76500	44477
4	-	-	-	+	-	-	-	-	-	-	-	62000	100	77000	100	100.0	0.0	69500	54377
5	-	-	-	-	+	-	-	-	-	-	-	66000	100	55000	100	100.0	0.0	60500	38820
6	-	-	-	-	-	+	-	-	-	-	-	73000	100	79000	100	100.0	0.0	76000	55791
7	-	-	-	-	-	-	+	-	-	-	-	60000	79.4	67000	100	89.7	14.6	63500	47313
8	-	-	-	-	-	-	-	+	-	-	-	71000	100	79000	80	90.0	14.1	75000	55798
9	-	-	-	-	-	-	-	-	+	-	-	68000	100	72000	100	100.0	0.0	70000	50841
10	-	-	-	-	-	-	-	-	-	+	-	56000	20.4	73000	21.3	20.9	0.6	64500	51604
11	+	-	-	-	-	-	-	-	-	-	+	52000	100	81000	100	100.0	0.0	66500	57205
12	+	+	+	+	+	+	+	+	+	+	-	34000	100	23000	100	100.0	0.0	28500	16193
13	+	+	+	+	+	+	+	+	+	+	+	36000	100	35000	100	100.0	0.0	35500	24678
Main Effects Accuracy	11.1	9.9	9.9	9.9	9.9	9.9	5.5	5.6	9.9	-24.4	9.0								
Main Effects Speed	-21.3	-25.2	-20.2	-23.3	-27.2	-20.4	-25.9	-20.9	-23.0	-25.4	-14.1								

Table 12: Experimental design and results of fractional factorial design limited to 13 experimental runs for Ensemble bagged trees classifier.

Factors												AWGN		AWGN + Rician		Mean Accuracy	STD Accuracy	Mean Speed	STD Speed
Run	A	B	C	D	E	F	G	H	I	J	K	Speed	Accuracy	Speed	Accuracy				
1	+	-	-	-	-	-	-	-	-	-	-	11000	100	11000	100	100.0	0.0	11000	7707
2	-	+	-	-	-	-	-	-	-	-	-	9400	100	9600	100	100.0	0.0	9500	6718
3	-	-	+	-	-	-	-	-	-	-	-	11000	100	12000	100	100.0	0.0	11500	8415
4	-	-	-	+	-	-	-	-	-	-	-	10000	100	11000	100	100.0	0.0	10500	7707
5	-	-	-	-	+	-	-	-	-	-	-	10000	100	10000	100	100.0	0.0	10000	7000
6	-	-	-	-	-	+	-	-	-	-	-	11000	100	10000	100	100.0	0.0	10500	7000
7	-	-	-	-	-	-	+	-	-	-	-	12000	76	7700	100	88.0	17.0	9850	5382
8	-	-	-	-	-	-	-	+	-	-	-	9800	100	8500	100	100.0	0.0	9150	5940
9	-	-	-	-	-	-	-	-	+	-	-	7600	100	7700	100	100.0	0.0	7650	5374
10	-	-	-	-	-	-	-	-	-	+	-	14000	8.9	1200	8.7	8.8	0.1	7600	842
11	+	-	-	-	-	-	-	-	-	-	+	12000	100	10000	100	100.0	0.0	11000	7000
12	+	+	+	+	+	+	+	+	+	+	-	8700	100	9900	100	100.0	0.0	9300	6930
13	+	+	+	+	+	+	+	+	+	+	+	11000	100	14000	100	100.0	0.0	12500	9829
Main Effects Accuracy	11.5	10.3	10.3	10.3	10.3	10.3	5.1	10.3	10.3	-29.2	9.4								
Main Effects Speed	1.4	0.6	1.4	1.0	0.8	1.0	0.7	0.4	-0.2	-0.3	2.1								

Table 13: Experimental design and results of fractional factorial design limited to 13 experimental runs for Ensemble subspace KNN classifier

Factors												AWGN		AWGN + Rician		Mean Accuracy	STD Accuracy	Mean Speed	STD Speed
Run	A	B	C	D	E	F	G	H	I	J	K	Speed	Accuracy	Speed	Accuracy				
1	+	-	-	-	-	-	-	-	-	-	-	6100	100	6500	100	100.0	0.0	6300	4525
2	-	+	-	-	-	-	-	-	-	-	-	6200	100	6200	100	100.0	0.0	6200	4313
3	-	-	+	-	-	-	-	-	-	-	-	6800	100	7100	100	100.0	0.0	6950	4950
4	-	-	-	+	-	-	-	-	-	-	-	5600	100	6400	100	100.0	0.0	6000	4455
5	-	-	-	-	+	-	-	-	-	-	-	6700	100	6800	100	100.0	0.0	6750	4738
6	-	-	-	-	-	+	-	-	-	-	-	6600	100	6200	100	100.0	0.0	6400	4313
7	-	-	-	-	-	-	+	-	-	-	-	6500	79.9	7800	100	90.0	14.2	7150	5452
8	-	-	-	-	-	-	-	+	-	-	-	6100	100	5900	80.6	90.3	13.7	6000	4108
9	-	-	-	-	-	-	-	-	+	-	-	6800	100	5500	100	100.0	0.0	6150	3818
10	-	-	-	-	-	-	-	-	-	+	-	6200	21.3	5300	21.3	21.3	0.0	5750	3733
11	+	-	-	-	-	-	-	-	-	-	+	6700	100	6100	100	100.0	0.0	6400	4243
12	+	+	+	+	+	+	+	+	+	+	-	4700	100	5200	100	100.0	0.0	4950	3606
13	+	+	+	+	+	+	+	+	+	+	+	4500	100	5300	100	100.0	0.0	4900	3677
Main Effects Accuracy	10.9	9.8	9.8	9.8	9.8	9.8	5.5	5.6	9.8	-24.3	9.0								
Main Effects Speed	-0.7	-1.0	-0.7	-1.1	-3.0	-0.9	-0.6	-1.1	-1.1	-1.2	-0.6								

Since the decision trees are the fastest classifier reaching an accuracy of 100%, the study of the interactions has been limited only to fine decision trees. Interaction between observation time and every of the spectral-based statistical feature (A-J) has been studied and plotted in Figure 6. Non-parallel lines corresponding to the main effects in Figure 6 indicate the presence of interaction between spectrum observation time and all the studied features. The study of interactions between extracted features has been discarded because they have been extracted from the same source: Haar transforms coefficients derived from frequency domain signal and therefore are not considered to be fully independent variables. The most prominent interactions are observed between 75th percentile and observation time, SNR as a classification feature and observation time, root-mean-square level and observation time, skewness and observation time, kurtoses, and observation time.

7. Conclusions and future work

In the scope of this work we have studied the possibility to classify the modulation type using instantaneous values of the time domain signal and SNR as inputs to the classifier. The main advantage of this approach is that it has potential to increase throughput by shortening the spectrum observation time and decreases the computational complexity: the raw values of in-phase and quadrature components of the signal are used as an input to the classifier and therefore there is no need for any preprocessing or feature extraction. This approach is applicable for binary classification between BPSK and 2FSK modulations, however it fails in case the classification task is extended to multiple classes containing higher order modulations. The possible reason for this is the missing data that is not included in the pair of in-phase and quadrature component. If the input values pair lies on the x axis the modulation could be classified as any of four types: BPSK, 2FSK, 8PSK or 16PSK. In case of the input pairs are located on the y axis it could be classified as any of three modulation classes: 2FSK, 8PSK, 16PSK. We have confirmed that to reach the classification accuracy of at least 85% or higher for classification between multiple high order modulations it is required to observe the received signal and perform feature extraction. In this study we have studied nine spectral/based statistical features derived from the wavelet coefficients obtained by applying the Haar wavelet transform to the frequency domain received signal. The highest classification speed of 170 000 objects/second and 100 % classification accuracy has been demonstrated by fine decision trees using as the single input kurtosis (C) derived from the wavelet coefficients derived from signal observed during 100 microseconds for AWGN channel. For line-of-sight fading Rician channel with AWGN demonstrated classification speed is slower 130 000 objects/s. Skewness and kurtosis have shown the highest main effect on classification accuracy for fine decision trees. The mean value demonstrated the highest main effects on classification speed for fine decision trees, fine KNN, weighted KNN, ensemble bagged trees, ensemble subspace KNN.

In the scope of the future work it is planned to implement the proposed fine decision trees classification algorithm on our target application hardware. Also, it is intended to extend this work to the multiple propagation channel models including Rayleigh and Rician for doppler shifts corresponding to target application embedded in the ground and aerial unmanned platforms moving with speeds 100-200 km/hour.

References

[1] I. Valieva, M. Björkman, J. Åkerberg, M. Ekström, I. Voitenko, "Multiple Machine Learning Algorithms Comparison for Modulation Type Classification for Efficient Cognitive Radio," in MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM), 318–323, 2019, doi:10.1109/MILCOM47813.2019.9020735.

[2] D. Rosker, D. M., "DARPA Spectrum Collaboration Challenge (SC2).," DARPA Spectrum Collaboration Challenge (SC2), 2020.

[3] C. Xin, Spectrum Sharing for Wireless Communications, 2015.

[4] O. Fink, Q. Wang, M. Svensén, P. Dersin, W.-J. Lee, M. Ducoffe, "Potential, challenges and future directions for deep learning in prognostics and health management applications," Engineering Applications of Artificial Intelligence, **92**, 103678, 2020, doi:https://doi.org/10.1016/j.engappai.2020.103678.

[5] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y.-D. Yao, "Modulation classification using convolutional Neural Network based deep learning model," 1–5, 2017, doi:10.1109/WOCC.2017.7929000.

[6] kaiyuan jiang, J. Zhang, H. Wu, A. Wang, Y. Iwahori, "A Novel Digital Modulation Recognition Algorithm Based on Deep Convolutional Neural Network," Applied Sciences, 2020, doi:10.3390/app10031166.

[7] Mathworks., Modulation Classification with Deep Learning., 2020.

[8] M. Véstias, R. Duarte, J. Sousa, H. Neto, "Fast Convolutional Neural

Networks in Low Density FPGAs Using Zero-Skipping and Weight Pruning," Electronics, **8**, 2019, doi:10.3390/electronics8111321.
[9] Z. Zhu, Automatic Modulation Classification: Principles, Algorithms and Applications, John Wiley & Sons, Incorporated, 2015.
[10] A. Kubankova, H. Atassi, A. Abilov, "Selection of optimal features for digital modulation recognition," 2011.
[11] W. Gardner, Cyclostationarity in Communications and Signal Processing, IEEE Press, New York, 1994.
[12] W.A. Gardner, C.M. Spooner, "Cyclic spectral analysis for signal detection and modulation recognition," in MILCOM 88, 21st Century Military Communications - What's Possible?'. Conference record. Military Communications Conference, **2**, 419–424, 1988, doi:10.1109/MILCOM.1988.13425.
[13] C. Cormio, K.R. Chowdhury, "A survey on MAC protocols for cognitive radio networks," Ad Hoc Networks, **7**(7), 1315–1329, 2009, doi:https://doi.org/10.1016/j.adhoc.2009.01.002.
[14] M. Raina, G. Aujla, "An Overview of Spectrum Sensing and its Techniques," IOSR Journal of Computer Engineering, **16**, 64–73, 2014, doi:10.9790/0661-16316473.
[15] K. Hinkelmann, Design and analysis of experiments., John Wiley & Sons, Inc., 2011.
[16] C. Rajan, Statistics for Scientists and Engineers, John Wiley & Sons, Incorporated, 2015.
[17] Mathworks, Mean or median absolute deviation, 2020.
[18] Mathworks, Root-mean-square level, 2020.
[19] Mathworks, Zero Crossing Rate, Nov. 2020.
[20] Mathworks, Interquartile range, 2020.
[21] Mathworks, 75th Percentile, 2020.
[22] Mathworks, Fading Channels, 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 660-668 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Issues in File Caching and Virtual Memory Paging with Fast SCM Storage

Yunjoo Park, Hyokyung Bahn*

Department of Computer Science & Engineering, Ewha Womans University, 03760, South Korea

ARTICLE INFO

Article history:
Received: 09 August, 2020
Accepted: 21 September, 2020
Online: 05 October, 2020

Keywords:
Storage-Class Memory (SCM)
File Caching
Virtual Memory Paging

ABSTRACT

Storage-Class Memory (SCM) like Optane™ has advanced as a fast storage medium, and conventional memory management systems designed for the hard disk storage need to be reconsidered. In this article, we revisit the memory management system that adopts SCM as the underlying storage medium and discuss the issues in two layers: file caching and virtual memory paging. Our first observation shows that file caching in the SCM storage is profitable only if the cached data is referenced more than once, which is different from the file caching in hard disks, where a single hit is also beneficial. Our second observation in virtual memory paging shows that the page size in the SCM storage is sensitive to the memory system performance due to the influence of memory address translation and storage access cost. Our simulation studies show that the performance of paging systems can be improved by adjusting the page size appropriately considering application characteristics, storage types, and available memory capacities. However, the page size will not be a significant issue in mobile platforms like Android, where applications are killed before the memory space is exhausted, making situations simpler. We expect that the analysis shown in this article will be useful in configuring file caches and paging systems with the emerging SCM storage.

1. Introduction

With the large performance gap between hard disk drive (HDD) and dynamic random-access memory (DRAM), the main purpose of memory management in computing systems has been the minimization of disk I/Os [1, 2]. The access latency of hard disks is more than tens of milliseconds, which is five to six orders of magnitude larger than DRAM's access latency. Meanwhile, due to the rapid improvement of storage access time by the adoption of flash-based solid state drive (SSD) and storage-class memory (SCM), the extremely large performance gap has been decreased [3-5]. The access latency of the flash storage is less than fifty milliseconds, and hence the performance gap of storage and memory becomes less than 3 orders of magnitude. Such trends have been speeded up by the commercialization of SCM whose access latency is just 1 or 2 orders of magnitude slower than DRAM [6, 7].

A lot of patents related to the detailed architectures and algorithms of SCM management have been suggested, and Intel manufactured the commercial product of SCM, called Optane™ [8, 9]. Owing to its desirable features like high performance, low energy consumption, and long write endurance, SCM is anticipated to be adopted in the storage systems like flash SSD and hard disks [10-13].

SCM can also be adopted in the main memory system because it allows byte-accesses like DRAM but consumes less energy because it is a non-volatile medium [14]. However, the access latency of SCM is longer than that of DRAM, and hence it is now considered as high-end storage or additional memory that can be used together with DRAM. Although SCM may be used as either memory or storage, this article focuses on storage. Since the performance gap of storage and memory becomes small by adopting SCM, memory management systems targeting at slow hard disk storage need to be revisited.

In this article, we quantify the performance of systems based on SCM storage and analyze a couple of issues in the management of main memory under the SCM-based storage. In particular, this article analyzes two memory management hierarchies affected by the acceleration of storage devices, *file caching* and *virtual memory paging*.

File caching preserves file data read from the storage to the main memory area called the file cache and services requests for the same data from the cache without accessing storage. The

*Corresponding Author: H. Bahn, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Republic of Korea, +82-2-3277-4247, bahn@ewha.ac.kr.

purpose of file caching is to minimize the storage access frequency. However, since the storage access latency is fast enough by making use of SCM, it is questionable whether file caching is still needed. To validate this, we investigate the condition that file caching is effective according as the storage access latency is varied. Our preliminary study exhibits that a storage access takes about 30% more latency than a cache access even though the same data is accessed and the access latency of storage and cache is identical. This is because there exists heavy software stack to passing through the storage media. By observing this result, our finding is that file caching is still needed for SCM-based storage if the data in the cache is requested more than once after inserted into the cache. That is, since file caching needs additional time for inserting the requested data into the cache, more than a certain cache hits are necessary for caching to be profitable, and we show that this is at least twice for SCM storage. Note that low-end storage media like hard disks require only a single hit for caching to gain.

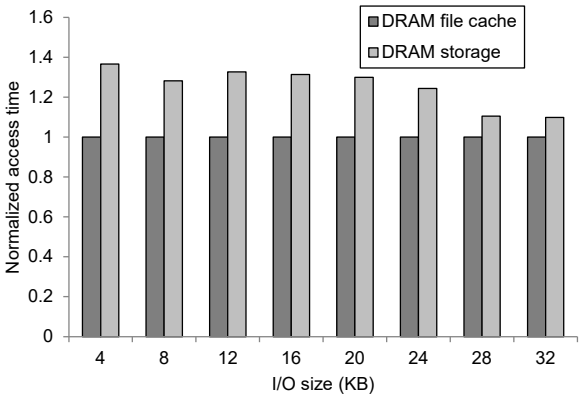


Figure 1: Latency of accessing data from the file cache and the DRAM-based storage.

Our second observation focuses on virtual memory paging if SCM is made use of as storage media. Similar to file caching, the purpose of virtual memory paging is the minimization of storage accesses, which is also called page faults, when storage is hard disk. Since the storage medium becomes fast enough, we observe that page fault handling may not be the main performance bottleneck of virtual memory paging. In particular, as the latency of storage access becomes small by adopting SCM, the bottleneck of the page access latency may be shifted to memory address translation. Note that to access a memory page, translating the address should be done first and then the data in the page is accessed from either memory or storage. As the data access is accelerated by the reduced storage access cost, the address translation process may be a new bottleneck by accessing page tables.

This article quantifies what will be the main bottleneck of virtual memory paging as the storage performance and the page size are changed. Based on our analysis, we discover the following two phenomena. First, a small page is not efficient in terms of the page fault rate but it performs well in terms of the data access latency. The reason is that a page fault handling latency is strongly related to the page size when the storage is SCM, which does not need seek movement. Second, even though a small page is efficient with respect to the latency of data access, it degrades the latency of address translation by increasing TLB miss counts. Due to this reason, deciding a page size needs to consider the trade-off relation of the data access latency and the address translation latency.

Our simulation results exhibit that the performance of virtual memory paging can be improved by adjusting the page size appropriately considering application characteristics, storage types, and available memory sizes. However, the page size will not be a significant issue in mobile platforms like Android, where applications are killed before memory space is exhausted, making situations simpler. We expect that the analysis shown in this article will be useful in configuring file caches and paging systems with the emerging SCM storage.

The remaining part of this article is organized as follows. In Section 2, we quantify the performance implication of file caching as the storage medium changes from hard disk to SCM. Section 3 anatomizes the virtual memory paging performances with high-performance SCM storage particularly focusing on page sizes. Section 4 presents the experimental results through conducting simulation experiments to observe the implications of SCM storage based memory management systems. Section 5 discusses the adoption of our model to the mobile application environments. Finally, we present the conclusion of this article in Section 6.

2. File Caching for SCM Storage

This section quantifies the efficiency of file caching as the storage medium changes from hard disk to SCM. To this end, we measure the latency of file system operations when DRAM is used as storage. Note that DRAM is volatile, but we use a certain area of DRAM as a storage partition just to see the effect of fast storage media. Note also that this situation implies the optimistic performance of SCM storage. We added a profiler to Ext4 for measuring the latency of directly accessing data from the DRAM storage and the latency of accessing data in the DRAM file cache.

Figure 1 depicts the access latency of DRAM storage and DRAM file cache as the size of data accesses changes. In this graph, we measure each case 10 times and plot their average. As can be seen in the graph, the access latency from DRAM storage is 30% longer than accessing the same data from DRAM file cache. Although the same DRAM is used for cache and storage, the performance gap occurs due to the existence of software I/O stack. When considering these results, file caching can be still necessary to buffer the latency gap between storage and memory although their access latencies are identical. However, as the gap is very small, some conditions need to be met for file caching to be beneficial.

When we use file caching, the accessed data should be stored into the cache, which requires additional latency. The access latency of DRAM storage shown in Figure 1 does not include this latency, and hence the accessed data is delivered directly to the user memory. If we use file caching, the accessed data is firstly stored in the file cache and then transferred to the user memory. The overhead of this additional copy operation in memory is not negligible when SCM storage is used because the time overhead to perform a memory copy is similar to that of a storage access. Thus, the gain of file caching is small or there are no profits at all because of this trade-off. Due to this reason, file caching is beneficial only when the benefit of subsequent cache hits is larger than the cost of

the additional memory copy operation. That is, cache hits are important for a cached data to gain, which differs from low-end storage media like hard disk drives where a single hit of data is sufficient for caching to gain. To quantify this, we measure the data access time t from the file cache, storage access time T, and time to access storage including the file caching cost T_m. Then, we evaluate the condition of caching to be beneficial with respect to the cache hit counts. The following equations represent the latency to access data with file cache T_c and without file cache T_x, respectively.

$$T_c = T_m + (n-1)\,t \quad (1)$$

$$T_x = n\,T \quad (2)$$

where n is the total access count for the data. File caching gains when T_c is less than T_x as follows.

$$n > (T_m - t) / (T - t) \quad (3)$$

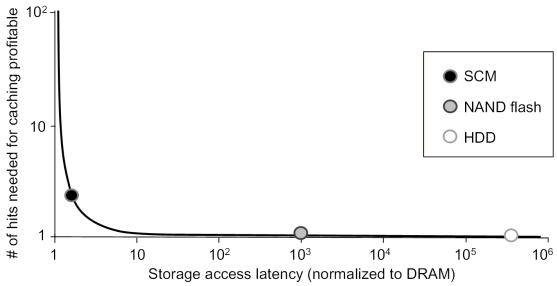
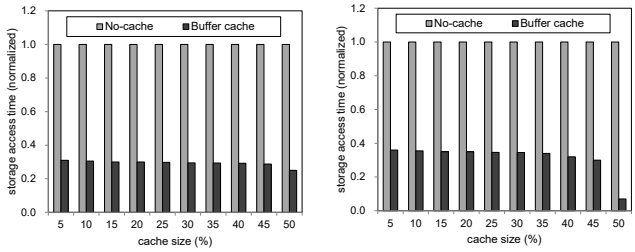


Figure 2: Hit counts necessary for a cached item to be beneficial as a function of the latency for accessing the storage media.

Based on Expression (3), we estimate the cache hit counts needed for caching to be beneficial. Figure 2 depicts the required hit counts when the storage media is flash memory, hard disk, and SCM. As plotted in the figure, the cache hit count necessary for caching to be profitable increases significantly as the performance gap of storage and memory is small. In the case of SCM storage whose access latency is only 30% slower than memory, the cache hit count necessary for caching to be profitable is two while resident in the cache. Hence, if a storage data is requested now and it will not be reused at least twice in the future, it would be better not to store it in the cache for performance improvement.

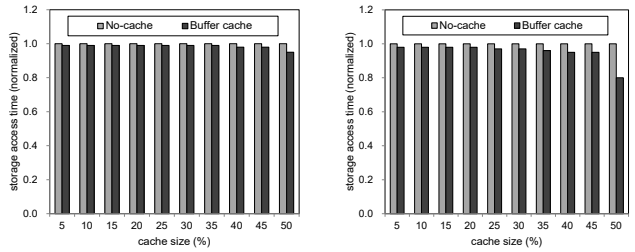
To quantify the influence of file caching in practical situations, we collect file request traces during the execution of two popular storage benchmarks, and replay them as the storage media is varied. The captured traces are web server and proxy server. We investigate the storage access latency with/without file caching under the two workload conditions. For the cache eviction algorithm, we use the least-recently-used (LRU), the most commonly adopted algorithm in file caching. LRU selects the data that was accessed the oldest among all data in the file cache and evicts it if there are no cache spaces to insert new data items.

Figure 3 depicts the storage access time with/without file caching when hard disk storage is used as the size of cache changes. Note that the cache size of 100% means that the cache size is equal to the total footprint of workloads, implying that the eviction algorithm is not necessary. This is not a realistic situation and in practical environments, the cache size is less than 50%. As can be seen from this figure, the effectiveness of file caching is significant in case of the hard disk storage. Specifically, file caching accelerates the storage performance by 70%, which is possible because hard disk is very slower than the file cache, and hence decreasing the disk I/O access counts by file caching is effective.



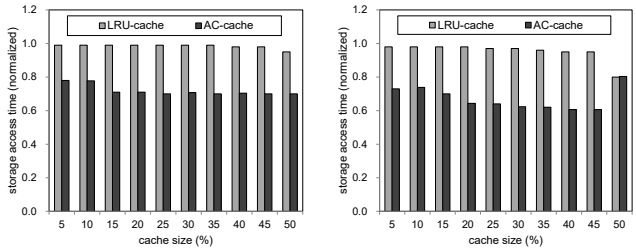
(a) Proxy (b) Web

Figure 3: Effectiveness of file caching under hard disk storage.



(a) Proxy (b) Web

Figure 4: Effectiveness of file caching under SCM storage.



(a) Proxy (b) Web

Figure 5: Performance of file caching with/without admission-controll (AC) under SCM.

Figure 4 depicts the storage access time without/with file caching when the storage medium is SCM. As can be seen from this figure, the effectiveness of file caching is insignificant in case of the SCM storage. In particular, the performance improvement by using file caching is less than 3% with the SCM storage. This apparently exhibits that the effectiveness of file caching decreases significantly as the performance gap of storage and memory becomes small. However, we conducted some new experiments, and show that file caching can be still effective by judicious management.

Figure 5 depicts the storage access time with file caching when the SCM storage is used, but we differentiate the result with two management policies. In particular, LRU-cache adopts the same configuration of Figure 4, which stores all data requested in the file cache and evicts the least-recently-used data if free cache space is necessary. The other graph, denoted by AC-cache (admission-controlled cache), is plotted by allowing the insertion of data into the file cache only if it is used more than once. This allows only

data used multiple times to be cached, thereby filtering out useless single-accessing data. Even though this is a simple technique, the effectiveness of AC-cache is significant as shown in Figure 5. The performance improvement ranges 20-40%. The implication of this result is that file caching is still useful when SCM storage is used, but management policies should be appropriately devised in order to get effective results.

3. Virtual Memory Paging for SCM Storage

This section presents the simulation results of virtual memory paging when we use SCM as storage.

3.1. Access Latency of a Virtual Memory Page

To use a memory page in virtual memory systems, translation of memory addresses between logical address and physical address needs to be conducted. This is done by referencing the page table, which is located at main memory. To improve the address translation performances, a certain part of the page table is cached in the Translation Look-aside Buffer (TLB), which is faster than main memory [15]. When a memory page is requested, address translation by TLB is tried first. The page table is referenced if address translation by TLB misses. Then, the memory data is accessed with the translated physical address. However, if the data does not reside in memory, storage should be accessed, which is called a page fault. Assume that the cache miss rate of TLB is R_T, and the page fault rate is R_F. Then the memory access time T_{TOTAL} can be denoted by

$$T_{TOTAL} = T_{ADDR} + T_{DATA} \quad (4)$$

$$T_{ADDR} = (1 - R_T) * t_\varepsilon + R_T * (t_\varepsilon + t_\tau) \quad (5)$$

$$T_{DATA} = (1 - R_F) * t_\tau + R_F * (t_\tau + T_{PF}) \quad (6)$$

where T_{ADDR} is the time required for performing address translation, T_{DATA} is the latency required for data access, t_ε is the latency required for TLB access, t_τ is the latency for accessing main memory, and T_{PF} is the latency for page fault handling including storage access time.

3.2. Eviction Policies

Since the size of TLB entries is fixed, adding an address translation data to TLB requires the eviction of a certain entry when no TLB entry is available. To choose an eviction victim, we adopt the least-recently-used (LRU) eviction policy, which is a representative algorithm adopted in TLB entry eviction [15].

Because the size of free pages in main memory is fixed, another eviction policy is necessary. Specifically, if the page requested is not in memory and should be loaded from storage, but there is no free page in memory, eviction of a page from memory is necessary. We use the second-chance algorithm, which is a popular eviction policy used in virtual memory systems, in our experiments [14].

The second-chance algorithm investigates if a page has been used recently or not by utilizing the reference bit of each page. When a page is used, the reference bit of that page becomes one. If a free page frame is needed, the second-chance algorithm investigates the reference bits of all pages in memory sequentially, and discards the page firstly found with its reference bit of zero. For each page whose reference bit is one while investigation processes, the second-chance algorithm resets the bit to zero, instead of discarding the page from memory. Thus, if a page is not accessed until the next investigation of the second-chance algorithm, it is evicted.

With this basic configurations, we conduct our simulations to quantify the efficiency of memory management systems when we use SCM-based storage.

3.3. Page Size and Prefetching

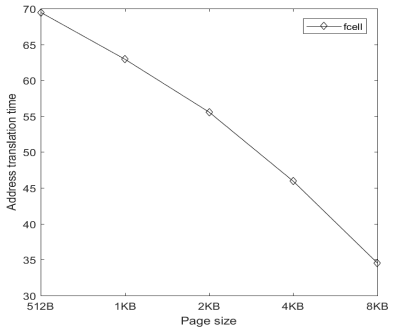
The current operating systems usually manage virtual memory by the unit of *page*. The size of a page is commonly set to 4 kilobytes, which is also the default page size of Linux. However, operating systems also allow the prefetching option, which loads maximum 128 pages together from the secondary storage when a page requested is not in memory. The rationale of this is to consider the characteristics of hard disk, which requires the basic cost for each access consisting of seek latency and rotational latency, accounting for the main part of storage access latency irrespective of the size of data loaded. This implies that loading large data in each request is efficient in case of hard disk based storage. Recent operating systems also support a huge page whose size is up to 4MB. This is also for hard disk storage systems, but it will not be efficient for SCM, in which the page size should be small because storage is fast and there are no seek time or rotational latency.

Nevertheless, it is not feasible to decrease the page size because memory address translation will be efficient with a large page size. That is, TLB with a large page size covers more memory address spaces, leading to improved address translation latency. Then, we need to decide the page size for SCM-based storage by considering the overall effect of storage access and address translation.

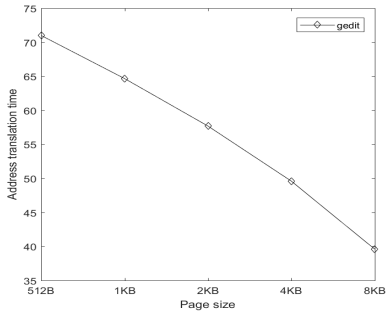
Also, the latency of SCM should be considered, which is optimistically as fast as DRAM, but can be up to 100 or 1000 times slower than DRAM. Due to the variance of performance gap between storage and memory, the relative impact of memory address translation compared to storage access is also varied. In particular, as the performance of SCM approaches that of DRAM, memory address translation incurs relatively more cost. On the contrary, as the performance gap of DRAM and SCM increases, the relative overhead of a storage access will be large. The relation of storage access and address translation will be analyzed in the next section as the performance of SCM is varied.

4. Performance Analysis

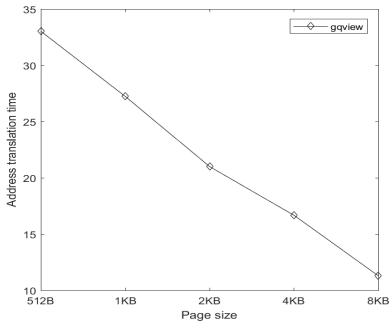
We conduct various simulations to investigate the efficiency of memory management subsystem as SCM storage is adopted. In particular, the impact of the page size on memory performance is investigated. The minimum page size is set to 512 bytes because the size of a page should be at least the block size of the last-level cache memory. We use the virtual memory reference traces captured by the Cachegrind tool of the Valgrind [16, 17]. We collect the virtual memory reference traces from five desktop workloads as listed in Table 1 [18-22].



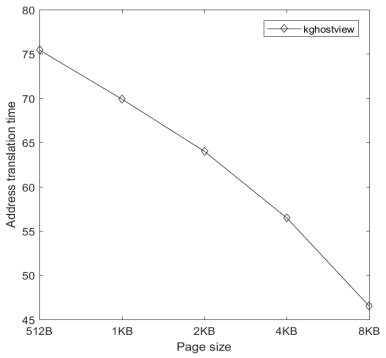
(a) freecell



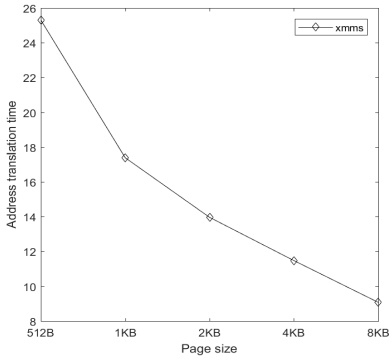
(b) gedit



(c) gqview

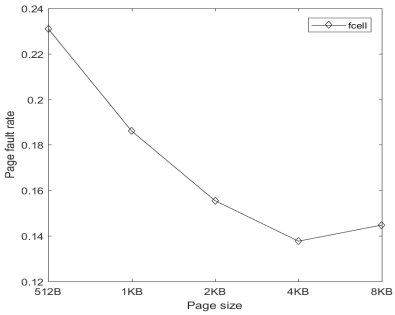


(d) kghostview

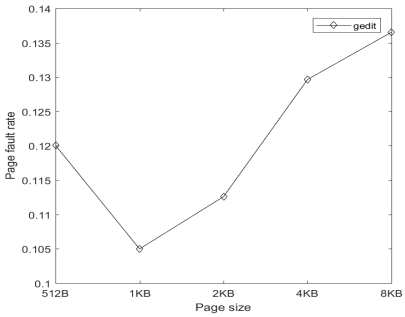


(e) xmms

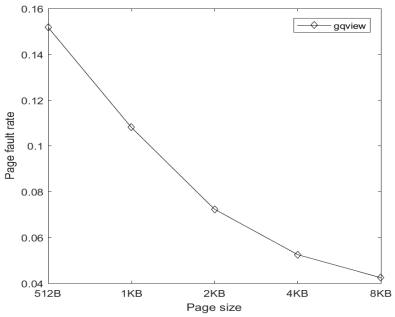
Figure 6: Memory address translation time of desktop workloads as the page size changes.



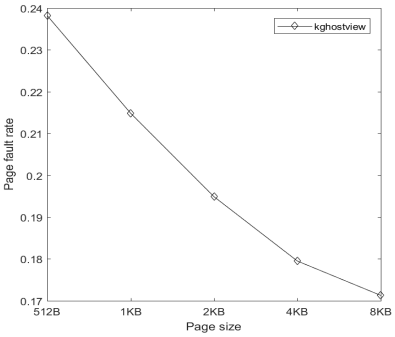
(a) freecell



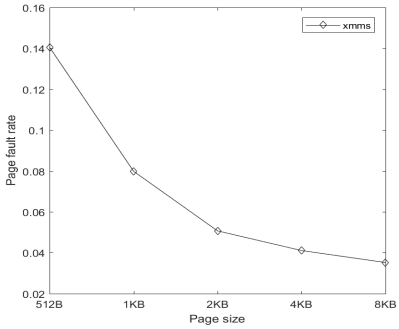
(b) gedit



(c) gqview

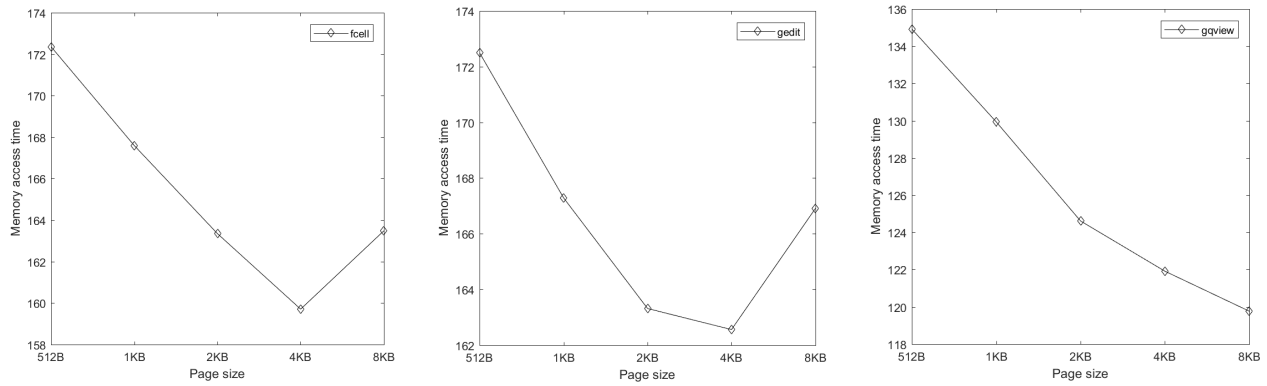


(d) kghostview

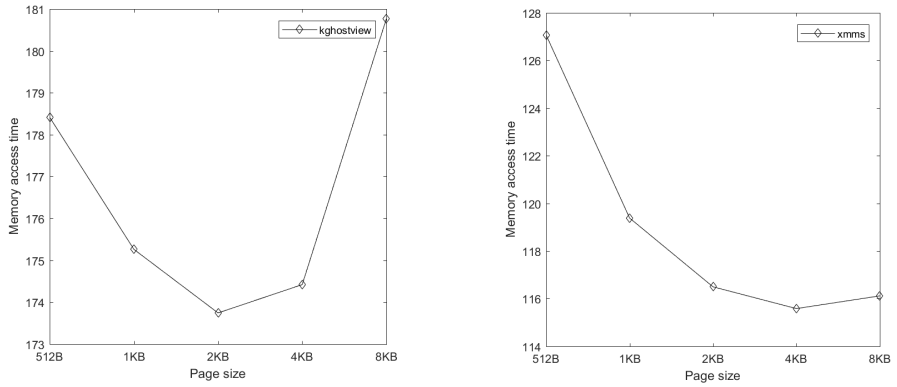


(e) xmms

Figure 7: Page fault rate of desktop workloads as the page size changes.

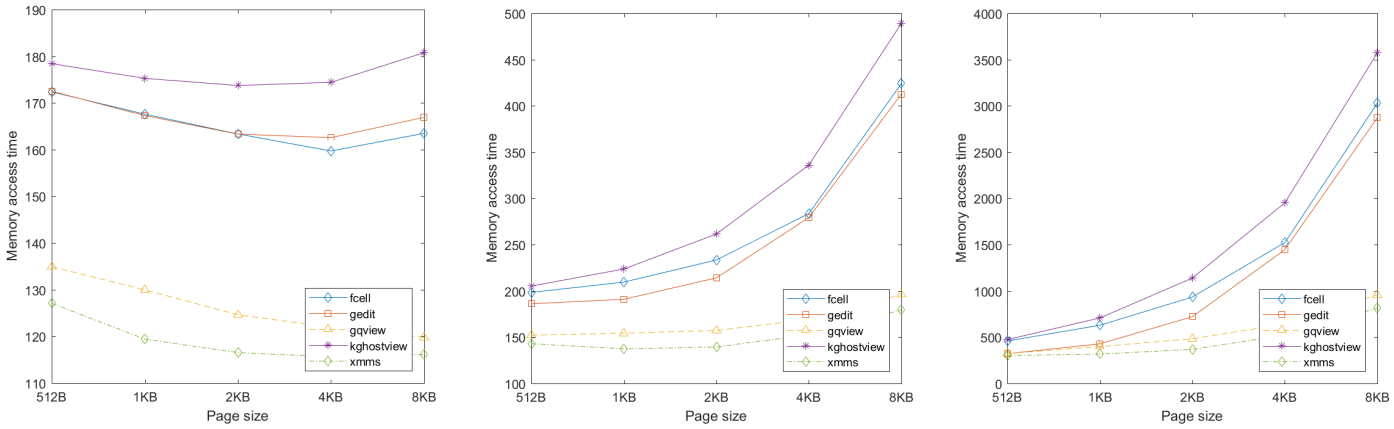


(a) freecell (b) gedit (c) gqview



(d) kghostview (e) xmms

Figure 8: Memory access time of desktop workloads as the page size changes.



(a) 1x of DRAM (b) 10x of DRAM (c) 100x of DRAM

Figure 9: Total memory latency of desktop workloads as the page size and storage access latency change.

4.1. Memory Address Translation

This section discusses the influence of the page size on the performance of memory address translation. Figure 6 depicts the memory address translation latency while running five desktop applications as the page size is changed. As the figure depicts, the latency of memory address translation decreases according as the page size grows. The reason is that the fixed TLB entries account for the address translations of wider memory spaces if the page size becomes large. An increased hit rate of TLB leads to the decreased latency of memory address translation by reducing the number of page table references. Nevertheless, it is known that the hit rate of TLB cannot increase any more when the page size is larger than a certain threshold [23]. Thus, the conclusion of this experiment is that a large page will perform well in terms of the memory address translation, but the page size does not need to be increased any longer after a certain large size.

Table 1: Characteristics of desktop workload

Application	Memory footprint (MB)	Memory access counts		
		Write	Read	Total
freecell [18]	9.84	60,040	430,135	490,175
gedit [19]	14.12	132,822	1,600,941	1,733,763
Gqview [20]	7.26	645,399	265,286	610,685
kghostview [21]	16.98	103,540	1,442,595	1,546,135
xmms [22]	7.86	978,242	190,697	1,168,939

4.2. Storage Access Frequency

This section discusses the impact of the page size on the storage access frequency. Figure 7 depicts the page fault rate of the five applications as the page size changes. As the figure shows, the best page size is not identical for different applications we considered. In most cases, a large page improves the page fault rate. In particular, the page size of 8 kilobytes exhibits the best results for gqview, kghostview, and xmms. On the contrary, the page size that exhibits the best result is 1 kilobytes for gedit and 4 kilobytes for freecell, respectively. The best page size relies not only on the application's characteristics but also on the free memory situation of the total system. When the memory size is insufficient for the given workload environment, decreasing the page size will perform well to retrieve only the data requested at that time. In contrast, when the available memory is sufficient, increasing the page size will be a better choice in terms of the storage access frequency. Also, when the workload is composed of sequential reference patterns, increasing the page size can reduce the number of storage accesses by retrieving large successive data together.

4.3. Total Memory access time

As hard disk requires large seek overhead for each storage access irrespective of the data size, decreasing the storage access frequency can lead to the improvement of data access latency. However, in case of SCM storage, as the data size becomes large, a storage access requires more time to load data. Assume a storage reference stream that is composed of sequential patterns and let us think of the storage access frequency for virtual memory paging that makes use of a large page size and a small page size. It is clear that a large page size is efficient in terms of the storage access frequency, but it does not essentially lead to the enhancement of data access latency because the storage access time becomes large for each I/O with SCM. Hence, instead of improving the storage access frequency, our aim is to reduce the data access latency. In other words, the page fault rate is not a fair performance index for SCM-based virtual memory paging, but the data access latency can be an alternative metric for SCM-based storage.

Unlike the page fault rate case, address translation latency has almost linear relation with the TLB miss rate because the access latency of a page table entry incurred by each TLB miss is identical though the page size is different.

Figure 8 depicts the total memory access latency of the desktop applications we considered as the page size changes. As the figures show, the best page size is not the same for different applications, and is also different from the results of Figure 7 that plots the page fault rate. From this result, we can see that the storage access frequency is not the primary factor of performance in case of SCM storage media. Also, this figure shows that the address translation latency can be another significant factor that influences the memory access time in case of fast SCM storage even though the effectiveness of address translation is smaller than the actual data access. Because a trade-off relation exists between data access and address translation in deciding an appropriate page size, we cannot simply conclude the page size as small or large but a judicious management is necessary for deciding the page size with given environments.

4.4. Relative Storage Performance

Figure 9 depicts the total memory latency for each application as the access latency of storage media changes. Specifically, Figures 9(a), 9(b), and 9(c) represent the total memory latency when the relative access latency of SCM is identical to DRAM, 10 times that of DRAM, and 100 times that of DRAM, respectively. As this figure shows, the page size that performs the best is not the same for each case as the access latency of storage media changes. If the performance of SCM is similar to DRAM as shown in Figure 9(a), the address translation latency becomes important, and hence increasing the page size performs relatively well. In other words, the role of TLB becomes significant in such environments, and the memory access time is improved with the page size of 4KB or 8KB. We cannot determine the best page size for all workloads as it depends on the total memory capacity as well as the workload characteristics. Figures 9(b) and 9(c) plot the total memory latency when the access latency of SCM media is 10 times and 100 times slower than DRAM, respectively. As these two figures show, the memory access time becomes better when we decrease the page size as small as possible. The reason is that the data access latency affects significantly if the storage becomes slow and thus the address translation procedure is less important. The best memory access time can be obtained if we set the page size to 512B for all cases.

Table 2: Characteristics of mobile workload.

Application	Memory footprint (MB)	Memory access counts		
		Write	Read	Total
facebook [24]	198.66	2,045,716	11,607,339	13,653,055
angrybirds [25]	76.94	3,822,479	14,368,068	18,201,717
youtube [26]	68.64	3,162,229	15,034,275	18,196,504
farmstory [27]	53.74	2,101,818	13,122,852	15,224,670
chrome [28]	259.86	4,104,436	16,895,563	20,999,999

When considering these overall situations, we can obtain the best memory access time not by fixing the page size to a constant but by varying relying on the application and storage characteristics. Hence, the common page size of 4KB will not be a good choice for all cases and needs to be changed appropriately if we use the SCM-based fast storage media.

5. Adopting the Model to Android Mobile Applications

In this section, we adopt our model to Android mobile applications to see the effect of the page size in SCM-based storage. To investigate a wide spectrum of Android applications, we collect virtual memory reference traces from five Android applications, namely, facebook a social network service [24], angrybirds a game [25], youtube an online streaming service [26], farmstory an online game [27], and chrome a web browser [28]. Table 2 lists the characteristics of the workloads.

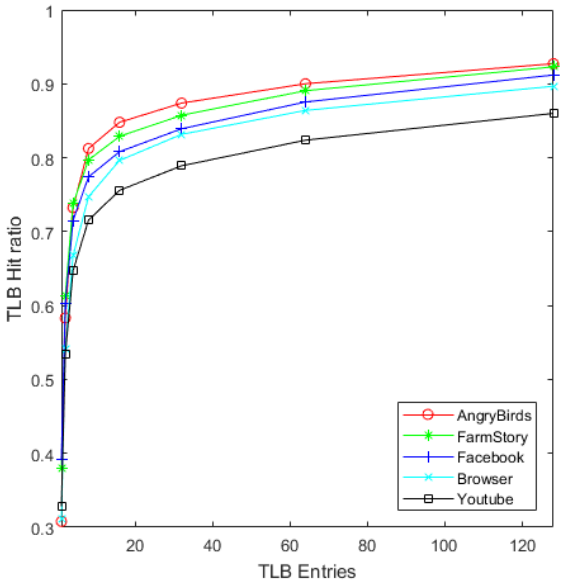


Figure 10: TLB hit rate of mobile workloads as the TLB size changes.

With these memory reference traces, we conduct extensive simulations, and analyze them similar to the desktop system cases in the previous section. Figure 10 depicts the TLB hit rate while executing five mobile applications as a function of the TLB size. As can be seen from the figure, the TLB hit rate is improved significantly as the TLB entries increase. The five applications show similar curve trends. Specifically, the TLB performance is improved sharply when the number of TLB entries is less than 20 and the slope of the curve becomes gradual after that point. Figure 11 depicts the TLB hit ratio as the page size changes. As can be seen from this figure, the TLB hit rate improves according as the page size increases. The reason is that the fixed number of TLB entries account for address translations of more memory capacity as the page size becomes large. An increased hit rate of TLB eventually leads to the decreased latency of memory address translation by reducing the number of page table references. However, the improvement of the TLB hit rate slows down when the page size is larger than a certain threshold. Thus, increasing the page size will perform well in terms of the performance of memory address translation, but it is not sensitive after a certain large page size.

Figure 12 depicts the page fault rate of the five Android applications as the page size changes. As can be seen from the figure, the page fault rate decreases as the page size increases for all applications we considered. Unlike desktop cases that do not exhibit the same trends for different applications, all mobile applications show similar results. The best page size relies on the application's characteristics in desktop cases, but a large page commonly improves the page fault rate of all mobile applications we considered. If the memory space is not sufficient for the given workload environment, large pages will not perform well due to retrieving the data not requested now. However, our Android application experiments show that the memory space is not exhausted in any applications. This is because Android does not use swap but kills applications if free memory space becomes small.

Figure 13 depicts the average memory access time (AMAT) of the five Android applications we considered as the page size changes. This graph separately represents the AMAT when the relative latency of SCM is identical to DRAM, 10 times that of DRAM, 100 times that of DRAM, and 1000 times that of DRAM, respectively. In the previous section, the page size that performs the best in desktop environments is not identical for each case as the performance of storage media is varied. If the performance of storage is similar to that of DRAM, increasing the page size performs relatively well, but if SCM becomes slow, decreasing the page size works well in desktop environments. However, as shown in Figure 13, the best results in mobile environments are obtained when the page size becomes large, and there are no significant differences in memory performances when the page size is larger than 1 kilobytes. This is also related to the memory management policies of Android, in which applications are killed before memory space is exhausted, and thus increasing the page size does not incur problems like desktop environments.

In summary, our conclusion is that the page size will not be a significant issue in mobile environments although fast SCM is used as the storage media.

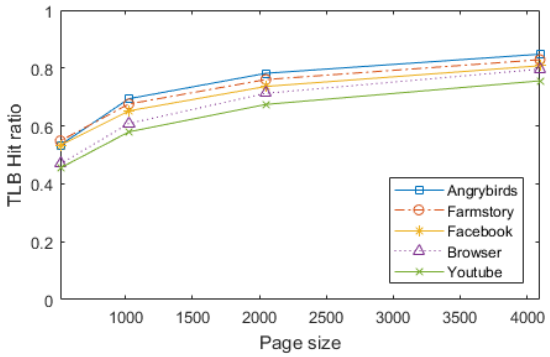


Figure 11: TLB hit rate of mobile workloads as the page size changes.

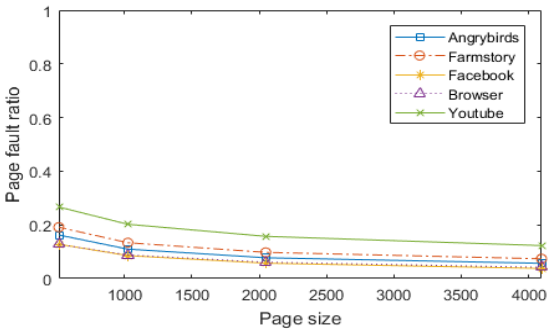


Figure 12: Page fault rate of mobile workloads as the page size changes.

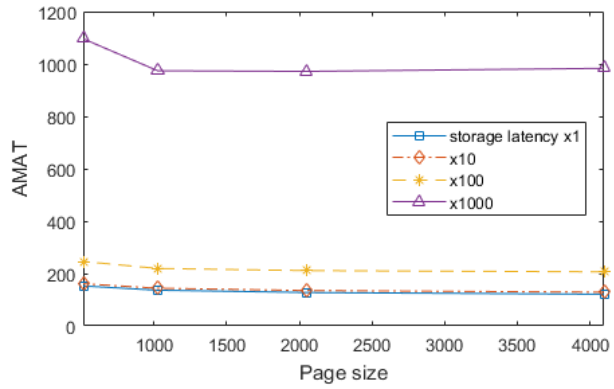


Figure 13: Average memory access time (AMAT) of mobile workloads as the page size and storage latency change.

6. Conclusion

In this article, we considered the effectiveness of traditional memory management systems and configurations if fast storage such as SCM is used as the storage media. Our study discussed the implication of SCM storage with respect to the two important memory/storage subsystems, file caching and virtual memory paging. The analysis results with respect to the file caching showed that caching file blocks is effective if the blocks retrieved from SCM storage is used more than once after loading to memory. Thus, the admission of file blocks to the cache should be controlled for managing the file cache appropriately in SCM-based storage media. In the case of virtual memory paging, our study showed that the number of storage accesses may not be the most important performance index, but the memory address translation process will also become important. In particular, our analysis showed that a common page size of 4 kilobytes is not a good choice for all cases when SCM storage is adopted, and determining the page size by decreasing or increasing it is needed because of the trade-off relation in storage access and memory address translation. However, we also showed that the page size is not a significant issue in mobile platforms like Android, where applications are killed before memory space is exhausted, making situations simpler. We expect that the findings in this article will be meaningful for emerging memory and storage management systems that make use of fast SCM media. In the future, we will perform the evaluation of our strategy with the complete system configurations consisting of the commercialized SCM products.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an extension of work originally presented in the 6th IEEE Int'l Conf. on Information Science and Control Engineering [1]. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2019R1A2C1009275) and also by the ICT R&D program of MSIP/IITP (2019-0-00074, developing system software technologies for emerging new memory that adaptively learn workload characteristics).

References

[1] Y. Park, K. Cho, and H. Bahn, "Challenges and Implications of Memory Management Systems under Fast SCM Storage," in 2019 IEEE International Conference on Information Science and Control Engineering, 190-194, 2019. https://doi.org/10.1109/ICISCE48695.2019.00046
[2] S. Ng, "Advances in disk technology: performance issues," Computer, **31**(5), 75-81, 1998. http://doi.org/10.1109/2.675641
[3] C. Evans, "Flash vs 3D Xpoint vs storage-class memory: which ones go where?" Computer Weekly, **3**(9), 23-26, 2018.
[4] M. Stanisavljevic et al., "Demonstration of Reliable Triple-Level-Cell (TLC) Phase-Change Memory," in 2016 IEEE International Memory Workshop, 2016. https://doi.org/10.1109/IMW.2016.7495263
[5] S. Hyun, H. Bahn, and K. Koh, "LeCramFS: an efficient compressed file system for flash-based portable consumer devices," IEEE Trans. Consumer Electron., **53**(2), 481-488, 2007. https://doi.org/10.1109/TCE.2007.381719
[6] M. K. Qureshi, V. Srinivasan, and J. A. Rivers, "Scalable high performance main memory system using phase-change memory technology," ACM SIGARCH Computer Architecture News, **37**(3), 24-33, 2009. https://doi.org/10.1145/1555815.1555760
[7] E. Lee, J. Jang, T. Kim, and H. Bahn, "On-demand snapshot: an efficient versioning file system for phase-change memory," IEEE Trans. Knowledge Data Engineering, **25**(12), 2841-2853, 2013. https://doi.org/10.1109/TKDE.2013.35
[8] R. K. Ramanujan, R. Agarwal, and G. J. Hinton, "Apparatus and Method for Implementing a Multi-level Memory Hierarchy Having Different Operating Modes," US Patent 20130268728 A1, Intel Corporation, 2013.
[9] B. Nale, R. Ramanujan, M. Swaminathan, and T. Thomas, "Memory channel that supports near memory and far memory access," PCT/US2011/054421, Intel Corporation, 2013.
[10] E. Lee, S. Yoo, and H. Bahn, "Design and implementation of a journaling file system for phase-change memory," IEEE Trans. Comput., **64**(5), 1349-1360, 2015. https://doi.org/10.1109/TC.2014.2329674
[11] S. Eilert, M. Leinwander, and G. Crisenza, "Phase change memory: A new memory enables new memory usage models," in 2009 IEEE International Memory Workshop, 2009.
[12] E. Lee, J. Kim, H. Bahn, S. Lee, and S. Noh, "Reducing Write Amplification of Flash Storage through Cooperative Data Management with NVM," ACM Transactions on Storage, **13**(2), 2017. https://doi.org/10.1145/3060146
[13] E. Lee, H. Bahn, S. Yoo, S. H. Noh, "Empirical study of NVM storage: an operating system's perspective and implications," in 2014 IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 405-410, 2014. https://doi.org/10.1109/MASCOTS.2014.56
[14] S. Lee, H. Bahn, and S. H. Noh, "CLOCK-DWF: a write-history-aware page replacement algorithm for hybrid PCM and DRAM memory architectures," IEEE Trans. Comput., **63**(9), 2187-2200, 2014. https://doi.org/10.1109/TC.2013.98
[15] B. Anita, J. B. Chen, and N. P. Jouppi, "A simulation based study of TLB performance," in 1992 Int. Symp. Computer Architecture, 114-123, 1992.
[16] Valgrind, http://valgrind.org/
[17] N. Nethercote and J. Seward, "Valgrind: a program supervision framework," Electronic Notes in Theoretical Computer Science, **89**(2), 2003. https://doi.org/10.1016/S1571-0661(04)81042-9
[18] Freecell, https://en.wikipedia.org/wiki/FreeCell
[19] Gedit, https://wiki.gnome.org/Apps/Gedit
[20] Gqview, http://gqview.sourceforge.net
[21] Kghostview, https://linuxappfinder.com/package/kghostview
[22] Xmms, https://www.xmms.org/
[23] P. Weisberg and Y. Wiseman, "Using 4KB page size for virtual memory is obsolete," in 2009 IEEE International Conference on Information Reuse & Integration, 262-265, 2009. http://doi.org/10.1109/IRI.2009.5211562
[24] Facebook, https://www.facebook.com/
[25] Angrybirds, https://www.angrybirds.com/
[26] Youtube, https://www.youtube.com/
[27] Farmstory, https://play.google.com/store/apps/details?id=com.teamlava.farmstory
[28] Chrome, https://www.google.com/chrome/

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 4, 667-675 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design of a Flapping Wings Butterfly Robot based on Aerodynamics Force

Kanjanapan Sukvichai*, Kan Yajai

Faculty of Enginerring, Kasetsart University, Bangkok, 10900, Thailand

ARTICLE INFO

Article history:
Received: 06 July, 2020
Accepted: 11 August, 2020
Online: 25 August, 2020

Keywords:
Flapping Wings
Butterfly Robot
Aerodynamics

ABSTRACT

Insect robots are always amazed by humans due to their ability to fly using a wing flapping mechanism. The butterfly robot was designed in this research based on aerodynamics and aeroelastic especially for designing a flapping mechanism due to its complexity. A butterfly wing structure was designed by considering aerodynamics forces based on assumptions. Aerodynamic equations were derived in order to obtain lift and thrust forces that acted on a small wing section. The wing was assumed to be in the Quasi-steady state when it was analyzed based on the thin airfoil theorem. Airflow was simulated in order to obtain air pressure and vertexes acting on the wing surface when it swings in the still air. By integrating the wing section's lift force for a flapping cycle motion trajectory, the average lift force was obtained. The robot wing structure was designed based on the average lift. The real butterfly wing was used as the guideline for designing the robot wing. Each wing was fabricated from a laminar plastic sheet. Carbon fiber robs were used as an internal reinforced support structures for wing frames. The reinforced wing achieved the wing's rigidity and was considered as a thin airfoil. The flapping mechanism was designed by using two separated servo motors because of its flexibility and performance. This mechanism enables the robot's rotation without an extra actuator. The butterfly robot body was manufactured from the 3D printer using PLA material. The experiments were conducted to identify the robot performance. The designed butterfly robot can take off from the support platform and fly up to a certain height.

1. Introduction

Flying robots has the ability to fly up into the air and amazes human because of their behavior. There are two major types of flyable animals which are insects and birds. They use different flying mechanisms. The flyable robot uses several aerodynamic properties benefits for creating aerodynamics forces. The robot that mimics the insect is called micro air vehicles. This manuscript is an extension of work originally presented in the International Conference on Electronics, Information, and Communication 2020 (ICEIC2020) [1]. The flapping winged type robot has an advantage over a regular drone robot that uses rotating propellers. In general, flapping motion produces a larger lift force compared to a regular fixed wing at a specific angle of attack (AoA) which is important for designing a tiny flying robot with a high efficiency. Furthermore, a flapping wing robot model has the advantages over other flying robot models which are a vertical taking off from and landing to a surface, hovering with a specific height, and maneuvering in a complex environment. The autonomous flying robots were interested because it is complex to analyze, model and design robot mechanisms and hard to determine control algorithms because of the nonlinearity term. Delfly [2] Nanohumming bird [3], Purdue Hummingbird robot [4] and RoboBee [5] were popular developed flying robots that can taking off and hovering based on flapping wing mechanisms. Delfy and Nanohumming bird robot's flapping mechanism were designed based on a crank mechanism that required one actuator to drive both wings while RoboBee and Purdue Hummingbird robot's flapping mechanism used two separated actuators for wings. The crank flapping mechanism is easy to develop but both wings share the same angle and angular velocity trajectories yield the less degree of freedom than the two separated actuators mechanism. A crank mechanism flying robot requires an extra actuator to rotate a robot. Although these robots were successful to taking-off and hovering but the robots are small and cannot carry any large payload and the wing flapping speed is very high that consumes a lot of energy.

*Corresponding Author: Kanjanapan Sukvichai, Faculty of Engineering, Kasetsart University, Bangkok, Thailand, Tel: +662-797-0999, Email: fengkpsc@ku.ac.th

The aerodynamics analysis of the butterfly wing section was explained and derived at the beginning of this manuscript in order to reveal the aerodynamics of the wing. The wing aerodynamics equations are derived based on assumptions that wing was in the Quasi-steady state with thin airfoil properties. Then, the derived equations were used to design the butterfly robot's wing. Navier-Stokes equations are used to obtain the numerical solution of the airflow against the wing. The flapping mechanism was designed based on two separate actuators mechanism which enables a rotation. The designed robot was based on a butterfly animal that has bigger wings and flaps its wings at a slower speed than other insects yield less energy consumption and more payload. Then, the designed robot can be used in the applications in Thailand such as the survey drone in the agriculture field and in the anti-terrorist application. Lastly, the butterfly wing structure, materials, flapping mechanism and the butterfly robot are designed. The robot parts were manufactured and constructed. The designed robot was experimented and the results were explained at the end of this manuscript. The overall robot design process is displayed in Figure 1.

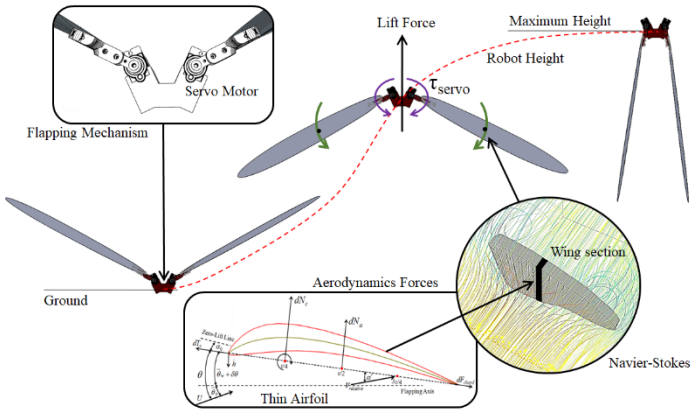


Figure 1: Overall robot design process.

2. Wing Aerodynamic Force Analysis

Aerodynamics of the butterfly wing can be analyzed by considering the small wing section along the wing span. A real butterfly wing consists of very thin membranes and a body wall made from thousands of scales and hairs, then, butterfly wing was considered as the thin airfoil. The butterfly robot model was modeled and aerodynamic forces that applied to the wing section are shown in Figure 2.

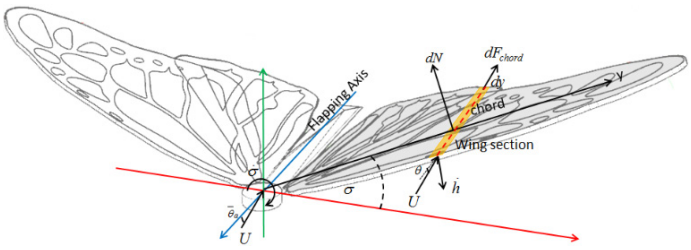


Figure 2: The butterfly robot model.

Due to the assumption that the wing's aspect ratio is large, the air flow over each section can be cosidered as a chordwised flow. The wing is in the Quasi-steady state when parameters and time-verying components are considered as time invariant constants. The wing section is shown in Figure 3. The aerodynamics and aeroelastic forces are orthogonal to the wing section's chord line and equal to (1). Aeroelastic happens in this calculation because the wing motion induces an added mass effect into the wing section.

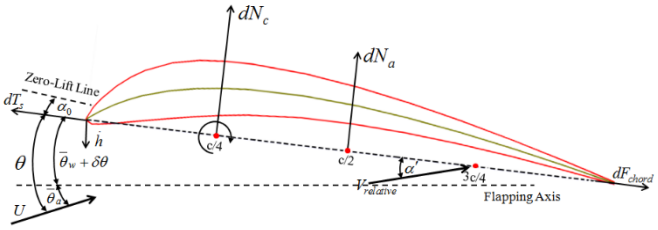


Figure 3: Wing Section

$$dN = dN_c + dN_a \tag{1}$$

where,

dN_c is the circulation normal force.

dN_a is the add mass normal force.

A plunging velocity at the leading edge $\dot{h}$, the local pitch angle θ are created by a swing motion specific for each section. For simplicity, the wing is considered as the fixed structure and the non-span-wised bending wing. A plunging motion is a pure sinusoid function in this research and it equals to (2)

$$h(y, t) = \Gamma \cdot y \cdot \cos(\omega t) \tag{2}$$

where,

Γ is the maximum magnitude of a flapping motion.

ω is the flapping frequency.

The section aerodynamic force, dN_c, acts at the one-quarter of the chord length. A real butterfly wing that consists of a very thin membrane with the support structure was used as the designed reference for the butterfly wing. Then, the butterfly wing is modeled as the thin airfoil because a ratio of the wing thickness to the chord length is small. The normal-force was derived as (3).

$$dN_c = \rho U V_{0.25} \pi \left(\alpha' + \left(\alpha_0 + \bar{\theta}_a + \bar{\theta}_w\right)\right) c\, dy \tag{3}$$

where,

U is the free stream velocity.

$V_{0.25}$ is the flow's relative velocity at a 0.5 chord.

ρ is the air density.

c is the wing chord length.

α' is the flow's relative angle of attack at the chord.

α_0 is the angle of the section's zero-lift line.

$\bar{\theta}_a$ is the angle of flapping axis respected to the free stream.

$\bar{\theta}_w$ is a mean pitch of chord respected to the flapping axis.

Due to the properties of the aeroelasticity, the added mass of the wing section generates an apparent normal force when a wing is flapping. This normal force acts at the middle of the chord length of the wing section and equals to (4).

$$dN_a = \frac{\rho\pi c^2}{4}\left(U\dot{\alpha} - 0.25c\ddot{\theta}\right)dy \tag{4}$$

where,

$V_{0.5}$ is the flow's relative velocity at half-chord.

α is the relative AoA at location of chord.

$\theta = \bar{\theta}_a + \bar{\theta}_w$

The relative angle of attack (AoA) was approximated by a diagram of velocities displayed in Figure 3 and equals to (4)

$$\alpha = \frac{\dot{h}\cos(\theta - \bar{\theta}_a) + \frac{3}{4}c\dot{\theta} + U\Delta\theta}{U} \tag{4}$$

where,

$\Delta\theta$ is the pitch angle change from a wing's motion.

Clearly, the pitch angle was changed according to time, flapping frequency and the twisted angle constant per unit distance along the wing span. By assumption, the wing is an untwisted and not bended wing, thus, the flow's relative AoA at the $3/4$ chord position was derived as (5)

$$\alpha' = C(k)_{Jones}\alpha - \frac{w_0}{U} \tag{5}$$

where,

w_0 is the downwash velocity at the chord.

$C(k)_{Jones}$ is the Jones's coefficient.

The Jones's coefficient [6] for the finite aspect ratio (AR) solution was derived based on the strip theory in case of the wing's finite span unsteady vortex wake [7]. AR equals to the wingspan^2 divided by the wing area. The dimensionless reduced frequency, k , was introduced for an unsteady aerodynamics and aeroelasticity condition. Reduced frequency was related to the flapping frequency as (6).

$$k = \frac{c\omega}{2U} \tag{6}$$

Jones's coefficient consists of the real part and imaginary parts, thus, coefficient is a complex function. The alternative approximated formulation was proposed [8] as (7)

$$C(k)_{Jones} = \left(\frac{AR}{2+AR}\right)\left(C'(k)\right) = \left(\frac{AR}{2+AR}\right)\left(F'(k) + iG'(k)\right) \tag{7}$$

The dimensionless constants $F'(k)$ and $G'(k)$ are depending on AR and k . Scherer proposed the approximated equations of $F'(k)$ and $G'(k)$ as (8) and (9) respectively.

$$F'(k) = 1 - \frac{C_1 k^2}{(k^2 + C_2^2)} \tag{8}$$

$$G'(k) = -\frac{C_1 C_2 k}{(k^2 + C_2^2)} \tag{9}$$

where,

$$C_1 = \frac{0.5AR}{2.32 + AR} \quad \text{and} \quad C_2 = 1.81 + \frac{0.772}{AR}$$

The flapping motion was designed as a sinusoidal equation (10)

$$\alpha = \Gamma e^{i\omega t} \tag{10}$$

By putting (6) to (10) into (5), the flow's relative AoA at the $3/4$ chord location can be approximated by (11)

$$\alpha' = -\frac{w_0}{U} + \left(\frac{AR}{2+AR}\right)\left(F'(k)\alpha + \frac{c}{2Uk}G'(k)\dot{\alpha}\right) \tag{11}$$

The untwisted elliptical-planform wing downwash and a free stream velocity have a relationship as shown in (12)

$$\frac{w_0}{U} = \frac{2(\alpha_0 + \bar{\theta})}{2 + AR} \tag{12}$$

Therefore, the flow's relative velocity at half-chord was calculated by (13)

$$V_{2.5} = \sqrt{\left(-\dot{h}\sin(\theta - \bar{\theta}_a) + U\cos(\theta)\right)^2 + \left(-\frac{1}{2}c\dot{\theta} + U(\alpha' + \bar{\theta})\right)^2} \tag{13}$$

The chordwise forces acts on the wing section. The total chordwise forces, dF_{chord} , was the combination of 3 forces which are the leading-edge suction, camber force, and chordwise viscosity friction drag was explained as (14).

$$dF_{chord} = dT_s - dF_{camber} - dF_{friction} \tag{14}$$

The leading-edge suction, dT_s , was given in [6] as (15)

$$dT_s = \eta_s 2\pi\left(\alpha' + \bar{\theta} - \frac{1}{4}c\frac{\dot{\theta}}{U}\right)^2 \frac{\rho U V_{2.5}}{2}cdy \tag{15}$$

where,

η_s is the leading edge suction efficiency factor

Two other chordwise forces were calculated by (16) and (17)

$$dF_{camber} = -2\pi\alpha_0\left(\alpha' + \bar{\theta}\right)\frac{\rho U V_{2.5}}{2}cdy \tag{16}$$

$$dF_{friction} = \frac{1}{2}\rho V_{chord}^2 C_{df}cdy \tag{17}$$

where,

V_{chord} is the section normal flow speed.

C_{df} is the skin friction drag coefficient.

$$V_{chord} = U\cos(\theta) + \dot{h}\sin(\bar{\theta}_a - \theta) \tag{18}$$

The skin friction drag coefficient C_{df} for a laminar flat plate skin was approximated as a function of the Reynolds number [9] which is calculated by (19) where μ is the air dynamic viscosity.

$$C_{df} = \frac{0.664}{\sqrt{\frac{\rho U c}{\mu}}} \tag{19}$$

The horizontal force dF_{chord} and the vertical force dN at each strip dy were altered into the normal lift and parallel thrust forces equations of the free stream velocity as (20) and (21) respectively.

$$dL = dF_{chord}\sin(\theta) + dN\cos(\theta) \tag{20}$$

$$dT_s = dF_{chord}\cos(\theta) - dN\sin(\theta) \tag{21}$$

Wing's lift and thrust at any particular instant time can be calculated by integrating (20) and (21) along the wing span. The instant section's dihedral angle $\sigma(t)$ in the swing motion is displayed in Figure 4 and it equals to $\Gamma\cos(\omega t)$. The average wing lift and thrust were calculated by integrating every instant wing's lift and thrust function over one flapping period ($1/f$), equals to (22) and (23).

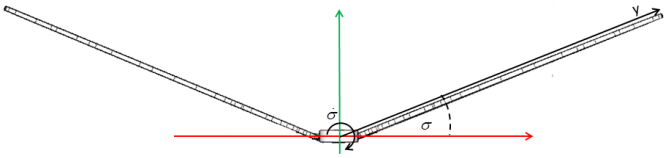


Figure 4: The section's dihedral angle at that instant in the flapping cycle.

$$\bar{L} = \frac{1}{1/f}\int_0^{1/f}\left(\int_{-b/2}^{b/2}\left(\cos(\sigma(t))dL\right)\right)dt \tag{22}$$

$$\bar{T}_s = \frac{1}{1/f}\int_0^{1/f}\left(\int_{-b/2}^{b/2}\left(dT_s\right)\right)dt \tag{23}$$

where,

$f = \omega/2\pi$ is the flapping frequency.

The average lift and trust shown in (22) and (23) are the function of the flapping frequency. Therefore, wing swings with the high flapping frequency, it produces large lift force for lifting robot body up to the air. Unfortunately, when wing flaps with symmetry magnitude for the upward and downward swing, wing generated down force and lift force in the motion. Unlike the bird's wing, the insect wing cannot be folded because the wing structure is fixed. To overcome this issue in this research, the nonsymmetric motion for the upward and downward swing is considered. The wing starts at position (a) and moves down to position (b) as shown in Figure 5. The wing has a high speed for the downward direction and a slow speed for the upward direction to generate the maximum possible lift force.

Butterfly robot will fly in the controlled wind-free indoor environment, then, the velocity of the air can be negligible. Thus, the wing motion can be considered as the laminar flat plate moves perpendicular to the air. In order to analyze the designed wing structure, the airflow against a wing is simulated. This flow is the non-reactive flow condition because the wing motion in this research was not generated from a fuel or heat which can alters the physical property of the air flow causing more shear layers in the flow. The governing Navier-Stokes (NS) equations were used as the basic flow equations. The conservation of air can be calculated from (24)

$$\frac{D(\rho)}{Dt} + \rho\nabla\cdot\upsilon = 0 \tag{24}$$

where,

$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \upsilon\cdot\nabla$ is material derivative.

υ is the velocity of air in x-, y- and z-axis.

∇ is the divergence in x-, y- and z-axis.

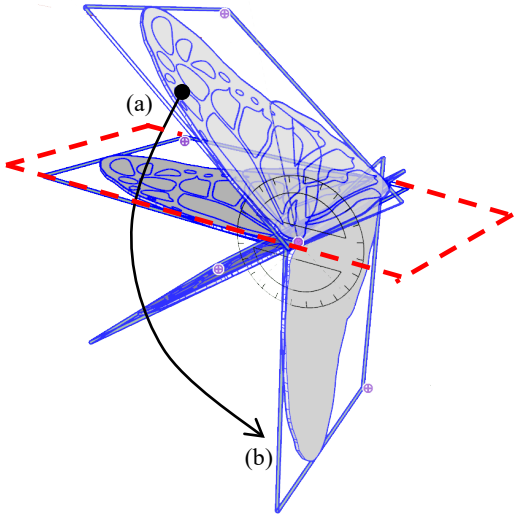


Figure 5: Butterfly wing motion

The Cauchy momentum equation of the air flow can be described as (25)

$$\rho\frac{D(\upsilon)}{Dt} = -\nabla P + \nabla\cdot\tau + \left(\rho - \rho_\infty\right)g \tag{25}$$

where,

P is a pressure.

τ is a viscous stress tensor.

g is an external force vector due to the buoyancy.

For simplicity, in this research the fluid was considered as Newtonian fluid and the fluid was incompressible. Therefore, the viscous stress tensor can be estimated as (26)

$$\tau = \mu(\nabla\upsilon + (\nabla\upsilon)^T) - \frac{2}{3}\mu(\nabla\upsilon)I \tag{26}$$

Navier-Stokes equations cannot be analytically solved due to the complexity [10]. The Semi-Implicit Method for Pressure Linked Equations Revised (SIMPLER) [11] method was selected

in order to obtain the numerical solution of a flow. The first step in this method is to calculate the divergence-free condition pseudo-time stepping by considering (26) with the pressure-free condition and equals to (27)

$$\rho\frac{D(\hat{\upsilon})}{Dt}=\left(\frac{\partial(\rho\hat{\upsilon})}{\partial t}+\hat{\upsilon}\cdot\nabla\left(\rho\hat{\upsilon}\right)\right)=\nabla\cdot\hat{\tau}+\left(\rho-\rho_{\infty}\right)g \qquad (27)$$

where,

$\hat{\upsilon}$ and $\hat{\tau}$ are the estimated values of υ and τ

Equation (27) is needed to be discretized for numerical method, yield (28)

$$\left(\frac{\rho(\hat{\upsilon}-\upsilon_0)}{\Delta t}+\hat{\upsilon}\cdot\nabla\left(\rho\hat{\upsilon}\right)\right)=\nabla\cdot\hat{\tau}+\left(\rho-\rho_{\infty}\right)g \qquad (28)$$

The next step, the estimated pressure is estimated from pressure Poisson equation as (29)

$$\rho\nabla\left(\frac{\hat{\upsilon}}{\Delta t}\right)=\nabla\cdot\nabla\hat{P}=\nabla^2\hat{P} \qquad (29)$$

where,

$\hat{P}$ is the estimated values of P

Then, value of $\hat{P}$ is used in the discrete version of equation (25) to obtain the updated velocity value υ^* and used this value to update $\hat{P}$ value to P^* value via (29). Finally, the next step velocity is obtained by (30)

$$\upsilon=\upsilon^*-\frac{\Delta t}{\rho}\nabla P^* \qquad (30)$$

There are several techniques to solve the numerical flow equations such as Euler, Lax-Wendroff, and 3rd – 4th orders Runge-Kutta method. In this research, the Runge-Kutta method was selected as the solver with proper boundary conditions such as flow inlet and outlet. For simplicity, the simulation is based on the Flowsquare+ flow simulator [12]. A designed wing STL mesh file was imported into the simulator, placed at the center of the airflow chamber and set perpendicularly to the x-axis. The number of grid points were 128, 64 and 64 for the x-axis, y-axis, and z-axis respectively. Air flowed in +x-direction with 5 m/s to make the simulation reaches the Quasi-steady state quickly. Dynamic viscosity was set at 1.8E-5 (kg/m/s) to simulate the atmospheric air at room temperature. The simulation results show in Figure 6 and Figure 7. Figure 6 shows that there are several vertexes happen at the back of the butterfly wing tip and the airflow simulation proves the assumption that when wing flap perpendicular to the air in the still air, it acks like a laminar flat plate. The pressure density is shown in Figure 7. From this result, the pressure condenses around the wing's center of mass. This information will be used for the analysis of dynamics based on masses system via Euler-Lagrange equation of motion. It is clear that if one wing flaps, the reaction force is perpendicular to the wing surface and has components in both x- and y-axis. This phenomenon generates pros and cons for controlling the butterfly robot. If two wings have a symmetrical synchronized motion, then, wings generate only the left force to the robot body but if wings motion is not symmetric or asynchronized then the robot will be rotated. The diagram of wings motion is displayed in Figure 8. The averaged reaction force acting point is estimated from the pressure simulation result.

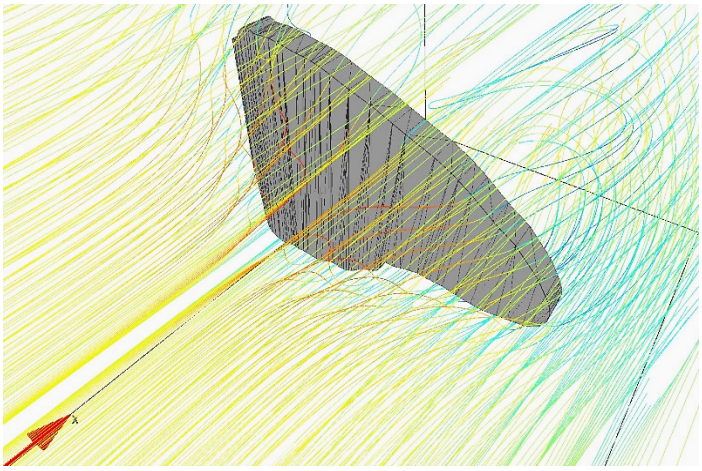


Figure 6: Air flow to a wing.

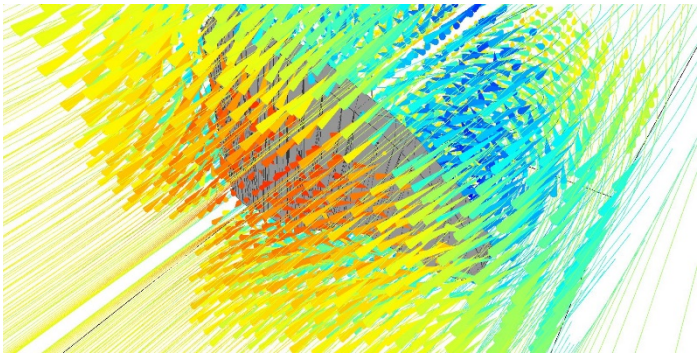


Figure 7: Pressure at a wing

3. Butterfly Robot Construction

In this section, the butterfly robot construction is explained. The first butterfly robot wing has been designed in such the way that it is 10 times larger than the real butterfly wing but keeps the wing physical properties such as a shape and the width to height proportion as shown in Figure 8. The first robot wing has a wingspan and area about 51 cm and 539 cm^2 respectively. Due to the assumption from the second topic, the wing was made from a reinforced carbon fiber rods structure with a transparent laminar plastic sheet. This reinforced structure creates the wing rigidity and ensures thin airfoil properties of a flat plate skin shape. Due to the thin wing, the chord line is approximately overlap to the wing camber line. Butterfly robot has one left-wing and one right-wing, unlike a real butterfly that has two left-wings and two right-wings which generate the 3 degrees of freedoms. The designed butterfly robot creates a rotation about the flapping axis while the real butterfly creates a rotation around a single rotation point. There are two major flapping mechanisms which are one actuator and two actuators driven mechanisms. A crank-and-piston is the one actuator flapping mechanism that creates the same motion for both left and right-wings while two separated actuators mechanism can create different motions for the left-wing and the right-wing. In this research, in order to achieve the advantages of unsymmetrical aerodynamics from two wings, two servos driven mechanism was selected in order to create mismatch motions and forces between

two wings. Servo motors were placed in the servos holder and attached to each wing by using a wing locker as shown in Figure 10. The complete butterfly robot is shown in Figure 11. The designed mechanism can enable robot rotation without an extra actuator such as a tail.

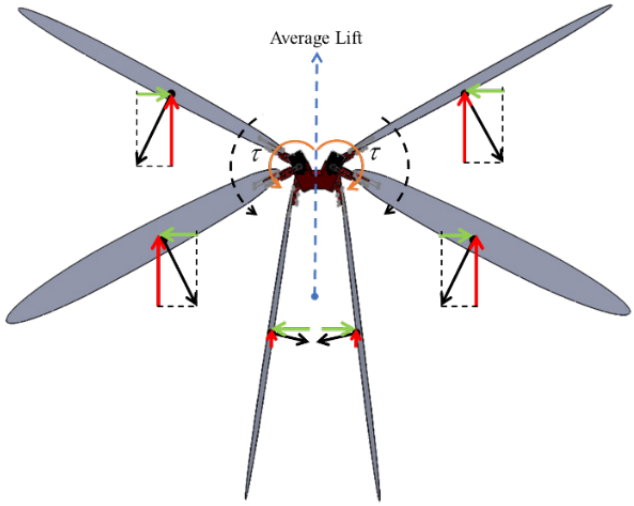


Figure 8: Two wings flapping downward motion.

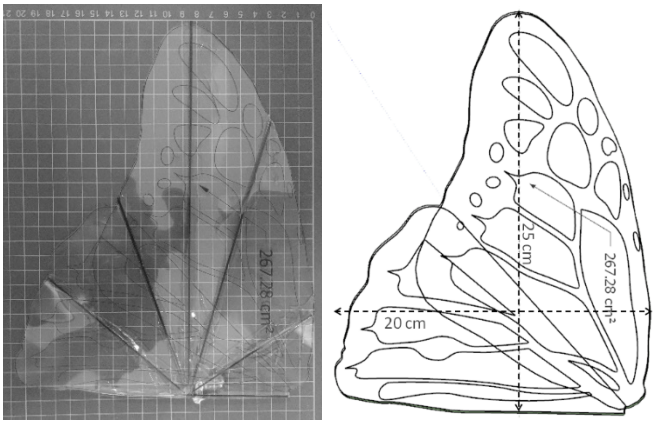


Figure 9: Prototype Butterfly Wings

The KST MS325 micro high-speed servo motors were used for the robot. The motor can generate 5.2 Kg.cm torque and achieve 60° per 0.07 second rotating speed when operated at 8.4 volts. Trust and lift forces were obtained by equations and simulations from the previous section. The completed butterfly robot dynamics was not analyzed in this manuscript because it requires masses system dynamics analysis via Euler-Lagrange or Newton-Euler methods. The completed butterfly robot mathematic model, then, will be explained in the future research. STM32 ARM Cortex-M3 is selected as the robot's main controller operated at 72 MHz frequency via a PLL module. Servo motor and ARM controller are shown in Figure 12. A servo motor shaft position can be controlled by using the RC control signal. This control signal has a specific pulse train that has a frequency of 50 Hz, and a duty cycle varies between 5 to 10 percent of the period depends on the desired shaft location. The selected servo motor consists of a local control circuit based on a PID control algorithm, a h-bridge motor driver and the shaft position sensor. Thus, servo motor's shaft can rotate to any desired position by using the RC control signal as long as a load torque doesn't greater than a maximum servo torque. If the maximum torque is much greater than a load torque, the servo shaft can rotate with a constant angular velocity. Therefore, in this research, the control algorithm in the ARM controller is used for the trajectory and RC control signal generators. The servo shaft will start at the start position and then rotates to the end position with the desired speed. In this research, the swing down speed is greater than the swing up speed which will be explained in the experiments.

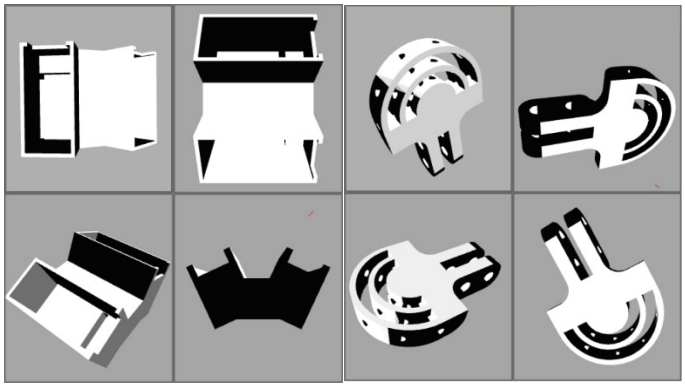


Figure 10: Servos holder (left) and Wing locker (right)

4. Experiments

The experiments on the designed butterfly robot were conducted. No power system was installed into the robot body in this research in order to reduce robot payload then the robot system was wired to the external power source. The butterfly robot was installed into a sliding rod with a support platform in order to make the robot move freely when the wings are flapped.

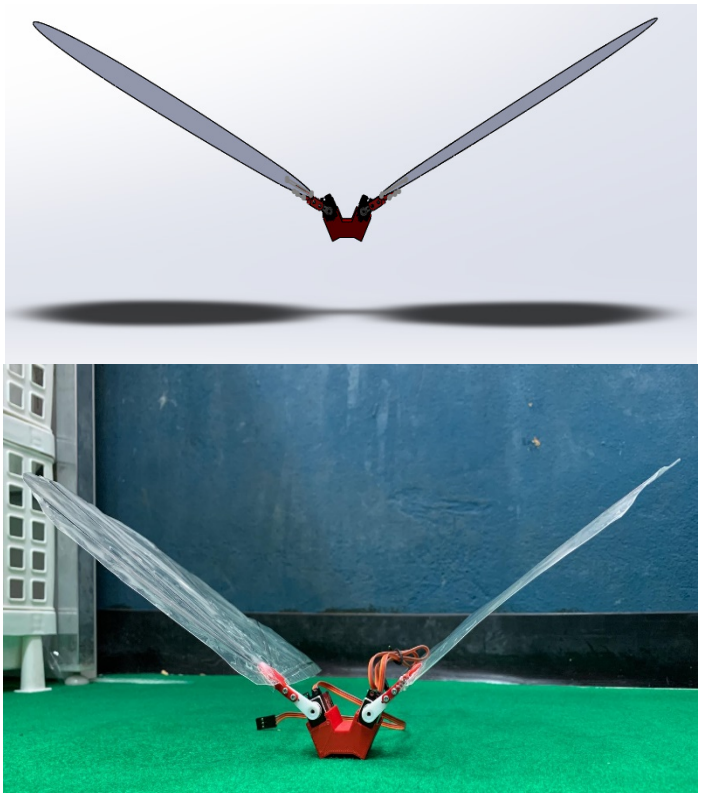


Figure 11: Butterfly robot model (top) and manufactured butterfly robot (bottom)

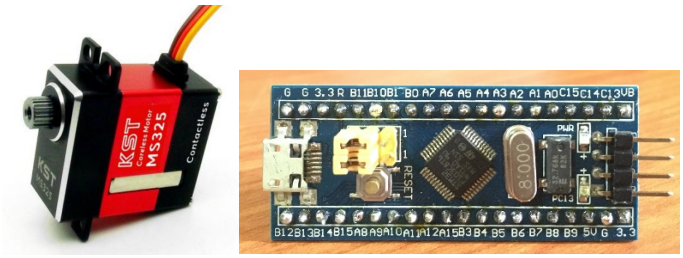


Figure 12: KST MS325 high speed servo (left) and ARM controller (right)

In the first experiment, the robot flapping mechanism was tested. The robot mechanism could flap the wings simultaneously as the desired motion but two problems occurred which were bended wing and robot rotation situation. The reinforced wings were bended due to servo torques and air pressure as shown in Figure 12. To understand this behavior each wing has to be model as a flexible robot arm which was not the aim of the wing design proposed in this research. Although, the wing was based on the real non-symmetry wing butterfly wing but the number of wings on each side was reduced from 4 to 2 wings, yield, the extra aerodynamics force and wake vertex rotate the robot body as shown in the experiment. Therefore, wings have to be further reinforced by another layer of a carbon fiber rods structure with duct tape to make a wing rigid as much as possible. The wing was redesigned to be a symmetry wing to make the air pressure condense around the center of the wing's mass and reduce the rotation as shown in Figure 13. The modified wing was simulated in order to confirm the location of the air pressure. Then, the butterfly robot was tested again. The experiment result shows that two problems are solved as shown in Figure 14.

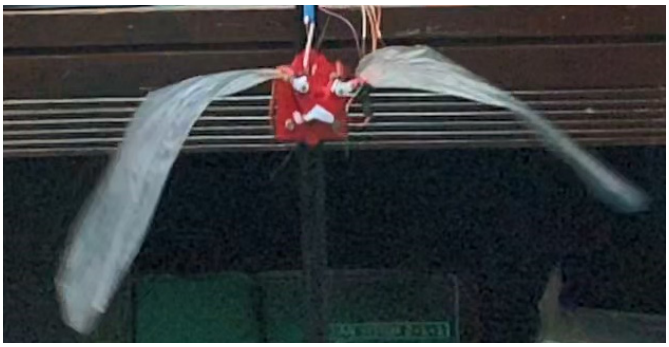


Figure 12: Bended wings situation

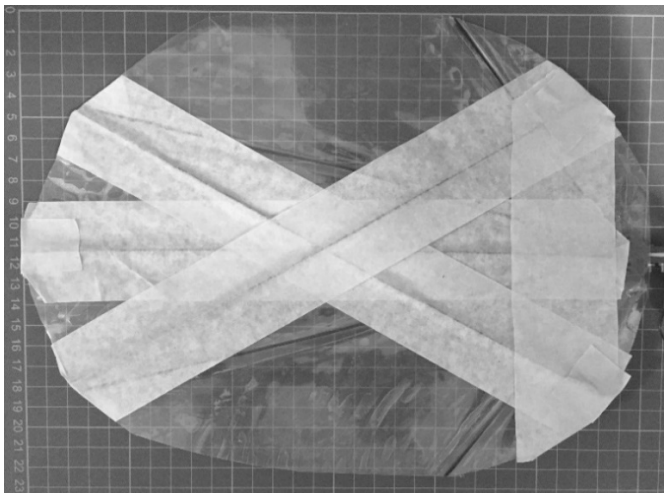


Figure 13: Redesigned and reinforced wing

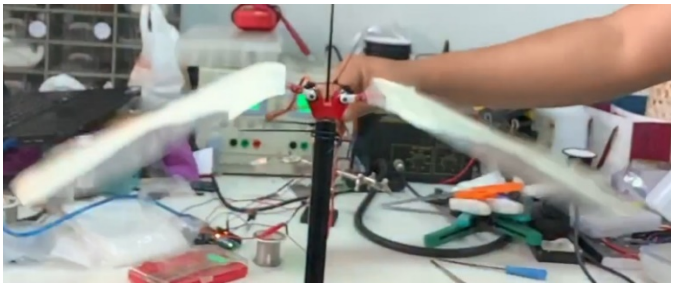


Figure 14: Unbended wings

Next, the butterfly robot was experimented. Servo motors were driven by the 50 Hz PWM signal with specified duty cycles depended on the wing locations. The driver circuit was installed to the robot power system since the ARM controller uses 3.3 volts logic while servo motors require 8.4 volts. In this experiment, the friction between the robot body and the sliding rod were neglected. The servo motor and wing lockers were printed by the 3D printer and the PLA plastic was selected as the printing material. Nuts and bolts were used to attach a wing locker to a servo horn. The total butterfly robot weight is 110 grams and the robot parameters are shown in table 1. The experiment was recorded by the Huawei P20 pro smartphone at 960 fps. In this research, the completed dynamics equation of the robot including the mass system was not considered. The coupling velocity effect and robot inertia also were neglected.

Table 1: Robot Parameters

	Part	weight
1	Wing with wing locker	25 g
2	KST servo motor with a plastic horn	23 g
3	Servos holder	9 g
4	Nuts and bolts	5 g

From the experiment, the butterfly robot can take-off from the support platform, fly up when wings swing down and fall back to the platform when wings swing up. It is clear that the robot wings can generate the average lift force is greater than 1.0791 N which is enough lift force to lift robot up from the support platform. The lift force can be simply estimated from the combination of the maximum torque generated from the servo motors acting at the center of the air pressure estimated from the simulation. The robot starts moving slowly from the platform then the speed is increasing and reaching the maximum acceleration when both wings are parallel to the ground and robot. Then, the robot starts moving slower until it stops at the maximum height as shown in Figure 15. The estimated lift force, if the coupling velocity effect and inertia is neglected, is about 5.4 N. At this point, the robot can reach the maximum body acceleration of 0.93 m/s. The robot can fly up to about 10 centimeters above the support platform in 1.02 seconds as shown in Figure 16. Although, in the experiment, the robot shows the success of take-off and fly up to certain height, but when the robot swings up wings, the robot drops down quickly because wing's downforce combined with the robot weight as shown in Figure 17. This situation makes the robot drops to the platform faster when wings swing up than the robot reaches the maximum height when wings swing down. This problem has a severe impact

if the butterfly carries payload such as battery and other electronic circuits. In order to overcome this problem, the wings motion is needed to be modified.

5. Conclusion and Future Work

A butterfly robot wing section's aerodynamics and aeroelastic are analyzed, estimated and derived in order to obtain a wing behavior. Lift and thrust are important for this research because they are critical of designing a robot wing. After wing section's lift and thrust forces were obtained, the average lift and thrust can be calculated by integrating all small wing section's forces in particular time over one flapping motion period. Navier-Stokes equations are used to simulated the air pressure acts on the wing when swings in the still air.

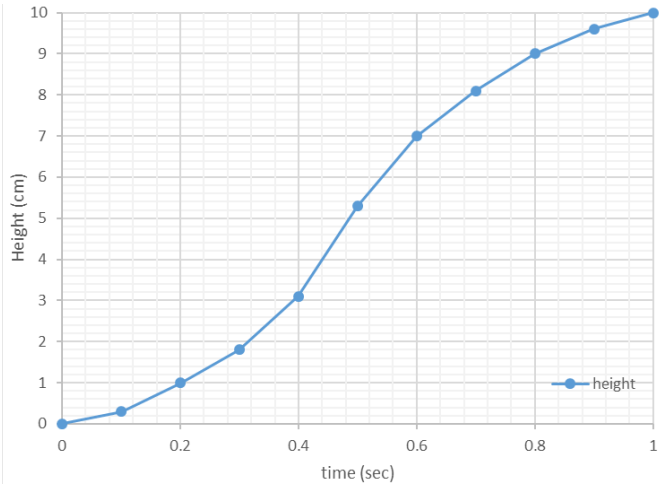


Figure 15: Height of the butterfly robot in the experiment.

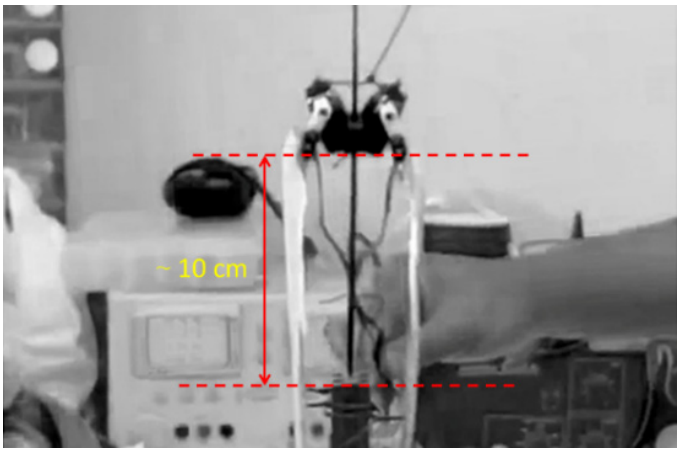


Figure 16: Maximum height.

The SIMPLER algorithm is used to obtain the numerical solution of Navier-Stockes equations. The real butterfly wing properties and dimension was used the reference for designing the desired robot wing. A laminar plastic sheet was used to mimic butterfly wing membranes properties while the reinforced fiber carbon rods was used to create wing wall structure like the scales and hairs in the butterfly wing. Thus, the designed wing was considered as a thin airfoil with laminar skin properties. A flapping mechanism was constructed from two servo motors to create flexible motions between two wings and enhance overall robot performance. The butterfly robot parts were designed and manufactured. In the first experiment, the designed wing was bended occurred from the servo torques and air pressure due to the wing material. Then, the wing was redesigned to be a symmetry wing and reinforced to reduce a bend. The experiment shows that the butterfly robot can take off from the platform and fly up to a certain height before it goes back down to the platform.

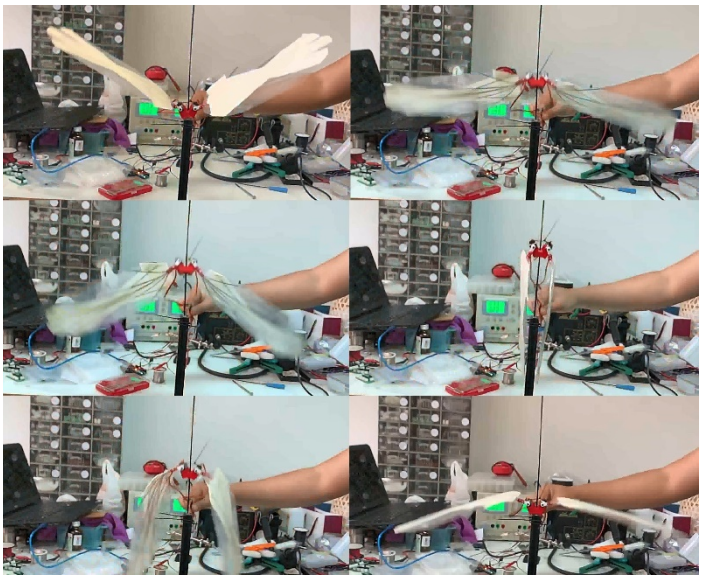


Figure 17: Experimental Result.

In this research, masses, inertias and the coupling velocity effect on the butterfly robot system were not considered. The standalone butterfly robot with payload will be designed and the completed butterfly robot dynamics will be derived using the Euler-Lagrange equation of motion. More experiments will be conducted in future work.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] K. Sukvichai, K. Yajai, "Aerodynamics Force Analysis for Designing a Flapping Butterfly Robot Wing" in International Conference on Electronics, Information, and Communication 2020 (ICEIC2020), Barcelona, Spain, 2020. https://doi.org/10.1109/ICEIC49074.2020.9051389

[2] G. De Croon, K. De Clercq, R. Ruijsink, B. Remes, C. de Wagter, "Design, aerodynamics, and vision-based control of the delfly" International Journal of Micro Air Vehicles, **1** (2), 71–97, 2009. https://doi.org/10.1260/175682909789498288

[3] M. Keennon, K. Klingebiel, H. Won, "Development of the nano hummingbird: A tailless flapping wing micro air vehicle," in 50th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition, Nashville, TN, USA, 2012. https://doi.org/10.2514/6.2012-588

[4] J. Zhang, F. Fei, Z. Tu, X. Deng, "Design optimization and system integration of robotic hummingbird," in 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, Singapore, 2017. https://doi.org/10.1109/ICRA.2017.7989639

[5] K. Y. Ma, P. Chirarattananon, S. B. Fuller, R. J. Wood, "Controlled flight of a biologically inspired, insect-scale robot" Science, **340**(6132), 603–607, 2013. https://doi.org/10.1126/science.1231806

[6] R. T. Jones, The Unsteady Lift of a Wing of Finite Aspect Ratio, NACA Report No. 681, 1940

[7] T. Theodorsen, General theory of aerodynamic instability and the mechanism of flutter, NACA Rept. 496, 1935.

[8] J. O. Scherer, Experimental and Theoretical Investigation of Large Amplitude Oscillating Foil Propulsion Systems, Hydronautics Incorporated,

Research in Hydrodynamics, Technical Report 662-1, Laurel, Maryland, USA, 1968.
[9] S. F. Hoerner, Fluid-Dynamic Drag: Theoretical, experimental and statistical information, Hoerner Fluid Dynamics; Second Edition Reprint edition , 1992.
[10] L. C. Fefferman, "Existence and Soothness of the Navier–Stokes equation", Millennium Prize Problems—Navier–Stokes Equation, Clay Mathematics Institute, March 27, 2017, retrieved 2017-04-02, [Online] Available: http://www.claymath.org/sites/default/files/navierstokes.pdf [Accessed: 7- Jul- 2020]
[11] S. Hosseinzadeh1, R. Ostadhossein, H.R. Mirshahvalad, J. Seraj, "Using Simpler Algorithm for Cavity Flow Problem" Mechatronics and Applications: An International Journal (MECHATROJ), **1**(1), 55–63, 2017. https://doi.org/10.2139/ssrn.3438116
[12] Flowsquare+, Available: https://fsp.norasci.com/ [Accessed: 7- Jul- 2020]

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 672-683 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Enhanced Artificial Intelligence-Based Approach Applied to Vehicular Traffic Signs Detection and Road Safety Enhancement

Anass Barodi*,[1], Abderrahim Bajit[1], Taoufiq El Harrouti[2], Ahmed Tamtaoui[3], Mohammed Benbrahim[1]

[1]Laboratory of Advanced Systems Engineering, National School of Applied Sciences, Ibn Tofail University, Kenitra, 14000, Morocco

[2]Laboratory of Engineering Sciences (SI), National School of Applied Sciences, Ibn Tofail University, Kenitra, 14000, Morocco

[3]National Institute of Posts and Telecommunications (INPT-Rabat), SC Department, Mohammed V University, Rabat, 10000, Morocco

ARTICLE INFO

Article history:
Received: 29 November, 2020
Accepted: 21 January, 2021
Online: 05 February, 2021

Keywords:
Object Detection
Traffic Road Signs
Optimized Detection
Artificial Intelligence
CNN, Classifications
Traffic Signs Recognition
Intelligent Transportation
Road Safety

ABSTRACT

The paper treats a problem for detection and recognition objects in computer vision sector, where researchers recommended OpenCV software and development tool, it's several and better-remembered library resource for isolating, detecting, and recognition of particular objects, what we would find an appropriate system for detecting and recognition traffic roads signs. The robustness with optimization is a necessary element in the computer vision algorithms, we can find a very large number of technics in object detection, for example, shapes transformation, color selection, a region of interest ROI, and edge detection, combined all these technics to reach high precision in animated video or still image processing. The system we are trying to develop, is in high demand in the automotive sector such as intelligent vehicles or autonomous driving assist systems ADAS, based on intelligent recognition, applying Artificial Intelligence, by using Deep Learning, exactly Convolutional Neural Network (CNN) architecture, our system improves the high accuracy of detection and recognition of traffic road signs with lower loss.

1. Introduction

Computer vision belongs to a large family of Artificial Intelligence-based systems, more precisely, it is to AI what eyes are to humans, an extremely powerful tool capable of analyzing, recognizing an image pattern, and even analyzing several images simultaneously, in other words, a video stream. To achieve this, computer vision relies on different technologies, firstly image processing, which applies operations and algorithms directly on the image, and secondly Deep Learning through neural networks using a large amount of data. In the automotive world, Industry 4.0 [1] has already chosen its IoT camp, the automotive groups have announced that their autonomous vehicles will embed an onboard camera, and not only electronic sensors, aiming to reduce a maximum accident's risk and also helping the driver to safely control his car to avoid dangers. The re-manufacturing sector has also already started to conquer computer vision, the major objective of the automotive industry, is to make non-polluting vehicles, e.g. electric or hybrid vehicles, to integrate intelligent options [2], to make vehicles easier, to move in and out of the city area, with road safety for both cars, users, and pedestrians. The procedure that we are going to follow in our system development, that integrates a series of automotive embedded systems, among the most famous of these systems, is the ADAS system that has accelerated the vehicle sector and improved its features, which has made the semi-automatic vehicles, to facilitate the tasks of conduction and avoid all possible dangers [3]. The human being is based essentially on his eyes, to make a description and a global analysis of his environment, because this description will allow him to manage these reactions, that he has to make either in the present or the future. Therefore, the equivalence of this behavior enters the field of computer vision or robotics, where it is called object detection [4, 5], which in our case are the traffic road signs, that help traffic management. Every year, researchers offer us a new advanced, highly precise technology, and slightly object detection error injection. Our traffic road signs system detection [5], which is based on the colors of the environment description, to detect the desired specific colors of the specified objects.

*Corresponding Author: Anass BARODI, Laboratory of Advanced Systems Engineering (ISA), National School of Applied Sciences, Ibn Tofail University. Kenitra, Morocco, barodi.anass@uit.ac.ma, https://orcid.org/0000-0003-3022-4761

The fact of making this description, therefore, required a test its effectiveness, based on onboard vehicle cameras, and then to integrate it into the developed algorithm. So detecting the colors of traffic roads signs with accuracy, is the necessary factor, so that we can recommend the approach, that can be applied in the future or improved, and applied to Industry 4.0 [6], but we're not going to stop only on the colors [7], we're also going to base it on the geometry of the traffic roads signs, to reduce detection errors. The general procedure of our system, is based on human characteristics, that it can easily detect objects, and avoids encountering difficulties due to light reflection. The performance will increase, if we have combined the colors and geometries of the traffic road signs, which will conduct us to reach an interesting and robust algorithm [8]. In computer vision and image processing, the detection or recognition of specific objects in movements, in unknown or a known environment, is commonly elaborate by different researches studies every year, the most technique based on the invariance properties of the interest objects, it can be the appearance of the objects, the geometry of the scene or objects, photometry or color, for example, space-time constraints.

Video is a media that poses complex problems, because of the large data volume to be processed, with the difficulty of extracting the representative information from its content. At present, indexing the content of images, and dividing the video into elementary unit shots, representative images only partially meets this growing need. But, the content of an image is often too rich in itself, and contains several objects of interest, i.e. semantic objects. When a target is located in a video frame, it is often useful to follow this object in the following images, each image target successfully tracked, can provide a lot of information or feature, about its activity with the identity, the visual characteristics most using are colors. Usually, the RGB represents color space, but this space doesn't a uniform perceptual, because the differences calculating between colors in this space by machines, do not correspond to the difference perceived by humans, besides, the RGB dimensions are highly correlated. But the HSV is an approximately uniform color space [9], with a high degree of sensitivity to the noise, for example, can be used to extract a characteristic for histogram-based appearance, for detected the contour-based edges of all or specific objects; that existed in images, in general, many tracking algorithms in real-time uses and recommend this space.

Object recognition is an important topic in computer vision, it has a huge evolution during these years, combined with Artificial Intelligence, which has been finalized by the appearance of Deep Learning [10], Deep Learning has touched different areas, such as automotive, robotics, industry. In computer vision, object recognition or object classification is a method for identifying the presence of an instance, or a class of objects in a digital image, for example, the application face detection and the detection of persons [11]. These methods often involve supervised learning, and have applications in multiple fields, such as, Biometric recognition using biometric technology, Surveillance, Industrial inspection, Robotics, Medical analysis, our case is applied for Intelligent Vehicle Systems, which are necessary for recognize and tracking the traffic road signs, as well as autonomous vehicle [12].

2. Methodology

The strategy is to develop a homogeneous system for detection and recognize traffic road signs, the system combines three essential parts, the first, detection of traffic road signs by methods that are proposed, and well described later on. Secondly, when the sign detected, strongly is going to be cropped with dimensions defined in real-time, finally, the last phase, is the recognition of the traffic road using Artificial Intelligence, establishment an architecture has the same concept of the human being analyzes with his brain, for classification or recognize the correct place different objects in the environment.

The paper organized as follows; section 1 contains an introduction to the essential for detection and recognition. Section 2 presents a system diagram, then section 3 describes the detection approach. Section 4, related work for different work of detection and our proposed approach, section 5 and section 6 for the discussion and implementation in real-time. Section 7 with section 8, to recognize road signs by deep learning and evaluation the performance of the system proposed.

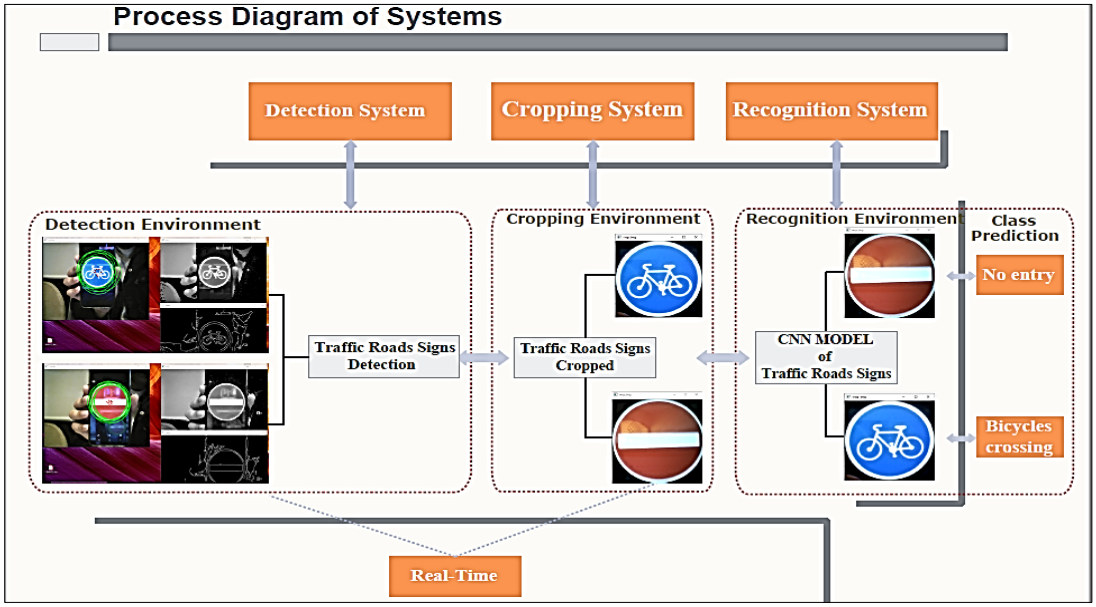


Figure 1: Traffic road signs detection and recognition System diagram

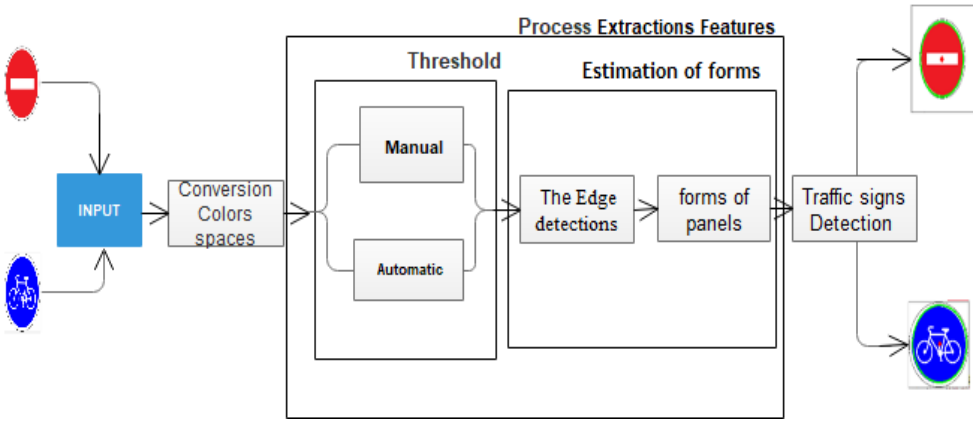


Figure 2: Traffic road sings detection diagram [5]

3. Road Signs Detection

In the figure. 2, the system tries to determine which road signs have a circular shape and specific colors.

The colors of the traffic road signs, are unique in the whole world, fixing the relevant colors or those that represent the majority, are blue and red. The second part, which is based essentially on the detection of the geometries of the traffic roads signs, which are also by default unique, can find either the geometry: rectangular, circular, triangular, and square, therefore for our study, based the circular, because it a percentage of 50% of the traffic roads signs, and at the same time contains the colors, to achieve our objective, it must to a play on the elements, that characterize the circular form, so the circle is described mathematically, by center position and radius.

4. Related Work

The paper has the main objective, to implement an application aiming to detect and recognize traffic signs in real-time, most of the techniques use the color detection presented in this paper [13], or based only on the shapes of signs treated in this paper [14], our approach combined between colors and shapes of traffic road signs, by applying the Canny matrix transformation. This approach uses the OpenCV library, because integrated powerful functions, to properly detect the edges in the image. Before the process detects shapes of signs, following sub-steps, firstly, the conversion input image from the RGB space to the HSV space for segmentation. secondly, applying the Gaussian smoothing filter, to minimize false detection results, as the color-based one just isolates red and blue colors of interest.

4.1. Color detection

The image can be represented by function G (X, Y), the X and Y are the spatial coordinates for two-dimension, besides this, the amplitude at any point G (X, Y) corresponding in general to gray level or mostly to intensity, and when the amplitude is discretized, the image is categorized to a digital image. The digital image also referred to as a scene, has a marker as an indication, each frame or image different from other mathematically, based mainly on its height (H) and width (W).

In image processing has several descriptions of images, for simplicity, can be a binary image, a rectangular matrix contains these values, 0 represented the blacks, or 1 represented the whites. Also, can be a grayscale image, the total number of pixels is 256, distributed from 0 (black) to 255 (white), these values means that each pixel encoded on 8 bits, grayscales are very powerful for recognition or detection of the most objects in a complexed scene [15]. Finally, the color image has 3 color channels, described respectively in this order by pixels number of red (R), green (G), and blue (B), it's coded or ranging between 0 and 255, which results in 255^3=16,777,216 possible colors, that it takes 24 bits to code a pixel. In the general hierarchy of video data present in a sequence, which means that contents one or more images groups, and ends with an end-of-sequence code, this grouping of one or more images facilitating direct access, thus, the image is the elementary unit for coding the video sequence [16].

As shown in figure 2, the first step based on color detection, to have a system capable of detecting; the colors of road signs with high accuracy, our approach proposes the conversion from RGB to HSV space, this conversion is essentially, based on the equations above [5], the first one is used to normalize each color channel, which makes the interval ordered [0,255] to [0,1], to calculate delta, C_{min} and C_{max} [17].

$$R_1 = \frac{R}{255}; G_1 = \frac{G}{255}; B_1 = \frac{B}{255} \quad (1)$$

$$C_{max} = \max(R_1, G_1, B_1) \quad (2)$$

$$C_{min} = \min(R_1, G_1, B_1) \quad (3)$$

$$\Delta = C_{max} - C_{min} \quad (4)$$

The HSV space is considered to be a natural representation model, which means has the same physiological color perception seeing by the eye of a human, in generally decomposing the colors consists of physiological criteria. This space decomposes the color, according to more intuitive characteristics close, to the current vocabulary to describe a color, based on pure color. Finally, these equations above serve to complete the conversion, to determine each channel: Hue, Saturation, and Value [5].

- The Hue Channel (H), corresponding to the perception of color:

$$H = \begin{cases} 0°, & \Delta = 0 \\ 60°x\left(\frac{G_1 - B_1}{\Delta} \bmod 6\right), & C_{max} = R_1 \\ 60°x\left(\frac{R_1 - B_1}{\Delta} + 2\right), & C_{max} = G_1 \\ 60°x\left(\frac{R_1 - G_1}{\Delta} + 4\right), & C_{max} = B_1 \end{cases} \quad (5)$$

- The Saturation Channel (S), describing the color purity, which can be bright or dull:

$$S = \begin{cases} 0, & C_{max} = 0 \\ \frac{\Delta}{C_{max}}, & C_{max} \neq 0 \end{cases} \quad (6)$$

- The Value Channel (V), indicating the light amount in the color, that can have a light or dark appearance:

$$V = C_{max} \quad (7)$$

4.2. Canny matrix transformation

The deployment between the Canny filter and the Hough transformation, to detect edges and locate shapes wants, a well-known algorithm for pattern detection, Hough transform do a parametric estimation method. The pattern recognition method, proposed by Hough are morphological study methods, but they require a large computation time, in the case of complex geometrical shapes or discontinuous contours, the Canny contour detection method, allowed us to highlight rather wide contours, that are perfectly suited to pair correlation localization, to achieve the work with higher detection accuracy.

4.3. Edge Detection

Canny a multi-step process for edge detection, First of all, it is a question of reducing the noise of the image to optimize contour detection, this reduces strong responses during the gradient calculation, leading to false positives, the solution using Gaussian Filter, very powerful filter, has a kernel of convolution operator in a 2D Gaussian plan [18], this operation due to Gaussian probability law, demonstrated in equation (8), σ^2 represent the variance of each variables couple (x, y) with the peak (μ). The main condition for using this type of filter, whether the image is in gray level so that it can be processed to standard deviation Gaussian σ.

$$G_0(x,y) = Ae^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}+\frac{-(y-\mu_y)^2}{2\sigma_y^2}} \quad (8)$$

A canny filter is based on Sobel's algorithm, it has additional steps that allow us to work on more complex images [19]. This method allows, by changing the intensity of edge detection, to eliminate linearization and noise defects and at the same time to have wide contours. The width is important for the correlation localization algorithm that follows this contour detection algorithm [20].

$$\text{Angle: } \theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (9)$$

$$\text{Edge: } G = \sqrt{G_x^2 + G_y^2} \quad (10)$$

The Angle and Edge algorithm-based, has been designed to be optimal according to three criteria:

1) Good detection: low error rate in contour signaling;
2) Good localization: minimizing detected and real contours gap;
3) Answer clarity: one answer per outline, no false positives.

After calculating the norm and the direction of the gradient, in each pixel of the image, these methods extract contours of a single-pixel, thickness by selecting the local maxima of the gradient norms. One can mention the detectors of Canny, Deriche, Marr - Hildreth, or the more recent one of Bao [20]. For the operators of Canny and Deriche, an additional variable allows us to adjust the detection sensitivity at the expense of the spatial precision of the contour and vice versa. Moreover, it can be noted that the contours detection in areas, with strong gradients does not necessarily, correspond to an object. In this case, post-processing can be effective, these useful methods yield good results for non-textured image analysis, high-contrast objects with marked contours.

4.4. Shape extraction

The international road safety regulations, are essentially based on laws, that must be respected by everyone, either drivers of automobiles or passengers, so the majority of its laws have been modeled by traffic road signs, that have unique shapes, for example triangular, rectangular, square, or circular, keeping only the circular shapes that exist in the images.

So, the objective of our approach, is to improve or increase the accuracy of the system, that based on artificial intelligence, will have accuracy to detect traffic road signs, we focus the studies on the circular shape, applying the Hough circles method of OpenCV [18] modeled by this equation:

$$(x - X_{center})^2 + (y - Y_{center})^2 = r^2 \quad (11)$$

Paul Hough invented a pattern recognition technique, named Hough transform, this technique allows us to recognize any shape in an image, for example, the circles or lines, this technique will serve, to identify circles presences in an image, it can be very easily extended to circle recognition, compared to other models, the Hough technique is a simple mathematical model and robust. This produces sinusoids curves, can be accumulated in a space called Hough space, its role is to draw all points set of the circles, to find out the places when the curves intersect, which means the presence of a circle, mathematically the circle is characterized, by (r) radius and center (X_{center},Y_{center}) in equation (11).

For forming a cone with a vertex ($X_{center} = x$, $Y_{center} = y$, r = 0), based essentially on the set number of circles, passing through a given point N (x, y) according to the axis r, and the best points correspond to the maximum intersection of cones. On the other hand, if the circle radius is specified, in a two-dimensional the (a, b) plane, in this plane, the set of circles passing through a given point N is described by the circle of center ($X_{center} = x$, $Y_{center} = y$) and radius r, also the best points that bring together the intersection of several circles. Finally, in accumulation, an M matrix is constructed, each element M (i, j) contains the number of circles passing through the point, but the method consists to construct a hyper-matrix for accumulation when the radius is unknown, which each value of cells M (i, j, k), corresponds to a cubic space (a, b, r), characterized by scanning between 1 pixel up and the dimension of the image.

5. Implementation

5.1. The manual threshold applied to images

We have tried to roughly summarize the procedure or essential steps in the traffic road signs detection process. Based essentially on the colors by setting the lower and upper limit markers for each color, and the geometry of the signs- see the evolution in Table. I.

5.2. Manual thresholding in Real-time

At this level, we try to apply the detection in real-time, for evaluation of our algorithm the detection of traffic road signs,

especially circular shape, which have the colors either Red or Blue, been using a web camera of Asus laptop.

Table 1: Detection of Color and Panel Form [5]

Color	HSV space	Canny Edge	Panel Detection
Red			
Blue			

The results obtained contains two essential windows: One window shows the color images conversions, that have several objects detected with Gaussian filtering to eliminate noise, to easily convert it into grayscale images, the other represents the Canny outlines that facilitate the detection of the desirable object outline, at each figure 4: (a) and (b). We took real-world images of two road signs of different colors: red for the "No-entry" sign and blue for the "Bicycle" sign, but the more relevant thing, the extraction or detection of objects were not perfectly detected. So, we ask several questions where this error comes from, but most probably the two HSV space filters are not perfectly established.

6. Discussion

The goal of our approach is to have or ensure high accuracy for the detection of traffic road signs, specific circular shapes, then we will be able to apply it to all geometric shapes. We have concluded from what we have obtained previously that the HSV space filter terminals did not prove our approach, indirectly the manual thresholding. In image processing, to have a high resolution or in-depth studies on images, we establish a histogram to have an idea of distributions pixels, because our objective is to reach automatic thresholding for the two test images used previously.

If we have visualized well or focused our attention on the two histograms, we interpret the results obtained, that we have a peak that seems to be the maximum of the pixels distributed for each histogram, in an interval from 250 to 255, it is the V channel in blue concerning the other channels H and S, for the figure. 4 (a) and (b). So, what we can conclude is to look for new values of the HSV space filters and make tests until we get high precision, it's hard because there are several factors among these factors is the luminance, the solution, we will unproven our work by using the automatic thresholding, we use the most robust is the OTSU method.

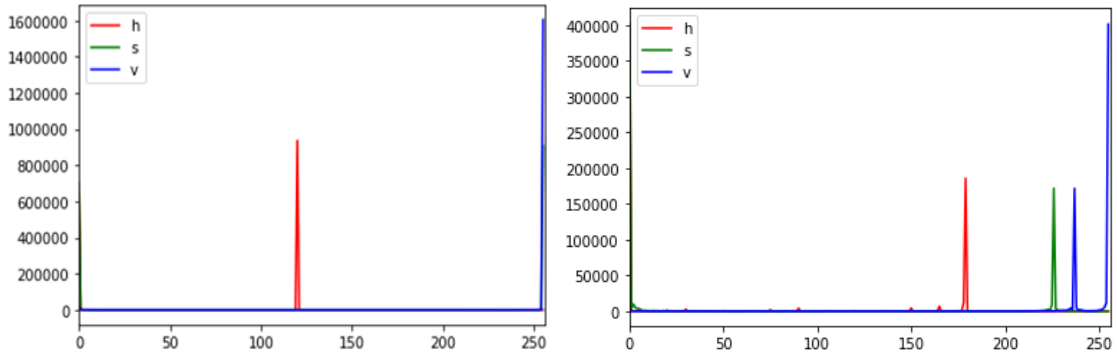
6.1. Automatic thresholding

The OTSU method based on a histogram of images, uses discriminant analysis to separate the text from the background, its principle is to divide the image into classes, and then look for the optimal threshold, by minimizing the variance function inter or intra classes. Otsu's method [21] tries to find or to get the optimal threshold, by separate the histogram into two segments, that maximize the inter-segment variance or minimizes a measure of intra-segment variance, the determination of these variances, based on the histogram normalization. The variances σ_i^2 for each threshold is given by equation (12), The ω_i designed weights of the probability for each class (i), which means the threshold separation.



(a) Red signs (b) Blue signs

Figure 3: Traffic Signs detection manual thresholding [5]



(a) Blue signs (b) Red signs

Figure 4: Traffic road signs histogram [5]

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \quad (12)$$

The Otsu method advises, for its technique to be effective, by minimizing the intra-class variance, due by maximizing the inter-class variance [20].

$$\omega_1(t) = \sum_{i=1}^{t} P(i) \quad \omega_2(t) = \sum_{i=t+1}^{I} P(i) \quad (13)$$

$$\mu_1(t) = \sum_{i=1}^{t} \frac{iP(i)}{\omega_1(t)} \quad \mu_2(t) = \sum_{i=t+1}^{I} \frac{iP(i)}{\omega_2(t)} \quad (14)$$

$$\sigma_1^2(t) = \sum_{i=1}^{t}[i-\mu_1(t)]^2 \frac{P(i)}{\omega_1(t)} \; \sigma_2^2(t) = \sum_{i=t+1}^{I}[i-\mu_2(t)]^2 \frac{P(i)}{\omega_1(t)} \quad (15)$$

The approach to detect regions of objects, therefore leads directly to a partition of the image, in this approach, we applied different methods consisting of manual or automatic thresholding, the simplest method for thresholding in image processing, through segmentation. From a grayscale image, image thresholding can be used to create an image with only two values, black or white, we have applied this method to the basic images that are shown below in Table. 1. The automatic thresholding applies a multi-level per Otsu's method, it consists to determine the number of regions, the optimal values of the different thresholds, based on the variance of the subdivisions. To have an algorithm for the detection of traffic roads signs with high efficiency and precision, from what we understood from the OTSU method, that is based on the classes average μ_1 and μ_2, also the class probabilities ω_i, which will be applied to the pixels of the road signs, either for blue or red, and we fly at the end of the binarized images.

Our strategy for the binary image was formed by the function of the OpenCV library, the bitwise function. Then we merged the three thresholds h, s, v resulting from the three binary images, to obtain at the end a single threshold value. Still, for the detection of traffic roads signs contours, we use Canny because it has proven its high precision in contour descriptions according to the previous tests, even when there is noise, it guarantees high detection and localization of objects for all two-dimensional applications. Before applying the Canny outline, it is preferable to apply a Gaussian filter. Table. 2 shows the manual thresholding values, which contain the bounds of the HSV space values and automatic thresholding.

Table 2: Identification signs for manual and automatic thresholding [5]

Threshold	HSV space	Canny Edge	Detection panels forms
Automatic			
Manual			

Starting from the fact that we are trying to detect the presence of the circle shape in an image, Firstly, the Hough transform consist of searching for the parameters of the structure to be detected, the second step aims to find the points of the image that may belong to the structure, typically the points where the gradient goes beyond a certain threshold. At each of these points, a representative of the structure will pass through and calculate its parameters, the values of these parameters correspond to a point in the parameter space, the third step consists of marking the curve represented by the calculated parameters as a possible representation of the structure being searched.

Table 3: Manual and automatic thresholding [5]

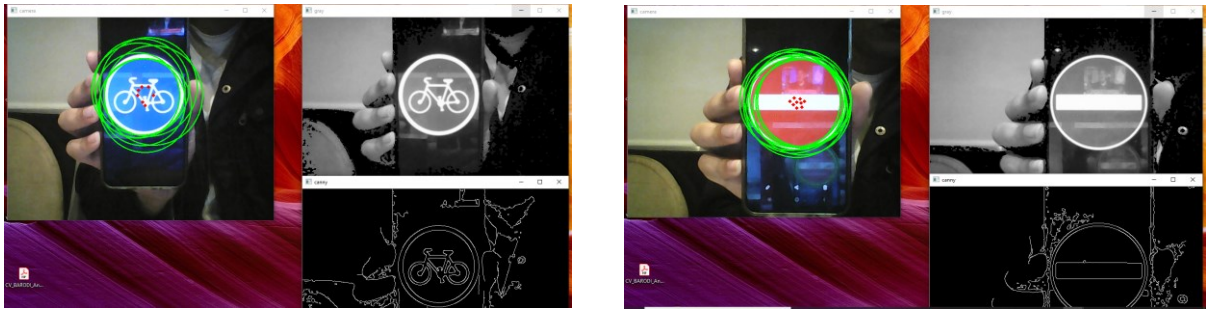
Threshold	Colors	Filter	Hue	Saturation	Value
Automatic	Red	Auto			
	Blue				
Manual	Red	Upper	180	255	255
		Lower	136	87	111
	Blue	Upper	130	255	255
		Lower	110	50	50

For the identification of colors in HSV space shown in the Table. 3, it is certainly necessary to specify very essential elements, for each color of the high and low filters. Our case is the colors that represent the majority of blue and red traffic roads signs, obtaining the values of channels H, S, and V after extensive experimentation and testing, but the trick is to note that these values can make a service or a work perfectly in an environment A and not sure certainly for an environment B. The solution is to apply automatic thresholding by applying the method of OTSU. OTSU adapts to each environment, to find optimal values for the HSV space that leads easily to the detection of traffic road signs. Finally, we are not only going to press on the color, but we are also going to combine the color and the detection of the circular geometry, it shows in table 3 we have an improvement that makes the algorithm more robust to detect the desired object with real-time optimization.

Practically, we are going to vote for parameters. The curve whose parameters obtain the most votes is therefore the one that best describes the structure. To collect the votes, we need an accumulator table, the size of this table is equal to the number of unknown parameters. The size of each dimension is equal to the maximum value of the corresponding parameter. Once the parameters have been calculated, they represent a cell of the accumulator table, we will increase the value of this cell by one unit. At the end of the process, we are going to explore the accumulator to find the boxes that have obtained the most votes, their corresponding parameters describe the occurrences of the structure we are looking at in table.2, which is the circular shape.

6.2. Real-time based Automatic thresholding

At this stage we try to make a real-time test for traffic road signs detection, we will produce the same scenario in an urban environment by using a webcam as shown in the two figures.5 (a) and (b). The relevant thing is that the detection of circular signs is made with high accuracy, so the automatic thresholding by the OTSU method has a great effect than the HSV space. For the three windows that make up the two figures, we have a very high image quality, which makes it easy for Canny to locate the outline of traffic road signs in a very powerful and robust. Otsu has a methodology able whatever the environments, to perform



(a) Blue signs (b) Red signs

Figure 5: Traffic road signs automatic thresholding [5]

automatic thresholding from the shape of the image histogram, so we deduce, that the conversion from gray-level space to binary space, makes it easier to calculate background and foreground classes, for optimal thresholding for each separate between classes, to comply with the condition, that intra-class variance to be minimal.

6.3. Cropping of the panels detected

The Hough feature descriptor used for road signs detection, calculated on a 32 × 32 patch of an image. Let an image of road signs with a size of 128 × 128. A patch with a size greater than 50 × 50 is selected for the calculation of the Hough descriptor. This patch is cropped from an image and resized to a size 32 × 32 in Table.4.

Table 4: Cropped traffic road signs

Cropped Blue Signs	Cropped Red signs
	

7. Recognizing Detected Road Signs

In the recognition part, applying model based on images of traffic road signs in Germany [22], the work is doing it with the Keras libraries for learning and classification, importing python framework, by using some simple and efficient techniques, such as data augmentation [23] and dropout [24].

7.1. Deep Learning

Figure.6 clarify the second part of the approach, for recognizing images, exactly the traffic roads signs [22] using Artificial intelligence, these images cropped from the first part of detection, so this second part, demonstrate how to apply Convolution Neural Networks (CNN) architectures in general, we are going to use in our approach the field of Deep Learning (DL), because this family of architecture, has made it possible to make significant progress in the fields of image classification and recognize different objects, the DL models contains multi-layer of perceptron described above, with more intermediate layers. Each intermediate layer, subdivided all images datasets, for getting sub-parts of features, and providing simpler results to the next layer, CNN currently the powerful models for classifying images [25].

In figure.6 demonstrate two distinct parts of the CNN model, training and performance evaluation, for the first part, in input color images exist, which consist of pre-treatment, the conversion of the color space from RGB to grayscale, for providing a matrix of pixels has two dimensions for a grayscale image, and normalization by value 255 for each image, forget the images ranging between 0 and 1. After the step of pre-treatment, starts training the model, then select the best training model, and then choose based essentially, the curve of accuracy function and loss function. For the second part, the best model passed by strict steps, to evaluate performance; so, we use confusion matrix and rapport classification each class image.

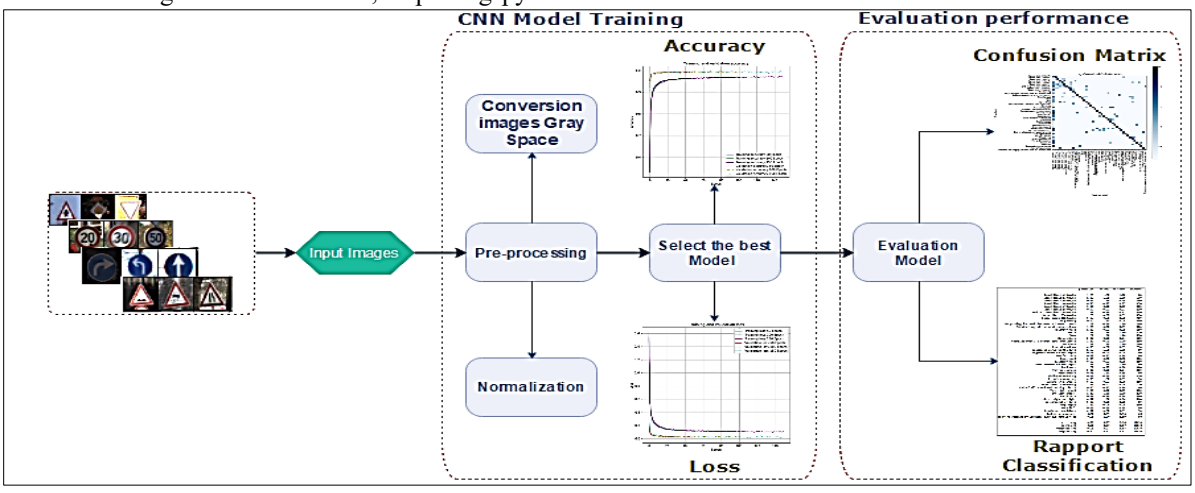
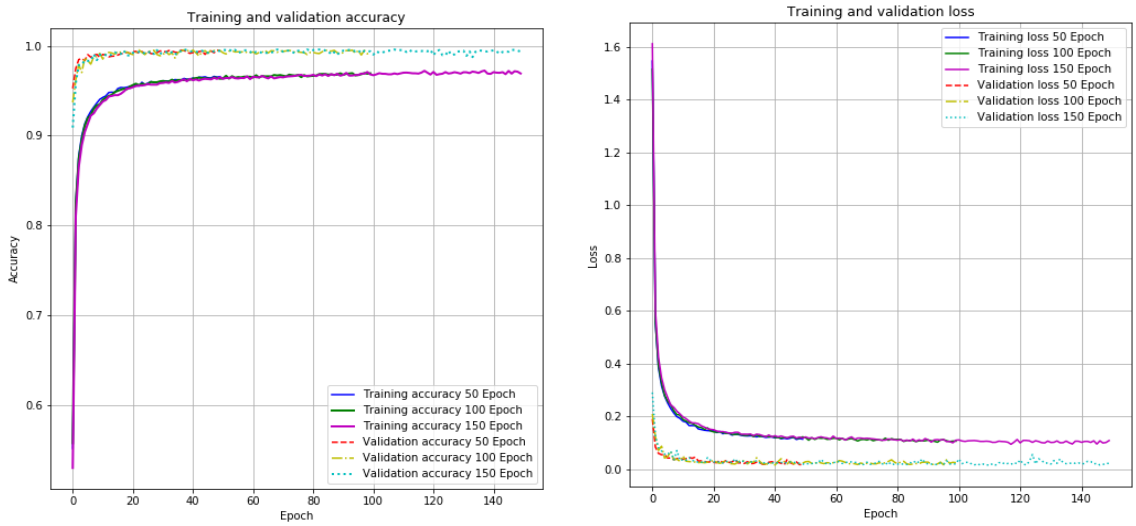


Figure 6: Traffic road signs CNN-based recognition model



(a) Accuracy curves (b) Loss curves

Figure 7: Results of training CNN Model proposed in Table.5

7.2. CNN-based Model Description

The first part of a CNN is the convolutional part, strictly speaking, functions as an extractor of the characteristics of the image, because the images pass through a succession of filters, or convolution cores, creating new images called convolution maps some broker filters reduce the images resolution, by using a local maximum of operation. After the convolution layers or maps, are flatteners and reassemble into a vector of features.

The CNN base on the MLP (Multi-Layer Perceptron) [26], and inspired by the behavior of the visual cortex, although effective for image processing, MLPs have great difficulty in handling large images, which are due to the exponential growth in the number of connections to the image size, for example, if taking an image size of 32x32x3, so the first hidden layer, need fully-connected neuron of the MLP, to have in inputs 3072 neuron, but for 224x224x3 size images, lead to being processed in inputs, need 150,528 neuron, which means when the images size are larger, the number becomes enormous of neurons.

A CNN architecture formed by an ensemble of different processing layers, the important layers are:

- ❖ The convolution layer (Conv2D) a receiver field the data, explained before.
- ❖ The pooling layer (MaxPooling2D) compresses or reduces the image size by subsampling.

Table 5: Architecture CNN Model

Layers	Output	Parameters
Conv2D	(None, 28, 28, 60)	1560
Conv2D	(None, 24, 24, 60)	90060
MaxPooling2D	(None, 12, 12, 60)	0
Dropout	(None, 12, 12, 60)	0
Conv2D	(None, 10, 10, 30)	16230
Conv2D	(None, 8, 8, 30)	8130
MaxPooling2D	(None, 4, 4, 30)	0
Dropout	(None, 4, 4, 30)	0
Flatten	(None, 480)	0
Dense	(None, 500)	240500
Dropout	(None, 500)	0
Dense	(None, 43)	21543
Total parameters: 378,023 Trainable parameters: 378,023 Non-trainable parameters: 0		

Table.5 is composed of four convolution layers, two layers of Max-pooling, and two layers of fully connected. The input image is 32*32 in size; goes to the first convolution layer, this layer composed of 60 filters of 5*5 size, with ReLU activation function forces neuron to return positive values, after this convolution, then 60 feature maps of size 32*32 are obtained, which represent the input of the second convolution layer, also composed of 60 filters. The Max pooling layer was applied afterward, to reduce the image size. At the output of this layer, we will have 60 feature maps of size 12*12. We repeat the same thing with the last convolution layers composed of 30 filters; the ReLU function always applies to each convolution, and Max-pooling layer valid all the time, at the output of these layers, the features vectors of sizes 480. Finally, using a neural network composed of two fully connected layers, the first layer composes of 500 neurons, where the ReLU activation function was used, the second layer uses softmax function, for calculating the probability distribution of the 43 classes of traffic roads signs.

7.3. Trained Model

According to Figure.7 (a), the curves of training accuracy increase when several epochs increases, the same observation for curves of validation accuracy, reflecting the fact of epochs, can conclude the model learns more information. If the validation

accuracy can't follow training accuracy, during the training model, this means that the CNN architecture, did not extract enough characteristics, for each class from the road sign images, called the underfitting, and when validation accuracy decreases during the training model, may result to overfitting, can conclude more features, allowing to model learn enough. Therefore, the curves of training loss decrease to achieve minimum error values, according to the number of epochs in the figure.7 (b), and the same for curves of validation loss. For all the graphs shown, the model does a very good job and powerful, according to the best scores shown in the figures previously during training, has reaches 11% in error rate and the totality for a 97% accuracy rate.

Table 6: Scores of accuracy and loss

Model	Epoch	Test Loss (%)	Test Accuracy (%)	Processor Type
CNN Model	50	06.93	97.91	NVIDIA GeForce 920M (GPU)
	100	11.64	97.05	
	150	11.31	97.11	

7.4. Training model description

Table.6 shows the results of scores accuracy and loss, using the test dataset, the execution time is too expensive, this comes back to the large size of the base which requires to use GPU that CPU [27]. The model presented the best results found, the number of epochs and convolution layers reflect these good results; however, the execution time was very expensive because of the number of epochs.

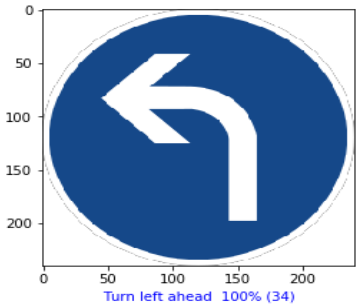
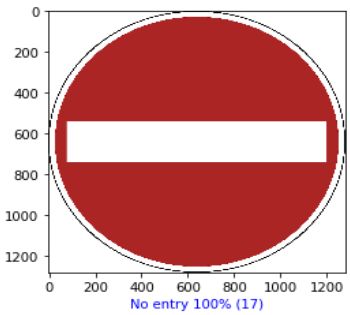
In general, a deep convolutional neural network gives good results and the performance of our network degrades, if the number of epochs increases. For example, according to the table, the increase leads to a loss of about 0.86% of the performance of the network. Therefore, depth is crucial to achieving good results. The results obtained have improved as we have deepened our network with the number of epochs. The learning base is also a crucial element in convolutional neural networks, you need a large learning base to achieve better results in the future.

8. System Performance Evaluation

8.1. Cropped images recognition

As can see in Table.7, after the detection of circular-shaped panels with high accuracy in a real-time environment. Then we cut the road signs, the next phase is the most important for our system, the recognition of the road signs, identifying it to which class or category belongs in an intelligent way using deep learning. The objects that contain both images have been classified correctly without any errors. The red sign has been classified or recognized with a prediction of 100% No-entry, the same for the blue sign with a percentage of 100% Turn left. So, we have good results that can be a basis for the improvement of ADAS systems in the automotive sector that will serve the evolution of industry 4.0.

Table 7: Recognition of TRAFFIC ROADS SIGNS cropped

Cropped Signs	Recognition Signs
	
	

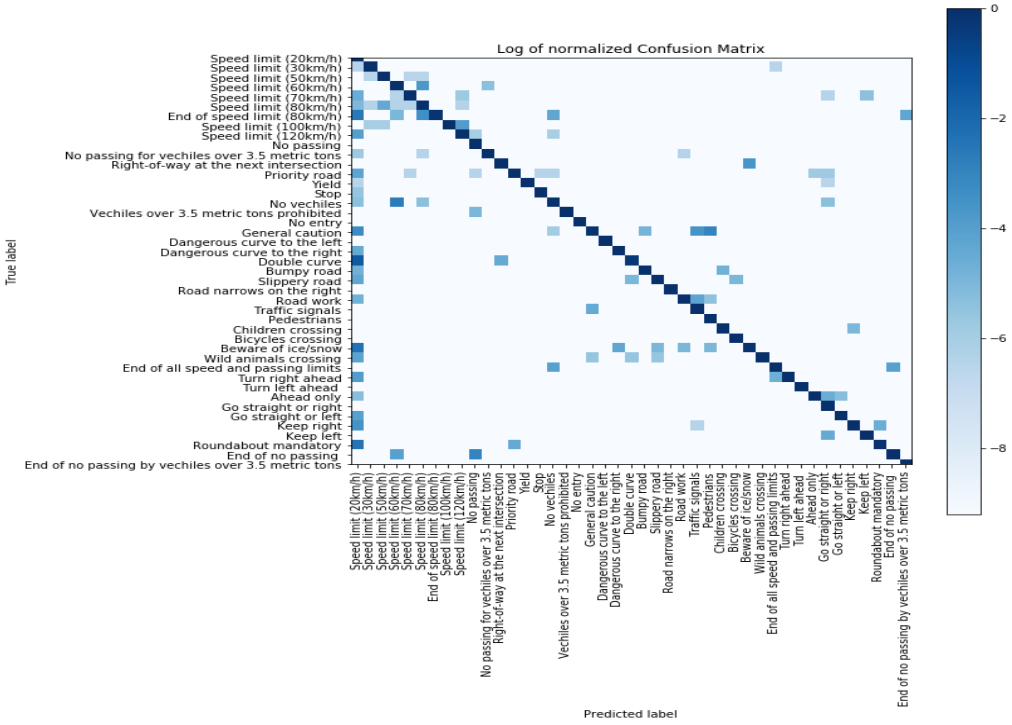


Figure 8: Confusion Matrix of CNN model

	precision	recall	f1-score	support
Speed limit (20km/h)	1.00	1.00	1.00	60
Speed limit (30km/h)	1.00	1.00	1.00	720
Speed limit (50km/h)	0.99	1.00	0.99	750
Speed limit (60km/h)	0.96	0.97	0.97	450
Speed limit (70km/h)	1.00	0.98	0.99	660
Speed limit (80km/h)	0.97	0.97	0.97	630
End of speed limit (80km/h)	1.00	0.86	0.92	150
Speed limit (100km/h)	1.00	0.98	0.99	450
Speed limit (120km/h)	0.97	0.98	0.98	450
No passing	0.99	1.00	0.99	480
No passing for vechiles over 3.5 metric tons	1.00	0.99	1.00	660
Right-of-way at the next intersection	1.00	0.97	0.99	420
Priority road	1.00	0.97	0.99	690
Yield	1.00	1.00	1.00	720
Stop	1.00	1.00	1.00	270
No vechiles	0.97	0.92	0.94	210
Vechiles over 3.5 metric tons prohibited	1.00	0.99	1.00	150
No entry	1.00	1.00	1.00	360
General caution	0.99	0.87	0.93	390
Dangerous curve to the left	1.00	1.00	1.00	60
Dangerous curve to the right	0.98	0.99	0.98	90
Double curve	0.97	0.77	0.86	90
Bumpy road	0.98	0.98	0.98	120
Slippery road	0.99	0.97	0.98	150
Road narrows on the right	1.00	1.00	1.00	90
Road work	1.00	0.97	0.98	480
Traffic signals	0.90	0.99	0.94	180
Pedestrians	0.73	1.00	0.85	60
Children crossing	0.99	0.99	0.99	150
Bicycles crossing	0.99	1.00	0.99	90
Beware of ice/snow	0.92	0.87	0.90	150
Wild animals crossing	1.00	0.97	0.99	270
End of all speed and passing limits	0.95	0.97	0.96	60
Turn right ahead	1.00	0.97	0.99	210
Turn left ahead	1.00	1.00	1.00	120
Ahead only	0.99	0.98	0.99	390
Go straight or right	0.92	1.00	0.96	120
Go straight or left	0.97	0.98	0.98	60
Keep right	1.00	0.96	0.98	690
Keep left	0.97	0.99	0.98	90
Roundabout mandatory	0.92	0.90	0.91	90
End of no passing	0.98	0.93	0.96	60
nd of no passing by vechiles over 3.5 metric tons	0.98	1.00	0.99	90
micro avg	0.99	0.97	0.98	12630
macro avg	0.98	0.97	0.97	12630
weighted avg	0.99	0.97	0.98	12630
samples avg	0.97	0.97	0.97	12630

Figure 9: The classification rapport of the CNN model

8.1. Confusion Matrix

The Confusion Matrix [28] is used to evaluate the performance of our model, since it reflects the metrics of True Positive, True Negative, False Positive, and False Negative. Figure.8 closely illustrates the position of these metrics for each class. For example, the model ranked the automotive images well and is ranked or confused by the images of the other classes. We notice that there is a little confusion in classification on everything for the first class, which reduced the classification rate a little.

The average classification rate, as shown in the figure. 9 is 98%, which indicates our good system, for precision and recall, and high classification values for all categories.

9. Conclusion

Object detection in images and more specifically the detection of road signs images is a problem addressed by artificial vision. Among its fields of application are video surveillance, road safety, detection assistance systems. In our work, we were interested in the detection of signs. A bibliographical study has been carried out to determine the possible detection adopted methods and the most used descriptors in this case. Image classification is an important task in the computer vision field, object recognition, and machine learning. Through Deep Learning, the future of artificial intelligence enhances large and very fast developing algorithms. We discussed how the detection process and image classification took by a camera device during the discussion of the concepts of the neural network. To realize our classification work we used deep learning, which has shown its performance in recent years and we chose the CNNs method as the classification method, this choice is justified by the method's simplicity and its efficiency. During the tests phase, the obtained result confirms our approach effectiveness. Our work is only in its initial version and remains open for comparison with other classification methods [29].

The objective was to create a system for the automotive sector that combines detection and recognition. Detection based on the color and geometry of road signs, recognition based on deep learning based on CNN architecture, we got a score of 100% for classifying the test images captured from a real environment. Our work to be complimentary for our next work of detected geometry: triangular and rectangular, and classified it using Artificial Intelligence. This approach can be applied to the development of embedded systems for the automobile, which will serve road safety for all road users, either the passenger or the driver.

References

[1] L. Li, "China's manufacturing locus in 2025: With a comparison of 'Made-in-China 2025' and 'Industry 4.0,'" Technological Forecasting and Social Change, **135**(February), 66–74, 2018, doi:10.1016/j.techfore.2017.05.028.

[2] K.M. Tan, V.K. Ramachandaramurthy, J.Y. Yong, "Integration of electric vehicles in smart grid: A review on vehicle to grid technologies and optimization techniques," Renewable and Sustainable Energy Reviews, **53**, 720–732, 2016, doi:10.1016/j.rser.2015.09.012.

[3] K. Armstrong, S. Das, J. Cresko, "The energy footprint of automotive electronic sensors," Sustainable Materials and Technologies, **25**, e00195, 2020, doi:10.1016/j.susmat.2020.e00195.

[4] S. Hossain, D.J. Lee, "Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices," Sensors (Switzerland), **19**(15), 2019, doi:10.3390/s19153371.

[5] A. Barodi, A. Bajit, M. Benbrahim, A. Tamtaoui, "An Enhanced Approach in Detecting Object Applied to Automotive Traffic Roads Signs", IEEE International Conference on Optimization and Applications (ICOA), 1–6 ,2020 , doi:10.1109/ICOA49421.2020.9094457.

[6] R.Y. Zhong, X. Xu, E. Klotz, S.T. Newman, "Intelligent Manufacturing in the Context of Industry 4.0: A Review," Engineering, **3**(5), 616–630, 2017, doi:10.1016/J.ENG.2017.05.015.

[7] K. Okuma, A. Taleghani, N. De Freitas, J.J. Little, D.G. Lowe, "A boosted particle filter: Multitarget detection and tracking," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **3021**, 28–39, 2004, doi:10.1007/978-3-540-24670-1_3.

[8] Y. Lu, Z. Zhou, J. Zhao, "Visual Object Tracking Using PCA Correlation Filters," 2544–2550, 2018, doi:10.2991/caai-18.2018.10.

[9] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, J. Jatakia, "Human skin detection using RGB, HSV and YCbCr color models," ArXiv, **137**, 324–332, 2017, doi:10.2991/iccasp-16.2017.51.

[10] A. Uçar, Y. Demir, C. Güzeliş, "Object recognition and detection with deep learning for autonomous driving applications," Simulation, **93**(9), 759–769, 2017, doi:10.1177/0037549717709932.

[11] R. Ranjan, V.M. Patel, R. Chellappa, "HyperFace: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition,"IEEE Transactions on Pattern Analysis and Machine Intelligence, **41**(1), 121–135, 2019, doi:10.1109/TPAMI.2017.2781233.

[12] M. Rout, J.K. Rout, H. Das, Correction to: Nature Inspired Computing for Data Science, 2020, doi:10.1007/978-3-030-33820-6_12.

[13] V. Balali, A. Jahangiri, S.G. Machiani, "Multi-class US traffic signs 3D recognition and localization via image-based point cloud model using color candidate extraction and texture-based recognition," Advanced Engineering Informatics, **32**, 263–274, 2017, doi:10.1016/j.aei.2017.03.006.

[14] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," Proceedings of the International Joint Conference on Neural Networks, 2013, doi:10.1109/IJCNN.2013.6706807.

[15] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, D. Metaxas, "StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks," Proceedings of the IEEE International Conference on Computer Vision, **2017-Octob**, 5908–5916, 2017, doi:10.1109/ICCV.2017.629.

[16] K. Muhammad, M. Sajjad, I. Mehmood, S. Rho, S.W. Baik, "Image steganography using uncorrelated color space and its application for security of visual contents in online social networks," Future Generation Computer Systems, **86**, 951–960, 2018, doi:10.1016/j.future.2016.11.029.

[17] G. Saravanan, G. Yamuna, S. Nandhini, "Real time implementation of RGB to HSV/HSI/HSL and its reverse color space models," International Conference on Communication and Signal Processing, ICCSP 2016, 462–466, 2016, doi:10.1109/ICCSP.2016.7754179.

[18] K. Preethi, K.S. Vishvaksenan, "Gaussian Filtering Implementation and Performance Analysis on GPU," Proceedings of the International Conference on Inventive Research in Computing Applications, ICIRCA 2018, (Icirca), 936–939, 2018, doi:10.1109/ICIRCA.2018.8597299.

[19] A. Barodi, A. Bajit, S. E. Aidi, M. Benbrahim, and A. Tamtaoui, "Applying Real-Time Object Shapes Detection To Automotive Traffic Roads Signs.", International Symposium on Advanced Electrical and Communication Technologies (ISAECT) , 1–6 ,2020-- Proceeding.

[20] B. Li, S.T. Acton, "Automatic active model initialization via poisson inverse gradient," IEEE Transactions on Image Processing, **17**(8), 1406–1420, 2008, doi:10.1109/TIP.2008.925375.

[21] M. Baygin, M. Karakose, A. Sarimaden, E. Akin, "An image processing based object counting approach for machine vision application," ArXiv, 18.

[22] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, "The German Traffic Sign Recognition Benchmark for the IJCNN'11 Competition," Proc. of the International Joint Conference on Neural Networks, 1453–1460, 2011.

[23] A. Mikołajczyk, M. Grochowski, "Data augmentation for improving deep learning in image classification problem', Inter. Interdisciplinary PhD Workshop (IIPhDW), 117–122, 19, doi10.1109/IIPHDW. 2018.8388338."

[24] G.H. de Rosa, J.P. Papa, X.S. Yang, "Handling dropout probability estimation in convolution neural networks using meta-heuristics," Soft Computing, **22**(18), 6147–6156, 2018, doi:10.1007/s00500-017-2678-4.

[25] H.C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Transactions on Medical Imaging, **35**(5), 1285–1298, 2016, doi:10.1109/TMI.2016.2528162.

[26] V.F. Kuzishchin, V.A. Dronov, "The specific features of the algorithm for

tuning controllers on the basis of the kvint firmware system," Thermal Engineering, **48**(10), 835–841, 2001.

[27] Y.D. Zhang, K. Muhammad, C. Tang, "Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification on GPU platform," Multimedia Tools and Applications, **77**(17), 22821–22839, 2018, doi:10.1007/s11042-018-5765-3.

[28] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," Neurocomputing, **321**, 321–331, 2018, doi:10.1016/j.neucom.2018.09.013.

[29] A. Barodi, A. Bajit, M. Benbrahim, and A. Tamtaoui,'Improving the Transfer Learning Performances in the Classification of the AutomotiveTraffic Roads Signs', E3S Web Conf., 2020 -- Proceeding.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 4, 676-685 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Clustering of Mindset towards Self-Regulated Learning of Undergraduate Students at the University of Phayao

Pratya Nuankaew*

School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand

ARTICLE INFO

Article history:
Received: 22 June, 2020
Accepted: 11 August, 2020
Online: 25 August, 2020

Keywords:
Educational Data Mining
Learning Analytics Model
Learning Styles
Self-Regulated Learning

ABSTRACT

The effects of Covid-19 severely affected the Thai higher education model. Therefore, there are three significant objectives in this research: (1) to cluster the mindsets and attitudes toward self-regulated learning styles of undergraduate students at the University of Phayao. (2) to construct a predictive model for recommending an appropriate student learning clusters. (3) to evaluate the predictive model that has been constructed. Samples collected a compilation of 472 student satisfaction with questionnaires from three schools, with seven disciplines at the University of Phayao, Thailand. Research tools consisted of statistical and machine learning techniques as follows: frequency, percentage, average, standard deviation, k-means clustering, decision tree techniques, cross-validation methods, confusion matrix performance, accuracy, precision, and recall measurement. Researcher found that the k-means model with the highest accuracy is the decision tree model that was classified into three clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of seven levels of the decision tree model and an accuracy of 98.73%. From the results and studies, it can be concluded that the developed model is effective and reasonable to be further developed as an application for further organizational development.

1. Introduction

Nowadays, the learning behavior of youth and the new generation has changed dramatically which have made educational system unable to cope with acts that wants to inquire more about the issue than just following what is instructed. The results have created learning that is more aligned with the new generation who demand to know more and expand their interests [1]–[5]. The learning styles of the new generation of children are highly identified as having a limited attention, known as ADHD: Attention Deficit Hyperactivity Disorder [6], [7]. In addition, the device addiction and mobile addiction symptoms are more intensifying and widespread among youth and the younger generation [8].

These type of behaviors often lead into internet addiction [9]. In 2009, a medical study found that the average age of internet addicts was 17.6 years (range: 12-27 years), as the use of internet was nine hours a day and increasing in proportion [9]. In 2016, there was a study of personality and positive orientation in Internet and Facebook addiction [10]. It has been found that age has a significant effect on factors distinguishing both Internet and Facebook addiction. Moreover, young people more often have problem with excessive use with the Internet and Facebook than adults [10]. Therefore, it can be summarized primarily that young people and the new generation are driven by the changes in technology, which has completely changed the learning behavior of students.

At the same time, the social patterns of the new generation have cause changes in the way of communication, making contact, and having a dialogue. Due to the addiction from the mobile phones and the internet, their habit has made them choose to be more associated with online communities rather than interacting with people in a normal society [11]. From the patterns and behaviors of the aforementioned young generation, the results of formal education, which is a basic education for everyone, are not consistent with the proper standards for student behavior of learning in a classroom environment [11], [12]. It is possible that the traditional teaching has been outdated for a long period of time thus presenting itself to be a problem in developing a healthy and

*Pratya Nuankaew, School of Information and Communication Technology, University of Phayao, Phayao, 56000, Thailand, +66 89 961 4832 & pratya.nu@up.ac.th

sound method for students to make any real progress in learning, and acquiring the necessary employable skills and abilities after graduation. The best solution is to change the style of knowledge management and teaching strategies to suit students.

The learning theory that is consistent and suitable for solving the above problem is the theory of Self-Regulated Learning (SRL), which is a widely accepted theory [13]–[16]. Self-regulated learning strategies can be applied to learning in the new era. It can also be applied to the promotion of Technology Enhanced Learning (TEL), which provides opportunities to increase the learning skills necessary for students [11], [17]. In addition, an important principle of self-regulated strategies is the development of learning that aims learning towards achieving the goals set by the learners themselves [15].

However, learning can happen anywhere and anytime, with each person learning differently, because each day in life presents itself with different situations for exploring something new. Moreover, each event may have similarities or differences with an experience that results in the behavior of the learners who are learning. In addition, when humans learn and achieve academic achievement, the result is a change in learning behavior that is the result from past events or situations. But changing human behavior may not always be learning, due to changes in a certain period of time whereas the person may have to find the self-motivation to get themselves to take part in the learning process.

From the benefits of self-regulated learning strategies, the researchers can use this theory to solve problems and design the learning processes that are appropriate for learners in the new normal education system. This is of vital interest and persuades the researcher to conduct research. The background of the researcher's past work is the study of the behavior of learners at the tertiary [5], [18], [19] and secondary [20], [21]levels. In addition, the researcher is also interested in developing educational models in order to create a learning model that is truly suitable for the learners [22], [23]. The success and achievement that researchers have found is to support learners to achieve learning success and to combine academic achievement. These are the forces that support and continually motivate researchers to pursue our research.

1.1. Research Objectives

There are three significant objectives: The first objective is to cluster the mindsets and attitudes toward self-regulated learning styles of undergraduate students at the University of Phayao. The second objective is to construct a predictive model for recommending an appropriate student learning clusters. Lastly, the third objective is to evaluate the predictive model that has been constructed.

The expectation of this research is to know the impact of the Covid-19 pandemic, which has had a severe impact on the education model by studying the perception and attitudes towards online and online learning styles of learners at the University of Phayao. In addition, the expected result is knowing the group of learners according to the attitude and self-learning style which will be used for developing the quality and potential of students in the future.

1.2. Research Approach

The research approach has been conducted according to the process of the CRISP-DM methodology [24], [25]. It consists of six steps as follows: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The details of the research process are presented on the topic of research methodology. Data collection was carried out in a compilation of 472 student satisfaction with questionnaires from three schools, with seven disciplines at the University of Phayao which is stored on the website: https://bit.ly/2BobB8l.

The research consisted of statistical tools and machine learning tools as follows: percentage, mean, average, standard deviation, k-means clustering, decision tree techniques, cross-validation methods, confusion matrix performance, accuracy, precision, and recall measurement.

1.3. Research Ethic

This research has requested permission from the School of Information and Communication Technology, the University of Phayao, and related agencies, by implementing the regulations of the university.

2. Literature Reviews

2.1. Self-Regulated Learning

Self-regulated learning refers to the process of setting goals, controlling, and managing the sources of knowledge based on the motivation of the learners to set learning goals and expect success in intellectual learning [15], [17], [26]–[29].

The composition of the self-regulated learning consists of three important phases [13], [27], [29]. The first phase is the forethought phase, which consists of two important processes as follows: (1) task analysis, and (2) self-motivation beliefs. The second phase is the performance phase, which consists of two important processes as follows: (1) self-control, and (2) self-observation. Finally, the third phase is the self-reflection phase, which consists of two important processes as follows: (1) self-judgment, and (2) self-reaction. Details of the components of the Self-regulated learning are shown in Figure 1.

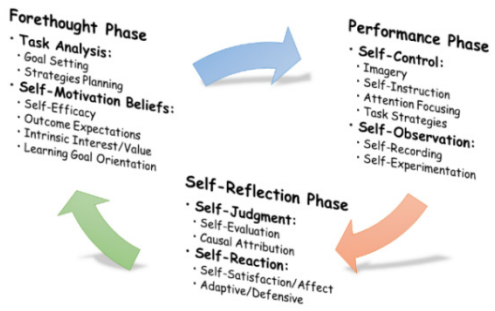


Figure 1: Phases and sub-process of self-regulated learning

From the theory of self-regulated learning, there is a lot of interest in research [13], [14], [17], [28]. It can therefore be concluded that this self-regulated learning theory is appropriate for this research.

2.2. Student Academic Performance

Student academic performance is a method of studying the effectiveness arising from graduation or learners receiving academic achievement. Many researchers have studied learning styles that encourage learners to receive high efficiency [5], [18], [30]–[32]. For example, there has been research works done to study the relationships of multiple variables that affect academic achievement using a method named "Structural Equation Modeling (SEM)" [30]. SEM is a type of statistical model that searches for and describes the relationships between multiple variables. The second example is research which studies the compatibility of different characteristics [32]. In the second example, they attempted to present the concept of the compatibility of the mentor and the receiver by comparing it to a jigsaw. The last example is research that studies the impact of unsuccessful studies or dropouts, in which these researches are discussed in many dimensions and perceptions of the researchers [19], [33].

There are also researchers who study the tools for applying the concepts for measuring and evaluating student performance [33]–[35]. Tools used in their studies include the use of statistical tools and data mining tools: basic statistical tools, decision tree techniques, k-means and k-medoids algorithms, confusion matrix performance, and cross-validation methods.

From the above example, it shows that various researchers clearly value the study of student academic performance. Therefore, it corresponds to the purpose of this research which aims to discover the pattern of relationships affecting graduation or non-graduation as specified by the curriculum.

2.3. Improving Academic Achievement

Generally, researchers in various fields have studied research and development in the science of educational quality development [5], [14], [36]–[38], in which the objectives of each research group are different in methods and perspectives. Some research groups want to study the factors that support and change learners' behavior and instructional methods [3], [13], [31], [33]. Some research groups want to study and develop tools that support the efficiency of teaching and learning [35], [37], [39]. Some research groups want to apply modern technology to be used in teaching and learning activities [40], [41]. Finally, all research groups have a similar objective, which is to encourage learners to have a positive impact with their studies toward graduation.

However, the concept of improving academic achievement is aimed on raising the level of student achievement performance values [5], [15], [38], [42]. The main goal of this research should focus on the three main components: The first target is to cluster the characteristics of mindset towards self-regulated learning styles of undergraduate students at the University of Phayao. The second target is to construct the predictive model for suggesting the appropriate student cluster. The third target is to evaluate the predictive model that has been developed.

2.4. Educational Data Mining

Educational data mining (EDM) is the science of combining the use of data science tools and educational technology for educational data analysis. It consists of machine learning applications, data mining tools and advanced statistics to carry out the process of educational system success [43]–[45]. In addition, the educational data mining refers to techniques, tools, and research designed to automate the definition of large data sources that are related to learning activities in the educational system [5], [19], [35], [45].

Examples of research in this field include research in application development to be recommended as suitable educational institutions for students [4], [35], [39]. They study and research on learners' factors, educational institutions' factors, educational data mining models that encourage learners to study in a suitable educational institution, and develop them into applications. In addition, there are other research studies such as Ahmad's research which analyzes the educational model to see where the best fit are for a particular program [36]. They study the behavior of learners occurring in online activities through a learning management system known as MOOC. Firdausiah Mansur's research [37] proposed a personalized learning model to find suitable learning methods based on a deep learning algorithm. Both of their results are impressive because they have applied machine learning applications, data mining tools, and advanced statistics which enabled them to discover some facts from the data.

Finally, it can be concluded that in the analysis of data mining for education it is necessary to find a new perspective that allows teachers to truly understand the learners.

3. Research Methodology

The research methodology was conducted according to the CRISP-DM method [24], [25]. It has six operations in accordance with the principles of CRISD-DM techniques, which include (1) business understanding, (2) data understanding, (3) data preparation, (4) modeling, (5) evaluation, and (6) deployment, as shown in Figure 2 below.

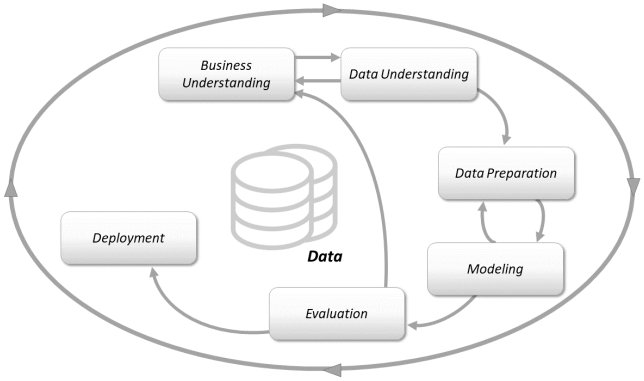


Figure 2: Research Methodology

3.1. Business Understanding

The business understanding phase focuses on understanding the audit objectives and needs from a case perspective and converting this knowledge into a definition of evidence mining problems and a preliminary plan designed to achieve the objectives [24], [25], [46]. For this reason, the researcher aimed

at the three objectives of the research as follows: The first objective is to cluster the characteristics of mindset towards self-regulated learning styles of undergraduate students at the University of Phayao. The second objective is to construct the predictive model for suggesting the appropriate student cluster. Finally, the third objective is to evaluate the predictive model that has been developed.

3.2. Data Understanding

The purpose of data understanding phase is to identify data quality problems, construct questions for finding patterns of data, and to examine interesting subset that create assumptions for hidden data [24], [25], [46]. The studied data was on the students of the University of Phayao, Phayao, Thailand. The data collected consisted of students' data from the School of Information and Communication Technology (ICT), the School of Management and Information Sciences (MIS), and the School of Law at the University of Phayao. In addition, the data collected were from seven disciplines of students majoring in the courses of accounting, business computer, law and accounting, management, marketing, and tourism. The purpose of the various data collection is to get information that covers the attitude of data providers.

3.3. Data Preparation

The data preparation phase aims to design activities that cover all activities to enable data collection for analysis and development of models. It details the sub-steps to prepare the data that will be fed into the research tool to create a model by initial management at raw data. It has five sub-steps to implement: selecting data, cleaning data, constructing data, integrating data, and formalizing data [24], [25], [46].

The data collected were on 472 students from the University of Phayao, Thailand. The data collection is divided into two categories according to the survey type. It consists of 319 students who responded from regular surveys and 153 people who answered from online surveys; details are shown in Table 1 to Table 6. In addition, the data collected in this research is defined in a digital format, which is stored on the website: https://bit.ly/2BobB8l.

Table 1: Data collected is classified by survey types

Survey Types	Absolute Count	Fraction (Percentage)
Regular Survey	319	67.58%
Online Survey	153	32.42%
Total:	*472*	*100.00%*

Table 1 shows that the regular survey has the largest number of respondents, with 319 students representing 67.58 percent of all respondents.

Table 2: data collected is classified by gender.

Gender	Absolute Count	Fraction (Percentage)
Male	112	23.73%
Female	360	76.27%
Total:	*472*	*100.00%*

Table 2 shows that Female have the largest number of respondents, with 360 students representing 76.27 percent of all respondents.

Table 3: Data collected is classified by disciplines.

Disciplines	Absolute Count	Fraction (Percentage)
Accounting	208	44.07%
Business Computer	73	15.47%
Law and Accounting	28	5.93%
Management	3	0.64%
Marketing	43	9.11%
Tourism	117	24.79%
Total:	*472*	*100.00%*

Table 3 shows that Accounting have the largest number of respondents, with 208 students representing 44.07 percent of all respondents. The second group that answered the most questionnaire was Tourism, with 117 students representing 24.79 percent.

Table 4: Data collected is classified by affiliation.

Affiliation	Absolute Count	Fraction (Percentage)
School of Information and Communication Technology (ICT)	73	15.47%
School of Law	28	5.93%
School of Management and Information Sciences (MIS)	371	78.60%
Total:	*472*	*100.00%*

Table 4 shows that the School of Management and Information Sciences (MIS) have the largest number of respondents, with 371 students representing 78.60 percent of all respondents.

Table 5: Data collected is classified by learning styles.

Learning Styles	Absolute Count	Fraction (Percentage)
Online Learning Style	81	17.16%
Offline Learning Style	391	82.84%
Total:	*472*	*100.00%*

Table 5 shows that the collected data is categorized according to the learning styles that students are interested in. Table 5 shows that students are interested in the traditional learning or in front of the classroom, with 391 students representing 82.84 percent of all respondents.

Table 6: Data collected is classified by acceptance of SRL style.

Acceptance of SRL style	Absolute Count	Fraction (Percentage)
Low Level (0-30%)	18	3.81%
Medium Level (31-70%)	418	88.56%
High Level (71-100%)	36	7.63%
Total:	*472*	*100.00%*

Table 6 shows that the level of acceptance to self-regulated learning style is at a medium level (31-70%), with 418 students representing 88.56 percent of all respondents.

3.4. Modeling

Various modeling techniques are selected and implemented, with their parameters being compared to the best values. In general, there are many techniques for the same type of data mining problem. Some techniques require a specific data format data [24], [25], [46]. In addition, modeling is the process of creating a suitable prototype. It consists of four important parts: selecting the modeling techniques, generating test design, building the model, and assessing the model [31], [35].

As mentioned above, the machine learning tools selected are k-means clustering and decision tree techniques. The benefit of the k-means is that it can be recommended for clustering with similar data patterns [31], while the benefit of the decision tree is that it is a structural decision consisting of nodes (features) and leaves (decisions) [35].

In this research, the analysis for clustering and charting of decision tree was based on data from questionnaires filled out by students who were assigned to take part in the research activities. The final result of the modeling is a set of variables (characteristics) that are important for predicting a reasonable cluster for the learners, which will be used to suggest activities for the learners in the next academic year.

3.5. Evaluation

The goal of the evaluation is to evaluate the performance of the results, which aims to construct the significant relationship models [19], [33]. The tools are used in the research, including the cross-validation techniques as shown in Figure 3, and the calculation of confusion matrix as mentioned in Figure 4.

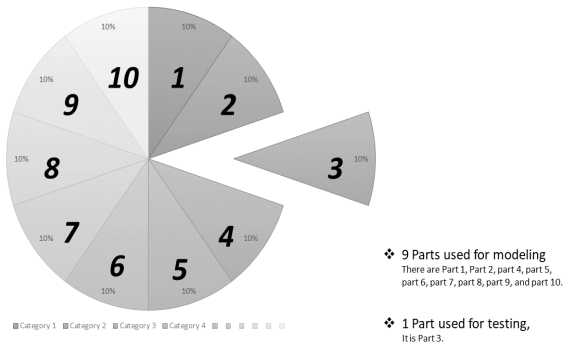


Figure 3: Cross-Validation Techniques

Figure 3 displays the separation of data for evaluating the model. The cross-validation method divides the data into two parts. The first part is used for modeling and the remainder is for testing the model. In addition, model evaluation requires a tool called a confusion matrix to test the model's performance, with the principles shown in Figure 4.

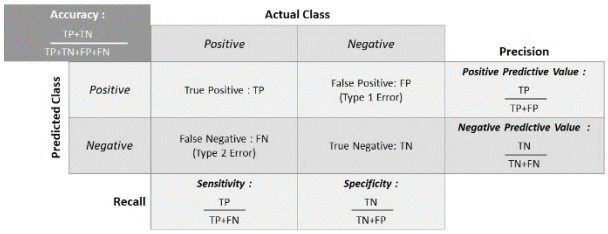


Figure 4: Confusion Matrix Calculation

Figure 4 presents the composition of the confusion matrix performance, which is composed of the actual class and the predicted class. An important benefit of the performance of the confusion matrix is the ability to determine the model's ability to predict results, such as the predictive ability or accuracy, model precision, model sensitivity, and model specificity (recall measurement). These values are used to determine the actual performance model. Moreover, Figure 4 also demonstrates the formulas and methods for calculating the various performance parameters in detail.

3.6. Deployment

The deployment is intended to bring results and discoveries in order to establish relationships and analyze the relationships that are discovered.

3.6.1. Testing Model Results

As mentioned early, the tools used to evaluate the model consist of two parts: cross-validation method and confusion matrix performance. This section presents the use of evaluation tools in research. The testing process divides the data into two parts according to the cross-validation method principles.

There are three methods of cross-validation in this research. The first method is 10-Fold cross-validation, which used 9-Fold for modeling and 1-Fold for testing. The second method is 50-Fold cross-validation, which used 49-Fold for modeling and 1-Fold for testing. The last method is leave-one-out cross-validation, which used 99 percent of data for modeling and 1 percent of data for testing. However, each time the cross-validation test is reported, the model results are also tested by using the confusion matrix performance.

3.6.2. Applying Model Results

The purpose of applying model results is to manage the developed models. It has four sections as follows: (1) decision tree model, (2) decision tree model applying results, (3) cluster model, and (4) number of members in each cluster.

4. Research Results

In the research results, the researchers classified the research report into four parts which are satisfaction level towards questionnaire, modeling results, model testing results, and model applying results.

4.1. Satisfaction level towards questionnaire

This section summarizes the satisfaction levels of the 472 students, which contain data from three schools and seven

disciplines at the University of Phayao. The details of the summary results are shown in Table 8.

In interpreting the data according to the characterization criteria, the interpretation is based on a five-level interpretation method by comparing it with the criteria that divides the level estimation into five equal levels, as followed in Equation (1). The result of the calculation is shown in Equation (2).

$$Width\ of\ the\ level\ = \frac{Maximum-Minimum}{Number\ of\ levels} \quad (1)$$

$$Width\ of\ the\ level\ = \frac{5-1}{5} = 0.8 \quad (2)$$

From the calculation results in Equation (2), the interpretation results can be specified as shown in Table 7.

Table 7: Interpretation and Meaning

Width of the level	Interpretation	Meaning
1.00 – 1.80	Highly Unacceptable	The lowest level of satisfaction
1.81 – 2.60	Unacceptable	Low level of satisfaction
2.61 – 3.40	Acceptable	Satisfaction
3.41 – 4.20	Highly Acceptable	High level of satisfaction
4.21 – 5.00	Maximum Acceptable	The highest level of satisfaction

Table 8: Satisfaction level towards questionnaire

Question	Average	S.D.	Interpretation
Question 1: What level of self-regulated learning can you apply to your studies?	3.38	0.64	Highly Acceptable
Question 2: What level do you agree that self-regulated learning is important for online learning?	3.58	0.78	Highly Acceptable
Question 3: What level do you agree that finding your own knowledge can be more useful than studying in the classroom?	3.45	0.81	Highly Acceptable
Question 4: What level do you agree that inspiration and motivation have more influence on learning than the advice of teachers?	3.73	0.56	Highly Acceptable
Total Average:	*3.54*	*0.79*	Highly Acceptable

Table 8 shows the level of satisfaction with the four questionnaires. It can be seen that the overall level of satisfaction is high (3.54). It can therefore be concluded that the students of the University of Phayao are satisfied with the high level of self-regulated learning styles.

4.2. Modeling Results

Modeling results are the various models on different criteria, such as defining the unequal depth of the decision tree, and determining the different types of cross-validation method tests, which have the results shown in Table 9 and Table 10.

Table 9: Modeling results from k-Means clustering

Cluster Number	Depth of Decision Tree	Cross-validation method		
		10-Fold	*50-Fold*	*Leave-one-out*
3 Clusters	Level 2	69.06%	69.09%	69.07%
	Level 3	89.38%	89.47%	89.41%
	Level 4	90.65%	91.60%	91.53%
	Level 5	92.56%	92.40%	92.37%
	Level 6	95.34%	95.73%	95.55%
	Level 7*	98.30%	98.69%	98.73%*
	Level 8	98.30%	98.69%	98.73%
	Level 9	98.30%	98.69%	98.73%
4 Clusters	Level 2	59.12%	59.22%	59.11%
	Level 3	78.40%	78.98%	78.81%
	Level 4	87.08%	84.22%	82.63%
	Level 5	91.52%	92.73%	91.10%
	Level 6	93.01%	92.07%	89.41%
	Level 7*	95.55%*	95.00%	92.58%
	Level 8	95.55%	95.00%	92.58%
	Level 9	95.55%	95.00%	92.58%
5 Clusters	Level 2	57.42%	57.51%	57.42%
	Level 3	80.93%	81.00%	80.93%
	Level 4	95.35%	95.62%	95.55%
	Level 5*	96.83%	97.07%*	97.03%
	Level 6	96.83%	97.07%	97.03%
	Level 7	96.83%	97.07%	97.03%
	Level 8	96.83%	97.07%	97.03%
	Level 9	96.83%	97.07%	97.03%
6 Clusters	Level 2	53.61%	53.62%	53.60%
	Level 3	56.34%	58.16%	58.05%
	Level 4	77.12%	77.84%	77.97%
	Level 5	86.24%	85.51%	84.75%
	Level 6	93.21%	93.71%	93.64%
	Level 7*	97.24%	97.71%	97.67%*
	Level 8	97.24%	97.71%	97.67%
	Level 9	97.24%	97.71%	97.67%
7 Clusters*	Level 2	45.77%	46.24%	46.19%
	Level 3	56.16%	56.18%	56.14%
	Level 4	71.42%	73.53%	77.12%
	Level 5	82.23%	85.69%	86.65%
	Level 6	88.58%	90.36%	89.83%
	Level 7	94.29%	94.13%	94.28%
	Level 8	96.20%	95.24%	95.13%
	Level 9*	97.90%*	97.00%	97.03%

From Table 9, it shows that the k-means model with the highest accuracy is the decision tree model that is classified into 3 clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of 7 levels of the decision tree model and has an accuracy of 98.73%. In Addition, the testing results classified by other clusters yield a lower accuracy. For example, the 4 clusters with the highest results are 95.55%, the 5 clusters with the highest results are 97.07%, the 6 clusters with the highest results are 97.67%, and the 7 clusters with the highest results are 97.90%.

4.3. Model Testing Results

From the results of the prototype model development, it can be concluded that the model with the highest accuracy is the development of the model from k-means clustering, with the appropriate number of 3 clusters and the leave-one-out cross-validation result with an accuracy of 98.73%. Details of the developed model are shown in Table 10.

Table 10: Model testing results

Precision Class		*97.37%*	*97.50%*	*100.00%*	*98.37%*	*97.96%*	*91.67%*	*97.30%*	
Actual Class	*True Cluster_6*	1	0	0	0	0	0	36	*97.30%*
	True Cluster_5	1	0	0	0	0	22	1	*91.67%*
	True Cluster_4	1	0	0	1	48	0	0	*96.00%*
	True Cluster_3	0	0	0	121	0	0	0	*100.00%*
	True Cluster_2	0	1	85	1	0	0	0	*97.70%*
	True Cluster_1	0	39	0	0	1	2	0	*92.86%*
	True Cluster_0	111	0	0	0	0	0	0	*100.00%*
Accuracy: 97.90% (+/- 2.41%)		*Pred. Cluster_0*	*Pred. Cluster_1*	*Pred. Cluster_2*	*Pred. Cluster_3*	*Pred. Cluster_4*	*Pred. Cluster_5*	*Pred. Cluster_6*	**Recall Class**
		Predicted Class							

Table 10 shows that the model performance testing is at the highest level, which can confirm the suitability of the model as well.

4.4. Model Applying Results

From the model that has been selected and demonstrated the performance, this section presents the decision tree model as shown in Table 11, the decision tree rules model for self-testing as shown in the test results in Table 12, the cluster model as shown in Table 13, and number of members in each cluster as shown in Table 14.

Table 11: Decision tree model

Decision Tree Model

```
Question 2 > 3.500
|  Question 4 > 2.500
|  |  Question 3 > 3.500
|  |  |  Question 4 > 3.500
|  |  |  |  Question 1 > 4.500
|  |  |  |  |  Question 4 > 4.500: cluster_1 {cluster_1=11}
|  |  |  |  |  Question 4 ≤ 4.500: cluster_5 {cluster_1=1, cluster_5=2}
|  |  |  |  Question 1 ≤ 4.500
|  |  |  |  |  Question 2 > 4.500
|  |  |  |  |  |  Question 1 > 3.500: cluster_1 {cluster_1=20}
|  |  |  |  |  |  Question 1 ≤ 3.500
|  |  |  |  |  |  |  Question 3 > 4.500: cluster_1 {cluster_1=3}
|  |  |  |  |  |  |  Question 3 ≤ 4.500: cluster_0 {cluster_0=7}
|  |  |  |  |  Question 2 ≤ 4.500
|  |  |  |  |  |  Question 3 > 4.500
|  |  |  |  |  |  |  Question 1 > 3.500: cluster_1 {cluster_1=6}
|  |  |  |  |  |  |  Question 1 ≤ 3.500: cluster_0 {cluster_0=6}
|  |  |  |  |  |  Question 3 ≤ 4.500: cluster_0 {cluster_0=71}
|  |  |  Question 4 ≤ 3.500
|  |  |  |  Question 3 > 4.500: cluster_4 {cluster_1=1, cluster_4=3}
|  |  |  |  Question 3 ≤ 4.500
|  |  |  |  |  Question 1 > 3.500: cluster_2 {cluster_2=10}
|  |  |  |  |  Question 1 ≤ 3.500: cluster_4 {cluster_2=1, cluster_4=3}
|  |  Question 3 ≤ 3.500
|  |  |  Question 4 > 4.500: cluster_0 {cluster_0=17, cluster_6=1,
cluster_5=1}
|  |  |  Question 4 ≤ 4.500: cluster_2 {cluster_2=75}
|  Question 4 ≤ 2.500
|  |  Question 3 > 3.500: cluster_2 {cluster_2=1, cluster_4=1}
|  |  Question 3 ≤ 3.500: cluster_3 {cluster_3=4}
Question 2 ≤ 3.500
|  Question 3 > 3.500
|  |  Question 1 > 3.500
|  |  |  Question 4 > 3.500: cluster_5 {cluster_5=20}
|  |  |  Question 4 ≤ 3.500: cluster_4 {cluster_4=7}
|  |  Question 1 ≤ 3.500
|  |  |  Question 4 > 4.500: cluster_0 {cluster_0=10, cluster_4=1}
|  |  |  Question 4 ≤ 4.500: cluster_4 {cluster_4=35}
|  Question 3 ≤ 3.500
|  |  Question 4 > 3.500: cluster_6 {cluster_6=36, cluster_5=1}
|  |  Question 4 ≤ 3.500: cluster_3 {cluster_3=117}
```

Table 11 shows the decision tree model. It can be developed into a decision tree rules model, as shown in Table 12. Decision tree rules model, it is used to test models using the data collected in the developed model tests.

Table 12: Decision tree rules model applying results

Rule	Condition (If)	Prediction (Then)
1	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 > 4.5 and Question 4 > 4.5	Then, suitable for cluster_1 = 100%
2	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 > 4.5 and Question 4 ≤ 4.5	then cluster_5 (0 / 1 / 0 / 0 / 0 / 2 / 0)
3	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 ≤ 4.5 and Question 2 > 4.5 and Question 1 > 3.5	Then, suitable for cluster_1 = 100%
4	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1	Then, suitable for cluster_1 = 100%

Rule	Condition (If)	Prediction (Then)
	≤ 4.5 and Question 2 > 4.5 and Question 1 ≤ 3.5 and Question 3 > 4.5	
5	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 ≤ 4.5 and Question 2 > 4.5 and Question 1 ≤ 3.5 and Question 3 ≤ 4.5	Then, suitable for cluster_0 = 100%
6	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 ≤ 4.5 and Question 2 ≤ 4.5 and Question 3 > 4.5 and Question 1 > 3.5	Then, suitable for cluster_1 = 100%
7	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 ≤ 4.5 and Question 2 ≤ 4.5 and Question 3 > 4.5 and Question 1 ≤ 3.5	Then, suitable for cluster_0 = 100%
8	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 > 3.5 and Question 1 ≤ 4.5 and Question 2 ≤ 4.5 and Question 3 ≤ 4.5	Then, suitable for cluster_0 = 100%
9	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 ≤ 3.5 and Question 3 > 4.5	Then, suitable for cluster_1 = 25%, and suitable for cluster_4 = 75%
10	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 ≤ 3.5 and Question 3 ≤ 4.5 and Question 1 > 3.5	Then, suitable for cluster_2 = 100%
11	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 > 3.5 and Question 4 ≤ 3.5 and Question 3 ≤ 4.5 and Question 1 ≤ 3.5	Then, suitable for cluster_2 = 25%, and suitable for cluster_4 = 75%
12	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 ≤ 3.5 and Question 4 > 4.5	Then, suitable for cluster_0 = 89.47%, suitable for cluster_5 = 5.26%, and suitable for cluster_6 = 5.26%
13	If Question 2 > 3.5 and Question 4 > 2.5 and Question 3 ≤ 3.5 and Question 4 ≤ 4.5	Then, suitable for cluster_2 = 100%
14	If Question 2 > 3.5 and Question 4 ≤ 2.5 and Question 3 > 3.5	Then, suitable for cluster_2 = 50%, and suitable for cluster_4 = 50%
15	If Question 2 > 3.5 and Question 4 ≤ 2.5 and Question 3 ≤ 3.5	Then, suitable for cluster_3 = 100%
16	if Question 2 ≤ 3.5 and Question 3 > 3.5 and Question 1 > 3.5 and Question 4 > 3.5	Then, suitable for cluster_5 = 100%
17	if Question 2 ≤ 3.5 and Question 3 > 3.5 and Question 1 > 3.5 and Question 4 ≤ 3.5	Then, suitable for cluster_4 = 100%
18	if Question 2 ≤ 3.5 and Question 3 > 3.5 and Question 1 ≤ 3.5 and Question 4 > 4.5	Then, suitable for cluster_0 = 90.90%, and suitable for cluster_4 = 9.10%
19	if Question 2 ≤ 3.5 and Question 3 > 3.5 and Question 1 ≤ 3.5 and Question 4 ≤ 4.5	Then, suitable for cluster_4 = 100%
20	if Question 2 ≤ 3.5 and Question 3 ≤ 3.5 and Question 4 > 3.5	Then, suitable for cluster_6 = 97.30% Then, suitable for cluster_5 = 3.70%
21	if Question 2 ≤ 3.5 and Question 3 ≤ 3.5 and Question 4 ≤ 3.5 then	Then, suitable for cluster_3 = 100%
		Correct: 464 out of 472 training examples (98.31%)

Table 12 shows the decision tree rules model. It consists of 21 rules. The test results show that the rules can be predicted at a high level, which can accurately predict 464 data sets from a total of 472 data (98.31%). It can therefore be concluded that the developed models are appropriate by showing details of each cluster in Table 13.

Table 13: Cluster model

Cluster / Attribute	Question 1	Question 2	Question 3	Question 4
Cluster_0	3.42	4.00	3.91	4.47
Cluster_1	4.24	4.76	4.67	4.71
Cluster_2	3.39	4.18	3.02	3.46
Cluster_3	3.00	2.90	2.72	2.81
Cluster_4	3.14	3.10	4.12	3.36
Cluster_5	4.25	3.08	4.04	4.33
Cluster_6	3.22	2.78	2.84	4.19

Table 13 shows details of each cluster. In addition, Table 14 shows the number of members in each cluster.

Table 14: Number of members in each cluster

Cluster	Absolute Count (Member)	Fraction (Percentage)
Cluster_0	111	23.52%
Cluster_1	42	8.90%
Cluster_2	87	18.43%
Cluster_3	121	25.64%
Cluster_4	50	10.59%
Cluster_5	24	5.08%
Total:	*472*	*100.00%*

From Table 9 to Table 10, it details of testing and selection of suitable models. Moreover, Tables 11 to Table 14 are showing the members and the centroid of each cluster. Finally, based on the data and the results of this research it can be concluded that the developed model is very suitable for this study.

5. Research Discussion

In this research, the researcher has divided the discussion process into two sections as follows: the first section is the discussion report of data collection, and the second section is the discussion of the model, testing results and model effectiveness.

5.1. Data Collection Discussion

Based on the summary of the research data collection report from the University of Phayao there were 472 students divided into two groups as follows: the first group is to collect data directly (regular surveys collection) with 319 students. The second group is 153 students from online data collection.

From the data gathered, it can be concluded that this data is small. The researchers should expand the results of the study and gather more data for further analysis in order to comply with data mining principles that require large amounts of data.

5.2. Identify the Headings

Table 9 to Table 14 show the model development process by presenting the analysis results for selecting the model in Table 9, which was done by showing the model performance test in Table 10 and testing the model by data collection in Table 12. It can be concluded that the model is effective and accepted by this research. In addition, Table 13 and Table 14 show the clustering and membership model for each cluster. It can be concluded that each cluster has members distributed appropriately.

6. Conclusion

This research achieved three objectives as follows: The first objective is to cluster the mindsets and attitudes toward self-regulated learning styles of undergraduate students at the University of Phayao. The second objective is to construct a predictive model for recommending an appropriate student learning clusters. Lastly, the third objective is to evaluate the predictive model that has been constructed. Data collection is a compilation of 472 student satisfaction with questionnaires from three schools, and majoring in seven disciplines at the University of Phayao, Thailand. The data consisted of students from the School of Information and Communication Technology (ICT), the School of Management and Information Sciences (MIS), and the School of Law.

The research consisted of statistical tools and machine learning tools as follows: percentage, mean, average, standard deviation, k-means clustering, decision tree techniques, cross-validation methods, confusion matrix performance, accuracy measurement, precision measurement, and recall measurement. The results of the research found that the k-means model with the highest accuracy is the decision tree model that is classified into three clusters by dividing the model testing into the leave-one-out cross-validation method with a depth of seven levels of the decision tree model and has an accuracy of 98.73%. From the results and studies, it can be concluded that the developed model is effective and reasonable to be further developed as an application for further organizational development.

For future studies, the researchers found that the results of this research could be clearly further enhanced. For example, the researcher could use results to cluster learners' instruction based on their perceptions and attitudes towards managing online instruction based on Self-Regulated Learning theory. In addition, the researcher can use the results of the research into a computer program to facilitate the learners and teachers to use the results of the research in a useful way.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is part of a study on “The Study of Trends and Impacts of Disruptive Technology on the Education System at the School of Information and Communication Technology, the University of Phayao”. This research is supported by the University of Phayao and Rajabhat Mahasarakham University. The authors would like to thank Mrs. Wongpanya Nuankaew, the advisor, lecturers, students, technicians, and all respondents for their entire support.

References

[1] P. Olivos, A. Santos, S. Martín, M. Cañas, E. Gómez-Lázaro, Y. Maya, “The relationship between learning styles, motivation to transfer of learning in a vocational training programme” Suma Psicológica, **23**(1), 25–32, 2016. http://doi.org/10.1016/j.sumpsi.2016.02.001

[2] A. Viloria, J. R. López, K. Payares, C. Vargas-Mercado, S. E. Duran, H. Hernández-Palma, M. A. David, “Determinating Student Interactions in a Virtual Learning Environment Using Data Mining” Procedia Comput. Sci., **155**, 587–592, 2019. http://doi.org/10.1016/j.procs.2019.08.082

[3] A. I. Adekitan, O. Salau, “The impact of engineering students’ performance in the first three years on their graduation result using educational data mining” Heliyon, **5**(2), 2019. http://doi.org/10.1016/j.heliyon.2019.e01250

[4] P. Nuankaew, W. Nuankaew, T. Thamma, “The Recommended System for the Relationship between Educational Programs, Students’ Interests” in International Conference on Digital Arts, Media, Technology, 2-3. 2016

[5] P. Nuankaew, W. Nuankaew, K. Phanniphong, S. Imwut, S. Bussaman, “Students Model in Different Learning Styles of Academic Achievement at the University of Phayao, Thailand” Int. J. Emerg. Technol. Learn. IJET, **14**(12), 133–157, 2019.

[6] D. P. Cantwell, L. Baker, “Association between attention deficit-hyperactivity disorder, learning disorders” J. Learn. Disabil., **24**(2), 88–95, 1991.

[7] P. N. Pastor, C. A. “Reuben, Diagnosed Attention Deficit Hyperactivity Disorder and Learning Disability: United States, 2004-2006. Data from the National Health Interview Survey” Vital and Health Statistics **10**(237), 1-14, 2008

[8] M. Salehan, A. Negahban, “Social networking on smartphones: When mobile phones become addictive” Comput. Hum. Behav., **29**(6), 2632–2639, 2013. http://doi.org/10.1016/j.chb.2013.07.003

[9] R. Tao, X. Huang, J. Wang, H. Zhang, Y. Zhang, M. Li, “Proposed diagnostic criteria for internet addiction” Addiction, **105**(3), 556–564, 2010. http://doi.org/10.1111/j.1360-0443.2009.02828.x

[10] A. Błachnio, A. Przepiorka, “Personality, positive orientation in Internet, Facebook addiction. An empirical report from Poland” Comput. Hum. Behav., **59**, 230–236, 2016. http://doi.org/10.1016/j.chb.2016.02.018

[11] S. K. Sharma, A. Joshi, H. Sharma, “A multi-analytical approach to predict the Facebook usage in higher education” Comput. Hum. Behav., **55**, 340–353, 2016. http://doi.org/10.1016/j.chb.2015.09.020

[12] M. Ozkan, B. Solmaz, “Mobile Addiction of Generation Z, its Effects on their Social Lifes: (An Application among University Students in the 18-23 Age Group)” Procedia - Soc. Behav. Sci., 205, 92–98, Oct. 2015. http://doi.org/10.1016/j.sbspro.2015.09.027

[13] W. S. Nuankaew, P. Nuankaew, D. Teeraputon, K. Phanniphong, S. Bussaman, “Perception, Attitude Toward Self-Regulated Learning of Thailand’s Students in Educational Data Mining Perspective” Int. J. Emerg. Technol. Learn. IJET, **14**(09), 2019.

[14] P. J. Gaskill, A. Woolfolk Hoy, “Self-efficacy and self-regulated learning: The dynamic duo in school performance. In J. Aronson (Ed.)” Improving academic achievement: Impact of psychological factors on education, 185–208, 2002. https://doi.org/10.1016/B978-012064455-1/50012-9

[15] B. J. Zimmerman, “Models of self-regulated learning, academic achievement” Self-regulated learning, academic achievement, Springer, 1–25, 1989.

[16] P. R. Pintrich, “Understanding self-regulated learning” New Dir. Teach. Learn., **63**, 3–12, 1995.

[17] H. C. Schmitz, A. Mikroyannidis, T. Connolly, L. C. E. Law, H. Vieritz, A. Nussbaumer, M. Berthold, C. Ullrich, A. Dhir, “Self–regulated learning in formal education: perceptions, challenges, opportunities” Int. J. Technol. Enhanc. Learn., **6**(2), 145–163, 2014.

[18] P. Nuankaew, W. Nuankaew, S. Bussaman, P. Jedeejit, “Education mining in the relationship between general knowledge, deep knowledge for lifelong learning” in 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, Information Technology (ECTI-CON), 694–697, 2017. http://doi.org/10.1109/ECTICon.2017.8096333

[19] P. Nuankaew, “Dropout Situation of Business Computer Students, University of Phayao” Int. J. Emerg. Technol. Learn. IJET, **14**(19), 2019.

[20] K. Pupara, W. Nuankaew, P. Nuankaew, “An institution recommender system based on student context, educational institution in a mobile environment” in 2016 International Computer Science, Engineering

Conference (ICSEC), 1–6, 2016. http://doi.org/10.1109/ICSEC.2016.7859877
[21] S. Bussaman, W. Nuankaew, P. Nuankaew, N. Rachata, K. Phanniphong, P. Jedeejit, "Prediction models of learning strategies, learning achievement for lifelong learning" in 2017 IEEE 6th International Conference on Teaching, Assessment, Learning for Engineering (TALE), 192–197, 2017. http://doi.org/10.1109/TALE.2017.8252332
[22] P. Nuankaew, W. Nuankaew, P. Temdee, "Institution recommendation using relationship optimisation between program, student context" Int. J. High. Educ. Sustain., **2**(4), 279–302, 2019.
[23] P. Nuankaew, W. Nuankaew, T. Thamma, "The Recommended System for the Relationship between Educational Programs, Stu-dents' Interests" in International Conference on Digital Arts, Media, Technology, 2-3, 2016.
[24] R. Wirth, J. Hipp, "CRISP-DM: Towards a standard process model for data mining" in Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery, Data Mining, 29–39, 2000.
[25] J. Venter, A. de Waal, C. Willers, "Specializing CRISP-DM for Evidence Mining" in Advances in Digital Forensics III, New York, USA, 303–315, 2007.
[26] B. L. McCombs, "Self-regulated learning, academic achievement: A phenomenological view" Self-regulated learning, academic achievement, Springer, 51–82, 1989.
[27] B. J. Zimmerman, "Self-regulated learning, academic achievement: An overview" Educ. Psychol., **25**(1), 3–17, 1990.
[28] B. J. Zimmerman, D. H. Schunk, "Self-regulated learning, academic achievement: Theoretical perspectives" Routledge, 2001.
[29] B. J. Zimmerman, "Becoming a self-regulated learner: An overview" Theory Pract., **41**(2), 64–70, 2002.
[30] E. Shahzadi, Z. Ahmad, "A study on academic performance of university students" Recent Adv. Stat., 255, 2011.
[31] P. Nuankaew, P. Temdee, "Matching of compatible different attributes for compatibility of members, groups" Int. J. Mob. Learn. Organ., **13**(1), 4–29, 2019.
[32] P. Nuankaew, P. Temdee, "Determining of compatible different attributes for online mentoring model" in 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory, Aerospace & Electronic Systems (VITAE), 1–5, 2014.
[33] P. Nuankaew, W. Nuankaew, K. Phanniphong, R. Fooprateepsiri, S. Bussaman, "Analysis Dropout Situation of Business Computer Students at University of Phayao" in The Impact of the 4th Industrial Revolution on Engineering Education, Cham, 2020, 419–432. http://10.1007/978-3-030-40274-7_42
[34] A. M. Shahiri, W. Husain, N. A. Rashid, "A Review on Predicting Student's Performance Using Data Mining Techniques" Procedia Comput. Sci., **72**, 414–422, 2015. http://doi.org/10.1016/j.procs.2015.12.157
[35] W. Nuankaew, P. Nuankaew, "The Study of the Factors, Development of Educational Model: The Relationship between the Learner Context, the Curriculum Context in Higher Education" Int. J. Emerg. Technol. Learn. IJET, **14**(21), 2019.
[36] N. B. Ahmad, U. F. Alias, N. Mohamad, N. Yusof, "Principal Component Analysis, Self-Organizing Map Clustering for Student Browsing Behaviour Analysis" Procedia Comput. Sci., **163**, 550–559, 2019. http://doi.org/10.1016/j.procs.2019.12.137
[37] A. B. Firdausiah Mansur, N. Yusof, A. H. Basori, "Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic" Procedia Comput. Sci., **163**, 125–133, 2019. http://doi.org/10.1016/j.procs.2019.12.094
[38] E. Aronson, "Chapter 10 - Building Empathy, Compassion, Achievement in the Jigsaw Classroom" Improving Academic Achievement, J. Aronson, Ed. San Diego: Academic Press, 209–225, 2002.
[39] K. Pupara, W. Nuankaew, P. Nuankaew, "An institution recommender system based on student context, educational institution in a mobile environment" in 2016 International Computer Science, Engineering Conference (ICSEC), 1–6, 2016.
[40] J. A. Meirink, P. C. Meijer, N. Verloop, T. C. Bergen, "Understanding teacher learning in secondary education: The relations of teacher activities to changed beliefs about teaching, learning" Teach. Teach. Educ., **25**(1), 89–100, 2009.
[41] G. Hampden-Thompson, J. Bennett, "Science teaching, learning activities, students' engagement in science" Int. J. Sci. Educ., **35**(8, 1325–1343, 2013.
[42] W. Nuankaew, P. Nuankaew, S. Bussaman, P. Tanasirathum, "Hidden academic relationship between academic achievement, higher education institutions" in 2017 International Conference on Digital Arts, Media, Technology (ICDAMT), 308–313, 2017. http://doi.org/10.1109/ICDAMT.2017.7904982
[43] C. Romero, S. Ventura, "Educational data mining: a review of the state of the art" IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., **40**(6), 601–618, 2010.
[44] S. K. Mohamad, Z. Tasir, "Educational data mining: A review" Procedia-Soc. Behav. Sci., **97**(2013), 320–324, 2013.
[45] S. Slater, S. Joksimović, V. Kovanovic, R. S. Baker, D. Gasevic, "Tools for educational data mining: A review" J. Educ. Behav. Stat., 42(1), 85–106, 2017.
[46] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, Reinartz, T., C. Shearer, R. Wirth, "CRISP-DM 1.0 Step-by-step data mining guide" SPSS inc, **9**, 13, 2000.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 684-691 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Dismantle Shilling Attacks in Recommendations Systems

Ossama Embarak*

Computer Information Science, Higher Colleges of Technology, Abu Dhabi, 51133 UAE

ARTICLE INFO

Article history:
Received: 18 December, 2020
Accepted: 24 January, 2021
Online: 05 February, 2021

Keywords:
Malicious users
Push & nuke attacks
Collaborative systems
Shilling attacks

ABSTRACT

Collaborative filtering of recommended systems (CFRSs) suffers from overrun false rating injections that diverge the system functions for creating accurate recommendations. In this paper, we propose a three-stage unsupervised approach. Starts by defining the mechanism(s) that makes recommendation vulnerable to attack. Second, find the maximum-paths or the associated related items valued by the user. We then rule out the two attacks; we will need to pull two different measures. (a) We will pull user ratings across all reviews and measure their centre variance. (b) We will then pull each individual user rating and measure them according to the original rating. Detected attack profiles are considered untrusted and, over time, if the same user is detected as untrusted, the profile is classified as completely untrusted and eliminated from being involved in the generation of recommendations. Thus, protect CFRS from creating tweaked recommendations. The experimental results of applying the algorithm to the Extensive MovieLens dataset explicitly and accurately filter users considering that a user could seem normal and slightly diverge towards attack behaviours. However, the algorithm used assumes that the framework has already begun and manages user accounts to manage the cold start scenario. The proposed method would abstractly protect users, irrespective of their identity, which is a positive side of the proposed approach, but if the same user reenters the system as a fresh one, the system will reapply algorithm processing for that user as a normal one.

1. Introduction

This paper is an extension of work originally presented in 2019 Sixth HCT Information Technology Trends (ITT) [1]. Personalization Collaborative Filter Recommending Systems (CFRSs) is becoming increasingly popular in well-known e-commerce services such as Amazon, eBay, Alibaba, etc. [2]. Most people are rating products or services without even realizing it is something they can do [3]. However, CFRSs are highly vulnerable to "profile injection" attacks; often referred to as "shilling" attacks [4]. It is common for attackers to pollute the recommended systems with malicious ratings. They either demote a target item with the lowest rating; called a nuke attack or promote a target item with the highest rating; called a push attack to reach their target or minimize the recommendation's accuracy. It is therefore necessary to develop an effective detection system for detecting and removing attackers before the recommendation [5–7].

Detection methods based on the attacks gained much coverage. Since the similarities between attackers are higher than the actual users, some have been presented based on the similarity between users. Traditional similarity metrics, including the Pearson Correlation Coefficient (PCC),[8,9]. However, the detection efficiency of these methods relies largely on the estimation of similarity. How to reduce the time consumption of the measurement of similarity is also a difficult problem, particularly when dealing with large datasets[10]. In addition, some attackers imitate the rating data of some legitimate users in order to increase their reliability. Only the use of similarities is difficult to discriminate incomplete [11]. In order to overcome these problems, a more efficient form of detection should be considered in the following aspects:

A. Applied algorithm complexity should be acceptable.

B. The proposed method should be able to defeat various kinds of "shilling" attacks.

This paper proposes an unsupervised detection method for detecting such attacks, consisting of three phases. The purpose of the proposed method is to filter out more legitimate users, and at the same time hold, all attackers step by step. Firstly, the recommendation system admin needs to understand the attacker's strategy based on the recommendations working mechanisms. Secondly, we constructed users' maximal paths that represent his/her loop less-visited nodes or rated items in an association, then calculate the maximal path's mean and standard deviation considering all ratings done by users visited that particular maximal path. Thirdly, we detected untrusted users by detecting the user's zone (normal, freeze or melt zone area) which reflect the gap between the user's ratings in the

*Corresponding Author: Ossama Embarak, oembarak@hct.ac.ae

https://dx.doi.org/10.25046/aj060174

maximal path (items) and the genuine profiles' ratings of the same path(items) considering the maximum expected calculated deviation. The proposed algorithm prevents untrusted users' maximal sessions should not be added to the system profiles used for generating recommendations, and the user is labelled as untrusted profiles in which the system should not consider such profiles towards generating recommendations.

Recommendation systems are massively used to support online systems provides services through the internet. As an alternative to passive search engines that simply accept as "true" the displayed results, these collaborative filtering systems attempt to take the perspective of the user's needs and what they are searching for and identify similar results based on that [12]. Decisions that depend on a user's previous demands can be made with a recommendation system; a recommendation system is a set of techniques used to process and generate recommendations. In order to inject the ratings of biased users into recommender systems, users attempt to influence the system by injecting biased ratings for a specific product they are interested or associated with. Through a process known as a rating bot attack, which can plough the ratings up or down, change the ratings of services or items to the highest or the lowest available levels, or launch attacks against them [13]. In order to avoid shilling attacks, the company has attempted to implement various suggestions, such as asking users to rate items by using certain code, or by only allowing ratings to occur after a certain amount of time, or by simply making profiles harder to create or by increasing the price of creating a profile? Because of the ways in which these methods can eliminate attacks, the participants in rating and evaluating services became much smaller than normal [14]. E-commerce web-systems suffer from attackers who push specific items to the recommendations list and promote the target item via manipulating recommendation systems [15]. The attackers create numerous profiles by injecting unusual identities and putting very high-ratings for their target items. An attacker's profile will camouflage itself by keeping track ratings for other non-target items, and the system will use that information to predict the attacker's targets. The manipulations of recommendations systems of injecting misleading and false data make a recommendation system's operation is negatively affected [16]. Recommendations systems powered by user-generated reviews will only be representative of individuals with very few reviews. Initial profiles will greatly affect future processing and output. Moreover, this system does not generate suitable recommendations, and occasionally, it breaks down the whole system; It causes the system to lose its credibility and respect by people in the community. Therefore, several research groups' main goal is to develop a system that can filter out fake profiles that could make those systems able to provide accurate recommendations [17].

This paper looks into a different type of attacks for recommendation systems and propose a solution to detect and isolate attackers. It contains the following sections: Section 1 introduction to recommendation systems and its attacks. Section 2, the background of recommendation systems different forms of attacks. Section 3, we discuss the previous researches and approaches for handling different attacks. Section 4 identify the study problem. In section 5, we explore the proposed solution. Section 6 demonstrate experimental results. Finally, The paper and the future work will be concluded.

2. Background

Numerous different sorts of recommenders are used on different sites. Studies in the past have shown dramatic growth in the methods used to improve recommendations. Recommendation systems can be broadly classified into content-based and collaboration-based [18]. Content-based filtering suggests products based on users browsing history. There are downsides to this type, such as over-specialization, which recommends only similar products to what was consumed before. Collaborative filtering takes into account past behaviour, considering that users will behave similarly. Shilling attacks manipulate recommendation systems by inserting malicious user profiles into the filter data [19]. The target item is either promoted or demoted. Attacks can be categorized as push or nuke strikes, a product may therefore be promoted or dismissed for advantages over competitors [20]. Over time, numerous attacks models have been developed. There are many detection techniques and algorithms to counter such attacks. All of the attack models are designed to generate vicious users. The discrepancies are recognized in how the strike profiles are established. Malicious users' domain items can be considered as $S = \{ I_s, I_p, I_n, I_{\emptyset}\}$, where I_s is the set of items that matched with the genuine users. *while* I_p is the set of items the malicious user promote, and I_n is the set of items he/she demote, and $I_{\emptyset}$ items that are not ranked yet by the trouble user.

Over the years, several attacks were developed. All such attacks can be classified either as a high-level attack or as a poor-knowledge attack based on their motivation and knowledge [21, 22]. Attacks with little knowledge are more practical and are more likely to impact the real world, but such attacks' effectiveness is also low. On the other hand, high-level attacks can have a huge impact on Recommender Systems' performance, but they are more difficult to pull out. *RandomBot attack* is the plainest form of a shilling attack. Items rated by the attackers selected in random, except for the target item. The rankings for such items are based on the overall system norm [22].

A maximum or minimum rating will be given on the basis that it is a push or a nuke attack. The purpose of a random attack is usually more effective in disrupting the Recommender Systems' performance than in promoting the target item. *Average Attack* is similar to Random Attack when selecting items. The randomized distribution of individual items is based on their distribution of ratings. The average rating of each filling item is assigned. This attack can only be carried out if the attacker knows the system's dataset. While random attack and average attack differ only from filler ratings, the average attack impact is much higher. *Bandwagon Attack* is the kind of attack where attacker profiles are filled with popular high rating items. Naturally, attack profiles are closer to many people. The highest value is given to the target item. Depending on the rating scheme used to provide the filler items, this attack can further be divided into random and bandwagon averages. Bandwagon is also classified as a low-knowledge attack since the attacker only needs public data. *Bandwagon Reverse Attack* is the exact reversal of an attack on a bandwagon. This attack is used to demote the target product by rating the items with high negative ratings low and giving the target item the least rating. *Segmented Attack* targets a particular group of users who can buy the target item in the e-commerce system. Segment attacks typically occur in collaborative item-based filtering [23, 24].

The ratings and items are based on knowledge of the segment from the attacker. The important benefit of this method is its ability to reach potential clients over other approaches. *Probe attack* occurs when an attacker provides genuine ratings for items based on their knowledge of the predicted rating scores. The attacker uses this detail to rate items, enabling it to appear similar to other ratings. The attacker composes the acceptable items list based on the rated items. This collaborative scheme allows attack profiles to remain close to their neighbours. It also facilitates an attacker in gaining more information about the system. *Love/Hate Attack*, where the attacker randomly chooses filler items and gives them the highest rating and lowest rating on the target item. It can also be known as a push attack by changing the highest ratings [25]. *Noise Injection attack* generates random numbers multiplied by a constant to each rating for a subset of rated profiles. The more obfuscation, the more constant. It can be effectively applied to obfuscate its signature to all standard methods. A slightly decreased ability to withstand an attack is observed in response to noise injection. *User-Shifting attack* where a subset of users revises ratings [26]. The attack subsets were selected randomly to reduce their similarities. Ratings are scaled differently for different subsets of the ratings [27]. *Target Shifting attacks* shift the targeted item rating to a lower level in push attacks than maximum. For nuke attacks, the minimum rating is shifted up by one. *The mixed attack* is made using the same proportions of random, average, bandwagon and segmented attacks. The detection technique should be able to detect all the standard attacks successfully. The various methods for attacking the same target item are used. It helps to avoid various methods of detection [28].

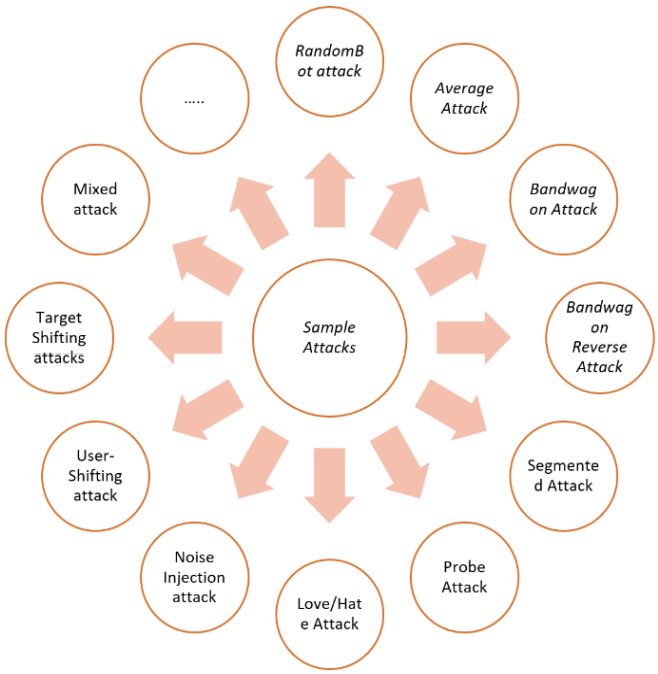


Figure 1: Sample Attacks Types

3. Related Works

A well-known class of attacks is called the Shilling attacks, which attempts to inject some profiles to influence the targeted system's performance [13]. This attack is divided into two attacks: push attack- which uses a malicious profile to rate a specific item highly (greater than most items) and nuke attack- which uses malicious profiles to lower rating for a specific item (s). Playbooks attacks reflect a series of actions undertaken to maximize a certain item's importance and increase its profits [29]. Some systems produce low-quality recommendations, so that once the users discover such attacks, the system loses its users' respect and loyalty [30]. Unfair rating attack occurs when a system uses trusted agents' ratings and compares them against each other for a service entity. This attack leads to predictions that are not actually correct [31]. Re-entry attacks occur when an agent is transferred between accounts using different names to avoid low ratings [29]. Sybil attacks reflect a fake identity to give different ratings. So far, a robust recommendation system is essential for any application of a recommendation system. Robustness measures a model's ability to produce good predictions with noisy data [32]. Whenever a user rates something, the system updates its databases, which means we have no way of knowing if the ratings are real or not. Ratings can sometimes be tedious, which could be caused by users' carelessness, and malicious users may attack. In robustness, there are two different aspects, *the first* of which is the accuracy of recommendation, which are the products that were recommended by the systems after the attack took place. *The second thing a system must do is be very stable*, which means won't recommend slightly different products if it detects an anomaly or attack.

Before storing user ratings into the system, the attack type must be detected to maintain its good reputation before affecting its performance. Injecting profiles another example of forged data being injected by malicious users into recommendations systems to manage recommendations [33]. Collaborative filters are susceptible to profile attacks by user-based and item-based recommendations. In this sort of attack, the attackers try to evaluate target and non-target items to make ratings appear normal by forging rating profiles and injecting them in rating systems [34]. Studies show that group attack profiles manipulate the target object ranking recommendation. Many attackers work together in a given time frame to attack certain target items to quickly put some items into a preferred list. The next section will provide a breakdown of the experiment and its results.

4. Identifying the Problem

Recommendation systems may provide health recommendations that reflect the preferences of users. However, injecting false ratings or trying to push or nuke ratings of the offered items differs in the system's performance. Many proposed solutions are almost implemented after attacking or generating recommendations that burden many computations and increase system time complexity. Furthermore, depending on the specific fixed average rating, the dynamic change in items' ratings over time is not reflected. Therefore, we propose a new approach to filter users into normal, fully trusted and untrusted users based on their maximal traversing path or items movable threshold.

5. The Proposed Solution

Details of the proposed method are implemented in three stages in this section. In the first phase, we need to understand the attacker's strategy based on recommendations generation mechanisms. In the second phase, we constructed users' maximal paths and calculated the maximal path's mean and variance, considering all users' ratings visited that particular

maximal path. In the third stage, we detected untrusted users by detecting the user's zone (normal, freeze or melt zone area), which reflect the gap between the user's ratings in the maximal path (items) and the genuine profiles ratings of the same path(items) considering the maximum expected calculated deviation. All detected untrusted users' maximal sessions should not be added to the system profiles used for generating recommendations, and the user is labelled as untrusted profiles in which the system should eliminate his future participations.

5.1. Users Profiles

Recommendation systems collect user ratings – items used to form an object rating matrix that includes three elements: users, items and ratings.

$$U = \{ u_1, u_2, \ldots., u_{m-1}, u_m \}$$

where *U* refers to set of *m* users.

$$I = \{ I_1, I_2, \ldots., I_{n-1}, I_n \}$$

where *I* refers to a set of n items.

All users can rate few or all items; a user-item rating matrix generated from all the ratings is used to find attackers and make many recommendations.

$$\begin{bmatrix} I_{u_1}^1 & \cdots & I_{u_1}^n \\ \vdots & \ddots & \vdots \\ I_{u_m}^1 & \cdots & I_{u_m}^n \end{bmatrix}$$

Figure 2: Items-users ratings matrix

where $I_{u_1}^1$ reflect the rating of the first item by the first user, while $I_{u_m}^n$ is the rating of the item number n by the user number m.

A regular user weight items in the overall round the norm. Similarly, both the push and nuke attackers aim to rate items extremely far from the item threshold. We call it a *freezing* zone, as the rates are very far from the item average rates measured, and the nuke attackers aim to rate items that are extremely low from the item thresholds, we call it a *melting* zone, as the rates are very far back from the item thresholds. We measure each user's ratings and can detect any change in weight in either "Freeze" or "Melting" zones by measuring the weighted mean of all ratings.

5.2. Attackers Background

The attackers usually deliberately attempt to change the efficiency of the recommendation system. They know how the system operates and use their expertise and understanding to affect the system's efficiency.

There are three types of systems in terms of user awareness of the Recommendation Mechanisms.

1. Systems that depend on the popularity threshold (like Insider) to produce recommendations are simply showing popularity to users. As a result, attackers know exactly what will be recommended, attempt to modify popularity by revisiting the same path(route) or increase or decrease other paths' ratings.
2. Systems that depend on each item's popularity to produce a recommendation, the popularity displayed to users with a list of recommendations can also use the association rule to show that items x are frequently selected with item y, such as YouTube videos. As a result, attackers have full knowledge of how ratings are allocated in the recommended device profiles. Attackers attempt to manipulate the specific item(s) by raising or decreasing the item(s) ratings or increasing or decreasing the related items' ratings.
3. Systems that show recommended items without a strong indication of popularity or affiliation between items, such as Amazon. As a result, attackers have no understanding of the system mechanisms.

Recommendation system developers must consider the attacker tactics and program the system to identify and remove any untrusted users from being part of the profiles used to produce recommendations.

5.3. Constructing Maximal User Path

As shown in the figure below, each user targets items and gives their ratings of those items. The ratings influence those items to be the top choices. Others will probably allocate very little to nuke individual objects at the lowest level far below the average weighted μ of the item. An average is normally relative to all items rated.

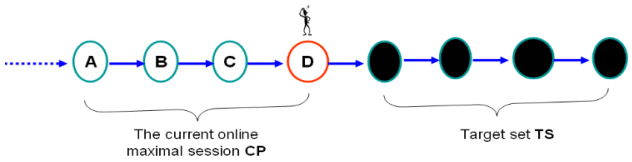


Figure 3: User-Targeted Items

We consider a maximal route occurred during a visit (session), as shown in the figure below, where there is no loop happened in a maximal path. The user may have several visits during which the rating occurred, as shown in the below figure, the user might be a genuine or an attacker. Whether a genuine person or an attacker, a user tries to check certain nodes, so we need to assess his traversing rates.

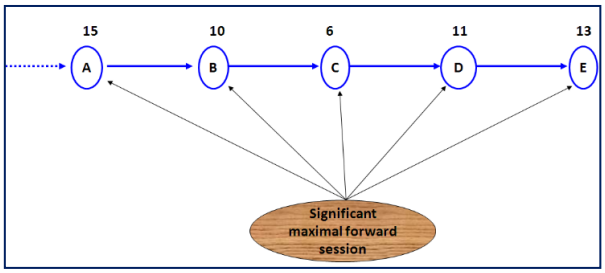


Figure 4: User Maximal Path

We begin with each user session, and gradually explore the data, moving from s_i. And as for all of the rates R_i, they do not exceed the limits. The user session is entered into the system of that session.

$$s_i = \{ R_1, R_2, \ldots., R_{n-1}, R_n \}$$

By measuring the session maximal path average rate, we can find the rate deviation of the user's rating of the node's visited and rate by looking at the other user's ratings of the user rated objects.

$$\mu_{s_i} = \frac{\sum_{i=1}^{n} R_i}{n} \quad (1)$$

$$\sigma(s_i) = \sqrt{\frac{\sum_{j=1}^{l}(|\mu_{s_i} - R_{u_j}^{i}|)^2}{n}} \quad (2)$$

where μ_{s_i} represents the mean weight of that path by all users and $\sigma(s_i)$ represents the calculated standard deviation of all rates happened on that maximal path.

5.4. Detect Untrusted Users

The two types of attacks are very severe from average users, but this does not mean that a normal user does not score an object at a low or high rate. However, it is uncommon for the same user to rate multiple items with very low or very high ratings relative to other users that rate the same item on similar maximum paths.

We then found the following rule to use every rate state (melting, freezing, or normal).

where F(s_i) is the evaluation function of each session as a collection of rankings made through a visit, $\Im_+$ represents the case when the user aims to push ratings far from the normal ratings (the freeze status), and $\Im$ reflects ratings happened within the threshold boundaries, while $\Im_-$ reflects the case when the user aims to downgrade target items (the melting status) as shown in the figure below.

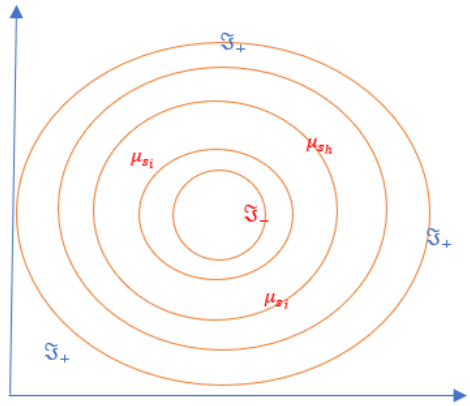


Figure 5: User Rates Zone

where,
Freeze case ($\Im_+$) = $\{R_i \mid R_i > \Im_+\}$
Meting case ($\Im_-$) = $\{R_i \mid R_i < \Im_-\}$
Normal case ($\Im$) = $\{R_i \mid R_i \cong \Im \; or \; R_i \cong \mu_{s_i}\}$

Therefore, $\Im_+$ & $\Im_-$ users should be eliminated from being part of recommendation generation processing.

Table 1: Attack Status

Attack model	Zone	Status
push attack	Freeze zone	$\Im_+$
No attack	Normal	$\Im$
Nuke attack	Melting zone	$\Im_-$

Table 1 shows that $\Im_+$ reflects the highest rates could be reached by attackers in the rating; for specific items or maximal path, far beyond the standard deviation. And $\Im_-$ reflect the customer's minimum prices relative to the measured average rate and the selected products' standard deviations. While $\Im$ reflects the normal rates of the trusted user obtained by the recommendation systems.

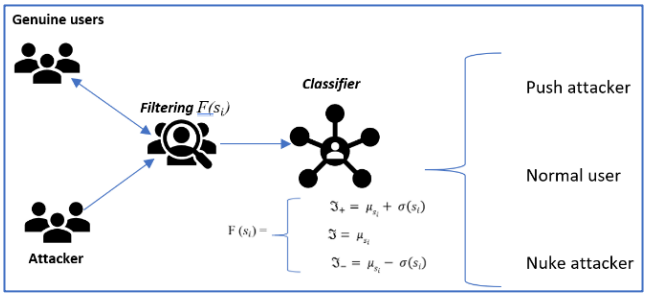


Figure 6: Suggested Model Stages

5.4.1. Algorithm for Detecting an Attack

The following algorithm demonstrates the followed steps to detect a user rating zone.

Input: users ratings or maximal path in a session.
Output: filter users into trusted, untrusted (melting, freeze).

1. Read a user u_i session s_i
2. $\forall_i \in s_i$, calculate μ_{s_i}, and $\sigma(s_i)$
3. Read ***UT***
4. ***While not UT***
   For R *in* s_i:
   a. *if* $R_i > \Im_+$
      i. *Detected Freeze case*
      ii. Refuse the whole session
      iii. Untrusted user is highlighted
      iv. ***Update UT*** $\leftarrow u_i$
   b. Else if $R_i < \Im_-$
      i. *Detected Melting case*
      ii. Refuse the whole session
      iii. Untrusted user is highlighted
      iv. ***Update UT*** $\leftarrow u_i$
   c. Else
      i. Detected Normal case
      ii. Accept the whole session ratings
      iii. Trusted user is highlighted
      iv. ***Update U*** $\leftarrow u_i$
5. Return zone

where R reflects items ratings in a session s, and ***UT*** reflects the untrusted users discovered and recorded on the system. The system would automatically maintain its functionality by upgrading any identified untrusted user to untrusted profiles that make sense of removing any ratings received from these users in the future. In contrast, the normal user ratings should be updated to the normal profiles used to produce recommendations.

5.5. The Contribution of the Manuscript

This paper's main contribution is implementing a new algorithm to filter user rating injections and item rating injections to prevent false similarities. A matrix of ratings is used in the processing to identify the trust of the user. Therefore, if the user ratings exceed a particular computed ceiling, the system routinely denies the user scores and moves

them in an untrustworthy category, and the user ratings are not considered in the recommendation creation.

6. Experimental Results

We use the available online Movie Lens 1M data set to apply and test the proposed algorithm, comprising 1,000,000,000 unidentified reviews of roughly 3,900 films received by 6,040 users of Movie Lens. All scores are 1 to 5 integers, with 1 being the lowest, and 5 being the highest. There are 18 separate domains in the data set and all users with at least 20 films rated. The proposed algorithm (see section 3.4.1) was able to detect cases near the freezing, and those were very close to the melting zone. Total cases were 11364): Normal ratings (2677), Malting near (3943), Freeze near (4611), found attacks (133) as demonstrated in the table below.

Table 2: Users- Malicious Detection

Domain name	Normal	Malting touch	Freeze touch	Detected Attacks
Action	240	158	154	3
Adventure	123	216	349	2
Animation	157	226	216	0
Children	125	152	173	9
Comedy	151	267	248	50
Crime	104	348	247	0
Documentary	177	326	536	4
Drama	173	158	195	15
Fantasy	123	234	165	0
Film-Noir	140	217	300	3
Horror	157	154	172	8
Musical	125	150	150	0
Mystery	141	148	141	11
Romance	177	178	272	0
Sci-Fi	173	353	467	12
Thriller	176	277	286	0
War	105	211	380	7
Western	110	170	160	9

In each movie domain, the following figure shows user patterns; it is clear that several freezes touch cases when user ratings exceed the normal ratings; the same is obvious in the melting zone. The experimental results showed that the algorithm detected 50 attacks from the comedy group, while many cases were under border conditions. The clearly reported attacks in the trained data showed that the attached percentage are as follows, in Action (2%) , Adventure(2%),Animation (0%), Children (7%), Comedy(38%), Crime(0%), Documentary(3%), Drama(11%), Fantasy(0%), Film Noir(2%), Horror(6%), Musical(0%), Mystery(8%), Romance(0%), Sci-Fi(9%), Thriller(0%), War(5%), Western(7%).

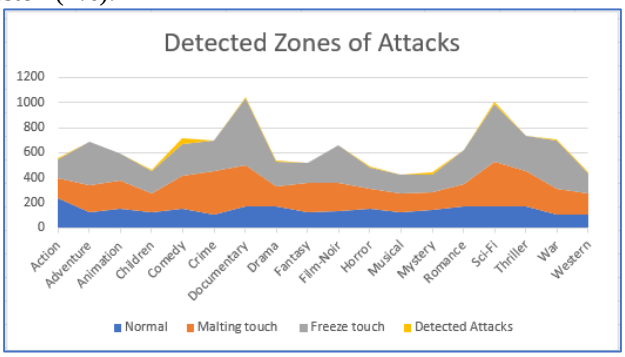


Figure 7: Attacks Detections For 300 Users

We expanded the data set and considered 450 users; the algorithm could detect more attacks, as seen in the figure below. This represents the incremental number of maximum paths that are being tested and the rise in the number of attacks observed by improvements in the slander of the deviation and the zones' size.

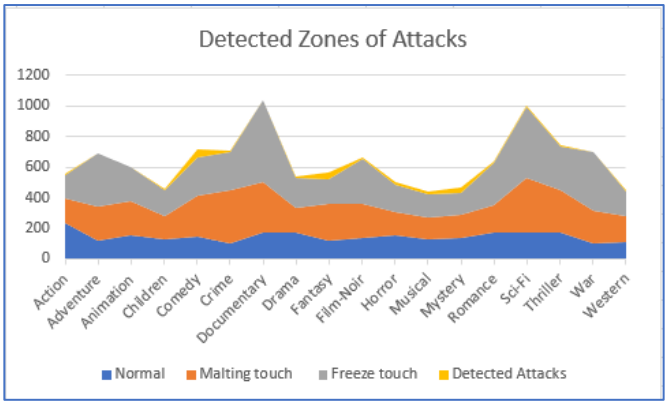


Figure 8: Attack Detection For 450 Users

The following table demonstrates a comparison between the proposed algorithm and another two algorithms were used to detect malicious attacks in recommendation systems.

Table 3: Algorithms Comparison

Points	Proposed algorithm	Rating behaviour	Profile similarity
Processing	Find the maximal path or sequence of items ratings	Pinpoint lower rating values considering that it has a minimal contribution to the systems	Discover profiles similarity
System behaviour	Classify users in three zones Normal, Melting and Freeze before updating their datasets	Detect all lower rating profiles.	Detect all identical profiles in ratings
Actions	Use trustworthy users to make recommendations	Ignore low rate profiles and only use moderately to highly qualified profiles	Ignore all identical profiles
Privacy concerns	Users' demographic data are not collected, only their maximum routes or sequence of items are collected. No privacy concerns.	The issue of privacy is valid.	Privacy is a valid issue.
Effective against	Various attack forms.	Random and Reverse	Average, Segment,

		Bandwagon attacks	and Bandwagon attacks
System reacts	The system will continue to serve untrusted users but will ignore their preferences for making recommendations.	The system eliminates profiles of an attacker	The system removes the profiles of attackers

Different attacks can be conducted using the proposed algorithm by finding the maximum path and measuring how far it is from the abstract user path. This helps classify users into different three zones; Normal, Melting (extreme below the threshold), and Freeze (extreme high from the threshold). The system then can ignore both Melting& Freeze cases and maintain only the normal cases for recommendation processing.

7. Conclusion and Future Works

Shilling attacks are a significant challenge to collaborative filtering recommendation systems. These attack profiles are likely to have similar rating information to many legitimate profiles to make them difficult to identify. This paper proposed an unsupervised detection method for detecting attacks (or irregular ratings) consisting of three phases. In the first phase, we need to understand the attacker's strategy based on the recommendation generation mechanisms. In the second phase, we built users' maximal paths and calculated the maximal path's mean and standard deviation considering all users' ratings visited that particular maximal path. In the third stage, we detected untrusted users by detecting the user's zone (normal, freeze or melt zone area) which reflect the gap between the user's ratings in the maximal path (items) and the genuine users' ratings of the same path(items) considering the maximum expected calculated deviation. All detected untrusted users' maximal sessions should not be added to the system profiles used for generating recommendations, and the user is labelled as untrusted profiles in which the system should eliminate any future participation. Experimental results showed that with a greater number of maximum routes used by the recommendation systems, the model could detect more attacks and thus discover untrusted users and prevent them from affecting their performance. The proposed method's main limitations include the following two aspects: (a) How the system could work in the launching phase when there are no ratings in the system collected from users. However, in the beginning, various methods can be used to overcome such systems' cold start, such as developing a suggestion considering the like-minded users when navigating the system. (b) The proposed approach abstractly handles users regardless of their identities, representing a good side of the proposed approach, but if the same user joins the system as a new user, the system will perform the calculations again for that user. One of our future objectives is to apply the suggested algorithm on a larger data set and apply it to live stream data in collaborative systems.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] O. Embarak, "Demolish falsy ratings in recommendation systems," in 2019 Sixth HCT Information Technology Trends (ITT), IEEE: 292–295, 2019.

[2] N. Sivaramakrishnan, V. Subramaniyaswamy, L. Ravi, V. Vijayakumar, X.-Z. Gao, S.L.R. J I.J. of B.-I.C. Sri, "An effective user clustering-based collaborative filtering recommender system with grey wolf optimisation," International Journal of Bio-Inspired Computation, **16**(1), 44–55, 2020. https://doi.org/10.1504/IJBIC.2020.108999

[3] Y. Pan, D. Wu, D.L. J D.S.S. Olson, "Online to offline (O2O) service recommendation method based on multi-dimensional similarity measurement," Decision Support Systems, **103**, 1–8, 2017. ttps://doi.org/10.1016/j.dss.2017.08.003

[4] Z. Yang, L. Xu, Z. Cai, Z. %J K.-B.S. Xu, "Re-scale AdaBoost for attack detection in collaborative filtering recommender systems," Knowledge-Based Systems **100**, 74–88, 2016.

[5] W. Zhou, J. Wen, Q. Xiong, M. Gao, J. J N. Zeng, "SVM-TIA a shilling attack detection method based on SVM and target item analysis in recommender systems," Neurocomputing **210**, 197–205, 2016.

[6] M. Si, Q. %J A.I.R. Li, "Shilling attacks against collaborative recommender systems: a review," Artificial Intelligence Review **53**(1), 291–319, 2020.

[7] S. Alonso, J. Bobadilla, F. Ortega, R. J I.A. Moya, "Robust model-based reliability approach to tackle shilling attacks in collaborative filtering recommender systems," EEE Access, **7**, 41782–41798, 2019. ttps://doi.org/10.1016/j.ds.2017.08.003

[8] Z. Yang, Z. Cai, X. %J K.-B.S. Guan, "Estimating user behavior toward detecting anomalous ratings in rating systems," Knowledge-Based Systems **111**, 144–158, 2016.

[9] K.G. Saranya, G.S. Sadasivam, M. %J I. journal of science Chandralekha, Technology, "Performance comparison of different similarity measures for collaborative filtering technique," Indian journal of science and Technology, **9**(29), 1–8, 2016.

[10] B. Li, Y. Wang, A. Singh, Y. Vorobeychik, "Data poisoning attacks on factorization-based collaborative filtering," in Advances in neural information processing systems, 1885–1893, 2016.

[11] M. Fang, G. Yang, N.Z. Gong, J. Liu, "Poisoning attacks to graph-based recommender systems," in Proceedings of the 34th Annual Computer Security Applications Conference, 381–392, 2018.

[12] M.N. TEKLEAB, RECOMMENDATION SYTEM ANALYSIS AND EVALUATION, NEAR EAST UNIVERSITY, 2019.

[13] A.M. Turk, A. %J E.S. with A. Bilge, "Robustness analysis of multi-criteria collaborative filtering algorithms against shilling attacks," **115**, 386–402, 2019.

[14] N. Nikzad–Khasmakhi, M.A. Balafar, M.R. %J E.A. of A.I. Feizi–Derakhshi, "The state-of-the-art in expert recommendation systems," **82**, 126–147, 2019.

[15] E. Çano, M. %J I.D.A. Morisio, "Hybrid recommender systems: A systematic literature review," **21**(6), 1487–1524, 2017.

[16] C. Wang, Y. Zheng, J. Jiang, K. %J E. Ren, "Toward privacy-preserving personalized recommendation services," **4**(1), 21–28, 2018.

[17] M. Sappelli, S. Verberne, W. %J J. of the A. for I.S. Kraaij, Technology, "Evaluation of context-aware recommendation systems for information re-finding," **68**(4), 895–910, 2017.

[18] J. Su, Content based recommendation system, 2017.

[19] L. Yang, W. Huang, X. %J I.E.T.I.S. Niu, "Defending shilling attacks in recommender systems using soft co-clustering," **11**(6), 319–325, 2017.

[20] V.W. Anelli, Y. Deldjoo, T. Di Noia, E. Di Sciascio, F.A. Merra, "Sasha: Semantic-aware shilling attacks on recommender systems exploiting knowledge graphs," in European Semantic Web Conference, Springer: 307–323, 2020.

[21] A.P. Sundar, F. Li, X. Zou, T. Gao, E.D. %J I.A. Russomanno, "Understanding Shilling Attacks and Their Detection Traits: A Comprehensive Survey," **8**, 171703–171715, 2020.

[22] P. Kaur, S.G. Goel, Shilling Attack Detection in Recommender Systems, 2016.

[23] X. Li, M. Gao, W. Rong, Q. Xiong, J. Wen, "Shilling attacks analysis in collaborative filtering based web service recommendation systems," in 2016 IEEE International Conference on Web Services (ICWS), IEEE: 538–545, 2016.

[24] M. Ebrahimian, R. Kashef, "Efficient Detection of Shilling's Attacks in Collaborative Filtering Recommendation Systems Using Deep Learning Models," in 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE: 460–464, 2020.

[25] I. Gunes, H. %J I.R.J. Polat, "Detecting shilling attacks in private environments," **19**(6), 547–572, 2016.

[26] Y. Hao, F. Zhang, J. Wang, Q. Zhao, J. %J S. Cao, C. Networks, "Detecting shilling attacks with automatic features from multiple views,"

2019, 2019.
[27] V. Mohammadi, A.M. Rahmani, A.M. Darwesh, A. %J H.C. Sahafi, I. Sciences, "Trust-based recommendation systems in Internet of Things: a systematic literature review," **9**(1), 1–61, 2019.
[28] S. Khusro, Z. Ali, I. Ullah, "Recommender Systems: Issues, Challenges, and Research Opportunities," Lecture Notes in Electrical Engineering, **376**, 1179–1189, 2016, doi:10.1007/978-981-10-0557-2_112.
[29] O. Embarak, M. Khaleifah, A. Ali, "An Approach to Discover Malicious Online Users in Collaborative Systems," in International Conference on Emerging Internetworking, Data & Web Technologies, Springer: 374–382, 2019.
[30] Y. Cai, D. %J D.S.S. Zhu, "Trustworthy and profit: A new value-based neighbor selection method in recommender systems under shilling attacks," **124**, 113112, 2019.
[31] D. Lee, P. Brusilovsky, Recommendations based on social links, Springer Verlag: 391–440, 2018, doi:10.1007/978-3-319-90092-6_11.
[32] K. Christakopoulou, A. Banerjee, "Adversarial attacks on an oblivious recommender," in Proceedings of the 13th ACM Conference on Recommender Systems, 322–330, 2019.
[33] C.-M. Lai, Attackers' Intention and Influence Analysis in Social Media, University of California, Davis, 2019.
[34] O.H. Embarak, "Like-minded detector to solve the cold start problem," in 2018 Fifth HCT Information Technology Trends (ITT), IEEE: 300–305, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 699-709 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Automatic License Plate Detection and Recognition for Jordanian Vehicles

Khalil Mustafa Ahmad Yousef *, Bassam Jamil Mohd, Yusra Abd-Al-Haleem Al-Khalaileh, Ahlam Hani Al-Hmeadat, Bushra Ibrahim El-Ziq

Department of Computer Engineering, Faculty of Engineering, The Hashemite University, P.O.Box 330127, Zarqa 13133, Jordan

ARTICLE INFO

Article history:
Received: 19 August, 2020
Accepted: 16 November, 2020
Online: 24 November, 2020

Keywords:
ANPR (automatic number plate recognition)
Jordanian license plates
OCR (optical character recognition)
rectangularity
make of the vehicle

ABSTRACT

Nowadays, automatic number plate recognition (ANPR) is very important especially in the era of smart cities and intelligent transport systems. Fully automated number plate detection and recognition system helps in reducing time, error, and cost for tracking of vehicles and for recording traffic violations. The main goal of this paper is to design a low cost fully automated number plate detection and recognition system targeting the Jordanian license plates. Several problems (e.g., cost, wasted efforts, manual intervention, and possible errors) were identified in the currently used Jordanian number plate recognition for recording traffic violations. We hope that the proposed system would mitigate such problems. The proposed system performs two main tasks. First, it automatically detects and recognizes the license plate number of a given Jordanian vehicle using a robust metric; the rectangularity measurement, and identifies the vehicle's type (e.g., governmental, private, visitor, public, etc.). Second, it recognizes a selected number of trained classes for the make of the vehicle whenever applicable. The experimentation results and the performance evaluations compared to other ANPR approaches show that proposed system achieves the best performance among the tested systems with a plate detection accuracy of 95%, OCR recognition accuracy of 94.68%, make recognition accuracy of 86.84%, and an overall ANPR accuracy of 89% excluding the make results.

1 Introduction

The system of automatic license plate number detection and recognition, known as ANPR (automatic number-plate recognition), is considered a great application and an example of the computer vision technology. ANPR is widely being used in, but not limited to, radar cameras, car park management and traffic management [1]. In Jordan, ANPR is being used in recording traffic violations (e.g., over-speeding), however it has several problems. In particular, some of the ANPR steps: cropping of the license plate and optical character recognition (OCR) are still being done manually [2]. This might be due to the high cost of a fully customized commercial ANPR system for detecting and recognizing the Jordanian license plates. Therefore, the main goal of this paper is to propose a low cost ANPR system that is customized for automatically detecting and recognizing the numbers of the Jordanian license plates. We aim to reduce and avoid human labor, and the possible human errors. Having human involved ANPR steps might result in misidentifying a vehicle plate numbers and data, which can cause erroneous tickets.

The main contributions of this work are as follows:

- Propose a low cost fully automated and customized Jordanian ANPR system.
- Propose and utilize an optimized calculation methodology of the rectangularity metric [3] to efficiently detect the Jordanian license plates.
- Build a database of Jordanian vehicles to be used as a benchmark for evaluating the performance of the proposed system.
- Compare the proposed system to other ANPR systems with the help of the built database for the Jordanian vehicles.
- Incorporate the recognition of the vehicle type and make into the system.

The rest of the paper is organized as follows. Section 2 presents the related work. Section 3 discusses in details the proposed system. Section 4 presents the experimentation results. Section 5 provides a discussion about the proposed system and the obtained results. Section 6 concludes the paper and provides some future work directions.

*Corresponding Author: Khalil Mustafa Ahmad Yousef, Department of Computer Engineering, Faculty of Engineering, The Hashemite University, P.O.Box 330127, Zarqa 13133, Jordan, Contact No: +962796669967 & Email: khalil@hu.edu.jo

2 Related Work

Generally speaking, any ANPR system consists of four main steps as listed below, which are also shown in Figure 1.

1. Capturing an image of a vehicle.
2. Detecting vehicle's license plate.
3. Segmenting out the license plate objects or characters.
4. Recognizing the segmented license plate characters using OCR (Optical Character Recognition).



Figure 1: ANPR steps

Based on the ANPR steps above, in the literature, several works have been proposed to address and optimize each step. Additionally, there is a significant work in combining the ANPR system with detecting and recognizing the make of the vehicle. In what follows, some of the related work regarding the ANPR systems and the vehicle make recognition systems are briefly presented.

In [4] and [5], the authors proposed an ANPR system, where mean shift segmentation is used to identify candidate plate regions. Then, the authors trained and used a Mahalanobis linear classifier with respect to three features to accurately detect number plate region. Namely, they used rectangularity [3], which is defined by the authors as "a measurement that represents how well an object fits its minimum enclosing rectangle", aspect ratio and edge density features.

In [6], the authors proposed an Iranian ANPR system, where edge detection and morphological operators are used to identify license plate candidate regions. The percentage of rectangularity and color filter of each candidate plate region are used to accurately detect the plate region. Multilayer Perception (MLP) neural network is finally used for recognizing the numbers of the detected plate.

In [7], the authors proposed an Indian ANPR system. License plate detection is based on using Sobel filter edge detector, morphological operations and connected component analysis. The segmentation and recognition of plate characters, which are in different Indian languages, are performed based on training a Support Vector Machine (SVM) classifier.

In [2], the authors proposed a Jordanian ANPR system that is based on SIFT features to detect Jordanian license plates and recognize them. Template matching technique is used to locate special marks in the license plate to segment out the license plate. Finally, OCR is used to recognize the characters or numbers from the segmented plate.

In [8], the authors surveyed several keypoint features under the feature-based automatic license plate detection techniques. They categorized the vehicle detection techniques into motion and appearance based, which included sophisticated probabilistic modeling and feature extraction. Also, the authors surveyed and discussed vehicle recognition utilizing vehicle attributes like license plate, color, type, and logo.

In [9], the authors applied a bag of features method to the car make and model recognition (VMMR). The sparse coding technique based on square mapped gradients (SMG) feature descriptors is used to learn a dictionary of codewords. The authors then apply supervised K-means or SVM with SMG to recognize the car make and model. Finally, the authors compared the performance of the two approaches.

In [10], the authors proposed an automatic license plate recognition system to trap traffic violations. For example, to trap offenders disobeying the traffic rules in an enclosed space (e.g., campus).

In [11], the authors proposed an ANPR-based vehicle tracking system, which monitors vehicle's Euclidean distance from a visible calibration line. License plate detection was based on the fact that license plate has an inherent property of having greater transitions in pixel values compared to that of its surroundings.

In [12], the authors proposed a number plate detection and localization system based on segmentation and morphological operators to segment out license plate from the captured image. Specifically, the authors used histogram checking, and vertical and horizontal edge detection methods for the segmentation of the license plate. The experimentation dataset consisted of 350 license plates from different countries including Jordan. The reported efficiency rate from the proposed system is 93.43%.

In [13], the authors proposed a Jordanian ANPR system, where the license plate was detected and localized based on a multistage algorithm using the aspect ratio and the edges ratio. Back propagation neural network was used for recognizing the license plate characters (OCR). The average success rate from the proposed system is 90% of 240 high resolution images of size 1188×960 pixel.

Very recently, in [14], the authors proposed an ANPR system to recognize the Bangladeshi characters in vehicle license plates. Morphological operation and histogram analysis with a dynamic threshold of vertical and horizontal projection profiles were applied in order to detect plate area. Character recognition is carried out using Support Vector Machine (SVM) classifier, where extracted Histogram of Oriented Gradient (HOG) features were used as input. The system was applied to 630 images of Bangla license plates of different categories of vehicles and achieved 91% accuracy in the extraction of plate ROI.

After surveying the literature and to the best of our knowledge, the work of [2, 12, 13] is the only published work that examines the detection and recognition of the license plates of the Jordanian vehicles. Therefore, the work presented in this paper extends the literature regarding the Jordanian license plates detection and recognition. In fact, the proposed work improves the automatic Jordanian license plate detection and recognition proposed in [2] by using another feature, which considers both the appearance shape of the license (rectangularity and aspect ratio) and the license plate content for detecting the license plate. Furthermore, the recognition of vehicle logo and make was also incorporated. It is worth mentioning the notes concerning the work by authors in [12] and [13]:

- The processing time for detecting the license plate and recognizing the plate numbers in both [12] and [13] was not reported at all, whereas in the proposed work it is being reported.
- The dataset of [13] consisted of high resolution images with-

out presenting any sample images. In the proposed work many sample images are provided from the created and employed dataset, which mostly consisted of mid to low resolution images, but better than 320 × 240.

- Neither the work in [12] nor in [13] presented a fully comprehensive ANPR system, where both the license plate type (private, public, government, etc.) and the vehicle's make are being recognized similar to what are being implemented in the proposed work.

Additionally, it is noteworthy to mention that the proposed work shares some similarity with the work presented in [4] and [5]. The proposed system utilizes the rectangularity and aspect ratio metrics in detecting the license plate (Jordanian plate in our case), where it achieved the best plate detection accuracy compared to using the rectangularity methodology proposed in [4] and [5]. This will be discussed in more details in the next section.

3 Proposed System

This section presents the details of the proposed Jordanian ANPR system. Taking into consideration the main ANPR steps presented in Figure 1, in the designed and implemented ANPR system, the following steps were considered as shown in Figure 2. The system was fully implemented in MATLAB R2018a [15], where the graphical user interface (GUI) is shown in Figure 3.

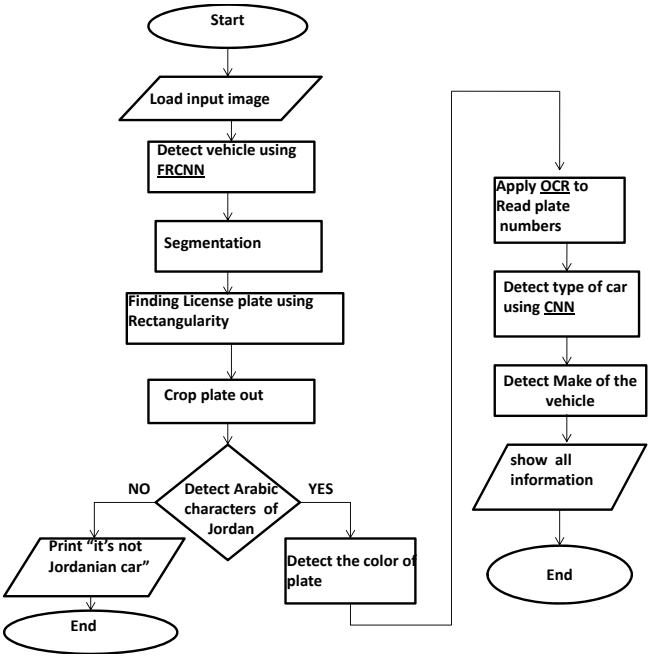


Figure 2: The proposed ANPR steps

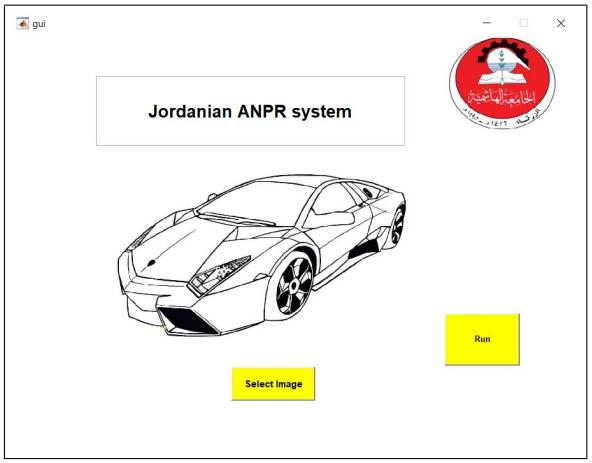


Figure 3: The graphical user interface of the proposed Jordanian ANPR system

3.1 Input image

The very first step in the system is to capture a vehicle image. An off-the-shelf webcam and a mobile phone camera were used to capture a set of database images to evaluate the system's performance. Additionally some of the images were obtained from the Jordanian Opensooq website [16]. In general, the captured images should be of good illumination conditions whenever possible, capture either the back or the front of the vehicle, and have a good resolution (e.g., better than 320 x 240). The database consisted of 100 images, which corresponds to different illumination conditions, different backgrounds, and different perspectives of the captured Jordanian vehicles (i.e., angle and distance between the camera and the vehicle location). Figure 4 shows 16 randomly selected input images chosen from the created dataset. Several preprocessing steps are then being performed on each input image that included histogram equalization to help in reducing varying day light effect and for contrast enhancement.

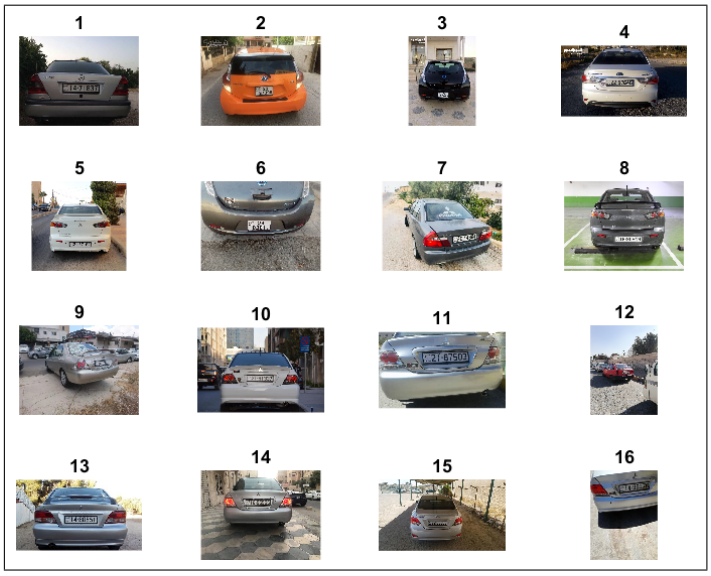


Figure 4: 16 randomly selected input images chosen from the created dataset

3.2 Vehicle detection

The second step is to detect the vehicle from the input image. Faster Regional Convolution Neural Network (FRCNN) [17], which is an objecte detection built-in MATLAB function, was trained to detect and determine the vehicle and its location in the image. FRCNN was chosen since it is significantly faster in training and testing sessions over other detection methods (e.g., R-CNN). Additionally, it is considered as a highly successful framework for generic object detection [18].

The process of using FRCNN is done iteratively among different zooming versions or scales of the input image. This iterative procedure stops, when a vehicle is found or the input image is zoomed to the maximum. In this case (the latter), it may be that no vehicle is found and thus the ANPR system will give an error message indicating such a state and halts. Otherwise (i.e., once the vehicle is found), a blue bounding box around the detected vehicle is drawn. The detected vehicle is then cropped and converted to gray scale. Figure 5 shows an example of this step.



Figure 5: An example of a given input image and the detected cropped vehicle

3.3 License plate detection

The third step is to detect the license plate from the cropped gray scale vehicle image, which is one of the most important ANPR steps. Initially, the fast marching segmentation method [19] is executed to find the best license plate candidate regions. Precisely, this segmentation method will segment the vehicle image into several connected components or parts. The segmented growing parts are initially initialized based on a uniformly chosen seeds within the cropped vehicle image as shown in Figure 6. One of those parts must be the license plate. Figure 7 shows some of the segmented parts as it was very difficult to visualize all candidate segmented parts for all seeds in the same image; as the segmented parts may overlap and therefore the segments will not clearly be viewed. To accurately identify the correct license plate region or part, the following analysis is performed:

- First, the shape of the Jordanian license plate is rectangle with two styles: European and American as shown in Figures 8 and 9, respectively. The European style places the numbers in one row. The American style places the numbers in two rows.
- Next, rectangularity measurement is used [3]–[5] to first identify the segmented parts whose shape is best described by a rectangle, but of certain aspect ratio. This is the aspect ratio, which is computed based on the ration between the height and width of either the American or European styles of the Jordanian license plate. Figure 10 correctly shows identified and detected license plate part based on the rectangularity measurement metric.
- Consequently, a red bounding box is drawn around the part, which satisfies the above constrain or criterion, and crop it as shown in Figure 11.

It is important to emphasize that one of the main contributions in this paper is the methodology used in computing the rectangularity measurement, and thus that for detecting the license plate more efficiently and precisely. Specifically, the presented methodology is quite different and efficient compared to that described in [4], [5] as explained next.



Figure 6: Uniformly chosen seeds in the vehicle cropped image to initialize the segmentation process for the plate detection step.



Figure 7: Some samples of the segmented vehicle parts after the fast marching segmentation during the license plate detection step.



Figure 8: European-style Jordanian plate.



Figure 9: American-style Jordanian plate.



Figure 10: Correctly identified and detected license plate part based on the rectangularity measurement metric.



Figure 11: Cropped license plate after being rotated as to maximize the rectangularity calculation

3.3.1 *Contribution in Calculating the Rectangularity Measurement*

The methodology for computing the rectangularity in [4], [5] requires that:

- each candidate license plate region is to be treated as a rigid object,
- computing its area A_0 by counting the total number of pixels in the region,
- finding its minimum enclosed rectangle (MER) fitted to its boundary by iteratively rotating this object in one degree steps in the range $[-15°, 15°]$ to account for possible tilted angles of the captured vehicles,
- computing the area of MER of this object at each incremental rotation,
- identifying the angle at which the MER goes through the minimum value (A_{MER}), and
- finally computing the rectangularity R at this angle as given in 1:

$$R = \frac{A_0}{A_{MER}} \tag{1}$$

However, the methodology does not require rotating each candidate plate region in a range of steps, which we believe is an important step to perform in order to efficiently detect the correct license plate with high confidence. However, performing this step iteratively and incrementally might be very time consuming and might not produce precise rectangularity calculation. Alternatively,the following steps are performed, which constitute one of the main contributions of this work:

- Each candidate license plate region or object is checked to make sure it has a convex shape as was proposed in [20]. In case it is not convex, it is converted to a convex shape so that spaces in the outwards lines of the object (if any) are connected and filled. This process uses a set of morphological operation (dilation, followed by filling, and lastly followed by erosion).
- Next, ellipse fitting is applied to each convex shaped object. This helps us determine the major and minor axis's of the fitted ellipse of the object.
- The object is then rotated as to remove the rotation angle, which is the angle between the elongation axis (major axis for the fitted ellipse) and the X-axis. Indeed, this is the angle at which the MER goes through the minimum value (A_{MER}).
- Then the area A_0 of the rotated object (the total number of pixels in the convex shape) is computed.
- Finally, the rectangularity R using Equation 1 is calculated.

3.3.2 *Rectangularity Analysis*

The methodology presented in [4], [5] was compared with the proposed methodology. It is important to mention to the best of our knowledge that we could not find any work other than [4], [5] which used the rectangularity metric as a major step in detecting and localizing the license plate similar to what is proposed in this paper. And, since this step is really the most critical step and the heart of any ANPR system, we believe that using the methodology of [4] and [5] as one of the baseline comparisons is still a fair choice especially with the proposed enhancement given in the previous subsection, which we believe shows an advancement to the literature in this direction. In the comparison, the dataset of 100 images is used to compute the detection accuracy of correctly finding the license plate region and the average rectangularity calculation time. The results are presented in Table 1. The results indicate that the proposed rectangularity calculation is more accurate, precise and faster. In fact, the analysis showed that proposed rectnagularity measurement values were very close to one for the candidate license plate objects compared to around 0.7 based on [4], [5]. Additionally, it is worth noting that the rectangularity measurement calculation under different $\Delta\theta$'s $\in [-15, \ldots, 15]$ as in [4], [5] might not be unique. The analysis shows that the same rectangularity value can be obtained under multiple different $\Delta\theta$'s, which can be problematic and result in misidentifying the correct license plate region.

Table 1: Rectangularity analysis

Rectangularity methodology	Average plate detection accuracy	Average rectangularity calculation time (sec)
[4], [5]	65%	0.2785
The proposed methodology	95%	0.0315

3.4 Jordanian license plate Verification

Upon the detection of a possible license plate region, it remains to verify that it is indeed a Jordanian plate. To do so, the label part is examined. It is located to the left side of the Jordanian license plate as shown in Figure 12. It contains the English and Arabic characters of the word Jordan. Therefore, it was decided to use template matching technique based on computing the score of the cross correlation coefficients between two binary windows; (the binary templates of the Arabic characters of the word Jordan and the binary connected components objects extracted from the anticipated label area of the identified license plate candidate). Following are the performed template matching steps:

- converting the label area into the binary scale using Otsu's method.
- segmenting out all possible characters or objects within the label area by performing connected component analysis on them.
- resizing these segmented objects to match the size of the stored binary templates of the Arabic characters of the word Jordan
- exhaustively computing the correlation coefficient between each object and the templates.
  - if the computed correlation coefficients are above heuristically computed threshold, then the detected license plate is declared to be a Jordanian one.
  - otherwise, it gives an error message that it could not find a Jordanian plate and halts.

3.5 Determine license plate type

The fifth step is to determine the type of the Jordanian vehicle. This information is color-coded (see Table 2) in the label area of the license plate. Thus, in the proposed ANPR system, we analyze the color of the detected license plate label area in the original input image. Consequently, based on the analyzed color (most frequently occurred color value) and the recognized code part from the next ANPR step (numbers in first line in American license plate and the digits before the dot in European license plate), we determine the type of the vehicle. Figure 13 shows some examples of the detected Jordanian license plate types from the ANPR system. Table 3 summarizes the types found in the created dataset.



Figure 12: Label image part.



(a) White Label Area: Private Plate



(b) Red Lable Area: Government Plate



(c) Yellow Label Area: Temporary plate



(d) Green Label Area: Public plate

Figure 13: Types of Jordanian license plate [21].



Figure 14: License plate area that excludes the label region

Table 2: Color coded Jordanian license plate types [22]

The two reserved digits indication: selected list		License plate color indication
1 – 5	Government vehicles	Red
6	Aqaba free zone vehicles	Red
7	Diplomatic vehicles	Yellow
8-9	Temporary entrance vehicles	Yellow
10 – 21	Passenger vehicles	White
41 – 42	Light goods vehicles	White
44	Tractors	Red
46	Motorcycles and Scooters	White
50	Taxi	Green
56	Small buses	White
58	Coaches	White
71	Trailer	Yellow

Table 3: Summary of the license plate types found in the dataset

Type of License Plate	Number of License Plates
Private, White	96
Diplomatic, Yellow	2
Taxi, Green	1
Temporary Entrance, Yellow	1
Total	100

3.6 Recognition of license plate numbers (OCR)

The sixth step is to recognize the license plate numbers. First, all possible characters within the detected license plate binary image

are extracted by performing connected component analysis within the plate region. Each extracted connected component or object is resized to match the size of the stored binary templates of the numeric plate numbers (from 0 to 9) that are graphically shown in Figure 15. The typeface used to generate such templates is known as Mandatory [23], which is the same typeface that is used for the United Kingdom vehicle registration plates.

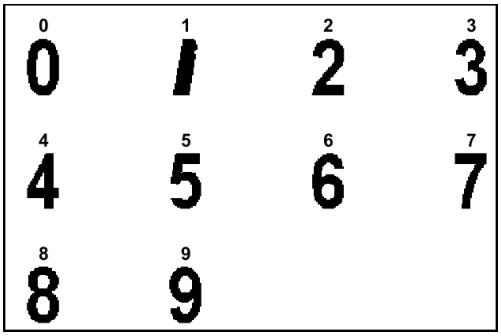


Figure 15: The stored binary templates (inverted in order to best be viewed) of the numeric plate numbers that are desired to be recognized (from 0 to 9)

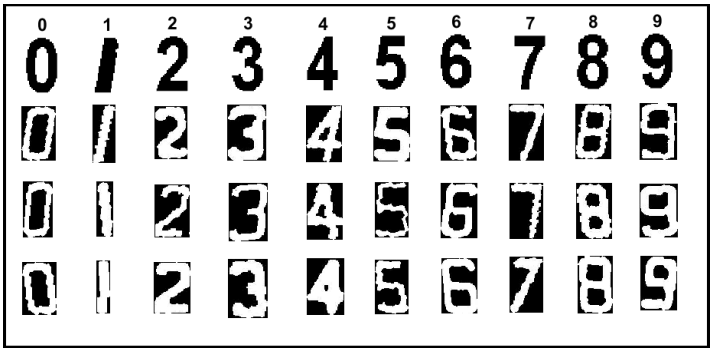


Figure 16: More binary templates of the numeric plate numbers that are desired to be recognized (from 0 to 9)

Finally, the correlation coefficient between each object and the templates is exhaustively computed to recognize the plate numbers. It is noteworthy to mention that, in order to increase the OCR recognition accuracy, the number of stored binary templates for the character set '0123456789' was increased as shown in Figure 16. This resulted in having better matching of the segmented objects and thus high confidence in the associated recognition results. Additionally, the license plate area that excludes the label region (Figure 14) is the OCR region of interest because the segmentation of possible numbers or objects is performed in it. As stated above, the Jordanian license plate splits into two different styles: European and American. Thus, the performed connected component analysis considers the properties and characteristics (distribution of the numbers) of the two types. It also considers the following facts:

- The width and length of all numbers are the same.
- The distance between numbers is nearly the same, except the code part (numbers in first line in American license plate and the digits before the dot in European license plate), which determines the type of vehicles such as: 50: taxi, 36: private bus, etc.).
- All numbers in the license plate have the same color, which is black. Additionally, the background color for the license plate (except label part) is white.

A green bounding box is drawn around each recognized number in the license plate region as shown in Figure 17. After applying the OCR, the code of the license plate is computed by calculating the distance between all numbers. The code part determines the type of the vehicle as we mentioned before and listed in Table 2.



Figure 17: OCR results: remove background from the license plate, segment-out the possible numbers, recognize them, and draw green bounding box around each recognized number.

Table 4: Summary of the vehicle makes found in the dataset

Vehicle Make Index	Vehicle Make	Number of Vehicles
1	Mercedes	11
2	Toyota	10
3	Nisan	7
4	Hyundai	10
5	Mitsubishi	39
6	Opel	3
7	Audi	1
8	Ford	8
9	Mazda	1
10	Lexus	2
11	Daewoo	1
12	Jaguar	1
13	Jeep	1
14	Kia	2
15	Honda	1
16	Peugeot	1
17	Chrysler	1
Total number of Images		100

3.7 Recognizing the make of the vehicle

The seventh and final step of the proposed ANPR system is to determine the make of the vehicle. In this step, the make of the vehicle is recognized by localizing and recognizing the vehicle's logo. Generally speaking, the fact that the vehicle's logo is approximately located in the region above the detected license plate can be relied on, which can be a weak assumption. However, in this presented step, it is desirable to prove the concept of being able to add such feature (i.e., recognizing the make of the vehicle) easily. Thus, the vehicle's logo in the segmented parts from the third step is searched for. The segmented parts are just above the detected license plate region. Once a logo was detected, a yellow bounding box is drawn around it. To help recognize those detected logos, several logo images from the Internet were collected. The images were specifically related to five auto makers MERCEDES, MITSUBISHI, KIA, NISSAN and TOYOTA. The list could be expanded easily. It should be mentioned that the dataset consisted of 17 vehicle makes or brands as summarized in Table 4, and the recognition focus was just put on five of them as indicated above. After that, using the collected logo dataset, a convolutional neural network (CNN) is trained with the help of MatConvNet MATLAB toolbox to recognize the five makes of the vehicles as listed above. While testing this ANPR step on the input images dataset, it is assumed the case where none of the logo images of the vehicles would match the makes from which the CNN classifier is trained on as an "UNKNOWN MAKE". Figure 18 shows the recognized make for the vehicle that was shown in Figure 10.



Figure 18: The recognized make for the vehicle that was shown in Figure 10

To conclude this section, Figure 19 shows and summarizes the output of all stages of the proposed ANPR system. As it is clear from Figure 19:

- The detected vehicle is identified by a blue bounding box.
- The detected plate is identified by red bounding box.
- The detected logo of the vehicle is identified by yellow bounding box.
- The detected license plate is a Jordanian one and its type is private; since it has a white color and its code is recognized as "24".
- The recognized license plate numbers are "12692", where each number is bounded by a green bounding box.
- The recognized make of the vehicle is "Mitsubishi".

4 Experimental results

This section presents the experimental results of the proposed ANPR system. A 100 images were tested as mentioned in the previous section. The average execution time for all tested images was found to be 8.0047 sec on average (ANPR steps 2-6) on a computer machine that has 16 GB of RAM and two quad core 2.5 GHz i7 processor. The system implementation was purely sequential and thus a single core was utilized. Table 5 presents the results (success rates or accuracies). The reported accuracy for each of the rows in Table 5 was computed for each step separately assuming correct results from previous step. As it can be seen from Table 5, the final license plate recognition accuracy from the OCR step was 94.68%. Consequently, the overall ANPR accuracy is 89% given the fact that the license plate detection step failed in 5 images, the license plate verification step failed in 1 image, and the OCR step failed in 6 images for a total of 11 images out of 100 (this is shown in the last row of Table 5). Finally, the accuracy for recognizing the make of the vehicles was 70.21% as given in the second row of Table 6 with an average execution time of 0.0096 sec. However the actual make accuracy was observed to be 86.84% as shown in the last row of Table 6. We should note that images without valid logos and unclear logos were manually excluded. Therefore, by excluding those images from the make evaluation dataset, the above accuracy was recomputed.

Table 5: The experimental results of the proposed ANPR system

ANPR Steps	Num. of input images	Num. of success images	Num. of fail images	Accuracy	Average execution time
Step 2: Vehicle detection	100	100	0	100%	0.3394 sec
Step 3: License plate detection	100	95	5	95%	3.6327 sec
Step 4: License plate verification	95	94	1	98.95%	3.5924 sec
Step 5: License plate type	94	94	0	100%	0.00089 sec
Step 6: OCR	94	89	5	94.68%	0.4393 sec
ANPR system with all Steps 2-6	100	89	11	89%	8.0047 sec

We have compared the proposed system with the following ANPR systems: the ANPR system, which is available online at [24] and the ANPR system proposed in [2] (our previous work). For this comparison, we only used the 94 images out of the created dataset for which the license plates were successfully detected and verified as listed in Table 5. The result of this comparison is shown in Table 7, which shows that the proposed system achieved the best performance in terms of the recognition accuracy.

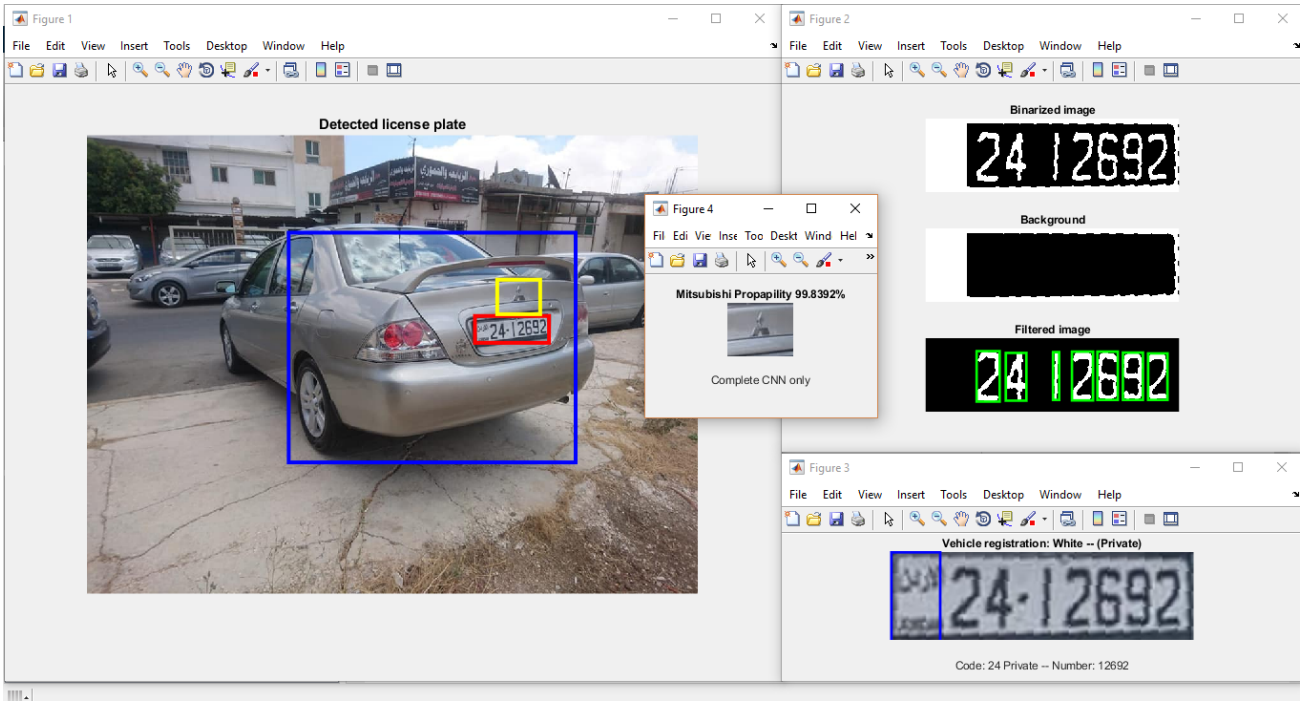


Figure 19: Summary of the output of all the steps of the proposed ANPR system.

Table 6: The make recognition results of the proposed ANPR system

ANPR Steps	**Num. of input images**	**Num. of success images**	**Num. of fail images**	**Accuracy**	**Average execution time**
Step 7: Make of the vehicle	94 (Successfully detected and verified license plate images)	66	28	70.21%	0.0096 sec
Step 7: Make of the vehicle	76 (Same as previous, but excluding the images that have unclear/missing logos)	66	10	86.84%	0.0096 sec

Table 7: Comparison results with other systems

ANPR Systems	License Plate Recognition Accuracy
`https://www.anpronline.net/demo.html`	56%
SIFT-based ANPR system [2]	75%
Our proposed system	94.68%

5 Discussion

As it can be inferred from Table 5, some of the steps need to be optimized either in terms of the average execution time or the accuracy of detection and recognition. For example, the average execution time of the license plate detection step was around 3.6327 sec and it was the largest among the rest of the steps. This is caused by the largest number of candidate segmented regions that were needed to check the license plate for; Rectangularity measurement and aspect ratio. Specifically, this is controlled by the number of the initialized seeds which is equal to $256 = 16 \times 16$ uniformly spaced grid as was shown in Figure 6. Despite that being said, it is important to highlight that our average calculation time of the rectangularity measurement per each candidate license plate part was better compared to the work of [4], [5] as was shown in Table 1 with a detection accuracy of 95%. Studying the factor of reducing the number of initialized seeds to reduce the license plate detection time is planned as a future work and how would that affect the detection accuracy.

Another observation from Table 5 is that the next highest average execution time was the license plate verification step. We believe that this is related to the way in which the verification process was performed as large number of small sized binary segmented windows were exhaustively compared against the binary templates of the English and Arabic characters of the word Jordan. Moreover, we carefully investigated the single failure case from this step, and it turned out to be caused by the bad resolution cropped license plate. The produced binary image of the cropped license plate is as shown in Figure 20.

Additionally, based on the results obtained from the OCR step, it was found that the main causes for the relatively low OCR recognition rate are related to one or more of the following factors:

- The numbers on some of license plates were not so much clear in the captured input images or may have bad illumination conditions; such as the example shown in Figure 21.
- Some of the numbers have certain painting defects that cause the OCR algorithm to not recognize them correctly such as the example shown in Figure 22 for the number ‘6’.

These kind of problems result in a bad segmentation of the plate characters and thus bad input to the OCR algorithm. Therefore, bad OCR recognition results.



Figure 20: The single case where the license plate verification step failed due to the bad resolution of the cropped license plate; the English and Arabic words of Jordan can not be found in the binary image.



Figure 21: A blurring problem in one of the numbers in the detected license plate (number 6) that caused the OCR step to fail

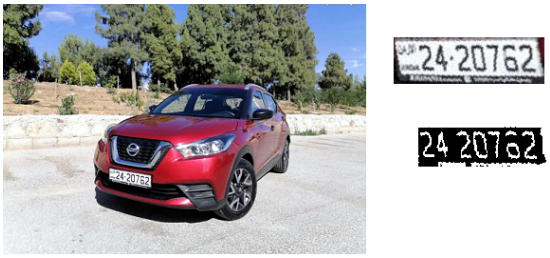


Figure 22: A painting defect problem in one of the numbers in the detected license plate (number 6) that caused the OCR step to fail

Moreover, the lowest recognition accuracy was for the step of recognizing the vehicle's make. This was due to the assumption in localizing the logo of vehicle, which was assumed to always be found in the upper regions above the license plate. We believe that the main challenge was that the location of the logo above the license plate is not indeed consistent among all the vehicles, especially the tested ones (there were too much variations). Therefore, as was mentioned in Section 4 and after analyzing the captured input images, it was found that there were several images without any logo as can be seen in Figure 23, and in some of the images the logo was not clear due to the bad illumination conditions or due to the shadow effect as seen in Figure 24. It turned out that the number of images in the created dataset that possess these kind of problems was found to be 18. As a result of this observation and in order to test the performance of the trained make recognition subsystem, we excluded those 18 images from the make evaluation dataset (94 images). Consequently, the make recognition acucray was observed to be 86.84% as was shown in last row of Table 6.

For the final cost of the proposed ANPR system in terms of the hardware needed to run, it is clear that it only needs a computer machine and a camera to capture input images of the vehicles. Thus, considering the fact that cameras are now almost available everywhere (e.g., mobile phone cameras and surveillance cameras) and that much of the processing needed can be done on the cloud, the estimated cost of the proposed system could be as low as zero.



Figure 23: Four sample vehicle images, where the logo of vehicle that indicates its make is missing



Figure 24: Two sample vehicle images, where there is bad illumination conditions around the logo of each vehicle

Last but not least, although the proposed Jordanian ANPR system in this paper was shown to provide very good recognition results, however it may not work or produce correct results for all cases of the input images. Indeed it needs to be further analyzed under bad illumination conditions, blurriness in image, dirtiness or painting or physical defects problems in the numbers and characters on the license plate, and under the bad weather conditions (e.g., foggy and

rainy conditions). About the speed performance of the proposed system, it is important to emphasis that the estimated time usually depends on the scene complexity of each input image, input image resolution, and the underlying speed of the running machine.

6 Conclusions and Future Work

The paper proposes a fully automated license plate recognition system targeting the Jordanian license plates as a case of study. The proposed system uses and optimizes the calculation of the rectangularity measurement to better identify the best license plate candidate region, which is considered one of the main contributions of the paper. The proposed Jordanian ANPR system was tested on a database that consisted of 100 Jordanian vehicles and compared it to two ANPR systems from the literature. Consequently, proposed system achieves and reports the best performance among the compared systems in terms of the OCR recognition rate, which was 94.68%. We hope that the presented system will be adopted in the very near future in Jordan.

The presented system still needs to be optimized in different aspects as was discussed in Section 5, which is slated for future work. For example, it is desirable to study the effect of the selected number of seeds during the license plate detection step on license plate detection accuracy and the average detection time. Additionally, it is useful to analyze and further experiment the proposed system under bad illumination and weather conditions, blurriness in image, and under physical defects problems in the numbers and characters of the license plate. Moreover, examining better methods for recognizing the make of the vehicle and increasing the number of trained classes for the recognition of other makes are another area of future work.

Conflict of Interest The authors declare no conflict of interest.

References

[1] L. Keilthy, "ANPR system performance," Parking trend international, 2008.

[2] K. M. Ahmad Yousef, M. Al-Tabanjah, E. Hudaib, M. Ikrai, "SIFT based automatic number plate recognition," in 2015 6th International Conference on Information and Communication Systems (ICICS), 124–129, IEEE, 2015, doi:10.1109/IACS.2015.7103214.

[3] K. R. Castleman, Digital Image Processing, Prentice Hall, 1996.

[4] W. Jia, H. Zhang, X. He, "Mean shift for accurate number plate detection," in Third International Conference on Information Technology and Applications, 2005. ICITA 2005., volume 1, 732–737, IEEE, 2005, doi: 10.1109/ICITA.2005.176.

[5] W. Jia, H. Zhang, X. He, "Region-based license plate detection," Journal of Network and computer Applications, **30**(4), 1324–1333, 2007, doi: 10.1016/j.jnca.2006.09.010.

[6] S. Rastegar, R. Ghaderi, G. Ardeshipr, N. Asadi, "An intelligent control system using an efficient License Plate Location and Recognition Approach," International Journal of Image Processing (IJIP) Volume (3), **3**(5), 252–264, 2009.

[7] M. Shidore, S. Narote, "Number plate recognition for indian vehicles," IJCSNS International Journal of Computer Science and Network Security, **11**(2), 143–146, 2011.

[8] R. A. S. Ma'moun Al-Smadi, Khairi Abdulrahimand, "Traffic Surveillance: A Review of Vision Based Vehicle Detection, Recognition and Tracking," International Journal of Applied Engineering Research, **11**(1), 713–726, 2016.

[9] I. Kamal, K. Housni, M. Y. Hadi, "Car Make and Model Recognition using Sparse Coding and Bag of Features," in Proceedings of the 2nd international Conference on Big Data, Cloud and Applications, 79, 1–5, ACM, 2017, doi: 10.1145/3090354.3090435.

[10] P. Agarwal, K. Chopra, M. Kashif, V. Kumari, "Implementing ALPR for detection of traffic violations: a step towards sustainability," Procedia Computer Science, **132**, 738–743, 2018, doi:10.1016/j.procs.2018.05.085.

[11] K. Indira, K. Mohan, T. Nikhilashwary, "Automatic License Plate Recognition," in Recent Trends in Signal and Image Processing, 67–77, Springer, Singapore, 2019, doi:10.1007/978-981-10-8863-6_8.

[12] A. A. Moustafa, M.-I. R. M. Jaradat, "A new approach for license plate detection and localization: Between reality and applicability," International Business Research, **8**(11), 13, 2015, doi:10.5539/ibr.v8n11p13.

[13] H. Alhaj Mustafa, S. Hassanin, M. Al-Yaman, "Automatic Jordanian License Plate Recognition System Using Multistage Detection," in 2018 15th International Multi-Conference on Systems, Signals Devices (SSD), 1228–1233, IEEE, 2018, doi:10.1109/ssd.2018.8570408.

[14] R. Islam, M. R. Islam, K. H. Talukder, "An efficient method for extraction and recognition of bangla characters from vehicle license plates," Multimedia Tools and Applications, 20107-20132, 2020, doi:10.1007/s11042-020-08629-8.

[15] The Mathworks, Inc., Natick, Massachusetts, MATLAB version 9.4.0.813654 (R2018a), 2018.

[16] Available online: `https://jo.opensooq.com/en/cars/cars-for-sale` (accessed on 3 Jan 2020).

[17] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, **39**(6), 1137 – 1149, 2017, doi: 10.1109/tpami.2016.2577031.

[18] Q.-x. Zhang, G.-h. Lin, Y.-m. Zhang, G. Xu, J.-j. Wang, "Wildland Forest Fire Smoke Detection Based on Faster R-CNN using Synthetic Smoke Images," Procedia engineering, **211**, 441–446, 2018, doi:10.1016/j.proeng.2017.12.034.

[19] J. A. Sethian, "A fast marching level set method for monotonically advancing fronts," Proceedings of the National Academy of Sciences, **93**(4), 1591–1595, 1996, doi:10.1073/pnas.93.4.1591.

[20] P. L. Rosin, "Measuring shape: ellipticity, rectangularity, and triangularity," Machine Vision and Applications, **14**(3), 172–184, 2003, doi:10.1007/s00138-002-0118-6.

[21] Available online: `http://www.plateshack.com/y2k/Jordan/jordany2k.htm` (accessed on 3 Jan 2020).

[22] Available online: `https://en.wikipedia.org/wiki/Vehicle_registration_plates_of_Jordan` (accessed on 3 Jan 2020).

[23] Available online: `https://en.wikipedia.org/wiki/Mandatory_(typeface)` (accessed on 3 Jan 2020).

[24] "ANPRonline," Available online: `https://anpronline.net/index.html` (accessed on 3 Jan 2020).

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 1029-1037 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Framework for Decentralising Municipal Solid Waste Management in Harare, Zimbabwe

Trust Nhubu[1,*], Edison Muzenda[1,2], Mohamed Belaid[1], Charles Mbohwa[3]

[1]*Department of Chemical Engineering Technology, University of Johannesburg, Johannesburg, 0027, South Africa*

[2]*Department of Chemical, Materials and Metallurgical Engineering, Botswana International University of Science and Technology, Palapye, 00267, Botswana*

[3]*Department of Quality and Operations Management, University of Johannesburg, Johannesburg, 0027, South Africa*

ARTICLE INFO

Article history:
Received: 22 December, 2020
Accepted: 12 April, 2021
Online: 22 April, 2021

Keywords:
Anaerobic digestion
Backyard composting
Community based municipal solid waste management
Composting
Municipal solid waste
Decentralisation
Incineration
Integrated municipal solid waste management
Recycling
Harare
Zimbabwe

ABSTRACT

Municipal Solid Waste Management (MSWM) decentralisation brings about reductions in the amount of municipal solid waste (MSW) earmarked for landfilling worse still under situations where MSW is being sent to dumpsites. It also reduces the distances MSW collection vehicles travel during MSW collection, maintenance and transport costs due to the establishment of local level decentralised MSWM and treatment facilities. Subsequently fuel use, greenhouse gas and other emissions together with MSWM associated environmental and human health risks are reduced. The Zimbabwe National Integrated Solid Waste Management Plan (ZNISWMP) provides for the need for decentralisation in MSWM. This study therefore assessed the framework along which MSWM decentralisation can be achieved in Harare. The study noted the presence of various opportunities for MSWM decentralisation in Harare namely household backyard composting, community level and industrial scale anaerobic digestion or composting of organic MSW fraction, anaerobic co-digestion of organic MSW fraction and dewatered sewage sludge, MSW source separation for material recovery, establishment of waste transfer stations, citizens drop off centers, buy back centers and thermal treatment facilities associated with energy recovery. Though the NISWMP plan provides for concrete actions for MSWM decentralisation, the study observed that almost all of the proposed actions have not been implemented hence the need for urgent review and subsequent operationalization and implementation of the review findings. MSWM has also been hindered by the lack of legislative and institutional reforms with ULAs remaining reluctant to devolve and delegate some of the MSWM responsibilities and functions to other players prompting the need for such reforms to be implemented as provided in goal 10 of plan. The Presidential national cleanup day proclamation needs to be complemented with other initiatives that will increase residents' interest in participation, allow for different types of participation such as provision of resources and equipment and above all the development of sustainable MSW disposal facilities unlike dumpsites.

1. Introduction

This paper is a revised and extended version of a paper entitled, "Options for decentralised municipal solid waste management in Harare, Zimbabwe", that was presented at the 7th International Renewable and Sustainable Energy Conference (IRSEC' 19), Sofitel Agadir Royal Bay, Agadir, Morocco, November 27-30, 2019.

In Zimbabwe, municipal solid waste management (MSWM) is a designated authority of urban local authorities (ULAs) including municipalities, town councils, local boards and rural district councils largely implemented under a conventional and centralised system [1-3]. Like most ULAs in Africa, the ULAs within and in the vicinity of the Harare metropolitan province in Zimbabwe are experiencing MSWM failures. These failures are largely due to increased municipal solid waste (MSW) generation emanating from increased urban population and improved lifestyles. The

*Corresponding Author: Trust Nhubu, Cr Siemert & Beit Streets, Doornfontein, Johannesburg, South Africa, 2094, +263777342295, nhubutrust@gmail.com

MSWM failures thus are a manifestation of mismatch between increased MSW generation and investment in MSWM and accompanying necessary infrastructural development. Other factors leading to MSWM failures are poor governance and inadequate capacity to fully adopt technology intensive MSWM and treatment methods being developed in the developed world.

The limited and or no external investment as well as the over reliance on rate payers characterizing the MSWM in the ULAs in and around Harare is largely unsustainable posing MSWM related human health and environmental hazards. Residents have developed general unwillingness to pay attitude and regard MSWM a sole responsibility of ULAs thus they resort to indiscriminate or irresponsible dumping of MSW in undesignated areas when over 50% of the MSW generated remain uncollected as reported by [4]. ULAs regard residents as MSWM rates or service payers who must only participate in MSWM through paying for MSWM services that would have been provided. The prevailing MSWM model thus lacks all the necessary partnership and cooperation ingredients amongst the stakeholders specifically between the residents served by MSWM systems and ULAs [2, 5]. The model does not provide adequate environment for residents as well as other stakeholders involvement in MSWM [6].

ULAs attributes the residents' inability and unwillingness to pay to the unsustainable, ineffective and inefficient MSWM systems currently prevailing within their jurisdictions. The residents on the other hand attribute their unwillingness and failure to pay for MSWM services to the poor MSWM services delivered, abuse of resources by the ULAs' executives and the general economic hardships emanating from the prevailing economic environment. Such contradictions shows the ineffectiveness, inefficiency and unsustainability of the conventional and centralised MSWM model inherited from the colonial era and being implemented in developing countries [7, 8]. Municipalities and or ULAs face financial constraints that hinder MSWM efficiency [9-11] leading to low or no collection of MSW generated [12-14]. The low or no MSW collection fuels illegal as well as indiscriminate open dumping, burying and burning of MSW in undesignated areas [4, 15, 16]. MSWM system failures, therefore, become inevitable and expected to worsen in Harare due to the residents' unwillingness and inability to pay as well as the absence of alternative financing sources. In addition MSWM system failures in developing countries has also been linked to the continued dependence on the linear MSW generation, collection, transportation and final disposal model or approach under the conventional and centralised system because of its rigidity that hinders the local level solutions acceptance and adoption of specific and unique local MSWM requirements [17]. Such MSWM approaches are not designed to bring solutions to MSWM associated complexity since one compartmentalized MSWM solution is likely to generate other MSWM problems [18]. Therefore, the linear (generation, collection, transportation and disposal), conventional and centralised model being implemented in developing countries and specifically in Harare is discredited for being a problem transfer (MSW transfer) from MSW sources to disposal sites especially land intensive dumpsites and landfills [19]. Zimbabwe thus realised and noted its lack of capacity to effectively provide MSWM services in its urban environments under the prevailing MSWM system or model and initiated the process of developing a national integrated solid waste management plan (NISWMP) in 2010 [20].

The Zimbabwe NISWMP amongst eight of the issues it sought to address noted in another study [21] speak on the limitations of the conventional and centralised MSWM system specifically the continued reliance on MSW generation, storage, collection and transportation and then final disposal by ULAs in Zimbabwe with major challenges being experienced under collection, transportation and disposal resulting from failures to meet the increased waste generation; lack of effort towards avoiding or preventing, reducing and controlling the increasing waste generation; and non-involvement, lack of participation and consultation of stakeholders especially the residents who are served by MSWM systems as well as the private sector since MSWM is largely a preserve of ULAs and their hired agents. In attempting to address the mentioned shortcomings the Zimbabwe NISWMP provides for decentralised MSWM options and the promotion of stakeholders' participation in MSWM within Zimbabwean urban environments. The Zimbabwe NISWMP's goal 10 of strategy C under action 5 (A5) provides for MSWM authority to be decentralised to the lowest appropriate level and enhancing ULAs' capacities to fulfil their obligations. Zimbabwe does not have a national waste management policy and Goal 10 provides for the need to develop one and review and assess the current legislative and institutional arrangements for the purpose of improving implementation [20]. This study, therefore, reviews MSWM models together with the associated available options for MSWM decentralisation in ULAs within and surrounding Harare City considering the MSWM decentralisation proposals in the Zimbabwe NISWMP as well as from other literature sources.

2. Decentralisation in MSWM

MSWM decentralisation encompasses relevant stakeholders' involvement and participation especially the residents and those who use and are served by the MSWM system. Stakeholder participation is integral in both community based municipal solid waste management (CBMSWM) and integrated municipal solid waste management (IMSWM) considering that MSWM sustainability is attained only when local authorities accept its corresponding management system. Such a MSWM system must be realistic and appropriate in providing solutions to the available and specific MSWM circumstances and challenges of an urban area (city, town, local board, growth point etc) and its surroundings. It must also be able to engage and capacitate all its stakeholders and users namely households, private businesses together with their employees, government departments (local, regional and national), non-governmental organisations including donors and international financing simultaneously recognizing and considering their socio-economic, political, cultural and environmental status.

2.1. Community Based Municipal Solid Waste Management (CBMSWM)

Communities are considered active role players under CBMSWM when it comes to the cleaning of their environments which has the potential of providing them with livelihoods in as far as earning an income from MSW [22, 23] and by taking advantage of the belief and understanding that people are likely to change their behaviours and attitudes towards something if they are part

of problem identification and solution design. Strong relationships are therefore built amongst the various MSWM stakeholders namely ULAs, residents and private business. It is under such CBMSWM that communities and residents bear the responsibility of providing clean and sanitary environments in the vicinity of their households and residences by discarding MSW in designated receptacles. Communities and residents are also potential agents for materials recovery (recycling and reuse) from MSW, backyard compositing of organic MSW fraction, MSW separation at source including at households and or other MSW generation sources which will be collected by respective MSW fractions collectors. They can as well resultantly deliver their source separated MSW fractions to any of the following: citizens drop of centers (CDOPs) or citizens convenient centers (CCCs), buy back centers (BBCs), waste transfer stations (WTS), materials recovery facilities, specific MSW fraction treatment facilities (anaerobic digestion and or composting for organic MSW fractions, incinerators, pyrolysis and gasification plants and landfills). In addition, they can as well perform other administrative MSWM functions such as recording MSW collection services like the daily number of premises and households served, quantity of MSW collected etc. within their localities [9, 14, 24].

CBMSWM provides communities and residents with the opportunity to be involved and participate in MSWM meetings, elect their local representatives and leaders responsible for the management of MSW collection and providing residents feedback and complaints to the ULAs. In addition to the role played by communities and residents in MSWM, CBMSWM recognizes the vital role of community-based organisations (CBOs) and that of local private businesses. Authors in [25] and [26] reported that CBMSWM has brought some improvements in MSWM in household or residential areas regardless of the other social challenges reported to have been encountered in some developing countries during its implementation [3, 27] as well as the negative perception regarding MSWM being considered a sole responsibility of ULAs [3, 28].

2.2. Integrated Municipal Solid Waste Management (IMSWM)

IMSWM model has received global acceptance as a system that could address the current MSWM challenges with full scale operationalisation and implementation having been achieved in the developed nations whereas developing nations are also making significant strides towards its operationalisation and implementation. IMSWM emerged from the need for MSWM shift from landfilling as the sole disposal method to a wider and holistic perspective that incorporated value extraction through materials and energy recovery from MSW that started in the 1990s. The need to shift from MSW landfilling as envisaged under the IMSWM model however is despite the continued reliance on dumpsites in some or majority of developing countries making the transition from dumpsites to landfilling inevitable, thus a transition that will not include landfilling which is to be avoided under the IMSWM for these developing countries will be too huge and difficult to implement. The IMSWM model is meant to bring about the necessary trade-offs regarding the effectiveness of the environmental sustainability, economic affordability and social acceptability dimensions of a MSWM system [29, 30]. The authors in [29] further pointed out that IMSWM focuses on integrating the numerous MSWM processes namely; MSW generation, storage, collection and transportation, treatment or processing methods, recovery of materials and energy, final disposal and entities (MSW producers and users, managers and policy formulators, governments, financiers and funders) that form a MSWM system to bring about continuous improvement together with the MSWM associated environmental impacts and costs reduction. The design of IMSWM systems is meant to achieve specific local level MSWM goals taking into consideration the prevailing MSW generation and characteristic together with the available and appropriate MSW prevention, reduction, reuse and recycle and final disposal methods, political, socio-economic, cultural, environmental and institutional systems factoring in stakeholders' perspectives and needs [29, 31, 32] which guarantees the needed MSWM systems decentralisation and sustainability since ignoring these social aspects and other local level priorities have led to systems failures [33, 34].

Effectiveness in a MSWM system is attained when both ULAs and residents fully embrace it thus departing from the traditional system where experts design and outline MSWM solutions prior to the involvement of residents meant to be serviced by the MSWM system. This is largely so since MSWM system acceptance by residents is anchored on their involvement and participation during planning, design, development, implementation and operationalisation with the evolving value systems, consumer attitudes, perceptions and behaviours being equally important as the technical and economic aspects [33, 34].

The traditional meaning of waste which is subjective in that it attaches a negative value on MSW by regarding it an unwanted item or material to the disposer is regarded only true in the IMSWM system only when the MSW waste has been utilised to its full potential leaving no further processing potential to recover materials and energy [18, 32]. IMSWM thus brings together the waste hierarchy elements factoring in the associated MSWM direct impacts from MSW generation, storage, collection and transportation, treatment and processing and final disposal together with indirect impacts from the use of recovered materials and energy [10]. Various players therefore, can take different roles along the different IMSWM system stages instead of having ULAs being the sole players responsible for the entire MSWM system functions. This brings about the much-needed decentralisation as well as specialisation. For instance, collection and transportation, treatment and processing may be delegated to other players whilst the final disposal be the duty of other players.

3. Support Pillars for MSWM Decentralisation in the NIMSWMP

The NISWMP provides several key strategies and action points under its goals 1 to 10 that support decentralisation of MSWM within ULAs in Zimbabwe. Such key actions include but not limited to the need for reduction in waste, source separation of MSW, reuse of biodegradable MSW for gardening and livestock feed at household level, establishment of viable markets for biodegradable MSW, production of manure from biodegradable MSW composting or vermi-composting, establishment of community and industrial recycling centers, development of recycling projects at institutions such as schools, tertiary education institutions and hotels, construction of AD facilities at established recycling centers and institutions like schools, universities, prisons

as well as markets and hotels, establish combined Heat and Power (CHP) generation facilities from AD derived biogas, putting in place a kerbside collection system with citizens' drop off or citizens' convenient centers where households or institutions bring waste as provided in Table 1.

All these key actions provided in the NISWMP with regards to enhancing decentralisation speaks on the need for stakeholder participation in MSWM which is further affirmed under goal 8 which seeks to promote cleanliness in Zimbabwe by involving the public, industry and government in efforts towards reducing, reusing and recycling all solid waste materials to manage and mitigate the MSW associated public health and environmental impacts. In addition, Goal 7 of the NISWMP supports decentralisation as it seeks to educate and raise awareness amongst Zimbabwean citizens to better understand the importance of participating in source separation; materials and resources recovery and integrated and sustainable solid waste management thus the high literacy level in Zimbabwe is believed will enhance citizens participation easing the introduction and implementation of decentralised waste management systems. Interestingly the tenth goal of the plan seeks to review and assess current legislation and institutional arrangements to improve implementation of ISWM amongst its key actions under its strategy C is the need to decentralize MSWM authority to the lowest appropriate level as well as strengthening the capacities of ULAs to meet their obligations with regards to decentralizing MSWM authorities.

4. Proposals for MSWM Decentralisation in Harare

The ULAs that constitutes Harare metropolitan province include Harare the Capital City of Zimbabwe, Chitungwiza municipality and Epworth Local Board. Other two ULAs that borders Harare city namely Ruwa Local Board to the east and Norton Local Boards to the west are generally considered part of ULAs that constitutes what is loosely termed Harare. The waste generation figures in these ULAs are as given in Table 2. Currently each ULA is responsible for the management of the MSW generated within their jurisdictions. These ULAs are indiscriminately collecting less than 50% of the MSW generated and dump it at open non engineered or sanitary dumpsites serve for Norton which has an engineered sanitary landfill whose environmental soundness is under questioning. Several options for decentralised management and treatment of the MSW fractions generated within a specific jurisdiction as shown in Table 2 can be pursued. The options for treating these MSW fractions generated should be focused on AD of biodegradable MSW fraction, recycling of at least 20% and incineration of at most 80% of the non-biodegradable fractions. The focus on the mentioned MSW treatment is informed by study findings from [35] which identified the option incorporating AD of organic fraction of MSW, at least 20% and at most 80 percent recycling and incineration of the non-organic MSW fraction generated in Harare respectively being the least environmentally impactful option. Composting of the biodegradable MSW fractions could also be considered after AD due to the better environmental performance and overall edge of AD over composting resulting from its renewable energy production capabilities and other associated benefits noted by [36].

4.1. Biological Treatment

The AD of the organic MSW fraction generated in Harare (136,612 tons), Chitungwiza (32, 822tons), Norton (6,21tons), Ruwa (5, 213tons) and Epworth (15,403tons) could be done at household, institutional, community or industrial scale. Households and institutions (schools, universities, hotels etc.) could source separate their biodegradable MSW they generate and feed it into purposely built small biogas digesters from which they can get biogas for their domestic cooking. Community level AD systems could therefore bring medium to large scale operation of up to 100,000 tons per annum where ULAs could partner with private sector, institutions and local residents. Further upgrading of biogas utilisation at institutional or community levels to install combined heat and power (CHP) generation units where the biogas produced from AD is burnt to generate renewable electricity and heat. The biodegradable MSW could also be composted at household and institutional level through backyard composting thereby reducing MSW amount transported for disposal at landfills in the face of heightened global concern against the landfilling of biodegradable MSW fractions. Other medium and large-scale composting facilities could also be established by the private sectors, ULAs and residents. The establishment of household and institutional AD and composting facilities as well as community or industrial scale AD and composting facilities within the neighborhood where the biodegradable MSW is generated brings about a reduction in MSW amount on a weight and volume basis to be transported to landfills. The number of trips and distance travelled by MSW collection vehicles are reduced subsequently reducing fuel consumption, MSW collection and transportation associated greenhouse gases (GHG) emissions, vehicles maintenance and overall costs. However, need for capacity building arises to avoid any unintended health and environmental impacts from the improper operation of household AD and backyard composting. Medium scale to large scale AD and composting facilities under decentralisation can be operated by the private operators with communities, institutions and ULAs facilitating source separation of organics to ensure the necessary AD and composting feedstock quality. ULAs will therefore collect and deposit their sorted organic waste to the composting and AD facilities at nominal agreed gate fees.

Alternatively, private players will be responsible for the collection of the source separated organics. The communities involved in source separation could benefit from the source separation associated incentives likely through MSWM fees reduction. Other benefits are largely the accompanying reduced environmental and human health risks from the AD and composting of biodegradable MSW compared to its open and indiscriminate dumping. Communities will also benefit from biogas which they could purchase and alternatively use for cooking. There is also envisaged increased availability of renewable electricity from the AD biogas associated CHP generation. The biogas potential production is as shown in Table 3.

Table 1: Goals, strategies and key actions in the Zimbabwe NISWMP promoting MSWM decentralisation

Goal	**Strategies**	**Key Actions**
Minimize SW generation.	Encourage the public to prevent SW generation.	i. Promote reduction of waste generation at household level.
Ensure source separation of solid waste.	Put in place appropriate source separation of SW systems at sources.	i. Separate waste into biodegradable and non-biodegradable fractions. ii. Introduce incentives for source separation to bring in public and industry buy in.
Promote the reuse of biodegradable SW to reduce its quantities earmarked for disposal sites.	Use biodegradable SW for gardening and livestock feed.	i. Use of biodegradable SW for gardening and livestock feed at household level. ii. Establish viable markets for biodegradable SW. iii. Packaged manure production from biodegradable SW composting or vermi-composting.
Maximize materials and resource recovery.	Support the development of a recycling economy.	i. Allocate land to enable informal waste traders/community groups to separate and sort recyclable SW. ii. Establish industrial sites or recycling centers where a multitude of recycling firms can rent space. iii. Establish waste recycling projects at institutions such as schools, tertiary education institutions and hotels. iv. Encourage the private sector, small and medium enterprises to develop and upgrade SW management facilities. v. Establish a supportive institutional framework for private recycling enterprises. vi. Establish research informed markets for recycled materials and resources.
	Promote SW value addition and innovation on recycled materials and resources.	i. Invest in the development of technologies for SW value addition and innovation on recycled materials and resources. ii. Use of indigenous knowledge systems on SW value addition like basket weaving, local crafts etc.
	Establish an enabling policy and legal framework for multi stakeholder participation in SW recycling.	i. Institute legislative, regulatory and institutional review to support recycling markets and promote SW recycling. ii. Introduce licensing system for SW recycling, incentives and subsidies for investment in SW recycling promotion. iii. Establish a Green Fund where individuals, banks and private sector contribute to support SW recycling.
	Explore and promote energy recovery from SW.	i. Construction of AD facilities at established recycling centers and institutions like schools, universities, prisons as well as markets and hotels. ii. Establish combined Heat and Power generation facilities from AD derived biogas.
Restructure and introduce an efficient source separated waste collection system.	Establish strong diversified and appropriate SW collection systems for source separated biodegradable (wet waste) and non-biodegradable waste (dry waste) fractions.	i. Introduce a collection system for recyclables where collection points include schools, office blocks and apartments. ii. Establish a kerbside collection system with citizens' drop off or citizens' convenient centers where households or institutions bring the recyclables. iii. Establish partnerships amongst stakeholders namely residents, ULAs, CBOs, individual waste collectors and private sector to optimize collection of sorted SW. iv. Integrate the residents, ULAs, private sector, informal collectors and CBOs into the IMSWM system within their local authorities.
	Pilot various SW management and collection systems in different residential areas.	i. Pilot ISWM systems in high density areas where individuals and CBOs collect, sort and recover, value add, compost and sell the products. ii. Pilot ISWM systems in high density areas characterised by a kerbside collection system with citizens' drop off or citizens' convenient centers where citizens bring in their source separated waste and subjected to a subsidy to reduce their individual household SW collection fee.
	Enhance ULAs capacity to provide efficient SW management services.	i. Review ULAs by laws to equip residents to provide appropriate SW receptacles at their households and have SW collection fees reduction incentive.

Invest in and build environmentally sustainable SW management infrastructure and systems.	Construct, upgrade and or rehabilitate SW collection, management, treatment and disposal infrastructure.	i. Construct, upgrade and or rehabilitate SW incineration, composting, treatment/neutralization and energy recovery facilities.

Table 2: Annual MSW and specific MSW fraction generation figures for Harare City and its surrounding ULAs [35]

ULA	Estimated annual MSW generation (ton)	Amount of MSW Fractions (ton)				
		Organic	Plastic	Metals	Paper	Glass
Harare	325,266	136,612	107,338	26,021	45,537	9,758
Chitungwiza	78,148	32,822	25,789	6,252	10,941	2,344
Norton	14,802	6,217	4,885	1,184	2,072	444
Ruwa	12,412	5,213	4,096	993	1,738	372
Epworth	36,674	15,403	12,102	2,934	5,134	1,100
Total	**467,303**	**196,267**	**154,210**	**37,384**	**65,422**	**14,019**

Table 3: Estimated biogas potential for Harare City and its surrounding ULAs [44]

ULA	Harare	Chitungwiza	Norton	Ruwa	Epworth	Total
Organic MSW generated (tons)	136,612	32,822	6,217	5,213	15,403	196,267
Estimated annual biogas potential (m^3)	1.59E+07	3.82E+06	7.24E+05	6.07E+05	1.79E+06	2.29E+07

4.2. Recycling

There is great emphasis on the need to increase the recycling levels in Zimbabwe. For example, in Harare only 9.6% of the MSW generated in Harare is recycled translating to only 23,000 tons of the over 467,000 tons generated. At least 20% of the non-biodegradable MSW equivalent to at least 56,000 tons has to be recycled [35] thus double the current recycling level. To meet this target therefore, recycling could be done at household, institutional and community level. It will have to be enhanced by the source separation of MSW generated. Source separation of MSW promotes stakeholder and citizens participation in any waste management system thus enhancing decentralisation. Citizens drop off or convenient centers could also be established under a kerbside waste collection system where those in the recycling business will come and separate the materials, they want for their recycling business. Recycling provides economic and employment opportunities to communities who would be involved either in source separation through incentives or those collecting recyclable materials to recycling facilities. The model for operation under a decentralised model will be to have recycling companies either collecting the source separated recyclables or having ULAs depositing the recyclables at a mutually agreed gate fee with waste pickers depositing their recyclable materials at an agreed cost. Those involved in source separation will benefit from reduced waste management fees they pay to ULAs through incentives for their participation in source separation.

4.3. Thermal Treatment

The incineration of MSW generated in Harare is likely to be done at institutional, community or industrial scale. In addition to the better environmental performance of incinerating MSW generated in Harare compared to its landfilling observed by [35], authors in [37] reported the suitability of MSW generated in Harare to be treated via incineration with a lower heating value of between 10100 kJ/kg and 9320 kJ/kg giving an estimated 3.8 ×106 GJ per annum. The lower heating value and incineration efficiency thus could be enhanced by the source separation of biodegradable MSW. The incineration of at most 80% of the non-biodegradable MSW generated translates to the incineration of at most 217,000 tons of non-biodegradable MSW estimated to give at most 1.96 ×106 GJ per annum. Like under AD and composting, ULAs could collect the source separated non-biodegradable MSW meant for incineration and deposit at incineration facilities at an agreed gate fee whereas incinerator operators can as well do the collection with residents benefiting in various forms from their participation in the decentralised MSWM system. Table 4 thus gives the estimates of potential energy generation from the incineration of the non-biodegradable MSW under each jurisdiction within and surrounding Harare.

5. Discussion

MSWM decentralisation promotes the backyard composting of organic MSW fraction at households and or the industrial scale AD and composting of source separated or mechanically separated organic MSW fraction. The industrial scale AD and composting of source separated organic MSW fraction leads to the production of a biofertilizer from the AD digestate and compost from composting thereby providing the safe utilisation or disposal of the digestate and compost as a nutrient rich biofertilizer devoid of biological, chemical and physical pollutants. Source separated

organic MSW fraction provides the AD and composting feedstock quality necessary for the production of biofertilizer from high quality AD digestate and compost that will be used for agricultural application [38]. The AD and composting associated nuisance in the form of noise from heavy MSW collection vehicles and facility operations, odours, bio-aerosols, environmental and human health impacts has however raised concerns from residents despite the undisputed associated benefits [39-42]. Solutions to these concerns could incorporate the installation of smaller decentralised AD and composting facilities of low operational capacities distributed within the ULAs as proposed by [43] unlike the large centralised facilities. These distributed and decentralised smaller AD and composting facilities of low capacities results in the reduction of distances travelled by MSW collection vehicles subsequently yielding corresponding reductions in fuel consumed, transport costs, GHG and other emissions namely traffic and noise. The renewable electricity generated from the CHP generation through the combustion of AD derived biogas will serve the local communities to bring about increased residents' acceptance of the AD and composting facilities in their neighbourhoods.

Opportunities for the anaerobic co-digestion of source separated organic MSW fraction and dewatered sewage sludge in the smaller decentralised facilities of low capacities distributed across the ULAs in and around Harare also exist. This will entail composting of the co-digestion digestate regarded ideal for localized and small communities' setups. The co-digestion of organic MSW fraction and sewage sludge increases the AD biogas yield and improves AD process stability [45, 46]. Decentralisation potentially facilitates for the participation of residents in material recovery or recycling activities which is anticipated to increase materials recycling levels beyond the current 9.6% of MSW generated reported by [4]. Decentralisation promotes the development and construction of decentralised community level WTS, CDOPs or CCCs as well as BBCs. All the proposed decentralised MSWM systems can reduce the amount of MSW on a volumetric and weight basis to be sent to landfills subsequently reduce land demand for MSW landfilling purposes. Thermal waste treatment methods to recover energy through incineration, pyrolysis and gasification facilities can potentially be integrated into the proposed MSWM decentralisation models.

A life cycle assessment (LCA) study by [47] on future likely least impactful MSWM systems for ULAs in and around Harare showed that incorporating the AD of organic MSW fraction generated in ULAs in and around Harare for CHP generation with 20% recovery of materials from the inorganic MSW fraction and incineration of 80% of the remaining inorganic MSW fraction for CHP generation with flue gas treatment and incinerator bottom ash use as road construction material yields the least environmental impacts (acidification, eutrophication, global warming and human health) potentials. Authors in [37] observed that 75% on a weight basis of MSW generated in ULAs in and around Harare is suitable for treatment via incineration without the need for supplementary fuel due to its lower heating value of 10.1MJ per kg with a potential annual energy yield of 3.8 ×106 GJ (112 GWh per annum) which will increase the national annual electricity share from MSW and bio-fuels by over 0.9% to 2.2% from the prevailing minimum of 1.3%. The incineration of the MSW generated in ULAs in and around Harare will potentially reduce the amount of MSW earmarked for landfilling by 40% [37]. Therefore, source separation of organic MSW fraction and subjecting it to either composting or AD is likely to increase the efficiency during the inorganic MSW incineration since organic biodegradable MSW fractions have been widely reported to compromise the incineration performance. Just like small scale distributed AD and composting facilities for managing and treating the organic MSW fraction generated in ULAs in and around Harare, small scale distributed incineration facilities can be deployed to manage and treat the inorganic and combustible MSW fractions.

MSWM decentralisation has largely failed to take off in Zimbabwe due to residents' attitudes and behaviors towards MSWM and MSW disposal. Authors in [3] noted that initiatives towards CBMSWM implementations in Bulawayo, Zimbabwe's second biggest City, failed since it failed to change the attitudes and behaviors of Bulawayo residents on MWSM and MSW disposal. The authors [3] further noted absence of different MSWM initiatives undertaken by CBOs that embraced and incorporated materials recycling. This is however different with Harare where the Zimbabwe Sunshine group has a MSWM system developed based on CBMSWM and IMSWM model in selected suburbs and at the Zimbabwe Exhibition Park. Interestingly the Zimbabwe Sunshine Group has already constructed CDOPs or CCCs where recyclable MSW fractions are brought and sorted having also started the collection of MSW fractions for their reuse and recycling from households and institutions. The success story of the Zimbabwe Sunshine Group should be used to scale up such initiatives across all the ULAs in and around Harare and possibly nationally.

The coming into effect of the NISWMP in July 2014 was viewed as a welcome development as it provided for decentralisation hence MSWM decentralisation initiatives were envisaged to emerge from its provisions. The plan provides several key strategies and action points that support the decentralisation of MSWM within ULAs in Zimbabwe. Key actions include education in waste reduction, source separation of MSW, reuse of biodegradable MSW for gardening and livestock feed at household level, establishment of viable markets for biodegradable SW, production of manure from biodegradable SW composting or vermi-composting, establishment of community and industrial recycling centers, development of recycling projects at institutions, construction of AD facilities at established recycling centers and institutions, establish combined Heat and Power (CHP) generation facilities for AD derived biogas, putting in place a MSW collection system with CDOPS or CCCS and BBCs where households or institutions bring in MSW etc. as provided in Table 1. Worryingly all of these key actions as provided in the plan have not taken off over six years since the plan came into effect hence the need for urgent review and subsequent operationalization and implementation of the review findings.

The need for stakeholder inclusion and participation in all these key actions provided in the NISWMP is evidently clear. This is affirmed under goal 8 which seeks to promote cleanliness in Zimbabwe by involving the public, industry and government in efforts towards reducing, reusing and recycling all solid waste materials to manage and mitigate the SW associated public health and environmental impacts. Goal 7 further supports decentralisation as it provides for the need to educate and raise awareness amongst Zimbabwean citizens to better understand the

Table 4: Estimated Energy potential from the incineration of the non-biodegradable MSW Harare City and its surrounding ULAs

ULA	Harare	Chitungwiza	Norton	Ruwa	Epworth	Total
Non-biodegradable MSW generated (tons)	188,654.28	45,325.84	8,585.16	7,198.96	21,270.92	271,035.16
Estimated annual biogas potential (GJ per annum)	1.36E+06	3.28E+05	6.21E+04	5.21E+04	1.54E+05	1.96E+06

importance of participating in source separation, materials and resources recovery and integrated and sustainable solid waste management. There is need for the reform of the available national legislation and ULAs' bylaws to ease the flexibility for entry and inclusion of other players in MSWM as ULAs are somehow reluctant to devolve and delegate some of the MSWM functions and roles despite their glaring failures to manage the MSW. Interestingly the tenth goal of the plan seeks to review and assess current legislation and institutional arrangements to improve implementation of ISWM amongst its key actions. Under strategy C of the plan is the need to decentralize MSWM authority to the lowest appropriate level as well as strengthening the capacities of ULAs to meet their obligations with regards to decentralizing MSWM authorities.

The Presidency, the highest office in Zimbabwe, made a proclamation that declared the first Friday of each calendar month being a national cleanup day for communities and citizens to take part in cleaning their surroundings. This is a significant milestone towards changing the citizens mindsets, attitudes and behaviors on MSWM, MSW disposal and general environmental stewardship. The Presidential proclamation is a significant and welcome development; however, it has to be complimented with the construction of safe MSW disposal facilities for the MSW that would have been collected during the national clean up since it is currently being dumped at dumpsites which are already classified as human health and environmental hazards thus the national clean up become another unsustainable problem transfer from communities to dumpsites. There is need for urgent development and construction of recycling, AD and composting, incineration and landfill facilities with only the inert materials and those from which materials and energy could not be further derived from sent to landfills. Zimbabwean citizens have not yet fully embraced the national cleanup day hence the need for extensive awareness raising together with other initiatives that will attract citizens to participate with the participation not only limited to sweeping and cleaning but allowing those who could provide the needed resources and equipment given the opportunity to do so. Above all the high literacy level in Zimbabwe is believed will enhance citizens participation easing the introduction and implementation of decentralised waste management systems.

6. Conclusion and Recommendations

The study has attempted to show that numerous MSWM decentralisation opportunities in ULAs in and around Harare exist anchored on the quantities of MSW generated and its characteristics, suitability for biological and thermal treatment methods as well as the available recycling potential. This is despite the prevailing MSWM challenges which have become potential human health and environmental hazards. The NISWMP thus has noted MSWM decentralisation potential and provides for the decentralisation of MSWM systems and functions. Another high-level commitment for decentralisation is noted from the Presidential proclamation that declared the first Friday of every calendar month a national cleanup day which provide for every stakeholder's participation in MSWM in their local communities. However, reluctance to devolve MSWM functions by ULAs has been noted hence the need for legislative and institutional reforms as provided for under goal 10 of the plan. All the key actions regarding MSWM decentralisation in the plan have not been implemented with no waste management policy in sight as provided six years after the plan came into effect thus the need for urgent review and subsequent operationalization and implementation of the review findings. The lack of review on the available legislations and institutions has hindered MSWM decentralisation. Despite the available opportunities for MSWM decentralisation in Harare, there is need for environmental sustainability, economic feasibility, social acceptability and material and energy recovery potential assessment studies of the likely decentralised IMSWM and CBMSWM models. Further studies that quantify the associated environmental impacts, materials and energy recovery potentials as well as the economic feasibility of local level decentralised MSWM need to be undertaken.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Special thanks go to University of Johannesburg postdoctoral fellowship that supported this study. The authors would like to thank the Life Cycle Initiative for the 2017 Life Cycle award with SimaPro software used for LCIA.

References

[1] A. M. Simon, Analysis of Activities of Community Cased Organizations Involved in Solid Waste Management, Investigating Modernized Mixtures Approach: The Case of Kinondoni Municipality, Dar es Salaam, Tanzania., Masters Thesis, Wageningen University, 2008.

[2] EMA, Hazardous Waste Management Regulations; Statutory Instrument (SI) 10 of 2007, 2007

[3] N. I. Sinthumule and S. H. Mkumbuzi, "Participation in community-based solid waste management in Nkulumane suburb, Bulawayo, Zimbabwe," Resources, **8**(1), 30-46, 2019, doi.org/10.3390/resources8010030.

[4] EMA, Waste generation and management in Harare, Zimbabwe: Residential areas, commercial areas and schools, 2016.

[5] J. Okot-Okumu and R. Nyenje, "Municipal solid waste management under decentralisation in Uganda," Habitat International, **35**(4), 537-543, 2011, doi.org/10.1016/j.habitatint.2011.03.003.

[6] C. Zurbrügg, S. Drescher, A. Patel, and H. C. Sharatchandra, "Decentralised composting of urban waste - an overview of community and private

initiatives in Indian cities," Waste management, **24**(7), 655-662, 2004, doi.org/10.1016/j.wasman.2004.01.003.
[7] A. Massarutto, A. de Carli, and M. Graffi, "Material and energy recovery in integrated waste management systems: A life-cycle costing approach," Waste management, **31**(9-10), 2102-2111, 2011, doi.org/10.1016/j.wasman.2011.05.017.
[8] F. H. Mudzengerere and A. Chigwenya, "Waste Management in Bulawayo city council in Zimbabwe: In search of Sustainable waste Management in the city," Journal of Sustainable Development in Africa, **14**(1), 228-244, 2012.
[9] N. K. A. Malik, S. H. Abdullah, and L. A. Manaf, "Community participation on solid waste segregation through recycling programmes in Putrajaya," Procedia Environmental Sciences, **30**, 10-14,2015, doi.org/10.1016/j.proenv.2015.10.002.
[10] J. K. Seadon, "Integrated waste management–Looking beyond the solid waste horizon," Waste management, **26**(12), 1327-1336, 2006, doi.org/10.1016/j.wasman.2006.04.009.
[11] L. A. Guerrero, G. Maas, and W. Hogland, "Solid waste management challenges for cities in developing countries," Waste management, **33**(1), 220-232, 2013, doi.org/10.1016/j.wasman.2012.09.008.
[12] S. A. Ahmed and S. M. Ali, "People as partners: Facilitating people's participation in public–private partnerships for solid waste management," Habitat International, **30**(4), 781-796, 2006, doi.org/10.1016/j.habitatint.2005.09.004.
[13] M. Sharholy, K. Ahmad, R. C. Vaishya, and R. D. Gupta, "Municipal solid waste characteristics and management in Allahabad, India," Waste management, **27**(4), 490-496, 2007, doi.org/10.1016/j.wasman.2006.03.001.
[14] A. A. Babaei, N. Alavi, G. Goudarzi, P. Teymouri, K. Ahmadi, and M. Rafiee, "Household recycling knowledge, attitudes and practices towards solid waste management," Resources, Conservation and Recycling, **102**, 94-100, 2015, /doi.org/10.1016/j.resconrec.2015.06.014.
[15] T. Nhubu, E. Muzenda, C. Mbohwa, and B. Patel, "Impacts of waste management practices on water resources in Harare," in WASTES 2019, Lisbon, Portugal, 1-6, 2019.
[16] A. Mubaiwa, Community Based Waste Management in Urban Areas, 2019
[17] D. C. Wilson, "Development drivers for waste management," Waste Management & Research, **25**(3), 198-207, 2007, doi.org/10.1177/0734242X07079149.
[18] G. P. J. Dijkema, M. A. Reuter, and E. V. Verhoef, "A new paradigm for waste management," Waste management, **20**(8), 633-638, 2000, doi.org/10.1016/S0956-053X(00)00052-0.
[19] J. T. Tukahirwa, A. P. Mol, and P. Oosterveer, "Civil society participation in urban sanitation and solid waste management in Uganda," Local envrironment, **15**(1), 1-14, 2010, doi.org/10.1080/13549830903406032.
[20] EMA, Zimbabwe's integrated solid waste management plan, 2014
[21] T. Nhubu, E. Muzenda, M. Belaid, and C. Mbohwa, "Implementation status assessment of the Zimbabwe national integrated solid waste management plan," in ICEESEN2020, 1-6, 2020.
[22] T. Sekito, T. B. Prayogo, Y. Dote, T. Yoshitake, and I. Bagus, "Influence of a community-based waste management system on people's behavior and waste reduction," Resources, Conservation and Recycling, **72**, 84-90, 2013, doi.org/10.1016/j.resconrec.2013.01.001.
[23] L. Xiao, G. Zhang, Y. Zhu, and T. Lin, "Promoting public participation in household waste management: A survey based method and case study in Xiamen city, China," Journal of Cleaner Production, **144**, 313-322, 2017, doi.org/10.1016/j.jclepro.2017.01.022.
[24] Y. A. Rigasa, A. G. Badamasi, N. Galadimawa, and G. U. Abubakar, "Community based solid waste management strategy: A case study of Kaduna metropolis," WIT Transactions on Ecology and the Environment, **210**, 761-772, 2017, doi.org/ 10.2495/SDP160641.
[25] M. Colon and B. Fawcett, "Community-based household waste management: Lessons learnt from EXNORA's zero waste management scheme in two South Indian cities," Habitat International, **30**(4), 916-931, 2006, doi.org/10.1016/j.habitatint.2005.04.006.
[26] J. Mongkolnchaiarunya, "Promoting a community-based solid-waste management initiative in local government: Yala municipality, Thailand," Habitat International, **29**(1), 27-40, 2005, doi.org/10.1016/S0197-3975(03)00060-2.
[27] W. Nchito and G. A. Myers, "Four caveats for participatory solid waste management in Lusaka, Zambia," Urban Forum, **15**, 109-133, 2004.
[28] D. Chikobvu and F. Makarati, "The challenges of solid waste disposal in rapidly urbanizing cities: A case of highfield suburb in Harare, Zimbabwe," Journal of Sustainable Development in Africa, **13**(7), 184-199, 2011.
[29] F. McDougall, P. R. White, M. Franke, and P. Hindle, Integrated Solid Waste Management: A Lifecycle Inventory, Blackwell Science, 2001.
[30] B. Thomas and F. McDougall, "International expert group on life cycle assessment for integrated waste management," Journal of Cleaner Production, **13**(3), 321-326, 2005, doi.org/10.1016/j.jclepro.2004.02.021.
[31] N. Kollikkathara, H. Feng, and E. Stern, "A purview of waste management evolution: special emphasis on USA," Waste management, **29**(2), 974-985, 2009, doi.org/10.1016/j.wasman.2008.06.032.
[32] A. van de Klundert and J. Anschutz, Integrated sustainable waste management – the concept: tools for decision-makers. Experiences from the Urban Waste Expertise Programme (1995–2001), 2001.
[33] R. K. Henry, Z. Yongsheng, and D. Jun, "Municipal solid waste management challenges in developing countries – Kenyan case study," Waste management, **26**(1), 92-100, 2006, doi.org/10.1016/j.wasman.2005.03.007.
[34] A. J. Morrissey and J. Browne, "Waste management models and their application to sustainable waste management," Waste management, **24**(3),297-308, 2004, doi.org/10.1016/j.wasman.2003.09.005.
[35] T. Nhubu and E. Muzenda, "Determination of the Least Impactful Municipal Solid Waste Management Option in Harare, Zimbabwe," Processes, **7**(11), 785, 2019, doi.org/10.3390/pr7110785.
[36] T. Nhubu, E. Muzenda, C. Mbohwa, and E. Agbenyeku, "Comparative assessment of compositing and anaerobic digestion of municipal biodegradable waste in Harare, Zimbabwe," Environmental Progress & Sustainable Energy, **39**(4), e13376, 2020, doi.org/10.1002/ep.13376.
[37] T. Nhubu, E. Muzenda, M. Belaid, and C. Mbohwa, Biogas potential from biomethanisation of biodegradable municipal solid waste generated in Harare, *in Handbook of solid waste management,* C. Baskar, S. Ramakrishna, S. Baskar, R. Sharma, and A. Chinnappan, Eds.: Springer, In Press.
[38] L. Makarichi, R. Kan, W. Jutidamrongphan, and K. A. Techato, "Suitability of municipal solid waste in African cities for thermochemical waste-to-energy conversion: The case of Harare Metropolitan City, Zimbabwe.," Waste Management & Research, **37**(1), 83-94, 2019, doi.org/10.1177/0734242X18804029 .
[39] T. Al Seadi, B. Drosg, W. Fuchs, D. Rutz, and R. Janssen, Biogas digestate quality and utilization, in *The biogas handbook*: Woodhead Publishing, 2013.
[40] S. Sironi, L. Capelli, P. Céntola, R. Del Rosso, and M. Il Grande, "Continuous monitoring of odours from a composting plant using electronic noses," Waste management, **27**(3), 389-397, 2007, doi.org/10.1016/j.wasman.2006.01.029 .
[41] J. L. Domingo and M. Nadal, "Domestic waste composting facilities: a review of human health risks," Environment International, **35**(2), 382-389, 2009, doi.org/10.1016/j.envint.2008.07.004.
[42] L. Giusti, "A review of waste management practices and their impact on human health," Waste management, **29**(8), 2227-2239, 2009, doi.org/10.1016/j.wasman.2009.03.028.
[43] EU Commission, Green Paper on the management of bio-waste in the European Union, 2008, .
[44] CalRecyle. Current Anaerobic Digestion Technologies Used for Treatment of Municipal Organic Solid Waste, 2008.
[45] R. Dereli, K, M. Ersahin, E, C. Gomec, Y, I. Ozturk, and O. Ozdemir, "Codigestion of the organic fraction of municipal solid waste with primary sludge at a municipal wastewater treatment plant in Turkey," Waste management, **28**(5), 404-410, 2010, doi.org/10.1177/0734242X09338227.
[46] J. C. Kabouris, U. Tezel, S. G. Pavlostathis, M. Engelmann, A. C. Todd, and R. A. Gillette, "The anaerobic biodegradability of municipal sludge and fat, oil, and grease at mesophilic conditions," Water Environment Research, **80**(3), 212-221, 2008, doi.org/10.2175/106143007X220699.
[47] T. Nhubu, E. Muzenda, C. Mbohwa, and E. Agbenyeku, Comparative environmental LCA of thermal and anaerobic treatment options for municipal solid waste in Zimbabwe, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 4, 700-709 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Deep Learning Approach for Automatic Topic Classification in an Online Submission System

Tran Thanh Dien, Nguyen Thanh-Hai, Nguyen Thai-Nghe*

College of Information and Communication Technology, Can Tho University, Can Tho city, 900000, Vietnam

ARTICLE INFO

Article history:
Received: 8 July, 2020
Accepted: 11 August, 2020
Online: 28 August, 2020

Keywords:
Article classification
Deep learning
Knowledge search
Online submission system

ABSTRACT

Topic classification is a crucial task where knowledge categories exist within hierarchical information systems designed to facilitate knowledge search and discovery. An application of topic classification is article (e.g., journal/conference paper) classification which is very useful for online submission systems. In fact, numerous online journals/magazine submission systems usually receive thousands of article submissions or even more for each month. This leads to a huge amount of time-consumption of editors to process and categorize the submissions aiming to look for and assign appropriate reviewers to the submitted articles. In this study, we propose an approach based on natural language processing techniques and machine learning algorithms (both classic machine learning and deep learning) to automatic classify the topics of articles in an online submission system. We show by promising performance collected from prediction tasks to present that the proposed method is a potential approach for applying to the real system.

1 Introduction

With the rapid development of data sources and computational algorithms, the classification tasks, especially in text classification, have revealed an important role in numerous fields [1]. Text classification is a supervised learning technique that has been deploying popularly in practical cases [2]. More specific, text classification tasks are classical text processing problems attempting to categorize a unseen text into a group of known texts based on its similarities to the considered group [3]. The authors in [4] stated that text classification tasks are the assignment or categorization of labels on a new text based on the similarities of the considered text to the labeled texts in the training set. The framework using text classification automatically allows information to be processed and searched easier. Furthermore, sorting each of them takes a lot of time and effort with a large number of texts, not to mention the possibility of inaccurate categorization due to the subjectivity of the people. From the proposed studies, there are numerous real applications of text classification tasks including news classification by topics in online newspapers, knowledge management, spam email filtering, and supportable tools for search engines on the Internet, etc. [1].

The text classification tasks are attracting numerous scientists with a vast of the studies proposed with different algorithms including machine learning as well as mathematical-statistical model. In recent years, machine learning has been widely implemented and investigated with numerous advancements. Many learning algorithms include k-nearest neighbors (kNN), Naïve Bayes, support vector machines (SVM), decision tree, and artificial neural network, etc. which are leveraged to solve text classification problem with the published studies in [5–12].

An important application of text classification is article/journal classification where not only the authors but also the editorial boards of the magazines/journals would like to know how to classify a document into a relevant topic of those magazine/journal to search or/and submit. More specifically, the submission system enables us to disseminate texts and extract relevant information automatically when a manuscript is submitted to the system. From the predicted category, editors can find appropriate reviewers for the submission faster and help to speed up the review process.

This paper is an extension of work originally presented in the International Conference on Advanced COMPuting and Applications (ACOMP) 2019 in Nha Trang, Vietnam [13]. We present an approach using machine learning algorithms to classify automatically the articles which were submitted via an online submission system. More specifically, when we submit an article (the extension can be doc(x) or pdf, etc.) to the online submission system, the system not only extract automatically the information on the author, title, and abstract but stratify and assign the topic to the submission. For this procedure, we can use the natural language processing techniques to pre-process the data before fetching them into a machine learning

*Corresponding Author: Nguyen Thai-Nghe, Can Tho city, Vietnam, Contact: +84918028402 & email: ntnghe@cit.ctu.edu.vn

algorithm to do the classification tasks. Comparing to the work in [13], this study provides contributions as follows:

- After pre-processing the data, we proposed to use Deep Learning (MLP) for the classification tasks. Experimental results show that by using this approach, the results even get better than using the SVM in our previous work.
- For experiments, we have collected 5 data sets (instead of 2 Vietnamese data sets as in [13]). These data sets represent for 3 languages (English, Turkey, and Vietnamese) and have multi-classes (2, 4, 6, 9, and 10 classes as presented in Table 1)
- In this work, since the topic classes are imbalanced, we have used AUC (Area Under the ROC Curve) as a metric for comparison. Previous works show that when the data sets are imbalanced, the AUC metric is a better measurement for evaluation [14, 15]. We will analyze this point in the result section.
- For faster experiments, the pre-processing data steps are combined to a single procedure as presented in Algorithm 1.
- We have reviewed more and up-to-date related works.

In the next sections of the paper, we present a literature review on robust machine learning algorithms used in text classification tasks in Section 2. In Section 3, we introduce the proposed framework including classification models and steps for pre-processing. Section 4 describes the empirical results and Section 5 provides insightful remarks of the study.

2 Related works

In the context of the enormous development of digital information in recent years, text mining techniques hold a crucial role in information and knowledge management and mining, attracting the attention of scientists [1]. Text classification through modern techniques is the division of a dataset including documents into two or more topics. The text classification purposes to assign a predefined label or category to a document. For example, a new article published on an online newswire system can be assigned to one of the given topics while each submission sent to an online journal system can be automatically stratified into its topics, etc.

2.1 Related works on text classification

Numerous papers have proposed methods and learning architectures on text classification to enhance performance and apply in practical cases. For example, authors in [16] employed the Chi-square feature selection (referred to, hereafter, as ImpCHI) to enhance the classification performance for Arabic documents classification. They also compared this improved chi-square with three traditional features selection metrics namely mutual information, information gain and Chi-square. Another study in [17] introduced a new firefly algorithm based feature selection method which achieved a precision value equals to 0.994 on an Open Source Arabic Corpora (OSAC) dataset.

Some techniques for data pre-processing phase have been introduced in numerous studies. Authors in [18] presented the maximum matching segmentation (MMSEG) algorithm to segment words. When the segmentation completed, the text was fetched into a vector form, with the vectorized TF*IDF. Next, data were classified using decision trees and Support Vector Machines with the Weka package. The experimental results of [18] were evaluated on the dataset with 7,842 texts categorized in 10 different topics. About 500 texts of each topic were selected randomly for training phase and the remaining were used to verify independence. As reported in the paper, the performance with Support Vector Machines algorithm was greater than the algorithm of decision tree. Besides, the authors deployed singular value decomposition to analyze and reduce the characteristic space dimension and hence, it can improve classification performance of Support Vector Machines algorithm.

Another work [19] studied the semantic relation extraction and classification in scientific paper abstracts. The authors presented the steps of setup procedures and experimental results of semantic relation extraction and classification on scientific papers datasets. The task included three sub-tasks: the first classification was performed on the clean data while the second one was on noisy data, and the final task combined extraction and classification scenario. Some datasets which were used in the challenges such as a subset of abstracts of published papers in the ACL Anthology Reference Corpus, annotated for domain specificentities and semantic relations were introduced in this study. Root Cause Analysis of Incidents using Text Clustering was proposed in [20]. The authors studied the use of two machine learning (ML) algorithms, namely random forest (RF), and support vector machine (SVM) and found that SVM performed best in classifying the accident narratives. Multinomial Naive Bayes (MNB), Logistic Regression (LR), Support Vector Machines were also deployed in [21] to prediction task on Twitter Data.

Approaches based on deep learning techniques are also carried out in numerous studies. The authors in [22] presented A Comprehensive Review on deep learning techniques for text classification. The authors in [23] proposed three fundamental architectures of deep learning models for the tasks of text classification: Deep Belief Neural (DBN), “Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). The work also introduced basic guidance about the deep learning models that which models is best for the task of text classification. The authors [24] presented a novel angle to further improve this representation learning, i.e., feature projection, and projected existing features into the orthogonal space of the common features. Another paper is [25] which illustrated a model of statistics like TF-IDF, to exploit pre-trained SOTA DL models (such as the Universal Sentence Encoder) without any need for traditional transfer learning or any other expensive training procedure on Text Classification tasks of UNGA Resolutions. The authors in [26] introduced a model with Gated recurrent unit (GRU) and support vector machine. The method implemented a linear support vector machine (SVM) as the replacement of Softmax in the final output layer of a GRU model to obtain comparative, remarkable results. The work in [27] introduced a new improved algorithm of the original Sine Cosine Algorithm for feature selection with the software written in Matlab2013 to achieve high performance on the datasets of Reuters-21578 collection [1]. Some studies [28, 29] have

[1]D.D., 2004. Reuters-21578. Retrieved November 6, 2013, http://www.daviddlewis.com/0Aresources/testcollections/reuters21578/

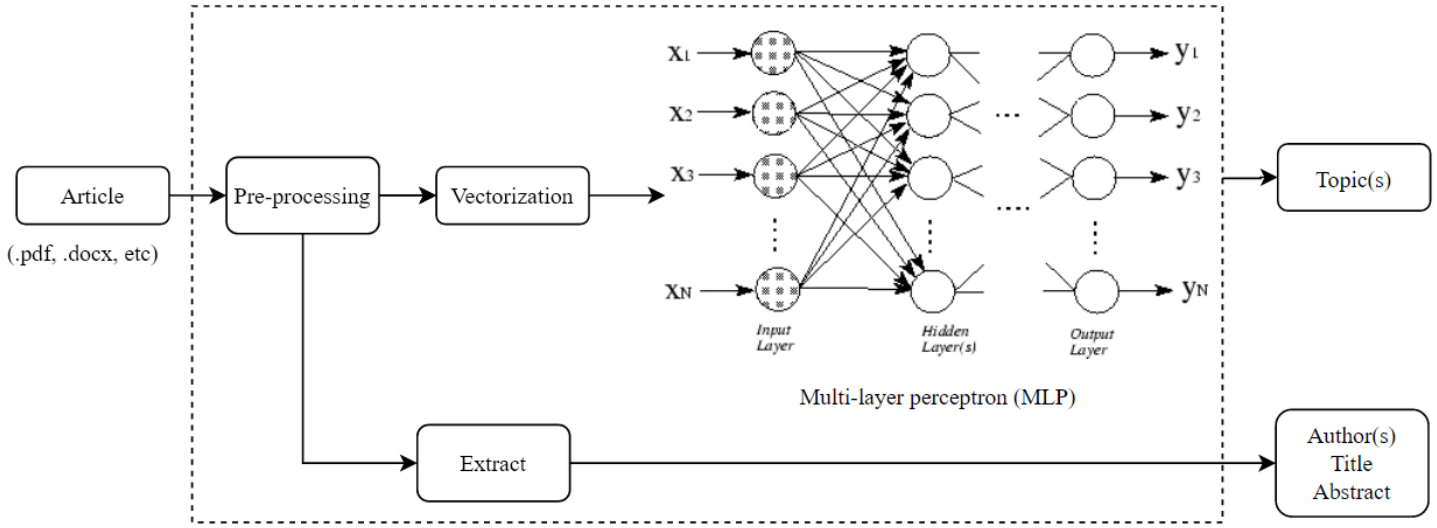


Figure 1: The proposed architecture for pre-processing and classification of articles

attempted to use Reinforcement Learning for text classification.

2.2 Text vectorization

In real word, some text representation models have been proposed including vector space model, bag of words model, and graph-based model. The vector space model [30] is deployed in this work because it can represent unformatted documents as simple and formulaic notation. Due to its advantages, many researches based on vector space model are implemented [31]. According to the details of this method, each document or article is represented as a vector while each component of such vector is a separated term that is assigned to a value namely "weight" of that term.

TF measures the frequency of a word in a document. It depends on the document length and the generality of word. For measuring the weight of a word, the number of occurrences of the word is divided by the length of the document (the number of words) as equation 1:

$$TF(t,d) = \frac{\text{number of occurrences of term } t \text{ in d}}{\text{Total number of terms in d}} \quad (1)$$

There is difference between TF and IDF. TF counts frequency of a term t in document d, where as DF counts occurrences of term t in the document set N. For calculating the TF, all the terms reveal the same importance. However, it is found that not all terms in a dataset are important, for instance, connecting terms, determiners; and prepositions. It is necessary to reduce the importance of such terms with computing IDF by the formula as Equation 2:

$$IDF(t,D) = log\frac{\text{Total document in D}}{\text{Number of document including } t} \quad (2)$$

We compute TF*IDF which integrate between TF and IDF. The method calculates the TF*IDF value of a term via its importance in a document belonging to a document set. Using such method, we are able to filter out common words and retain high value words as equation 3.

$$TF*IDF(t,d,D) = TF(t,d) \times IDF(t,D) \quad (3)$$

In this study, we propose an approach of automated classification of articles submitted via an online submission system. When the submission (with extension such as *.doc(x), *.pdf, etc.) is uploaded, the system, then, extracts the author's information and abstract to stratify the submission.

3 Proposed method

The proposed overall system including extracting information, pre-processing data methods and categorizing articles is exhibited in Fig. 1.

In proposed model, when a new article under formatting of .docx, .pdf, etc. is submitted to the system, its information consisting of abstract, author(s), title can be automatically extracted, and especially categorized into an appropriate topic based on the previous data trained by machine learning models. Due to article's pre-formatted template, article's information extraction is easy to get. Therefore, this work only focuses on how to classify article's topics when it is submitted to the system.

In the following sections, we will present how to pre-process the data and setting up the classification models.

3.1 Data pre-processing

Data pre-processing described as Algorithm 1. This algorithm receives a document as input, then the document is converted and normalized with techniques such as changing to lower cases, removing blanks, etc. The algorithm also separates documents into words and eliminate noises then transform to the vector. The details are in Algorithm 1.

File format conversion and word standardization: If the documents of datasets have the extension of .doc(x), they will be converted to plain text (.txt) for easy use in most of classification mod-

els. After converting file format, word standardization performs to convert all text characters into lowercase while spaces are also eliminated.

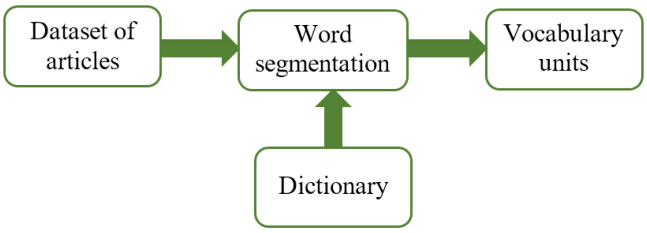


Figure 2: Process of word segmentation

Algorithm 1: Pre-processing for documents

Data: InputDocument d
Result: Vectorized-Documents D
1 $Convertion(d)$: convert the input document (word/pdf) to text
2 $WordNormalization(d)$: changed to lower cases, removing blanks
3 $WordSegmention(d)$: separate document to words
4 $RemovingStopWords(d)$: remove noise words
5 $Vectorization(d)$: convert documents to respectively vectors
6 **Return** sets of Vectorized-Documents D

Word segmentation: In some languages including Vietnamese, spaces do not segment words but separate syllables so the segmentation is an important phase in NLP. There are a lot of tools that have been successfully developed to segment words with relatively high accuracy. For instance, the *VnTokenizer* segmentation tool [2] was used for Vietnamese word segmentation. It was based on the integrated methods of maximum matching, weighted finite-state transducer and regular expression parsing, running on the dataset of Vietnamese syllabary and Vietnamese vocabulary dictionary. This tool segments Vietnamese documents into vocabulary units (words, names, numbers, dates and other regular expressions) with over 95% accuracy described in Fig. 2

Removing stop words: As mentioned in [32] that stop words are the words that widely appear in all documents of the considered dataset , or the words that appear only in one and several documents. Therefore, they do not contain useful information or make sense. For tasks of text classification, the appearance of such types of words not only do not help examine the classification but also cause noises leading to a decrease in performance of the classification process.

Text vectorization In this study, we use TF-IDF (term frequency-inverse document frequency) as a statistical measure that evaluates how relevant a term is to a document in a collection of documents.

Text classification algorithms From a set of documents $\{d_1 ... d_n\}$ called a training set, in which document d_i is labeled under c_j belonging to set of categories C = $\{c_1 ... c_m\}$, classification model is determined for classifying any document d_k into an appropriate category of set C.

In our experiments, some text classification algorithms which are implemented for comparison include SVM algorithms, tree decision and deep learning techniques.

3.2 Deep learning model

The proposed model is presented in Fig. 1 where the input attributes are selected from Table 1 and the output (prediction) of the model including classes depending the selected dataset. The proposed MultiLayer Perceptron (MLP) architecture includes one hidden layer with 16 neurons (see an illustration in Figure 3) which is conducted from fine-tune hyperparameters experiments (see Section 4.2) running n hidden layer(s) with various (m) numbers of neurons.

In order to investigate the difference in the performance, starting at one hidden layer, we increase the number of neurons from 2 to 128. Each multiplies two times the preceding one (2^n with $n = 1..7$). When we obtain the best number of neurons, let say k, for example, we begin to increase the number of hidden layers from 2 to 5 with k neurons for each hidden layer to observe the changes in the prediction results. The experiments for hyperparameters search are done on Scientific_Articles dataset. As exhibited from Figure 3, the network receives 3431 attributes of Scientific_Articles dataset as the input following by a hidden layer including 16 neurons and produces 9 outputs corresponding to the predicted probabilities of 9 topics for the classification. After selecting hyperparameters from the experiments, we keep the number of neurons and one hidden layer for prediction tasks on five datasets while the number of nodes of the input layer and output layer can be vary depending on the considered dataset.

The MLP models which perform the binary classification are implemented Sigmoid function to do prediction tasks. The Sigmoid function [33] usually appears in the output layers of Deep learning architectures. It transforms the input values which lie in the domain $\mathbb{R}$ to outputs have the domain in [0,1]. The Sigmoid function is also called “squashing" because this function squashes any input in the range of (-inf,-inf) to the range of [0,1]. When we shifted to gradient-based learning, the Sigmoid was considered as a natural selection due to its smooth and differentiable approximation to a thresholding unit. The Sigmoid function is given by the formula:

$$Sigmoid(x) = \frac{1}{1+e^{-x}} \tag{4}$$

where, x denotes data after being computed by the preceded neural layer.

For multi-classification problems, we use softmax function (Equation 5) with k classes. The Softmax function normalizes an input value into a vector of values that follows a probability distribution whose total sums up to 1.

$$Softmax(x_i) = \frac{e^{x_i}}{\sum_j^k e^{x_j}} \tag{5}$$

The activation function namely, ReLU [34], is also implemented in our architecture. ReLU follows the formula:

$$f(x) = max(0,x) \tag{6}$$

where, x denotes data after being processed by the preceded neural layer.

[2]N. T. M. Huyen, V. X. Luong and L. H. Phuong, “VnTokenizer”, 2010. http://vntokenizer.sourceforge.net/

Table 1: Five considered datasets descriptions

Data set	#Instances	#Attributes	#Classes	Language
Reuters_Corn	2,158	1,503	2	English
School_Text_Books	1,786	2,566	4	English
Turkish_News_Articles	3,600	5,693	6	Turkish
Scientific_Articles	650	3,431	9	Vietnamese
VnExpress_Newsletters	10,000	3,266	10	Vietnamese

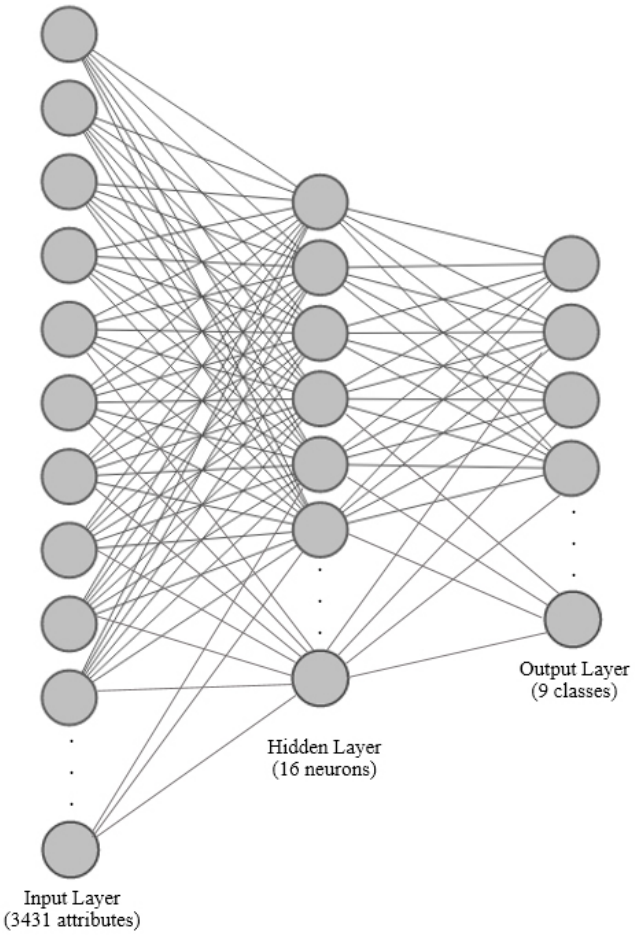


Figure 3: The proposed MLP architecture conducted from tune parameters experiments on Scientific_Articles dataset

ReLU is the most widely used activation function for deep learning architectures with state-of-the-art results to date. ReLU helps models to produce better performance and generalization in deep learning compared to the Sigmoid and Tanh activation functions. It represents a nearly linear function, so this activation function preserves the properties of linear models that made them easy to optimize, with the gradient-descent method [35, 36].

ReLu holds a role as an activation function to transform the output of the preceded hidden neural layer. ReLU helps to improve neural networks by speeding up training. Gradients of logistic and hyperbolic tangent models are lower than the positive portion of the ReLU. That means the positive portion is updated more rapidly as training progresses.

In order to reduce overfitting issues, we consider implementing Early Stopping technique with a patience epoch of 5. This means that the loss cannot be improved after 5 consecutive epochs, the learning will be stopped. Otherwise, the learning will be continued to run to 10 epochs. The network is implemented with Adam optimizer function, use a batch size of 100 and a default learning rate of 0.001.

4 Experimental results

4.1 Data Description

This study used five experimental datasets in three various languages (including English, Turkish, and Vietnamese) as described in Table 1. The reuters corn is available at UCI repository[3] for binary classification tasks. The School text books of 11^{th} and 12^{th} grade which is available at Kaggle Website[4] with four topics. A collection of Turkish news and articles dataset can be downloadable at UCI repository [5] including 3600 samples on 6 categories. The Scientific articles of a university and VnExpress Newsletters in Vietnam were used in our previous work at ACOMP 2019 [13] include 650 samples, 3431 features and 10000 samples, 3266 features, respectively. The considered numbers of classes also vary from binary classification to 10-class classification. The largest dataset is VnExpress_Newsletters with 10000 samples on 10 different topics.

We also face imbalanced datasets issues where the number of samples of some classes is much more than other classes. For example, a class of Reuters_corn dataset occupies to 97% while only 3% is for the other.

The performances of classifiers are examined by average AUC on 3-fold cross-validation. The folds are the same for all classifiers, i.e. training and test sets were identical for each classifier. AUC is a reliable metric for evaluating classifiers where data are not balanced. AUC is widely used in numerous studies to examine the performance of prediction tasks and it is reliable metric to measure the performance of prediction.

4.2 Hyper-parameter turning

In order to select appropriate parameters for the MLP models, we run the experiments with various configurations of MLP architec-

[3] https://storm.cis.fordham.edu/ gweiss/data-mining/datasets.html
[4] https://www.kaggle.com/deepak711/4-subject-data-text-classification
[5]https://archive.ics.uci.edu/ml/datasets/TTC-3600:%20Benchmark%20dataset%20for%20Turkish %20text%20categorization

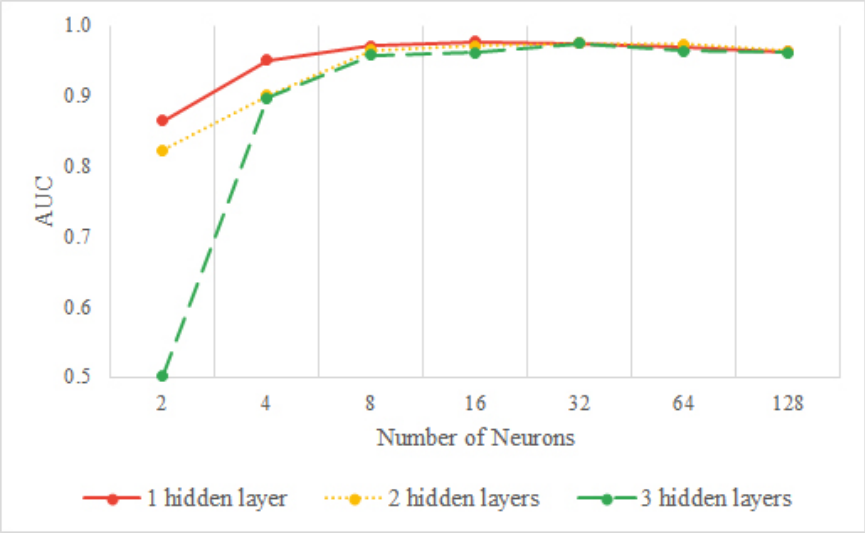


Figure 4: Hyper-parameter search for the Number of Hidden Layers

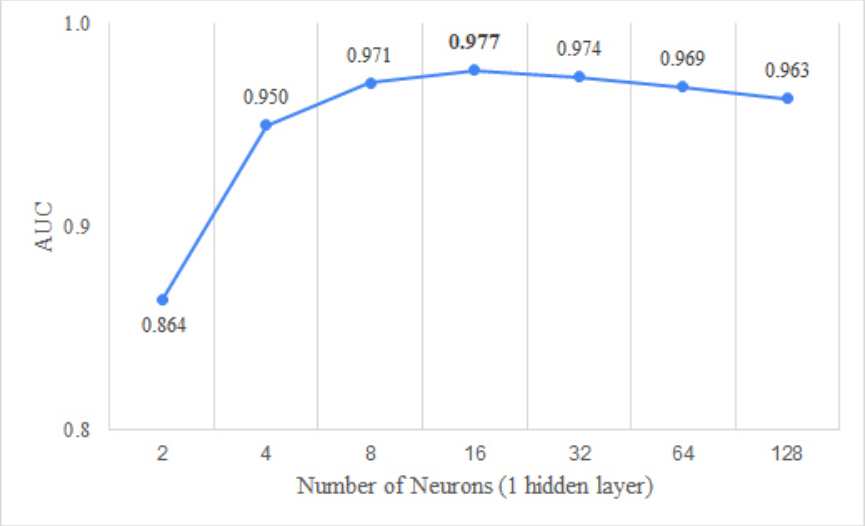


Figure 5: Hyper-parameter search for the Number of Neurons

tures.

In Figure 4 and Figure 5, we reveal the performance of different configurations of MLP on Scientific_Articles dataset. The results show the performance is rising according to the width of the MLP. However, the performances reach peaks which are vary depending on the number of hidden layers used. After reaching the peak, the performance tends to go down when we keep increasing the number of neurons. As seen from Figure 4, with one hidden layer, we obtain the best with 16 neurons while using two or three hidden layers achieve the highest performance with 32 neurons per hidden layer. We noted that the performance decreases when we add more layers due to overfitting, but these differences are trivial when the number of neurons per hidden layer is high.

In most of the cases, the architectures of MLP with one neural layer obtain greater performance than the architectures with more hidden layers (see details performance of the number of neurons with one hidden layer in Figure 5). However, with a large number of neurons, we can see the performance of various numbers of neural layers are nearly the same. Our results exhibit that the shallow architecture of MLP including one hidden layer with 16 neurons reaches the best performance. The number of hidden layers and the number of neurons conducted from the experiments in hyperparameter tuning, we implement an MLP architecture with one hidden layer including 16 neurons to run the learning and validation on the other 4 datasets.

The experiments of five considered datasets are presented and compared with various machine learning methods in Section 4.3.

4.3 Topic classification Results of various machine learning algorithms on five considered datasets

Previous work [13] we showed that SVM works very well for automatic topic classification in an online submission system, however, in this work we have continued to improve and showed that using Deep Learning approach the results even better. Since the data sets are imbalanced, we report the AUC instead of the accuracy metric as

Table 2: Performance Comparison in AUC (Area under the ROC Curve) of various machine learning approaches on five considered datasets

Data set	**Classifier**	**AUC**
Reuters_Corn	MLP	**0.991**
	SVM	0.811
	Decision Tree	0.813
School_Text_Books	MLP	**0.999**
	SVM	0.991
	Decision Tree	0.928
Turkish_News_Articles	MLP	**0.962**
	SVM	0.949
	Decision Tree	0.871
Scientific_Articles	MLP	**0.977**
	SVM	0.965
	Decision Tree	0.819
VnExpress_Newsletters	MLP	**0.990**
	SVM	0.985
	Decision Tree	0.876

reported in Table 2. We have also used Decision Tree as a baseline for comparison.

Table 2 reveals the topic classification performances in AUC of three different machine learning algorithms on five considered datasets. It is easy to see that MLP outperforms other algorithms. In most cases, Decision Tree algorithm gives the worst result while SVM holds the second place. The classification performances of MLP are promising results which all achieve over 0.960 in AUC. Three datasets of them reach over 0.990 while the article classification in Turkish reveals the lowest performance but this result is still high with an AUC of 0.962. An example of visualization for the AUC and precision-recall are presented in Figures 6 and 7, other datasets are quite similar.

We also present the results in confusion matrices of School_Text_Books dataset to observe how different in the performances between MLP and SVM are in Table 3.

As shown from the results above, MLP outperforms SVM and we might like to know how different between them. We can see the difference is that SVM performs worse than MLP on the class where owns the minimum number of samples among the considered classes (these numbers are formatted with blue and bold, revealed in Table 3). With 98 samples for the class of “geography", this number is compared to other classes to see that we face imbalanced issues in data. However, MLP achieves a promising classification result comparing to SVM on the class with much fewer samples. Similar results are also revealed in other datasets. This is expected to bring to a reliable result in practical cases where we usually meet imbalanced dataset issues.

Binary classification tasks on Reuters_Corn dataset reveal the same results where Deep learning approach also reaches a better performance on the class with fewer samples (see the results in Table 4). The Class indicates whether the entry is related to corn ($b = 1$) or not ($a = 0$). As observed from the results, we collected fewer samples which related to “corn" so we also face imbalanced issue in the dataset. In this case, the class of 1 only occupies 3% compared to 97% samples belonging to the class of 0. The same result with multi-classes classification tasks, we obtain better performance for the binary classification tasks with MLP on the class with fewer samples (these numbers are formatted with blue and bold, revealed in Table 4) when we compare to SVM algorithm.

Experimental results in this work showed that the MLP is more suitable than the SVM in case of imbalanced data where the minority class is more interesting to predict. This is also the reason why we have selected the AUC as a measure instead of the Accuracy [14,15].

For the training times, the MLP was completed the training stage in a couple of minutes for the datasets using in this study, so it does not a matter for an online system where we can automatically set-up a training schedule after a time interval (e.g., the model can be automatically updated after one day or other intervals depending on the real number of submissions).

5 Conclusion

Leveraging techniques of natural language processing and machine learning algorithms, we presented a solution to the automated classification of articles to support authors/editors saving their efforts and time for processing articles on the system. Data pre-processing techniques with steps introduced in this work are significantly improved to make the dataset in a standardized format for learning with the three considered algorithms of Multilayer Perceptron, Support Vector Machines and Decision Tree. The experiments are done on five various datasets. The data used vary in the number of features, attributes as well as the number of classes. All datasets reach the performance of over 0.960 in AUC with deep learning models.

As shown from experimental results, deep learning algorithm

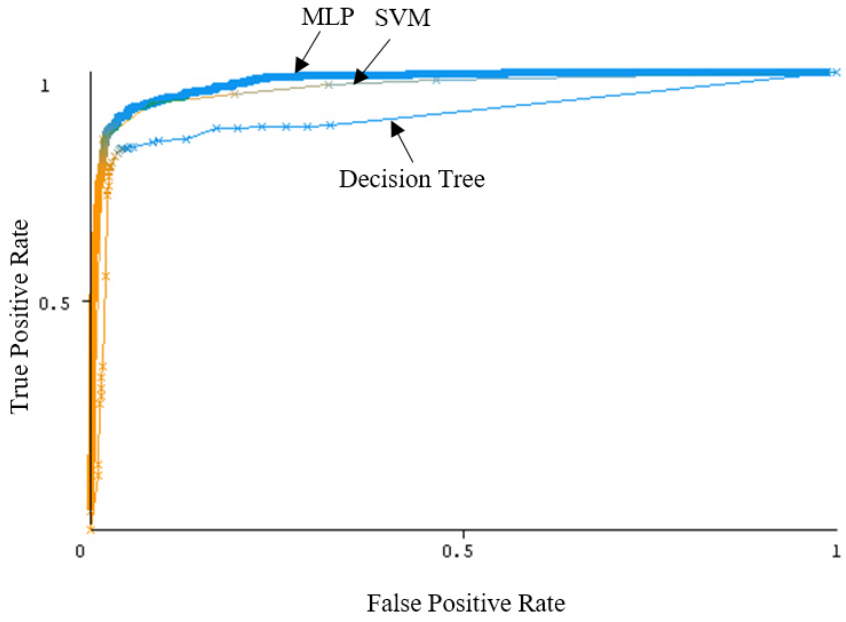


Figure 6: The Area under the ROC (AUC) for Turkish_News_Articles dataset

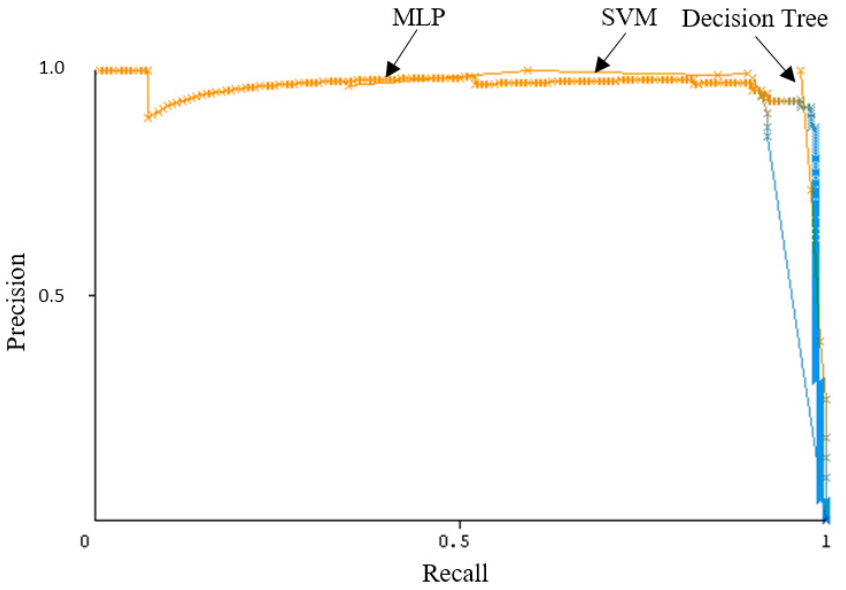


Figure 7: The Precision-Recall for Turkish_News_Articles dataset

Table 3: Confusion matrix comparison of predictions between Multilayer Perceptron and Support Vector Machines on School_Text_Books dataset

	Deep learning					**SVM**			
	Predicted classes					Predicted classes			
Actual classes	a	b	c	d		a	b	c	d
	283	0	0	1	a = accounts	281	1	0	2
	0	626	7	2	b = biology	0	632	3	0
	0	8	**87**	3	c = geography	0	16	**78**	4
	0	4	3	762	d = physics	0	4	1	764

with Multilayer perceptron exhibits better classification performance than the classic machine learning such as Support Vector Machines.

Table 4: Confusion matrix comparison of predictions between Multilayer Perceptron and Support Vector Machines on Reuters_Corn dataset with binary classification

	Deep learning			**SVM**	
	Predicted classes			Predicted classes	
Actual classes	a	b		a	b
	2064	25	a = 0	2088	1
	13	**56**	b = 1	26	**43**

Some parameters are also evaluated to reach promising results in classification tasks. The results show that the proposed model is feasible to extract information and stratify articles automatically whenever a document is submitted to the system. We continue to find a solution for larger datasets in further research.

The proposed architecture of Multilayer perceptron is rather shallow with one neural layer. A vast of hyper-parameters of MLP are evaluated and we see that the performance tends to be saturated when we increase both the number of hidden layers as well as the number of neurons. Further studies should take into sophisticated architectures to improve performance in document categorization tasks.

Conflict of Interest The authors declare no conflict of interest.

References

[1] K. Thaoroijam, "A Study on Document Classification using Machine Learning Techniques," International Journal of Computer Science Issues, **11**(1), 217–222, 2014.

[2] Y. Li, L. Zhang, Y. Xu, Y. Yao, R. Y. K. Lau, Y. Wu, "Enhancing Binary Classification by Modeling Uncertain Boundary in Three-Way Decisions," IEEE Transactions on Knowledge and Data Engineering, **29**(7), 1438–1451, 2017, doi:https://doi.org/10.1109/TKDE.2017.2681671.

[3] F. Sebastiani, "Machine Learning in Automated Text Categorization," **34**(1), 1–47, 2002, doi:https://doi.org/10.1145/505282.505283.

[4] Y. Yang, J. O. Pedersen, "A Comparative Study on Feature Selection in Text Categorization," in Proceedings of the Fourteenth International Conference on Machine Learning, ICML '97, 412–420, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997, doi:https://dl.acm.org/doi/10.5555/645526.657137.

[5] C. C. Aggarwal, C. Zhai, A Survey of Text Classification Algorithms, 163–222, Springer US, Boston, MA, 2012, doi:10.1007/978-1-4614-3223-4_6.

[6] M. A. Bijaksana, Y. Li, A. Algarni, "A Pattern Based Two-Stage Text Classifier," in P. Perner, editor, Machine Learning and Data Mining in Pattern Recognition, 169–182, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, doi:https://doi.org/10.1007/978-3-642-39712-7_13.

[7] J. Chen, H. Huang, S. Tian, Y. Qu, "Feature selection for text classification with Naïve Bayes," Expert Systems with Applications, **36**(3, Part 1), 5432 – 5435, 2009, doi:https://doi.org/10.1016/j.eswa.2008.06.054.

[8] M. Haddoud, A. Mokhtari, T. Lecroq, S. Abdeddaïm, "Combining supervised term-weighting metrics for SVM text classification with extended term representation," Knowledge and Information Systems, **49**, 909–931, 2016, doi:https://doi.org/10.1007/s10115-016-0924-1.

[9] A. Chouchoulas, Q. Shen, "A Rough Set-Based Approach to Text Classification," in N. Zhong, A. Skowron, S. Ohsuga, editors, New Directions in Rough Sets, Data Mining, and Granular-Soft Computing, 118–127, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999, doi:https://doi.org/10.1007/978-3-540-48061-7_16.

[10] A. P. Gopi, R. N. S. Jyothi, V. L. Narayana, K. S. Sandeep, "Classification of tweets data based on polarity using improved RBF kernel of SVM," International Journal of Information Technology, 2020, doi:https://doi.org/10.1007/s41870-019-00409-4.

[11] N. Thai-Nghe, Q. D. Truong, "An Approach for Building a Semi-automatic Online Consultancy System," in 2015 International Conference on Advanced Computing and Applications (ACOMP), 51–58, 2015, doi:https://doi.org/10.1109/ACOMP.2015.11.

[12] B. E. Boser, I. M. Guyon, V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," in Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, 144–152, ACM Press, 1992, doi:https://doi.org/10.1145/130385.130401.

[13] T. T. Dien, B. H. Loc, N. Thai-Nghe, "Article Classification using Natural Language Processing and Machine Learning," in 2019 International Conference on Advanced Computing and Applications (ACOMP), 78–84, 2019, doi:https://doi.org/10.1109/ACOMP.2019.00019.

[14] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, G. Bing, "Learning from class-imbalanced data: Review of methods and applications," Expert Systems with Applications, **73**, 220 – 239, 2017, doi:https://doi.org/10.1016/j.eswa.2016.12.035.

[15] N. Thai-Nghe, Z. Gantner, L. Schmidt-Thieme, "A new evaluation measure for learning from imbalanced data," in The 2011 International Joint Conference on Neural Networks, 537–542, 2011, doi:https://doi.org/10.1109/IJCNN.2011.6033267.

[16] S. Bahassine, A. Madani, M. Al-Sarem, M. Kissi, "Feature selection using an improved Chi-square for Arabic text classification," Journal of King Saud University - Computer and Information Sciences, **32**(2), 225 – 231, 2020, doi:https://doi.org/10.1016/j.jksuci.2018.05.010.

[17] S. Larabi Marie-Sainte, N. Alalyani, "Firefly Algorithm based Feature Selection for Arabic Text Classification," Journal of King Saud University - Computer and Information Sciences, **32**(3), 320–328, 2020, doi:https://doi.org/10.1016/j.jksuci.2018.06.004.

[18] C.-H. Tsai, "MMSEG: A word identification system for Mandarin Chinese text based on two variants of the maximum matching algorithm," Avaible on internet at http://www. geocities. com/hao510/mmseg, 2000.

[19] K. Gábor, D. Buscaldi, A.-K. Schumann, B. QasemiZadeh, H. Zargayouna, T. Charnois, "SemEval-2018 Task 7: Semantic Relation Extraction and Classification in Scientific Papers," in Proceedings of The 12th International Workshop on Semantic Evaluation, 679–688, Association for Computational Linguistics, New Orleans, Louisiana, 2018, doi:http://dx.doi.org/10.18653/v1/S18-1111.

[20] S. Sarkar, N. Ejaz, M. Kumar, J. Maiti, "Root Cause Analysis of Incidents Using Text Clustering and Classification Algorithms," in P. K. Singh, B. K. Panigrahi, N. K. Suryadevara, S. K. Sharma, A. P. Singh, editors, Proceedings of ICETIT 2019, 707–718, Springer International Publishing, Cham, 2020, doi:https://doi.org/10.1007/978-3-030-30577-263.

[21] P. Harjule, A. Gurjar, H. Seth, P. Thakur, "Text Classification on Twitter Data," in 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), 160–164, 2020, doi:https://doi.org/10.1109/ICETCE48199.2020.9091774.

[22] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. A. Chenaghlu, J. Gao, "Deep Learning Based Text Classification: A Comprehensive Review," ArXiv, **abs/2004.03705**, 2020.

[23] M. Zulqarnain, R. Ghazali, Y. M. M. Hassim, M. Rehan, "A comparative review on deep learning models for text classification," Indonesian Journal of Electrical Engineering and Computer Science, **19**(1), 325–335, 2020, doi:https://doi.org/10.11591/ijeecs.v19.i1.pp325-335.

[24] Q. Qin, W. Hu, B. Liu, "Feature Projection for Improved Text Classification," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 8161–8171, Association for Computational Linguistics, Online, 2020, doi:http://dx.doi.org/10.18653/v1/2020.acl-main.726.

[25] F. Sovrano, M. Palmirani, F. Vitali, "Deep Learning Based Multi-Label Text Classification of UNGA Resolutions," ArXiv, **abs/2004.03455**, 2020.

[26] M. Zulqarnain, R. Ghazali, Y. M. M. Hassim, M. Rehan, "Text classification based on gated recurrent unit combines with support vector machine," International Journal of Electrical and Computer Engineering, **10**(4), 3734–3742, 2020, doi:http://doi.org/10.11591/ijece.v10i4.pp3734-3742.

[27] M. Belazzoug, M. Touahria, F. Nouioua, M. Brahimi, "An improved sine cosine algorithm to select features for text categorization," Journal of King Saud University - Computer and Information Sciences, **32**(4), 454 – 464, 2020, doi:https://doi.org/10.1016/j.jksuci.2019.07.003.

[28] E. Lin, Q. Chen, X. Qi, "Deep reinforcement learning for imbalanced classification," Applied Intelligence, **50**(8), 2488–2502, 2020, doi:10.1007/s10489-020-01637-z.

[29] D. Chai, W. Wu, Q. Han, F. Wu, J. Li, "Description Based Text Classification with Reinforcement Learning," ArXiv, **abs/2002.03067**, 2020.

[30] C. S. Perone, "Machine learning: Cosine similarity for vector space models (Part III)," URL:http://blog.christianperone.com, 2019.

[31] J.-Y. Chang, I.-M. Kim, "Analysis and Evaluation of Current Graph-Based Text Mining Researches," Advanced Science and Technology Letter, **42**, 100–103, 2013, doi:http://dx.doi.org/10.14257/astl.2013.42.2.

[32] H. Saif, M. Fernandez, Y. He, H. Alani, "On Stopwords, Filtering and Data Sparsity for Sentiment Analysis of Twitter," in Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), 810–817, European Language Resources Association (ELRA), Reykjavik, Iceland, 2014.

[33] C. Nwankpa, W. Ijomah, A. Gachagan, S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning," ArXiv, **abs/1811.03378**, 2018.

[34] V. Nair, G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in J. Fürnkranz, T. Joachims, editors, Proceedings of the 27th International Conference on Machine Learning (ICML-10), June 21-24, 2010, Haifa, Israel, 807–814, Omnipress, 2010.

[35] M. D. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. V. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean, G. E. Hinton, "On rectified linear units for speech processing," in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 3517–3521, 2013, doi:https://doi.org/10.1109/ICASSP.2013.6638312.

[36] G. E. Dahl, T. N. Sainath, G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 8609–8613, 2013, doi:https://doi.org/10.1109/ICASSP.2013.6639346.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 744-761 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Analysis of Long-term Equilibrium Relationship Between KRW, RMB, JPY Exchange Rates and International Financial Market Variables: Comparative Analysis of KRW, RMB, JPY

Moon-Kyum Kim*, Woong Ryeol Kim

Graduate School of Soongsil University, Seoul, 06978, Republic of Korea

ARTICLE INFO

Article history:
Received: 06 October, 2020
Accepted: 31 October, 2020
Online: 24 November, 2020

Keywords:
Currency
RMB
Yuan
CNY
CNH
KRW
JPY
NDF
VECM

ABSTRACT

This study used 15 VECM analysis models to analyze the relationship among exchange rates of South Korea, China and Japan; and between exchange rates of each country and international financial market variables. The analysis variables are the Won, Yuan, Yen spot exchange rates and international financial market variables. The results of the analysis are as follows: First, there was a long-term cointegration relationship between RMB, KRW, JPY and international financial market variables. Second, the analysis results of the VECM showed that explanatory power of Korea's offshore Won-Dollar accounts for 50% of the onshore Won-Dollar. The results of the analysis of the Yuan's VECM showed that the onshore Yuan (CNY) and offshore Yuan (CNH) exchange influence, but each has its own independent characteristics. Overall, the Won is more integrated with the international financial market than the Yuan. The Yen's relationship with variables in the international financial market was stronger than that of the Won and the Yuan. Third, offshore Won-Dollar, onshore KRW, and JPY had similar Granger causality relationship and impulse responses with international financial market variables. However, CNH and CNY indicated weaker than that of Won and Yen. The onshore Won-Dollar is shown to partially offset the shock from the onshore Yuan and offshore Yuan. The implications of this study are as follows: First, it is important to look at the exchange rate from the perspective of the international financial market. Second, Yuan investment and risk management are necessary considering the characteristics of the onshore and offshore Yuan which are interrelated but also distinctly unique markets. Third, it is necessary to manage the exchange rate position, taking into account the long-term equilibrium relationship among the Won, Yuan, and Yen currencies. Fourth, Korean companies should find a way to actively utilize the Won which shows the characteristics of partial internationalization.

1. Introduction

International credit rating agencies note changes in exchange rates in relation to country credit ratings. The exchange rate represents the credit risk of the currency country viewed in the financial market. In [1], the author argued that the credit rating announcement by the credit rating agency has a significant impact on the exchange rate and its influence varies depending on emerging markets or developed countries. In [2], the author argued that the depreciation of the exchange rate is a negative factor for the financial system of the country. In [3], the author argued that large differences in interest rates between local and foreign currencies would increase dollar loans with low interest rates. In [4], the author that entities in East Asian countries use foreign currency loans to reduce borrowing costs. They argued that company selectively use foreign currency loans in consideration of low foreign currency borrowing interest rate benefits and foreign exchange hedge costs arising from foreign currency borrowing. As they argument, the mid- to long-term exchange outlook, adjusted according to interest rates, expected currency values, macro-economic variables, etc., are important for the parties to the foreign exchange transaction. Therefore, it is important to understand the mid- to long-term equilibrium relationship between each currency and with variables in each currency and international financial markets variables. However, the exchange rate is generally highly volatile in the short term, as the following arguments suggest. In [5], the author argued that the short-term exchange rate determinant model by macroeconomic variables is almost

*Corresponding Author: Moon-Kyum Kim, Professor, Graduate School of Soongsil University. mgkim@ssu.ac.kr

indescribable. In [6], the author argued that the empirical analysis of advanced countries' currencies generally showed that the exchange rate was close to random walking, and that the won-dollar exchange rate was also random walking after the introduction of the floating exchange rate system in Korea. In [7], the author argued that the sensitivity of the trade-in between the local currency and the foreign currency was greater under the floating exchange rate system. He argued that the floating exchange rate system also limits the ability of monetary authorities to carry out their own policies. The following studies have shown that emerging economies' exchange rate-setting mechanisms and dollarization increase exchange rate volatility. In [8], the author studied the factors affecting the transactions of local and foreign currencies in emerging countries. He argued that the holding of the local currency or foreign currency depends on the level of interest rate on the foreign currency against the local currency and the expected depreciation of the local currency. Therefore, it was argued that the foreign currency trade was determined in accordance with reasonable expectations for the local currency and the foreign currency. In [9], the author argued that the dollarization generally increases in emerging economies. The dollarization is the use of the dollar (foreign currency) as a means of storing value instead of the local currency. The IMF defined the dollar phenomenon in emerging economies as the proportion of foreign currency deposits to broad cash. They used the dollarization level as the ratio of foreign currency deposits to M2. In [10], the author argued that the dollarization was a reasonable response from the investors to future uncertainties in emerging economies' currencies. The dollarization also affects macroeconomic policies, foreign exchange risk management and exchange rate systems in emerging economies. In [11], the author analyzed the Turkish currency using a VAR model. He argued that the dollarization is caused by instability in three macroeconomic variables: currency devaluation volatility, inflation volatility, and uncertainty forecast volatility. In [12], the author argued that the loss of confidence in economic management would reduce confidence in the local currency, thereby promoting the retention of foreign currencies. Bank plays a pivotal role in developing emerging economies' financial markets. Therefore, deepening the dollarization of banks' assets and liabilities increases foreign currency exchange risk. He argued that the dollarization is not significant in countries where the financial industry has grown to a considerable level, but in countries with small financial industries, the dollarization is relatively large. The above studies suggest that the medium- to long-term relationship of exchange rates is important for parties to foreign exchange transactions. The demand and supply of foreign exchange trading arise from the import-export transactions of enterprises and the hedging of foreign exchange positions held and the capital markets transactions invested in stocks and bonds. Thus, long-term exchange rate relationships between currencies and key variables in the international financial market are important for businesses and investors. South-Korea, China and Japan are the top 10 global import-export trading partners. In addition, China and Japan are Korea's largest export-import trading partners. The Korean won is a non-international currency, but the Chinese yuan and Japanese yen are SDR international currencies. As a non-international currency country, like most ASEAN countries, Korea has to do import-export transactions and capital market transactions with international currency countries. Therefore, it is an important issue for Korean import-export companies and pension fund investors to find out the long-term equilibrium relationship between the won dollar, the yuan, and the yen. Regarding exchange rate relationship between Korea, China and Japan. In [13], the author analyzed the non-international currency of the Won, the SDR international currency of the Yuan, and the Yen by applying CIP theory. The CIP theory is that there is an equilibrium relationship between the difference between the risk-free interest rates of the two currencies and the difference between the spot exchange rate and the forward exchange rate. Therefore, the country's foreign exchange market maintains equilibrium relationship under the CIP theory. In general, international currency is not regulated and there are no restrictions to trade, so the CIP theory is established. Therefore, it is in principle impossible to trade risk-free arbitrage transaction that arises from the collapse of the CIP equilibrium relationship. They argued that the Yen, which is an international currency, is impossible for risk-free arbitrage transaction when considering the transaction cost. However, they argued that the Yuan, which is in the process of internationalization, and the Won, a non-international currency, imply arbitrage opportunities through SWAP transactions. In [14], the author used the VECM (Vector Error Correction Model) model to analyze the relationship between the country CDS Premiums of Korea, China and Japan, the exchange rate and the variables of the international financial market. They argued that while the country CDS Premiums of Korea and Japan were more closely related to the variables in the international financial market, the Chinese CDS Premiums had a weak relationship. In [15], the author used the VECM model to analyze the relationship between the international financial market variables and the spot exchange rate of Latin America and the country CDS Premium. He argued that although variables in the international financial market affect the spot exchange rate of Latin America countries, their impact varies from country to country.

This study analyzes the long-term equilibrium between the Korean won (onshore KRW and offshore Won-Dollar, herein “NDF KRW” where NDF is non deliverable forward), the Chinese yuan (onshore CNY and offshore CNH), and the Japanese yen JPY currency. And analyze the long-term equilibrium relationship between these currency and international financial market variables. The analysis results compared the characteristics of a total of five currency markets by dividing them into won (KRW and NDF KRW) and yuan (CNY and CNH) and yen JPY. This study can be differentiated from the existent researches based on the following; First, we have identified the long-term equilibrium relationship between the exchange rates of Korea won, China yuan, and Japan yen from the perspective of the international financial market. Second, the degree of integration with the international financial market was compared by comparing and analyzing the non-international currency, the won, the yuan and the yen, the international currency. Third, it was found that although the onshore and offshore yuan, which is being promoted for internationalization, is related to each other, there is a unique market characteristic that is distinguished from each other. Therefore, we have provided suggestions that company and international investors need yuan trading and yuan position management in consideration of the onshore and offshore yuan characteristics. The composition of the study is as follows; Chapter 1 describes the purpose and background of the research. Chapter 2 describes prior research and the variables used in analysis. Chapter

3 described the research model and the research methodology. Chapter 4 basic statistics, unit root test, and co-integration test of the analysis data. Chapter 5 presents the analysis results of the VECM model. Chapter 6 presents summaries and conclusions.

2. Previous research

2.1. Dollarization

The dollarization is the use of the dollar (foreign currency) as a means of saving value instead of the local currency. However, there is no consensus as to the various hypotheses in previous studies about the dollar phenomenon and the cause of currency exchange trading. In [16], the author put together four types of research on the dollarization. First, the study that the dollarization increases when the policy and institutional financial system of the local currency finance is poor. In [17], the author and in [18], the author, and in [19], the author argued that financial markets are not naturally fostered, but are developed through the process of maintaining stable asset values through their own economic growth and continued policy and institutional legislation. In [20], the author and in [19], the author argued that the dollarization is caused by increased risk of default, expected currency devaluation and a lack of dollar liability mismatch information. They argued that market failures were caused by problems between financial industry instability and related legal systems, and the dollarization was also caused by poor policy results and financial market imperfections. Second, the study that devaluation of exchange rates causes capital outflow and aggravates effect on the financial statements of domestic companies. In [21], the author argued that, in the event of a crisis, a significant foreign debt repayment burden reduces aggregate demand and, therefore, further depreciation of the exchange rate occurs. They argued that this would worsen the company financial statements, making borrowing conditions more unfavorable. Third, the study of currency substitution in relation to the foreign currency hedging. If investors experience high inflation in assets held in their local currencies, they will replace their holdings with foreign currencies to maintain future asset values. In [22], the author outlined the theory of trade-in between two currencies. In [23], the author argued in empirical studies that inflation and currency exchange trading are strongly interrelated. In [24], the author argued that holding foreign currencies in emerging countries played a role in hedging macroeconomic risks in emerging country economies. He argued that if investors were concerned about inflation in their currencies, the demand for foreign currency holdings would continue to increase. Fourth, the study that portfolio investment manager will have foreign currency assets at an optimal rate to minimize the risk of their investment portfolios. In [25], the author used the Minimum Variance Portfolio (MVP) method to study with five Latin America countries. They confirmed that assets and liabilities of financial institutions are readjusted in accordance with changes in exchange rates to hedge the inflation. In [26], the author argued that the proportion of dollar assets in MVPs and the dollarization are positive correlations.

2.2. Internationalization of Yuan

In 2008, when the global financial crisis recurred, China proceeded the internationalization of the yuan, mainly in Hong Kong. The yuan was newly incorporated into the SDR international currency in October 2016. The internationalizing of the yuan was a major issue for non-international currency countries. It was because currency internationalization could be an alternative to addressing the exchange rate problems common to emerging Asian countries. However, it is necessary for companies and financial institutions in non-international currency countries that do business with China to understand the characteristics of the new international currency, the renminbi, to manage exchange rate risk. In [27], the author asserted four backgrounds for the internationalization of the yuan. First, currency internationalization brings about the development of the direct financial market, so it has the effect of increasing the efficiency of monetary policy. Second, internationalization of currency has a good function to reduce the difference in interest rates and exchange rates between currencies by enabling arbitrage trading when changes in monetary policy are expected. Third, it argued that currency internationalization could further expand the wealth effect of the currency by enticing investors in the international financial market. Fourth, it argued that currency internationalization could lead to an increase in yuan deposits abroad, thereby promoting the growth of the yuan's international financial market. They summed up the requirements for the internationalization of the yuan as follows: The increase in China's economic weight in the global economy, the convenient exchange function of the yuan, the development of financial market and the liberalization of the Chinese financial market, the development of the offshore yuan financial market and the expansion of the flexibility of the yuan exchange rate. In [28], the author and in [29], the author cited five common factors in currency internationalization: political, military power, economic scale, national financial soundness and financial market competitiveness. In [30], the author argued that the higher the proportion of GDP, which represents the size of the economy, and the higher the proportion of trade in the global economy, the higher the possibility of internationalization of currencies. Trade dependence and inflation, on the other hand, have a negative impact on currency internationalization. They also argued that the greater the share of the stock market, the higher the proportion of foreign bonds issued in the local currency, and the greater the financial opening, the greater the possibility of internationalization of currencies. In [31], the author studied the international demand for yuan as a means of value storage. He used the explanatory variables of existing research to analyze the yuan. He estimated that the demand for the yuan would be 1 to 2 percent of the global currency volume in 2015 and 3 to 8 percent in 2020. However, he said that if China actively liberalizes its capital market and opens up its financial markets, the yuan's demand may rise more than expected as the yuan's trading volume increases further in the international financial market. The previous study of internationalization of currency were summarized in Table 1 below.

Table 1: Pre-study of the internationalization of currency

		Description of variable			
[Dependent variable]		Macroeconomic variables	Currency market variables	Financial market variables	Others
[Requirements to be used as foreign	Chinn & Frankel	GDP portion Inflation rate	Exchange rate volatility	Foreign exchange	time lag

exchange reserves)			depreciation rate	transaction weight	
	Chen et al	GDP portion, trade size portion, inflation rate	Exchange rate volatility		time lag
	Li Daokui et al	GDP portion Inflation rate	Exchange rate volatility depreciation rate		
	Dae Won Oh	GDP portion, trade size portion, inflation rate	Exchange rate volatility	M2 / GDP Openness	
[Currency internationalization requirements]	Seok Hyeon, Sang Heon Lee	GDP portion, Trade portion, Trade dependency, Inflation		Stock Market Openness	Financial openness
[RMB internationalization requirements]	Haihong & Yongding Yu	China's Economy portion in the Global Market	Flexibility of exchange rate Convenient exchange	Financial liberalization Financial Market Development Offshore Market Development	

In [32], the author analyzed the impact of the onshore Chinese yuan (CNY) on the offshore yuan (CNH) market in Hong Kong. According to their analysis, the change in the implied interest rate of the forward exchange rate in China immediately affects the renminbi forward exchange rate market in Hong Kong. In [33], the author analyzed the effect of interest rate volatility transferred from the Chinese market to the Hong Kong offshore renminbi market through a variance decomposition analysis. They argued that volatility was transferred from the Chinese market to the Hong Kong offshore market at an average of 60.9%. However, the effect of transferring volatility from the Hong Kong offshore market to the Chinese local market was insignificant at 13.8%. In [34], the author analyzed the CNH and CNY markets from September 2010 to August 2013. They argued that CNY had a strong influence on CNH at the beginning of the analysis period, but argued that CNH's influence on CNY increased in the later period. Therefore, they argued that CNH Hong Kong's offshore yuan has predictive of CNY's Chinese onshore yuan. In [35], the author studied the difference in price between CNH offshore renminbi and CNY onshore currency. They argued that the yuan liquidity and risk aversion of international financial market investors were the main causes of the price difference between CNY and CNH and incurring arbitrage trading. They argued that an increase in the inflow of yuan from the CNY onshore market to the CNH Hong Kong offshore market narrowed the CNH-CNY price difference, and thus the yuan exchange rate was adjusted based on the available liquidity in the market.

2.3. Analysis of variable

We describe the variables used in the fifteen VECM analysis models. The first variable is the spot exchange rate. The exchange rate represents the credit risk of the currency country viewed in the financial market. In [2], the author argued that the depreciation of the exchange rate is a negative factor for the financial system of the country. In [36], the author argued that the depreciation of the exchange rate is a problem with the repayment of national debt. In [37], the author argued that a sharp depreciation of the exchange rate lowers the value of the assets denominated in the local currency, resulting in a lack of foreign currency liquidity, which increases the occurrence of a default risk. The second variable is the 10-year US treasury interest rate. In [38], the author argued that the 10-year U.S. government bond interest rate is the benchmark for the level of interest expected by investors in international financial markets. In [39], the author argued that the determinants of investment capital flowing into emerging economies are largely attributable to supply factors, such as the level of U.S. treasury bond interest rates and preference for safe assets in the international financial market. It argued that a 50bp drops in interest rates on 10-year U.S. treasury bonds would increase the total inflow of capital into emerging economies by 13 percent. The third variable is credit risk spread. In [40], the author argued that credit risk spreads represent investors' risk aversion and are explained by the spread difference between the 10-year U.S. treasury bond and credit risky bonds. In [41], the author argued that an increase in the hedge propensity of investors in the international financial market increases the credit risk spread, and that the credit risk spread is significant in explaining the EMBI index of emerging economies in 31 countries. Alsaka and ap Gwilym [1] argued that the announcement by the credit rating agency had a significant impact on exchange rate fluctuations. The fourth variable is the VIX index. VIX is Chicago Board Options Exchange Volatility Index and is a variability in S&P100 index option transactions. The index represents the level of hedge propensity among international financial market investors. The VIX index is to the extent that investors feel an event risk, and it is the risk compensation rate expected by investors who expect high returns in the international financial market. In [42], the author argued that the VIX index is an important explanatory variable that determines the rate of return on treasury bonds. In [43], the author argued that it is an important variable that explained the premium on government bonds in emerging countries. In [44], the author argued that VIX is an important variable that explains the CDS premium in Mexico, Turkey, and South Korea. In [38], the author argued that VIX is a factor that explains the premium on short-term government bonds. In [39], the author stated that the VIX index represents a preference for safe assets in the international financial market and that if the VIX index fall halves, total capital inflows from emerging economies would increase by 11%. The fifth variable is TED spread. TED spread is the difference between the 90-day yield on U.S. treasury bonds and the three-month interest rate on Euro-dollar Libor. TED spread represents the average level of risk felt by the counterparty in inter-bank money market transactions in the international financial market. Higher cross-bank counterparty risk burdens are considered to have increased TED spreads and increased hedging tendencies in the international financial market, worsening the funding situation. The change in TED spread implies changes in overall capital liquidity in the international financial market or changes in credit risk levels and bond rates. In [45], the author used the TED spread as a measure of the liquidity level of hedge funds. In [41], the author used TED spread as a measure of the liquidity of government-issued government bonds in the international financial market.

3. Research model

3.1. Research Analysis

In [46], the author analyzed the mid- to long-term relationship of the Turkish exchange rate using the VECM analysis model. He

constructed a nominal exchange rate and market exchange rate model and analyzed it using the VECM analysis method. They argued that the market exchange rate model has better predictability of the future mid-term exchange rate than a random walk with and without drift. In [15], the author also used the VECM analysis methodology to analyze the relationship between the Latin America spot exchange rate and the international financial market variables (the 10-year US treasury interest rate, the credit risk spread between the AAA credit rating and the BBB credit rating, VIX index, TED spread). This study also analyzed the Granger Causality, the impulse response, and the variance decomposition analysis using VECM analysis methodology that studied the mid- to long-term exchange rate relationship. The analysis variables are Korea won onshore (KRW) and offshore won, NDF KRW (KWD1m), Chinese onshore (CNY) and offshore yuan (CNH), Japanese yen (JPY), 10-year US treasury bond interest rate (T10Y), credit risk spread, VIX index, TED spread. The Chinese yuan is analyzed by dividing the Chinese onshore yuan (CNY) market, the Hong Kong offshore yuan (CNH) market, and the Seoul offshore yuan market (herein we refer as "CNK"). The Korean won is divided into the Seoul Foreign Exchange onshore (KRW) Market and the offshore NDF KRW market (KWD1M). The offshore won NDF KRW, has the shortest maturity of one month's forward rate. In [47], the author also analyzed the causal relationship between the spot exchange rate of the onshore won dollar and the offshore NDF KRW exchange rate (the same one-month forward rate). NDF KRW is a forward exchange contract in which only the difference between the forward exchange rate and the spot exchange rate is settled in dollars without having to exchange the principal of the contract at maturity. As of 2018, the Korea won's NDF market is worth about $9.5 billion. The won is non-international, but there is no limit to profit-taking arbitrage transactions in which investors take advantage of the price gap between the Seoul foreign exchange KRW market and the Seoul bond market and the offshore NDF KRW market. Therefore, it is assumed that the won has become substantially and effectively internationalized. The Japanese yen is International currency. Therefore, it analyzes the yen traded in international financial markets as there is no distinction between onshore and offshore markets. A total of fifteen VECM analysis models are used to verify the long-term cointegration relationship between variables and apply the VECM analysis methodology only if the long-term cointegration relationship is valid at statistical confidence level. The data analysis period is from December 1, 2014, when the yuan (CNK) began trading on the Seoul foreign exchange market to November 30 2017. The daily average trading volume of the Seoul offshore yuan (CNK) currency market is $1.5 billion to $2 billion, ranking third to fourth among the global offshore yuan market. There is no restriction on transactions because CNK is the same offshore yuan as the Hong Kong offshore yuan (CNH). It also means that the cross rate between the currencies can also affect the won-dollar KRW exchange rate. Therefore, the data analysis period was set for three years from December 1, 2014 to November 30, 2017 to compare fifteen analysis models.

3.2. Analysis Model

This study first tests the long-term equilibrium relationship between the analysis variables using the cointegration analysis method proposed by [48] the author. In general, time series data are non-stationary and therefore have a unit root. However, if there is a stable linear combination between these variables, it can be considered that there is a cointegration relationship between the variables, so that a long-term, stationary relationship exists. Therefore, regression analysis can be used when a cointegration relationship is recognized even among unstable time series data. When the cointegration relationship is confirmed, the relationship between the variables is analyzed using the VECM model, which includes the error correction term (ECT) in the VAR model. The VECM model can measure the effect of the time lag between endogenous variables and analyze the long-term equilibrium relation between variables as well as the short-term dynamic relations simultaneously. Nevertheless, since the VAR model is not constrained, the results may vary depending on the variable settings. Therefore, this study performed Granger causality analysis in addition. The impulse response analysis was used to analyze how external impacts were transferred between variables. The study also add variance decomposition analysis to analyze the magnitude of the influence of unexpected impacts on variables. It is necessary to verify the cointegration relationship between research variables in order to apply the VECM analysis methodology. Thus, this study conducted Johansson's cointegration analysis. The methods for verifying Johansen's cointegration vector are trace (λ trace) and maximum eigenvalue (λ max) tests. The test statistic is as follows:

$$\lambda_{trace} = -\ T\sum_{i=r+1}^{n} \ln\left(1-\hat{\lambda_i}\right) \quad (1)$$

$$\lambda_{max}(r, r+1) = -T\ln\left(1-\hat{\lambda_{r+1}}\right) \quad (2)$$

here, λ_i indicates the matrix eigenvalue and T is the observed number. The null hypothesis of λ trace is: "there is no cointegration"; if this is rejected, there exist at least r number of cointegration relations, and the variables become stable *I (0)* time series. Only after the cointegration verification is confirmed, the relationship between the endogenous variables is analyzed by applying the VECM analysis methodology. The cointegration test show the long-term equilibrium relationship between variables, but it does not suggest its direction. To overcome this problem, this study analysis the relationship between variables using VECM model. The use of the VECM model has the advantage of being able to identify both short- and long-term causal relationships, as it can identify not only the effects of the differential term of the independent variables on the dependent variables, but also the effect of the variation of the error correction term on the dependent variables. The VECM model is as follows:

$$\Delta[LKRW_t] = [\alpha_1] + \sum_{i=1}^{p} \alpha_i \Delta[LKRW_{t-i}]$$
$$+ \sum_{j=1}^{p} \alpha_j \Delta\left[LKRW1m_{t-j}\right] \ldots + [\theta_1][ECT_{t-1}] + \left[\varepsilon_{1,t}\right] \quad (3)$$

The extension of (3) of two variables into six variables shall be expressed as follow:

$$\begin{bmatrix} LKRW_t \\ LKWD1m_t \\ T10Y_t \\ CREDIT_t \\ LVIX_t \\ TED_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix} + \sum_{i=1}^{p} \Delta \begin{bmatrix} \beta_{11i}\beta_{12i}\beta_{13i}\beta_{14i}\beta_{15i}\beta_{16i}\,\beta_{17i} \\ \beta_{21i}\beta_{22i}\beta_{23i}\beta_{24i}\beta_{25i}\beta_{26i}\,\beta_{27i} \\ \beta_{31i}\beta_{32i}\beta_{33i}\beta_{34i}\beta_{35i}\beta_{36i}\,\beta_{37i} \\ \beta_{41i}\beta_{42i}\beta_{43i}\beta_{44i}\beta_{45i}\beta_{46i}\,\beta_{47i} \\ \beta_{51i}\beta_{52i}\beta_{53i}\beta_{54i}\beta_{55i}\beta_{56i}\,\beta_{57i} \\ \beta_{61i}\beta_{62i}\beta_{63i}\beta_{64i}\beta_{65i}\beta_{66i}\,\beta_{67i} \end{bmatrix} \times \begin{bmatrix} LKRW_{t-j} \\ LKWD1m_{t-j} \\ T10Y_{t-j} \\ CREDIT_{t-j} \\ LVIX_{t-j} \\ TED_{t-j} \end{bmatrix} + \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \end{bmatrix} [ECT_{t-1}] + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \\ \varepsilon_{6t} \end{bmatrix} \quad (4)$$

where $\triangle$ is the difference coefficient, and $ECT\ t\text{-}1$ is the term describing the long-term equilibrium relation of the estimation coefficient as an error correction term, where $\alpha\ \beta\ \theta$ is the measurement coefficient value, p is the time difference and ε_t is the error term. The Granger causality test verifies the significance of the coefficient values measured in relation to the lagged variable *j* and the variable *i. If "H_{null} : j does not granger cause i*" is rejected, j shall become a Granger cause to i. Impulse response verification measures how quickly information is passed between variables when impacts are delivered to the VECM model with time lag. In other words, this study analyzed how the variables in the model react with time when a certain amount of impact is delivered. A variance decomposition analysis shows the relative importance of variables, as the estimation error of VECM shows the proportion that is explained by its own and other variables. Therefore, if there is no decomposition value of the variable due to the impact of the residual, this variable is considered as exogenous.

4. Analysis data

4.1. data

The analysis variables are summarized in Table 2. The analysis period is from December 1, 2014 to November 30, 2017, and is daily closing data. The closing price is determined by the accumulation of both new information for the day and trading activities between the parties. The data was obtained through Bloomberg or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Table 2: Analysis variables

	Market	Variables
	Global	US Treasury 10 Year
		Credit spread
		VIX index
		TED spread
Korea	onshore	KRW
	offshore	KWD1m
China	onshore	CNY
	offshore	CNH*, CNK**
Japan	Global	JPY

The exchange rate is the spot exchange rate. SPOT of the Chinese yuan, the Japanese yen, and the Korean won, against U.S. dollar. The Chinese yuan is analyzed by dividing it into onshore yuan (CNY) and offshore yuan in Hongkong (CNH). For the Korea won dollar, there is no offshore spot exchange rate. Therefore, the shortest term of one-month term in offshore market was used as a won-dollar offshore exchange rate (KWD1m). The analysis variables for the international financial market are the 10-year U.S. treasury bond interest rate, Credit risk spread, VIX index and TED spread, which are daily closing data. During the analysis period, there were no fundamental shift in the international financial market, such as the 1998 Asian financial market crisis or the 2008 international financial market crisis. Therefore, dummy variables were not considered. The analysis of basic statistics of yen, yuan, and won dollar is summarized in Table 3. The spot exchange rate and the VIX index were calculated as the difference between today's closing price taken on the log value and the previous day's closing price taken on the log value. The 10-year treasury interest rate, credit risk spread and TED spread are the difference between today's and the previous day's closing price of data.

Table 3: Raw data statistic

	onshore market		Global market							
	Korea	China	Korea	China		Japan	global market variable			
Variables	KRW	CNY	KWD1m	CNH	CNK	JPY	T10Y	CREDIT	LVIX	TED
Mean	1141.52	6.55	1142.34	6.56	174.38	114.17	2.09	1.70	14.72	34.53
Median	1136.95	6.58	1137.69	6.58	174.61	113.62	2.17	1.66	13.72	31.38
Maximum	1240.90	6.96	1246.00	6.98	189.53	125.63	2.63	2.30	40.74	68.93
Minimum	1068.60	6.15	1064.95	6.15	162.01	100.03	1.36	1.29	9.14	14.52
Std. Dev.	35.29009	0.24247	34.99482	0.23624	7.00772	6.40875	0.28092	0.23559	4.25084	10.46365
Observations	784	784	784	784	784	784	784	784	784	784
ADF	-2.1636	-1.6404	-2.1636	-1.7592	-1.5864	-1.5845	-2.1857	-0.3129	-4.5982***	-2.2357
PP	-2.2105	-1.6339	-2.2105	-1.7090	-1.3730	-1.5983	-2.0749	-0.6730	-4.4872***	-2.1803

* 10% significance level, ** 5% significance level, *** 1% significance level. KRW, CNY, and JPY are Korean, Chinese and Japanese spot exchange rates. CNH is the Chinese offshore yuan spot exchange rate, and KWD1m is the 1-month NDF KRW exchange rate. T10Y is the US treasury 10-year interest rate (%). CREDIT is the difference (%) of the spread between T10y and BBB credit rating bonds. VIX is the chicago options market index (%). TED is the difference between the libor 3-month interest rate and the T-bill rate (basis point)

Table 4: Stationary test

	onshore market		Global market							
	Korea	China	Korea	China		Japan	global market variable			
	KRW	CNY	KWD1m	CNH	CNK	JPY	T10Y	CREDIT	LVIX	TED
ADF	-2.1636	-1.6404	-2.1636	-1.7592	-1.5864	-1.5845	-2.1857	-0.3129	-4.59***	-2.2357
PP	-2.2105	-1.6339	-2.2105	-1.7090	-1.3730	-1.5983	-2.0749	-0.6730	-4.48***	-2.1803
ADF	-27.09***	-24.50***	-28.91***	-20.61***	-25.75***	-27.44***	-30.12***	-27.41***	-27.71***	-11.79***
PP	-27.07***	-24.43***	-29.03***	-24.56***	-25.73***	-27.44***	-30.13***	-42.64***	-30.09***	-27.92***

4.2. Data

The Stationary of the data was confirmed prior to the analysis. The study used ADF (Augmented Dickey Fuller) and PP (Phillips-Perron) unit root analysis method to verify the data. The analysis result of the yen, the yuan, and the won and international financial market variables are summarized in Table 4.

The first difference of the natural logarithm of spot exchange rate and VIX index data rejected the null hypothesis that "there is a unit root". Therefore, time series data satisfy the stationary condition. The first differenced data of 10-year US treasury interest rate (T10Y), Credit risk spread (CREDIT) and TED spread (TED) showed that the data satisfied the stationary without the unit root. Therefore, this study analyzed by using the log conversion for spot exchange rate, and VIX index data, and by using the level data for the US treasury 10-year interest rate, Credit risk spread, and TED spread.

4.3. Cointegration Test

The optimal time lag of the fifteen VECM model among the aforementioned six variables was determined prior to the cointegration test between the analytic variables. The optimal time lag was selected based on the time differences where AIC and SC values are minimized in the Akaike Information Criteria (AIC) or Schwarz Criteria (SC) analysis. The results of the time lag selection of the VECM (p) model are summarized in Table 6 below. The optimal time lag of the yen and Korea won is the time lag 3. The optimal time lag of the Chinese yuan is time lag 6. In the case of the yuan, the long-term equilibrium cointegration relationship was found to lag behind the yen and the won dollar. The time lag determined for each VECM model was applied to both the cointegration verification and the VECM analysis. The following includes the examination of whether there is at least one cointegration vector, which is a long-term equilibrium relationship between level variables of each time-series data, through cointegration analysis commonly used in financial time series data analysis. Table 5 summarizes the results of the number of the cointegration analysis of fifteen VECM analysis model between the analytic variables.

As summarized in Table 5, the null hypothesis of "no cointegration vector" was rejected at the 5% level for the all fifteen analysis model cases. Therefore, more than one cointegration relation between time series variables used in fifteen analysis model was revealed. The existence of the cointegration indicates that the individual variables are non-stationary, but when they are linked by a cointegration vector relation it indicates the stationary. Therefore, this study analyzed the lead-lag causality between variables using the VECM analysis model that includes the error correction term in the VAR model analysis.

Table 5: Cointegration vector

Currency		market sector (model No)		Cointegration vectors							
						dependent	independent variable				
KRW	on-shore KRW	[CNK] KWD1m, KRW, CNH, CNY, JPY	(1)		constant	LCNK	LKWD1m	LKRW	LCNH	LCNY	LJPY
				Coint EQ (1)	-4.49	1	0	0	-25.97316***	25.92838***	-0.118494
				Coint EQ (2)	-4.62	0	1	0	-26.39582***	25.37661***	-0.099479
				Coint EQ (3)	-4.55	0	0	1	-26.64894***	25.60423***	-0.104134
		[KRW] KWD1m, CNK, CNH, CNY, JPY	(2)		constant	LKRW	LKWD1m	LCNK	LCNH	LCNY	LJPY
				Coint EQ (1)	-4.55	1	0	0	-26.64894***	25.60423***	-0.104134
				Coint EQ (2)	-4.62	0	1	0	-26.39582***	25.37661***	-0.099479
				Coint EQ (3)	-4.49	0	0	1	-25.97316***	25.92838***	-0.118494
		[KRW] KWD1m	(3)		constant	LKRW	LKWD1m				
				Coint EQ (1)	0.11	1	-1.015389***				
		[KRW] KWD1m, T10Y, Credit, VIX, TED	(4)		constant	LKRW	LKWD1M	T10Y	CREDIT	LVIX	TED
				Coint EQ (1)	-8.52	1	0	0.081849	-0.895473***	1.074763***	-0.000894
				Coint EQ (2)	-8.47	0	1	0.079503	-0.872705***	1.044728***	-0.000904
	off-shore KWD1m	[KWD1m] KRW, CNK, CNH, CNY, JPY	(5)		constant	LKWD1m	LKRW	LCNK	LCNH	LCNY	LJPY
				Coint EQ (1)	-4.62	1	0	0	-26.39582***	25.37661***	-0.099479
				Coint EQ (2)	-4.55	0	1	0	-26.64894***	25.60423***	-0.104134
				Coint EQ (3)	-4.49	0	0	1	-25.97316***	25.92838***	-0.118494
		[KWD1m] KRW	(6)		constant	LKWD1m	LKRW				
				Coint EQ (1)	-0.11	1	-0.984844***				
		[KWD1m] KRW, T10Y, Credit, VIX, TED	(7)		constant	LKWD1m	LKRW	T10Y	CREDIT	LVIX	TED
				Coint EQ (1)	-9.44	1	0	0.153340	-1.253755***	1.587332***	-0.000418
				Coint EQ (2)	-9.52	0	1	0.158203	-1.290331***	1.636855***	-0.000389
	on-shore	[CNY] CNH, CNK,	(8)		constant	LCNY	LCNH	LCNK	LKWD1m	LKRW	LJPY
				Coint EQ (1)	-7.33	1	0	0	64.16184***	-63.55239***	0.235192

RMB	CNY	KWD1m, KRW, JPY		Coint EQ (2)	-6.87	0	1	0	61.64652***	-61.09849***	0.229879
				Coint EQ (3)	7.09	0	0	1	-62.4577***	60.8898***	-0.245945
		[CNY] CNH	(9)		constant	LCNY	LCNH				
				Coint EQ (1)	0.06	1	-1.030136***				
		[CNY} CNH, T10Y, Credit, VIX, TED	(10)		constant	LCNY	LCNH	T10Y	CREDIT	LVIX	TED
				Coint EQ (1)	-12.80	1	0	0.773481	-4.477629***	6.285092***	0.005925
				Coint EQ (2)	-12.96	0	1	0.786428	-4.551355***	6.38005***	0.006126
	off-shore CNH	[CNH] CNY, CNK	(11)		constant	LCNH	LCNY	LCNK			
				Coint EQ (1)	0.23	1	-1.004402***	-0.04332***			
		[CNH] CNY, CNK KWD1m, KRW, JPY	(12)		constant	LCNH	LCNY	LCNK	LKWD1M	LKRW	LJPY
				Coint EQ (1)	-6.87	1	0	0	61.64652***	-61.09849***	0.229879
				Coint EQ (2)	-7.33	0	1	0	64.16184***	-63.55239***	0.235192
				Coint EQ (3)	7.09	0	0	1	-62.4577***	60.8898***	-0.245945
		[CNH] CNY	(13)		constant	LCNH	LCNY				
				Coint EQ (1)	-0.06	1	-0.969678***				
		[CNH] CNY, T10Y, Credit VIX, TED	(14)		constant	LCNH	LCNY	T10Y	CREDIT	LVIX	TED
				Coint EQ (1)	-5.27	1	0	0.194119	-1.489189***	2.081043*	-0.000288
				Coint EQ (2)	-5.25	0	1	0.192055	-1.473122***	2.066396*	-0.000370
JPY	[JPY] T10Y, Credit, VIX, TED		(15)		constant	JPY		T10Y	CREDIT	LVIX	TED
				Coint EQ (1)	-2.87	1		-0.22638***	0.524198***	-0.882197***	0.002019

5. Analysis Result

5.1. VECM Analysis

Table 6 shows the relationship between the rate of change of the analytic variables by applying the time lags to the fifteen VECM model using the long-term cointegrating equilibrium equations. The fifteen VECM analysis models' error correction term, ECT values are expressed as Coint EQ (1). In the VECM model analysis, if there are two cointegrations, it is expressed as Coint EQ (1) and (2). In the VECM analysis model, the ECT, error correction term must be negative (-) and statistically significant at the same time. Table 6 analysis results showed that 12 of the total 15 analysis models had negative (-) error correction terms and were statistically significant at the same time. However, in analysis models 5, 6 and 12, the error correction term was negative (-), but it was not significant.

* Korea won dollar and Seoul offshore yuan

We found four onshore won-dollar models, three offshore won-dollar models, and a total of seven VECM analysis models. The following Table 6 shows that all four VECM models of the onshore won-dollar were significant. The long-term equilibrium relationship explanation power of the VECM model was found to be around 50 percent to 68 percent. However, the explanation power of the three offshore won-dollar VECM models was about 3 percent to 6 percent. The error correction term, Coint EQ of the five VECM analysis models was negative (-) and statistically significant. However, analysis models No. 5 and No. 6 showed that the error correction term were negative (-), but were not statistically significant. The VECM Model No. 1, CNK, shows a long-term equilibrium relationship between the won-dollar and the yuan and the yen. Since CNK is an offshore yuan traded in the Seoul foreign exchange market, there are no restrictions on its transactions. Therefore, it has been shown that the Seoul offshore yuan (CNK) has a long-term equilibrium relationship with onshore yuan (CNY) and offshore yuan (CNH) and Japanese yen (JPY) either. These results suggest that although the yuan's liquidity in the Seoul foreign exchange market is relatively insufficient compared to Hong Kong. Therefore, it secures insufficient liquidity through cross-rate transactions between currencies of KRW↔USD↔CNH, CNY, and JPY. The amount of liquidity of a currency used in transactions turned out to be the influence factor of these currencies. It gives impact on the Seoul offshore yuan (CNK) exchange rate either. VECM model No. 2, onshore won-dollar (KRW) exchange rate, indicated that there was a long-term equilibrium relationship with the offshore won-dollar exchange rate, the yuan and the yen. This result implied that the won-dollar exchange rate received impact by the yuan and yen flowing through the export/import trade and inter-bank money markets transactions between Korea, China, and Japan. VECM Model No. 4 is an analysis of the relationship between the onshore won-dollar (KRW) and the variables in the international financial market. The analysis results indicated that the won-dollar exchange rate has a long-term equilibrium relationship with variables in the international financial market. VECM Analysis Models No. 1, No. 2, No. 3 and 4 show that the error correction factor of the offshore won-dollar exchange rate (KWD1m) is greater than that of the other variables. VECM model No. 3 analyzed the relationship between two variables, the onshore won-dollar exchange rate (KRW) and the offshore won-dollar (KWD1m). According to the analysis of the long-term equilibrium relationship between the two variables, the offshore won-dollar exchange rate explains the onshore won-dollar exchange rate about 50%. These results are similar to [47] the author's argument. They studied the Granger causal relationship between the KRW exchange rate and the offshore NDF KRW market. They argued that the offshore one-month forward exchange market was strongly causative in the onshore spot exchange market and that the volatility shock was exchanged between the two markets. In [34] the author analyzed the relationship between offshore yuan (CNH) and onshore yuan (CNY) for the period from September 2010 to August 2013. They argued that the CNH Hong Kong yuan has a predictive power for the CNY yuan in China. The Korean won has not yet been internationalized, but like the yuan, the onshore won-dollar exchange rate has been consistently and strongly influenced by the offshore market won-dollar exchange rate. In addition, the offshore won-dollar exchange rate, KWD1m is found to have a long-term equilibrium relationship with four variables in the international financial market, and as a result, the Korean won-dollar exchange rate (KRW) is affected by variables in the international financial market either. The results suggest that Korea onshore won-dollar exchange market and the Korea won-

Table 6: VECM analysis

	KRW							Yuan							JPY
[dependent variable]	[onshore KRW]				[offshore KWD1m]			[onshore CNY]			[off-shore CNH]				
	[CNK]	[KRW]	[KRW]	[KRW]	[KWD1m]	[KWD1m]	[KWD1m]	[CNY]	[CNY]	[CNY]	[CNH]	[CNH]	[CNH]	[CNH]	[JPY]
variable 1	LKWD1m	LKRW1m	LKWD1m	LKRW1m	LKRW	LKRW	LKRW	LCNH	LCNH	LCNH	LCNY	LCNY	LCNY	LCNY	
variable 2	LKRW	LCNK		T10Y	LCNK		T10Y	LCNK		T10Y	LCNK	LCNK		T10Y	T10Y
variable 3	LCNH	LCNH		CREDIT	LCNH		CREDIT	LKRW1m		CREDIT		LKWD1m		CREDIT	CREDIT
variable 4	LCNY	LCNY		LVIX	LCNY		LVIX	LKRW		LVIX		LKRW		LVIX	LVIX
variable 5	LJPY	LJPY		TED	LJPY		TED	LJPY		TED		LJPY		TED	TED
Coint EQ (1)	-0.43***	-0.584***	-0.60***	-0.617***	*-0.1303*	*-0.0646*	-0.1747**	-0.056**	-0.035*	-0.058**	-0.064**	*-0.0455*	-0.051*	-0.063*	-0.002*
Coint EQ (2)	-0.1237	0.2444***		0.16879**	*0.6307****	*0.6114*	0.6353***	0.1095*		0.05109		*0.0261*		0.065***	
Coint EQ (3)	-0.0416	0.0911**			*0.3452****			-0.191***				*-0.235****			
time lag	3	3	2	3	*3*	*2*	3	3	4	3	4	*3*	6	2	3
adjusted $R^{\wedge 2}$	68.48%	50.94%	49.97%	51.29%	*3.21%*	*0.66%*	6.47%	7.92%	7.34%	8.63%	4.78%	*5.25%*	5.46%	4.93%	9.42%
model number	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)

government bond market and the inter-bank money market are substantially and effectively open, so this increase the link to the international financial market. In [49] the author and in [50] the announcement argued that certain currency shocks are transferred through the interbank money market. It also argued that the impact from the US dollar currency affects the interbank money markets in Hong Kong (HKD), Australia (AUD), Japan (JPY), and Korea (KRW). However, they argued that the impact of the US dollar currency impact on China (Yuan), Thailand (THB) and Malaysia (RHB) is limited. The fact that the Korean won dollar exchange rate is strongly related to the international financial market is in line with [49] the author's argument and with [50] announcement. The VECM model relationships model No. 5 and model No. 6 were not significant. The analysis shows that offshore won-dollar exchange rates are more affected by international financial market factors than by regional factors between Korea, China and Japan in Northeast Asia.

*China Yuan

We found three onshore analysis model and four offshore analysis models, and a total of seven VECM analysis models. The three VECM models of the onshore yuan (CNY) were shown to be statistically significant. The three long-term equilibrium relationship VECM models of the onshore yuan had the explanatory power of between 7% and 8%. The explanatory power of the four VECM models of the Hong Kong offshore yuan (CNH) was between 4% and 5%. Although the error correction term of the six VECM analysis models was significant, analysis model 12 showed that the error correction term was not significant. The VECM model No. 8 Chinese onshore yuan (CNY) was found to have a long-term equilibrium relationship with the onshore won dollar (KRW) and the yen (JPY). The analysis results of VECM Model No. 10 and No. 14 show that onshore and offshore yuan have long-term equilibrium relationship with international financial market variables. The VECM model No. 11 Hong Kong offshore yuan (CNH) has a long-term equilibrium relationship with the Chinese onshore yuan (CNY) and the Korean offshore yuan (CNK). The unusual was that, VECM analysis models No. 9 and 13, show that China's long-term equilibrium relationship between onshore and offshore yuan is weaker than Korea's onshore and offshore won-dollar relations. In the previous model No. 3 Korean Won-dollar VECM analysis model, offshore Won-dollar's explanation power on the onshore KRW was about 50%, at 1% significant level. However, the explanatory power of the VECM No. 9 and No. 13 analytical models analyzed the Chinese yuan was 7.3% and 5.4%, respectively, and it was found to be significant at the 10% level. It is unusual for the long-term equilibrium relationship between onshore and offshore yuan markets to be weaker than the non-international currency of Korea won. This analysis result is interpreted by the fact that the yuan onshore-offshore trade related cross border transactions are liberalized, but capital market transactions are limited. Overall, the yuan's long-term equilibrium relationship with other currencies was weaker than that of Korea won. And these results are in line with [49] the author's argument that the impact of the shock from the US dollar currency on Chinese yuan is limited.

*Japan Yen

VECM Analysis Model No. 15 of Table 6 is the result of an analysis of the long-term equilibrium relationship between the Japanese yen and variables in the international financial market. The analysis indicated that the yen exchange rate is more related to the variables in the international financial market than that of the won and the yuan. This result suggest that the yen is international currency and the Japanese bond markets are fully open to international financial market investors, so, the greater association with the variables of the international financial market than the non-international currency, the Korea won and the yuan, which is on the way of internationalized. Analysis models 1, 2, and 8 show that the onshore won dollar (KRW) and the onshore yuan (CNY) have a long-term equilibrium relationship with the yen (JPY). This result show that since most of the demand and supply of foreign exchange transactions occurred through import-export transactions and were supplied to the foreign exchange markets in Korea and China, so the currency market result in long term equilibrium relationship. However, looking at analysis models Nos. 5, 6, and 12, it was found that the offshore won (KWD1m) and the offshore yuan (CNH) had no long-term equilibrium relationship with the yen (JPY). This result is interpreted as that, in the offshore international financial market, since investors trade currencies in consideration of the country risk of the currency, a long-term equilibrium relationship between the offshore won and the offshore yuan is not established with the international currency, the yen.

5.2. Granger Causality Analysis

The VECM model has no restriction on the analytical model and may cause a fictitious spurious regression analysis problem. As such, this study further conducted a widely used in time series

analysis, Granger causality test. The null hypothesis of Granger causality test is that "the offshore won, KWD1m has no effect on the spot exchange rate of the onshore won, KRW", and the direction of the causal relationship in Table 7 was written as "KWD1m→KRW". If the null hypothesis is rejected, the offshore won shall have an impact on onshore KRW exchange rates. Table 7 summarizes the results of the Granger causality analysis between the exchange rates of Japanese yen, Chinese renminbi, and Korean won, and between the exchange rates of the three currencies of Korea, China, and Japan and variables in the international financial market.

Table 7: Granger Causality analysis of won, yuan, yen and international financial market variables

Currency		market sector (model No)		Granger Causality relation			Time lag (1)	Time lag (2)	Time lag (3)	Time lag (4)	Time lag (5)
							Chi-square	Chi-square	Chi-square	Chi-square	Chi-square
KRW	on-shore KRW	[CNK] KWD1m, KRW, CNH, CNY, JPY	(1)	KWD1m	→	CNK	14.9383	56.2589	214.79	165.636	135.221
				CNH	→	CNK	*na*	*na*	*na*	2.94628**	2.39355**
				CNY	→	CNK	3.33834*	*na*	*na*	2.05321*	*na*
				KRW	→	CNK	*na*	77.2157***	58.6483***	45.4459***	36.4794***
				JPY	→	CNK	6.40453**	3.18475**	3.45293**	3.06304**	2.74506**
				CNK	→	*KWD1m*	*na*	*na*	*na*	*na*	*na*
				KRW	→	*KWD1m*	*na*	*na*	*na*	*na*	*na*
				CNH	→	*KWD1m*	*na*	*na*	*na*	*na*	*na*
				CNY	→	*KWD1m*	*na*	*na*	*na*	*na*	*na*
				JPY	→	*KWD1m*	*na*	*na*	*na*	*na*	*na*
				KWD1m	→	KRW	732.949***	379.449***	254.111***	189.808***	152.381***
				CNK	→	KRW	*na*		2.49953*	2.63884**	2.13671*
				CNH	→	KRW	*na*	16.3348***	13.3646***	10.049***	9.04209***
				CNY	→	KRW	*na*	12.1681***	10.988***	8.23484***	6.57248***
				JPY	→	KRW	*na*	5.78389***	6.56095***	5.00083***	4.69926***
				CNY	→	CNH	4.22086**	6.22332***	3.6091**	2.61292**	2.23464**
				CNK	→	*CNH*	*na*	*na*	*na*	*na*	*na*
				LKWD1m	→	*CNH*	*na*	*na*	*na*	*na*	*na*
				KRW	→	*CNH*	*na*	*na*	*na*	*na*	*na*
				JPY	→	*CNH*	*na*	*na*	*na*	*na*	*na*
				CNK	→	*CNY*	*na*	*na*	*na*	*na*	*na*
				LKWD1m	→	*CNY*	*na*	*na*	*na*	*na*	*na*
				KRW	→	*CNY*	*na*	*na*	*na*	*na*	*na*
				CNH	→	CNY	13.5758***	14.8769***	11.4008**	8.35465***	6.67746***
				JPY	→	*CNY*	*na*	*na*	*na*	*na*	*na*
				KWD1m	→	JPY	2.99698*	*na*	*na*	*na*	*na*
				KRW	→	JPY	3.07157***	*na*	*na*	2.07901*	2.3439**
				CNK	→	*JPY*	*na*	*na*	*na*	*na*	*na*
				CNH	→	*JPY*	*na*	*na*	*na*	*na*	*na*
				CNY	→	*JPY*	*na*	*na*	*na*	*na*	*na*
			(2)	*[KRW] KWD1m, CNK, CNH, CNY, JPY*							
			(3)	*[KRW] KWD1m*							
		[KRW] KWD1m, T10Y, Credit, VIX, TED	(4)	T10Y	→	KRW	*na*	*na*	4.50778***	3.50753***	3.53952***
				Credit	→	KRW	3.33958**	*na*	*na*	*na*	*na*
				VIX	→	KRW	5.16845**	4.82868***	5.68474***	4.39125***	3.57154***
				T10Y	→	KWD1m	*na*	2.86013*	2.92025**	3.17319**	3.53952***
				Credit	→	KWD1m	4.01793**	*na*	*na*	*na*	*na*
				TED	→	KWD1m	*na*	4.19474**	2.73643**	3.00317**	2.4705**
				KWD1m	→	T10Y	*na*	3.16114**	2.02383*	2.48292**	1.8993*
	off-shore KWD1m		(5)	*[KWD1m] KRW, CNK, CNH, CNY, JPY*							
			(6)	*[KWD1m] KRW*							
			(7)	*[KWD1M] KRW, T10Y, Credit, VIX, TED*							
RMB	on-shore CNY		(8)	*[CNY] CNH, CNK, KWD1m, KRW, JPY*							
			(9)	*[CNY] CNH*							
		[CNY] CNH, T10Y, Credit, VIX, TED	(10)	TED	→	CNY	3.57194**	3.87336**	2.81656**	2.17497*	1.95067*
				CNY	→	Credit	3.76118*	2.47433*	*na*	*na*	*na*
				CNY	→	VIX	5.01105**	3.03027**	*na*	*na*	*na*
				CNY	→	TED	*na*	*na*	*na*	*na*	1.86105*
	off-shore CNH		(11)	*[CNH] CNY, CNK,*							
			(12)	*[CNH] CNY, CNK, KWD1m, KRW, JPY*							
			(13)	*[CNH] CNY*							
		[CNH] CNY, T10Y, Credit VIX, TED	(14)	TED	→	CNH	*na*	3.64619**	2.49048*	*na*	*na*
				CNH	→	Credit	3.52004*	3.13441**	3.03439**	2.40199**	2.44063**
				CNH	→	VIX	3.4716*	2.36669*	*na*	*na*	*na*
JPY	[JPY] T10Y, Credit, VIX, TED		(15)	T10Y	→	JPY	*na*	6.90707***	5.24848**	4.46118***	3.72127***
				Credit	→	JPY	*na*	14.591***	13.7139***	10.8005***	8.83545***
				VIX	→	JPY	*na*	3.52556**	2.28807*	*na*	*na*
				TED	→	JPY	3.00953*	6.51429***	4.74023***	4.46049***	4.40631***
				JPY	→	T10Y	*na*	*na*	*na*	*na*	*na*

	JPY	→	Credit	7.08832***	10.0327***	6.64887***	4.78381***	4.11865***
	JPY	→	VIX	6.24512**	4.23995**	2.92458**	2.29145*	1.9398*
	JPY	→	TED	*na*	3.43594**	2.43067*	*na*	2.01376*

* (KWD1m → KRW) the direction of the arrow is the direction of the granger causal relationship. 'na' is no the granger causal relationship does not exist at the significance level. The Granger causal relationship analysis results of analysis models 2, 3, 5 to 9, 11 to 13 are overlapped with model 1

*Korea won dollar and Seoul offshore yuan

Models No. 1 through No. 4 in Table 7 summarizes the variables affecting the Granger causal relationship between the Korean won (KRW) and the Seoul offshore yuan (CNH). The causality analysis among the variables showed that offshore won (KWD1m) was the biggest cause-and-effectiveness for Korean onshore won (KRW). After KWD1m, the onshore won exchange rate was shown to be causal in the order of Hong Kong's offshore yuan (CNH), China's onshore yuan (CNY) and Japan's yen (JPY) at 1% significant level. These results suggest that the offshore won-dollar exchange rate had the greatest impact on the onshore won-dollar exchange rate. And it is in line with the findings of [34] the author who claimed that the influence of the Chinese offshore yuan on the onshore yuan has increased. The unusual was that the Chinese yuan causes more on the onshore won dollar exchange rate than the yen. These results are interpreted as reflecting the largest export-import trade volume of Korea with China. Model 4 analyzed the Granger causal relationship between the won dollar and international financial market variables. The analysis results show that 10-year US treasury bonds, Credit risk spread, and VIX index causes Korean onshore won dollars (KRW) at 1%, 5%, and 1% significance levels, respectively. Models No. 5 through 7 analyzed and summarized the variables affecting causality in offshore won dollars. According to the Granger causal analysis among the variables, the offshore won-dollar exchange rate (KWD1m) has no causal relationship with the onshore dollar, the Hong Kong Overseas Yuan (CNH), the Chinese onshore yuan (CNY) and yen (JPY). Similar to this analysis, the analysis results of VECM model No. 5 showed that offshore won has no long-term equilibrium with the yuan and yen. In [51] the author argued that the depreciation of the exchange rate is an increase in the sovereign credit risk. The analysis results of this paper and in [51] the author reconfirm that when international financial market investors trade offshore won (KWD1m), they trade based on the expected rate of return versus risk considering the country-specific factor only. In addition, offshore dollars are freely traded in overseas markets without being restricted by Korea foreign exchange transaction regulations, and the difference in price at maturity is settled in US dollars. Therefore, the absence of a liquidity problem in offshore won have affected the analysis results. This analysis result is reconfirmed in the Granger causality relationship analysis result of the model 7 variables. International financial market variables such as US 10-year treasury bonds, Credit risk spread, and TED spread were found to cause offshore won dollars (KWD1m) at 1%, 5%, and 5% significant level, respectively. The results of the Granger causality relationship analysis of Models 5 and 7 show that the offshore won-dollar exchange rate has no causal relationship with the yuan and yen but has a causal relationship with international financial market variables. These results provide policy implications that the analysis of future won-dollar exchange rate fluctuations should be approached from the perspective of global investment portfolio investor in international market.

*China Yuan

Models 8 to 10 in Table 7 summarize variables that have a Granger causal effect on the Chinese onshore yuan (CNY). Since the variables used in the Granger causal relationship analysis are the same as those of models 1 to 4, we need to refer the analysis result of model 1. The analysis results of this study showed that the Hong Kong offshore yuan, CNH was causal to the Chinese onshore yuan (CNY) at a significance level of 1%. However, it was found that the Seoul offshore yuan (CNK), the offshore won dollar (KWD1m), the onshore won dollar (KRW), and the Japanese yen (JPY) do not cause the Chinese onshore yuan (CNY). Model No. 10, the analysis of Granger causality between the yuan and the variables in the international financial market showed that only the TED spread Granger causal the yuan at a 5% significance level. However, 10-year US treasury interest rate, the Credit risk spread, the VIX index did not cause onshore yuan. Models 11-14 summarize the variables that have a Granger causal effect on the Hong Kong offshore yuan (CNH). Since the variables used in the Granger causal relationship analysis are the same as those of models 1 to 4, we refer to the ones summarized in model 1. According to the results of Granger causal relationship analysis between variables, it was found that the Chinese onshore yuan (CNY) caused the Hong Kong offshore yuan (CNH) at a 1% significance level. In the model 14, the analysis of the Granger causal relationship between the offshore yuan and the variables of the international financial market variables, it was found that only TED spread caused Hong Kong offshore yuan (CNH) at a 5% significance level. However, similar to the onshore yuan (CNY), the 10-year US treasury interest rate, Credit risk spread, VIX index, did not appear to be Granger causal to the Hong Kong offshore yuan (CNH). These results show that since dollar liquidity required for currency exchange transactions is secured through the interbank money market, the TED spread, which represents the level of liquidity in the interbank money market, has an impact. In the case of Korean won, the offshore won dollar market caused onshore won, but not in the opposite case. In the case of the yuan, however, it was found that the onshore-offshore yuan exchanged its causal relationship. In [35] the author argued that yuan liquidity and risk aversion of investors in international financial markets are the main reasons for creating price differences and generating profit-taking arbitrage transaction between CNY and CNH. They argued that increased inflow of yuan from the onshore CNY market to the Hong Kong CNH offshore market would reduce the difference in CNH-CNY prices, so that the yuan exchange rate would be adjusted based on the onshore CNY, which has more liquidity. According to [50] the statement, China began supplying yuan liquidity to Hong Kong's offshore markets in cooperation with HKMA on November 10, 2014. Therefore, the onshore yuan's Granger causal of offshore yuan is showed. Offshore yuan have been influenced by the ample liquidity of the onshore yuan market. In [34] the author analyzed CNH and CNY markets for the period from September 2010 to August 2013. They argued that CNH's influence on CNY has increased, and argued that CNH Hong Kong offshore yuan has

predictive power over CNY China onshore yuan. The onshore yuan market is limited to investors in the international financial market, while the Hong Kong currency market has no trading limitation. Therefore, as argued by [34] the author, it has been shown that the price-discovery effect of the Hong Kong offshore yuan market may have affected the onshore yuan market. The yuan (CNY and CNH), the SDR currency, is the same yuan, but the onshore yuan and the offshore yuan exchange each other's independent causal relationship and each has a distinct characteristic. These analysis results provide important implications for investment and risk-hedge transactions for companies that use the yuan to trade in import-exports and for investors in the international financial market.

*Japan Yen

Model 15 of Table 7 summarizes the Granger causal relationship of international financial market variables to the Japanese yen. The analysis results showed that all four variables (the 10-year US treasury bond, credit risk spread, VIX index, and TED spread) were responsible for the change in the yen's exchange rate at 1%, 1%, 5%, and 1%, significant level respectively. The aforementioned analysis of the Korean won dollar showed that three variables in the international financial market affect the won-dollar exchange rate. In yuan case, only TED spread affects the yuan's exchange rate. It was expected that the yen, the international currency, would be the strongest in the strength of Granger causality relationship among the variables in the international financial market, Chinese the yuan that is in the process of internationalization, would be the next and the won-dollar, the non-international currency would be the weakest. However, the analysis result showed that the non-international currency, the won, has a stronger Granger causal relationship with variables in the international financial market than that of Chinese the yuan. In [14] the author analyzed the effects of international financial market variables and currency exchange rates on the country CDS premium. They argued that the Granger causal relationship of the international financial market variables and the won-dollar exchange rate that affect the Korea country CDS premium is stronger than that of the China country CDS premium. Therefore, when a shock occurs in the international financial market, the shock could be dispersed and absorbed into the Korea CDS premium and the onshore and offshore won-dollar currency market. However, for China the shock could be concentrated in the CDS premium only. The analysis results that the Korean won dollar has a stronger Granger causal relationship with international financial market variables than the Chinese yuan is in line with their arguments.

5.3. Impulse Response Analysis

The Granger causality test informs the direction of the causal relationship between variables. However, the Granger causality test cannot determine whether the direction of the causal relationship is positive (+) or negative (-). Therefore, the impulse response analysis was performed to measure the impact of a specific variable on other variables. The Impulse response results are summarized in Table 8 for the results of the 1 to 10th and 20th, and 30th days.

Table 8: Impulse response analysis of won, yuan, yen and international financial market variables

Curr-ency		model number		Response of dependents variable to one standard deviation of innovations						
				days	depen-dent	independent variables				
KRW	on-shore KRW	[CNK] KWD 1m, KRW CNH CNY JPY	(1)		LCNK	LKRW1m	LKRW	LCNH	LCNY	LJPY
				1 day	0.00261	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00110	0.00163	0.00062	(0.00087)	(0.00004)	(0.00002)
				3 day	0.00074	0.00412	0.00031	(0.00175)	(0.00052)	(0.00010)
				4 day	0.00034	0.00399	0.00024	(0.00207)	(0.00077)	0.00039
				5 day	0.00020	0.00400	0.00022	(0.00196)	(0.00055)	0.00062
				6 day	0.00002	0.00399	0.00054	(0.00184)	(0.00067)	0.00072
				7 day	(0.00005)	0.00394	0.00081	(0.00176)	(0.00094)	0.00073
				8 day	(0.00011)	0.00399	0.00087	(0.00161)	(0.00107)	0.00071
				9 day	(0.00015)	0.00408	0.00084	(0.00147)	(0.00117)	0.00068
				10 day	(0.00017)	0.00414	0.00080	(0.00138)	(0.00127)	0.00067
				20 day	0.00014	0.00434	0.00060	(0.00061)	(0.00182)	0.00056
				30 day	0.00033	0.00437	0.00051	(0.00032)	(0.00201)	0.00049
		[KRW] KWD 1m, CNK CNH CNY JPY	(2)		LKRW	LKRW1m	LCNK	LCNH	LCNY	LJPY
				1 day	0.00372	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00236	0.00292	0.00006	0.00003	(0.00014)	(0.00007)
				3 day	0.00201	0.00468	0.00009	(0.00001)	(0.00003)	0.00009
				4 day	0.00195	0.00471	(0.00031)	(0.00010)	(0.00001)	0.00027
				5 day	0.00213	0.00479	(0.00022)	(0.00013)	(0.00004)	0.00053
				6 day	0.00227	0.00455	(0.00029)	0.00004	(0.00014)	0.00064
				7 day	0.00240	0.00453	(0.00032)	0.00022	(0.00027)	0.00059
				8 day	0.00245	0.00457	(0.00036)	0.00035	(0.00042)	0.00055
				9 day	0.00244	0.00461	(0.00034)	0.00047	(0.00052)	0.00052
				10 day	0.00241	0.00466	(0.00031)	0.00057	(0.00059)	0.00050
				20 day	0.00236	0.00481	0.00007	0.00125	(0.00106)	0.00038
				30 day	0.00233	0.00484	0.00025	0.00149	(0.00121)	0.00032
		[KRW] KWD 1m	(3)		LKRW	LKRW1m				
				1 day	0.00372	0.00000				
				2 day	0.00235	0.00292				
				3 day	0.00203	0.00474				
				4 day	0.00197	0.00485				
				5 day	0.00194	0.00482				
				6 day	0.00194	0.00481				
				7 day	0.00195	0.00481				

Curr-ency		model number		Response of dependents variable to one standard deviation of innovations						
				days	depen-dent	independent variables				
RMB	on-shore CNY	[CNY] CNH CNK KWD 1m, KRW JPY	(8)		LCNY	LCNH	LCNK	LKRW1M	LKRW	LJPY
				1 day	0.00182	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00210	0.00038	(0.00006)	(0.00005)	0.00000	(0.00004)
				3 day	0.00195	0.00044	(0.00003)	(0.00008)	0.00007	(0.00003)
				4 day	0.00206	0.00041	(0.00006)	(0.00011)	0.00014	(0.00004)
				5 day	0.00209	0.00043	(0.00004)	(0.00009)	0.00011	(0.00005)
				6 day	0.00206	0.00045	(0.00004)	(0.00009)	0.00008	(0.00005)
				7 day	0.00208	0.00046	(0.00004)	(0.00009)	0.00008	(0.00005)
				8 day	0.00209	0.00048	(0.00004)	(0.00008)	0.00008	(0.00005)
				9 day	0.00209	0.00049	(0.00004)	(0.00008)	0.00007	(0.00005)
				10 day	0.00209	0.00050	(0.00003)	(0.00009)	0.00007	(0.00005)
				20 day	0.00211	0.00056	(0.00001)	(0.00011)	0.00006	(0.00006)
				30 day	0.00212	0.00058	0.00001	(0.00011)	0.00005	(0.00007)
		[CNY] CNH	(9)		LCNY	LCNH				
				1 day	0.00182	0.00000				
				2 day	0.00209	0.00038				
				3 day	0.00194	0.00044				
				4 day	0.00205	0.00041				
				5 day	0.00209	0.00044				
				6 day	0.00208	0.00045				
				7 day	0.00208	0.00046				
				8 day	0.00209	0.00048				
				9 day	0.00210	0.00050				
				10 day	0.00210	0.00051				
				20 day	0.00212	0.00061				
				30 day	0.00213	0.00065				
		[CNY] CNH T10Y Credit VIX TED	(10)		LCNY	LCNH	T10Y	CREDIT	LVIX	TED
				1 day	0.00182	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00208	0.00038	0.00009	(0.00005)	0.00005	(0.00011)
				3 day	0.00193	0.00045	0.00012	0.00001	(0.00001)	(0.00017)
				4 day	0.00204	0.00043	0.00010	0.00000	(0.00001)	(0.00019)
				5 day	0.00206	0.00047	0.00011	0.00000	(0.00001)	(0.00015)
				6 day	0.00203	0.00051	0.00013	0.00000	(0.00003)	(0.00014)
				7 day	0.00205	0.00054	0.00013	0.00000	(0.00006)	(0.00014)

				8 day	0.00194	0.00481				
				9 day	0.00194	0.00482				
				10 day	0.00194	0.00482				
				20 day	0.00194	0.00482				
				30 day	0.00194	0.00482				
		[KRW] KWD1m T10Y Credit VIX TED	(4)		LKRW	LKRW1m	T10Y	CREDIT	LVIX	TED
				1 day	0.00371	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00233	0.00284	(0.00000)	0.00003	(0.00004)	(0.00030)
				3 day	0.00201	0.00461	0.00055	0.00036	0.00037	(0.00049)
				4 day	0.00192	0.00460	0.00073	0.00052	0.00047	(0.00040)
				5 day	0.00203	0.00460	0.00112	0.00070	0.00038	(0.00071)
				6 day	0.00215	0.00446	0.00136	0.00063	0.00037	(0.00090)
				7 day	0.00216	0.00431	0.00137	0.00055	0.00032	(0.00093)
				8 day	0.00217	0.00428	0.00138	0.00054	0.00024	(0.00089)
				9 day	0.00217	0.00423	0.00137	0.00053	0.00016	(0.00086)
				10 day	0.00216	0.00418	0.00136	0.00052	0.00008	(0.00084)
				20 day	0.00212	0.00411	0.00142	0.00048	(0.00026)	(0.00083)
				30 day	0.00211	0.00409	0.00144	0.00046	(0.00036)	(0.00083)
	off-shore KWD1m		(5)	na	LKWD1m	LKRW	LCNK	LCNH	LCNY	LJPY
			(6)	na	LKWD1m	LKRW				
		[KWD1m] KRW T10Y Credit VIX TED	(7)		LKWD1m	LKRW	T10Y	CREDIT	LVIX	TED
				1 day	0.00584	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00546	0.00003	0.00062	0.00037	0.00048	(0.00062)
				3 day	0.00509	0.00017	0.00101	0.00081	0.00056	(0.00062)
				4 day	0.00505	0.00047	0.00134	0.00082	0.00057	(0.00092)
				5 day	0.00473	0.00061	0.00141	0.00056	0.00054	(0.00091)
				6 day	0.00462	0.00067	0.00134	0.00058	0.00044	(0.00085)
				7 day	0.00460	0.00067	0.00131	0.00056	0.00033	(0.00079)
				8 day	0.00454	0.00067	0.00128	0.00055	0.00023	(0.00076)
				9 day	0.00452	0.00067	0.00129	0.00054	0.00016	(0.00074)
				10 day	0.00451	0.00067	0.00130	0.00054	0.00009	(0.00075)
				20 day	0.00445	0.00065	0.00137	0.00050	(0.00024)	(0.00075)
				30 day	0.00442	0.00065	0.00140	0.00049	(0.00034)	(0.00075)
JPY		[JPY] T10Y Credit VIX TED	(15)		JPY		T10Y	CREDIT	LVIX	TED
				1 day	0.00597		0.00000	0.00000	0.00000	0.00000
				2 day	0.00593		0.00087	(0.00131)	(0.00036)	(0.00071)
				3 day	0.00583		0.00115	(0.00156)	(0.00027)	(0.00098)
				4 day	0.00589		0.00137	(0.00139)	(0.00011)	(0.00136)
				5 day	0.00580		0.00124	(0.00143)	(0.00003)	(0.00134)
				6 day	0.00578		0.00121	(0.00144)	0.00009	(0.00130)
				7 day	0.00575		0.00124	(0.00141)	0.00021	(0.00127)
				8 day	0.00571		0.00122	(0.00140)	0.00032	(0.00128)
				9 day	0.00568		0.00122	(0.00138)	0.00042	(0.00129)
				10 day	0.00566		0.00122	(0.00136)	0.00050	(0.00130)
				20 day	0.00552		0.00121	(0.00129)	0.00093	(0.00131)
				30 day	0.00549		0.00121	(0.00127)	0.00104	(0.00131)

				8 day	0.00206	0.00057	0.00013	0.00000	(0.00007)	(0.00014)
				9 day	0.00206	0.00060	0.00014	0.00000	(0.00009)	(0.00014)
				10 day	0.00207	0.00062	0.00014	0.00000	(0.00010)	(0.00013)
				20 day	0.00210	0.00073	0.00017	(0.00000)	(0.00020)	(0.00012)
				30 day	0.00210	0.00075	0.00017	(0.00001)	(0.00023)	(0.00011)
		[CNH] CNY CNK	(11)		LCNH	LCNY	LCNK			
				1 day	0.00261	0.00000	0.00000			
				2 day	0.00293	0.00025	(0.00008)			
				3 day	0.00269	0.00032	(0.00014)			
				4 day	0.00248	0.00040	(0.00004)			
				5 day	0.00244	0.00049	(0.00016)			
				6 day	0.00247	0.00049	(0.00017)			
				7 day	0.00245	0.00050	(0.00016)			
				8 day	0.00243	0.00055	(0.00013)			
				9 day	0.00240	0.00056	(0.00011)			
				10 day	0.00237	0.00059	(0.00011)			
				20 day	0.00219	0.00074	(0.00004)			
				30 day	0.00213	0.00079	(0.00001)			
off-shore CNH			(12)	na	LCNH	LCNY	LCNK	LKRW1M	LKRW	LJPY
		[CNH] CNY	(13)		LCNH	LCNY				
				1 day	0.00260	0.00000				
				2 day	0.00291	0.00028				
				3 day	0.00266	0.00035				
				4 day	0.00245	0.00046				
				5 day	0.00242	0.00059				
				6 day	0.00235	0.00080				
				7 day	0.00235	0.00069				
				8 day	0.00237	0.00062				
				9 day	0.00241	0.00064				
				10 day	0.00240	0.00064				
				20 day	0.00231	0.00070				
				30 day	0.00227	0.00072				
		[CNH] CNY T10Y Credit VIX TED	(14)		LCNH	LCNY	T10Y	CREDIT	LVIX	TED
				1 day	0.00260	0.00000	0.00000	0.00000	0.00000	0.00000
				2 day	0.00293	0.00026	0.00002	(0.00001)	0.00006	(0.00019)
				3 day	0.00268	0.00028	0.00004	0.00012	0.00003	(0.00025)
				4 day	0.00255	0.00027	0.00007	0.00009	(0.00003)	(0.00025)
				5 day	0.00251	0.00033	0.00008	0.00007	(0.00007)	(0.00024)
				6 day	0.00248	0.00037	0.00008	0.00008	(0.00010)	(0.00024)
				7 day	0.00244	0.00040	0.00008	0.00008	(0.00012)	(0.00024)
				8 day	0.00240	0.00044	0.00009	0.00008	(0.00014)	(0.00024)
				9 day	0.00236	0.00047	0.00009	0.00008	(0.00016)	(0.00024)
				10 day	0.00233	0.00049	0.00009	0.00008	(0.00017)	(0.00025)
				20 day	0.00211	0.00067	0.00010	0.00008	(0.00021)	(0.00026)
				30 day	0.00203	0.00072	0.00010	0.00008	(0.00020)	(0.00026)

* a negative number in parentheses

*Korea won dollar and Seoul offshore yuan

Table 8 summarizes the impulse response results of 15 analysis models. The analysis results of this study showed that Model No. 1, the Seoul offshore yuan (CNK) reacts positively to the shock from the offshore dollar. It responds positively to shocks from the onshore dollar and yen. However, it reacts negatively to the shock of the yuan offshore and onshore in China. It was found that the Seoul offshore yuan currency (CNK) is most strongly affected by the impulse shock from the offshore won-dollars (KWD1m). It was found that the influence of the offshore won dollar on the Seoul offshore yuan (CNK) was greater than that of the CNH, and CNY. These results are found to be due to the fact that the lack of yuan liquidity in the Seoul offshore yuan market compared to the Hong Kong and Chinese yuan markets. It was found that the Seoul offshore yuan (CNK) market financed insufficient yuan liquidity through offshore won dollar cross rate arbitrage transaction. In addition, the impact from onshore won (KRW) and offshore won (KWD1m), Chinese onshore yuan (CNY), Hong Kong offshore yuan (CNH), and yen (JPY) to the Seoul offshore yuan (CNK) was mostly reflected within 1 to 6 time lag. This result is interpreted as the Seoul offshore yuan (CNK) exchange rate being adjusted over time through arbitrage trading between the foreign exchange markets. Model No. 2 onshore won dollar (KRW) was found to react positively to the shock from offshore won dollars. It reacts positively to shocks from the Hong Kong offshore yuan and yen. However, it reacts negatively from the shock of the Chinese onshore yuan (CNY). The onshore won dollar (KRW) has a positive response to the shock from the offshore won dollar (KWD1m) and the yen and remains constant. The onshore won dollar showed adjustment response to the impact of the offshore yuan, but a constant negative (-) response to the impact from the onshore yuan. The result is interpreted as the fact that it takes time for the onshore yuan to be delivered to the Seoul offshore yuan market through import-export transactions and limited capital market transactions. This can also be seen in the VECM analysis model No. 9 and No. 13 in Table 6 which has optimal analysis time lag was 4 and 6 days. It is also unusual for the onshore won-dollar, KRW to have different impulse reactions of the onshore yuan and offshore yuan. The onshore won-dollar showed an adjustment reaction that reverse the positive sign to negative to the impact from the Hong Kong offshore yuan (CNH). These results are interpreted that since onshore won-dollar markets and Hong Kong offshore yuan markets have no restrictions on foreign exchange trading in general, so the two markets adjust to the external shocks with cross rate arbitrage transactions. However, onshore won-dollar continuously responded negatively to the shock from the Chinese onshore yuan. The reason for such results can be interpreted that because foreign exchange transactions

between the Seoul and China are only export-import related transactions are liberalized but capital market transactions are still limited, so the adjustment responses to the shocks from the outside are different. In [49] the author argued that the impact of the shock from the US dollar currency on Chinese yuan was limited. The onshore won-dollar's negative reaction to the shock of the onshore yuan is in the same line as his argument. However, the direction of onshore and offshore yuan shock transfer to the onshore Korean won is partly opposite. This is interesting in that the impact of the yuan on the onshore won-dollar exchange rate is partially offset by offshore yuan. And these results can be interpreted that the different characteristics of the onshore and the offshore yuan contributed in part to the stabilization of the onshore won-dollar exchange rate. Models 4 and 7 onshore and offshore won dollars responded almost identically to shocks from international financial market variables. The onshore and offshore won dollar responded positively to the impact of the 10-year US treasury bond and credit risk spread. In the VIX index shock, it reacted to positive (+) and then converted to negative (-), and the shock disappeared. It responded continuously negative to the impact from the TED spread. Overall, the Korean onshore and offshore won dollar and the Japanese yen had similar shock reactions to the impact from the international financial market. However, the onshore and offshore yuan showed relatively weaker responses to shocks from variables in the international financial market than that of the won and yen. In [39] the research report argues that the determinants of investment capital inflow to emerging countries are largely influenced by factors on the supplier side, such as the level of US treasury bond rates and preference for safe assets in the international financial market. They argued that a 50bp fall in the 10-year US treasury bond yield would increase the total capital inflow to emerging countries by 13%. Korean onshore and offshore won dollars and Chinese onshore and offshore yuan responded positively to the shock arising from 10-year US treasury bonds. Rising risk free interest rates on 10-year U.S. treasury bonds increase demand for the U.S. dollar and lead to capital outflows from emerging markets, which in turn could weaken emerging economies' currencies. This study also confirmed the results consistent with the IMF research. In [39] the research report argued that the VIX index represents the preference for safe assets in international financial markets, and that if the VIX index falls in half, the total capital inflow of emerging countries increases by 11%. However, the onshore and offshore won dollars were slightly different from the IMF argument. The impact from the VIX index was weaker than US 10-year treasury and credit spread. The reason for such results can be interpreted that, because the impact from the VIX index can be adjusted through the NDF KWD1m transaction and the offshore won dollar (KWD1m) readjusts the onshore won dollar (KRW). This result is also explained as the volume of offshore won-dollar, NDF transactions in the international financial market accounts for about 20% of the total global NDF transactions. The won dollar, the yuan, and the yen all reacted in the same negative direction to the shock from the TED spread. Changes in the TED spread implies changes in the overall liquidity level in the international financial market, or changes in credit risk levels and bond yields. In [45] the author used the TED spread as a measure of hedge funds' liquidity level. In [49] the author argued that the impact of a particular currency is spill over through the interbank money market. As he claimed, this study also reconfirmed that the shock generated by TED spread, a variable in the international financial market, was similarly transmitted to Korea won, China yuan and Japan yen. The reason for such results can be interpreted that because banks in Korea, Japan, and China, which are counterparties in money market transactions, do not have a large difference in financing capacity in the international interbank market, and therefore all respond similarly to the shock from the TED Spread change.

*China Yuan

Table 8 summarizes the analysis results of Model No. 8, Chinese onshore yuan (CNY). CNY was found to react negatively to shocks from offshore won dollars (KWD1m), Seoul offshore yuan (CNK) and yen (JPY). On the other hand, Chinese onshore yuan responded positively to shocks from Hong Kong offshore yuan and onshore won dollar. The Chinese onshore and offshore yuan reacts positive (+) to the shocks generated from each other because the two markets are closely interconnected. The Chinese onshore yuan (CNY) reacted positively to the shock from the Korean onshore won (KRW). This phenomenon is attributed to the fact that the 250 billion dollars import-export transactions volume between South Korea and China. Analysis of models 10 and 14, onshore and offshore yuan, and international financial market variables, showed that the onshore and offshore yuan responded similarly to shocks from international financial market variables. The unusual was that the yuan is the SDR international currency, but the response to the impact from international financial market variables was relatively weak compared to the non-international currencies Korean won and international currency Japanese yen. This phenomenon once again support [49] the author's argument that the impact of the US dollar currency on Chinese yuan is limited.

*Japan Yen

The analysis result of model No. 15, yen, JPY summarized in Table 8 shows that JPY consistently responds to the shocks from international financial markets. The yen responded positively to shocks from 10-year US treasury bonds and credit risk spreads, and negatively (-) to shocks from the VIX index and TED spreads.

5.4. Variance Decomposition Analysis

Variance decomposition analysis is to measure the variance of the prediction error of the VECM analysis model and divide the weight for each variable. Table 9 summarized analysis result.

Table 9: Variance decomposition analysis of won, yuan, yen and international financial market variables

Curr-ency	model number		Variance decompositions of dependent variables								
			days	S. E	depen-dent	independent variables					
on-sho	[CNK]	(1)	10 day	S. E	LCNK	LKRW1m	LKRW	LCNH	LCNY	LJPY	
				0.0134	4.8881	73.8031	2.0152	13.946	3.6724	1.6745	

Curr-ency	model number		Variance decompositions of dependent variables								
			days	S. E	depen-dent	independent variables					
on-sho	[CNY]	(8)	10 day	S. E	LCNY	LCNH	LCNK	LKRW1m	LKRW	LJPY	
				0.0066	95.4783	4.1412	0.0407	0.1492	0.1486	0.0420	

KRW	re KRW	[KRW]	(2)	20 day	0.0202	2.1805	77.7151	2.0285	8.3040	8.1100	1.6619
				30 day	0.0254	1.4867	78.7531	1.7485	5.5444	11.006	1.4613
					S. E	LKRW	LKWD1m	LCNK	LCNH	LCNY	LJPY
				10 day	0.0157	24.6279	73.6158	0.2814	0.3046	0.3747	0.7956
				20 day	0.0234	21.2813	74.2448	0.1611	1.9899	1.6258	0.6969
				30 day	0.0295	19.7039	73.5180	0.1423	3.5161	2.5471	0.5725
		[KRW]	(3)		S. E	LKRW	LKWD1m				
				10 day	0.0156	20.5450	79.4550				
				20 day	0.0227	17.1122	82.8878				
				30 day	0.0280	16.0438	83.9562				
		[KRW]	(4)		S. E	LKRW	LKRW1m	T10Y	CREDIT	LVIX	TED
				10 day	0.0154	22.8658	68.8931	4.8228	1.0169	0.3475	2.0539
				20 day	0.0220	20.6905	69.2345	6.4061	0.9974	0.2285	2.4430
				30 day	0.0270	19.9320	69.1256	7.0975	0.9654	0.2990	2.5806
	off-shore		(5)			LKWD1m	LKRW	LCNK	LCNH	LCNY	LJPY
			(6)			LKWD1m	LKRW				
		[KWD 1m]	(7)		S. E	LKWD1m	LKRW	T10Y	CREDIT	LVIX	TED
				10 day	0.0164	89.9893	1.0681	5.0957	1.2291	0.5770	2.0409
				20 day	0.0223	88.4382	1.4175	6.3739	1.1847	0.3566	2.2291
				30 day	0.0270	87.6089	1.5451	7.0189	1.1424	0.3758	2.3090
JPY	[JPY] T10Y Credit VIX TED		(15)		S. E	JPY		T10Y	CREDIT	LVIX	TED
				10 day	0.0195	88.1914		3.4020	4.6873	0.2073	3.5121
				20 day	0.0273	86.5664		3.7050	4.7074	0.9522	4.0690
				30 day	0.0333	85.6938		3.8121	4.6436	1.5557	4.2948

RMB	re CNY	[CNY]	(9)	20 day	0.0095	94.4583	5.1621	0.0231	0.1807	0.1165	0.0592
				30 day	0.0118	93.8791	5.7306	0.0152	0.2071	0.0970	0.0711
					S. E	LCNY	LCNH				
				10 day	0.0066	95.7357	4.2643				
				20 day	0.0096	94.4312	5.5688				
				30 day	0.0119	93.5308	6.4692				
		[CNY]	(10)		S. E	LCNY	LCNH	T10Y	CREDIT	LVIX	TED
				10 day	0.0066	93.7528	5.4277	0.3049	0.0065	0.0673	0.4408
				20 day	0.0096	91.2001	7.7013	0.4060	0.0032	0.3224	0.3671
				30 day	0.0120	89.8291	8.8248	0.4644	0.0023	0.5514	0.3279
	off-shore CNH	[CNH]	(11)		S. E	LCNH	LCNY	LCNK			
				10 day	0.0081	96.7287	3.0479	0.2234			
				20 day	0.0110	94.3654	5.4780	0.1566			
				30 day	0.0132	92.6297	7.2585	0.1118			
			(12)			LCNH	LCNY	LCNK	LKRW1M	LKRW	LJPY
		[CNH]	(13)		S. E	LCNH	LCNY				
				10 day	0.0081	95.2963	4.7037				
				20 day	0.0112	93.8654	6.1346				
				30 day	0.0135	93.0042	6.9958				
		[CNH]	(14)		S. E	LCNH	LCNY	T10Y	CREDIT	LVIX	TED
				10 day	0.0081	96.9695	1.9251	0.0760	0.0863	0.1644	0.7787
				20 day	0.0109	94.2397	4.1311	0.1265	0.0981	0.4395	0.9650
				30 day	0.0129	92.2013	5.8894	0.1510	0.1067	0.5628	1.0888

*Korea won dollar and Seoul offshore yuan

In Table 9, it was found that the offshore won dollar (KWD1m) accounts for more than 70% of the variance of the onshore won dollar (KRW) and the Seoul offshore yuan (CNK). In [52] the author argued that the volatility of the offshore yuan had a great influence on the exchange rate of the onshore yuan. The results of this study also showed similar results to their argument. The analysis results of this study showed that the onshore won dollar exchange rate is greatly affected by the offshore won dollar. The reason for such results can be interpreted that, First, the won dollar is effectively and practically internationalized in general. Korean won, transactions between onshore and offshore markets are processed through the cross rates arbitrage transactions. Various SWAP transactions using interest rate differential between the Seoul foreign exchange market and the international financial market support this finding. Therefore, the Korea won may be considered as partly internationalized currency. Second, offshore won-dollar (KWD1m) transactions, which have absolute influence on the onshore won-dollar (KRW), have no liquidity problems. This is because offshore NDF won-dollar payments are settled in U.S. dollars at the maturity, so they are not subject to KRW liquidity constraints. Third, it is because investors in the international financial market actively utilize the offshore won-dollar market. This phenomenon is reaffirmed that about 20 percent of the world's NDF currency market transactions are DNF Korea won- dollar. Fourth, Korea is the top 10 global exporters and importers. Therefore, the demand for won-dollar transactions related to import-export settlement and capital market transactions are sizeable in onshore and offshore markets. The analysis results of this study showed that the variance decomposition analysis results between the onshore-offshore won dollar and the variables of the international financial market (analysis models 4 and 7) are similar. In addition, the analysis of variance decomposition of the Korean won, the Japanese yen and the international financial market variables show similarity either. This phenomenon is attributed to the fact that the Korea won has both the characteristics of emerging market currencies and international currencies either.

*China Yuan

The results of analysis of models No. 8-14 in Table 9 show that the onshore-offshore yuan variable explains 92%-95% of the total variance of yuan itself. Compared to the won dollar and yen, the onshore-offshore yuan showed relatively weak explanatory power of other currencies and international financial markets variables. The results of this analysis also reaffirm of [49] the author's argument that the impact of the US dollar is limited in the yuan.

*Japan Yen

The results of the analysis of variance decomposition of the Japanese yen (JPY) of model No. 15 summarized in Table 9 are similar to the results of analysis of Korean onshore and offshore won dollars in analysis models No. 4 and No. 7. The explanatory power of the variables in the international financial market for the won dollar and the yen were similar.

6. Summary and Conclusion

This study analyzed the relationship between renminbi (onshore and offshore), won dollar (onshore and offshore), yen spot exchange rate, and international financial market variables using the VECM model. The analysis period is 3 years from December 1, 2014 to November 30, 2017, and daily closing price data were used. Analysis variables are US dollar-based Chinese yuan (onshore and offshore), Korean won dollar (onshore and offshore), Japanese yen spot exchange rate, and international financial market variables. During the analysis, there were no unusual situations in the international financial market, such as the 1998 Asian financial market crisis or the 2008 global financial market crisis. Therefore, dummy variables were not considered. We used Bloomberg data. The results of the analysis are as follows: First, there was a long-term cointegration equilibrium relationship between yuan (onshore and offshore), won dollar (onshore and offshore), yen spot exchange rate, and international

financial market variables. Second, In the results of VECM analysis of won dollar, yen and yuan, it was found that there is a long-term equilibrium relationship between the Seoul offshore yuan (CNK), Hong Kong offshore yuan (CNH) and the China onshore yuan, (CNY). The result is interpreted as the yuan's liquidity in Seoul's offshore foreign exchange market is relatively insufficient compared to Hong Kong's, and the lack of the liquidity is financed through cross rate transactions between currencies of KRW↔USD↔CNH, CNY. The onshore won dollar (KRW) has a long-term equilibrium relationship in the VECM model from the relationship between the offshore won dollar, the onshore and offshore yuan, and the yen. The explanatory power of offshore won dollars (KWD1m) to onshore won dollars (KRW) using the VECM model was found to be about 50%. And it was found that the offshore won dollar exchange rate has a long-term equilibrium relationship with four variables in the international financial market. As a result, Korea's onshore won-dollar exchange rate was found to be largely influenced by international financial market variables. China's Onshore yuan (CNY) is found to have a long-term equilibrium relationship with the Hong Kong currency (CNH), the onshore and offshore won dollar, and the yen. CNY has also a long-term equilibrium relationship with four variables in the international financial market. Hong Kong's offshore yuan (CNH) has a long-term equilibrium relationship between China's onshore yuan (CNY) and Korea's offshore yuan (CNK). There is also a long-term equilibrium relationship with four variables in the international financial market. However, Hong Kong's offshore yuan has not had a long-term equilibrium with the onshore and offshore won and the yen. The unusual was that although there was a long-term equilibrium relationship between the onshore and offshore yuan, the relationship was weaker than that of the onshore and offshore won-dollar. Overall, the yuan has a weaker long-term equilibrium relationship with other currencies than that of the won dollar. This phenomenon can be interpreted as the yuan is a currency in the progress of internationalization, and still foreign exchange transactions are limited. Commercial export-import cross border yuan transactions are liberalized but capital market transactions are still limited. VECM model analysis of onshore won dollar and onshore yuan have a long-term equilibrium relationship with yen. The reason for such results can be interpreted that the supply of the won dollar, yuan, and yen currency positions are exchanged in the foreign exchange markets of Korea, China and Japan through export-import transactions between Korea, China and Japan. However, the offshore won dollar (NDF, KWD1m) and offshore yuan (CNH) does not show long-term equilibrium relationship with the yen (JPY). This result is interpreted to be because investors in the international financial market, which do not have transaction restrictions, take into account country specific factors only when they trade in currency market. The results of the analysis of the yen VECM model showed that the yen is more correlated with international financial market variables than that of the won dollar and yuan, as expected. Third, the results of the Granger causality relationship between the won dollar, yen, yuan and international financial market variables. It was found that the offshore won dollar (KWD1m) was the biggest causative (Granger causality) to the Korean onshore won dollar (KRW). Subsequently, in the order of the Hong Kong offshore yuan (CNH), the Chinese onshore yuan (CNY) and the Japanese yen (JPY). It was found that the Chinese yuan has slightly stronger Granger causality the Korean won than that of the Japanese yen. The offshore won dollar exchange rate (KWD1m) showed no Granger causal relationship from the onshore won dollar (KRW), Hong Kong offshore yuan (CNH), Chinese onshore yuan (CNY), and Japanese yen (JPY). However, the international financial market variables such as 10-year US treasury bonds, Credit risk spread, and TED spread were found to Granger cause offshore won (KWD1m). Chinese onshore yuan (CNY) was found to Granger cause Hong Kong offshore yuan (CNH) at a 1% significance level. And the reverse case was also significant at the 1% level. As for yuan, onshore and offshore yuan were found to exchange Granger causality relationship. However, it was found that international financial market variables do not Granger cause onshore and offshore yuan. In addition, onshore yuan (CNY) and offshore yuan (CNH) each independently showed a Granger causal relationship with the won dollar and the yen. Therefore, it was found that onshore and offshore yuan have distinct characteristics. This phenomenon is interpreted as the fact that although the yuan is an internationalized currency, only export-import commercial transactions are liberalized but capital market transactions are still limited. The Japanese yen (JPY) has a stronger Granger causal relationship than that of the won dollar and yuan. It was found that three out of four variables in the international financial market have a Granger causal effect on the Korean won dollar exchange rate. However, it was found that only one TED spread has a Granger causal effect on the yuan exchange rate. It was found that international financial market variables had a stronger Granger causal relationship on the non-international currency, the Korean won dollar, than that of the SDR international currency, the Chinese yuan. Fourth, the results of Impulse response analysis. The onshore won dollar (KRW) and the Seoul offshore yuan (CNK) were found to be most strongly affected by the impact from the offshore won dollar (KWD1m). The onshore won dollar (KRW) consistently reacts negatively to the onshore yuan (CNY) shock, but reverse response of positive (+) to negative (-) with different time lag of the offshore yuan (CNH) shock. Therefore, some of the onshore and offshore yuan shocks delivered to the onshore won-dollar (KRW) exchange rate were found to be mutually offset. Overall, Korean onshore (KRW) and offshore won dollars (KWD1m) and Japanese yen (JPY) showed similar impulse responses to shocks from international financial markets. However, the impulse response of the onshore (CNY) and offshore yuan (CNH) to international financial market variables was weaker than that of the won dollar and yen. China's onshore and offshore yuan responded positively to the mutual shock. These results are interpreted as the two markets are closely related to each other as China supplies yuan liquidity to the Hong Kong offshore (CNH) market. Fifth, the result of Variance decomposition analysis. The offshore won dollar (KWD1m) was found to account for more than 70% of the variance of Korean onshore won dollar (KRW) and the Seoul offshore yuan (CNK). International financial market variables have similar degrees of explanation power for the variance of Korean onshore and offshore won and Japanese yen. On the other hand, international financial market variables showed weaker explanatory power for Chinese onshore and offshore yuan compared to Korean won and Japanese yen. These results suggest that the Korea won dollar is closely related to international financial market.

The implications of this study are as follows: First, from the perspective of the international financial market, we compared and

analyzed the won dollar (onshore and offshore), yuan (onshore and offshore), and yen. As a result of analyzing the long-term equilibrium relationship with the VECM model, it was found that the offshore won-dollar exchange rate and the international financial market variables explained more than 50% of the onshore won-dollar exchange rate. Therefore, it was revealed that it is important to monitor the international financial market for changes in the onshore won-dollar exchange rate. Second, overall, the Korean won dollar was found to have greater integration with the international financial market than the Chinese yuan, the SDR international currency. Therefore, it is necessary to consider the characteristics of the won, which is practically and partially similar to the international currency. It also revealed that a new perspective on the Korea won-dollar exchange rate is needed from the perspective of global portfolio investors. Third, the SDR currency, the yuan, is closely related to the onshore yuan (CNY) and the offshore yuan (CNH), but at the same time, each has its own unique market characteristics. Therefore, it was revealed that companies and investors that trade in yuan need to manage currency risks in consideration of the characteristics of the onshore and offshore yuan market. Fourth, it was revealed that for the development of the Korean financial market, it is necessary to prepare a way to actively utilize the won, which has already partially internationalized, in areas where possible, such as SME import-export transaction, investment of sovereign wealth funds. Under the modern monetary economy, where integration with global financial markets is deepening and the role of finance is important, the Korean won will need a partial and managed internationalization process.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

I would like to thank you Graduate School of Software, Open Source lab, Soongsil University.

References

[1] R. Alsakka, O. ap Gwilym, "Foreign exchange market reactions to sovereign credit news," Journal of International Money and Finance, 2012, doi:10.1016/j.jimonfin.2012.01.007.
[2] D.J. Jüttner, D. Chung, W. Leung, "Emerging market bond returns - An investor perspective," Journal of Multinational Financial Management, 2006, doi:10.1016/j.mulfin.2005.05.005.
[3] H.S. Basso, O. Calvo-Gonzalez, M. Jurgilas, "Financial dollarization: The role of foreign-owned banks and interest rates," Journal of Banking and Finance, 2011, doi:10.1016/j.jbankfin.2010.11.018.
[4] M. Brown, S. Ongena, P. Yeşin, "Foreign currency borrowing by small firms in the transition economies," Journal of Financial Intermediation, 2011, doi:10.1016/j.jfi.2010.12.001.
[5] R.A. Meese, K. Rogoff, "Empirical exchange rate models of the seventies. Do they fit out of sample?," Journal of International Economics, 1983, doi:10.1016/0022-1996(83)90017-X.
[6] R. Sangdai, "Analysis of exchange rate characteristics before and after the introduction of floating exchange rates," Bank of Korea, 3–24, 2000.
[7] G. Ortiz, Dollarization in Mexico: Causes and consequences, 1983.
[8] P. Agenor, M.S. Khan, "Foreign currency deposits and the demand for money in developing countries," **50**, 101–118, 1996.
[9] C.M. Reinhart, E. Carmen, K.S. Rogoff, M.A. Savastano, "Addicted to Dollars *," **50**, 1–50, 2014.
[10] M. Lebre de Freitas, "The dynamics of inflation and currency substitution in a small open economy," Journal of International Money and Finance, 2004, doi:10.1016/j.jimonfin.2003.10.003.
[11] K. Metin-Özcan, V. Us, "Dedollarization in Turkey after decades of dollarization: A myth or reality?," Physica A: Statistical Mechanics and Its Applications, 2007, doi:10.1016/j.physa.2007.06.018.
[12] I. Asel, "Financial Sector Development and Dollarization in the Economies of Central Asia," Unpublished Paper (London: European Bank for …, 2010.
[13] Woong Ryeol Kim, "Analysis of RMB Covered Interest Rate Parity and its Implication," The Academic Society of Global Business Administration, **16**(3), 129–159, 2019.
[14] W.R. Kim, M.K. Kim, "Empirical Analysis of RMB Price Discovery Effects between China Sovereign CDS spread, offshore RMB and Global Financial Market Variables: Comparative analysis of China, Japan and South Korea," in Proceedings - 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD 2019, 2019, doi:10.1109/SNPD.2019.8935719.
[15] A.T. Wang, S.Y. Yang, N.T. Yang, "Information transmission between sovereign debt CDS and other financial factors - The case of Latin America," North American Journal of Economics and Finance, 2013, doi:10.1016/j.najef.2013.02.023.
[16] I. Marcelin, I. Mathur, "Financial sector development and dollarization in emerging economies," International Review of Financial Analysis, 2016, doi:10.1016/j.irfa.2016.04.010.
[17] I. Marcelin, I. Mathur, "Financial development, institutions and banks," International Review of Financial Analysis, 2014, doi:10.1016/j.irfa.2013.09.003.
[18] I. Marcelin, I. Mathur, "Privatization, financial development, property rights and growth," Journal of Banking and Finance, 2015, doi:10.1016/j.jbankfin.2014.03.034.
[19] I. Mathur, I. Marcelin, "Institutional failure or market failure?," Journal of Banking and Finance, 2015, doi:10.1016/j.jbankfin.2014.12.018.
[20] E.L. Yeyati, "Financial dollarization: Evaluating the consequences," Economic Policy, 2006, doi:10.1111/j.1468-0327.2006.00154.x.
[21] A. Korinek, "Excessive Dollar Borrowing in Emerging Markets: Balance Sheet Effects and Macroeconomic Externalities," SSRN Electronic Journal, 2011, doi:10.2139/ssrn.967524.
[22] P.E. Guidotti, C.A. Rodriguez, "Dollarization in Latin America: Gresham's Law in Reverse?," Staff Papers - International Monetary Fund, 1992, doi:10.2307/3867472.
[23] A. Kokenyne, J. Ley, R. Veyrune, "Dedollarization," IMF Working Papers, **10**(188), 1, 2010, doi:10.5089/9781455202225.001.
[24] A. Berg, E. Borensztein, Full Dollarization - The Pros and Cons, Internatiol Monetary Fund, 2000.
[25] A. Ize, E.L. Yeyati, "Financial dollarization," Journal of International Economics, 2003, doi:10.1016/S0022-1996(02)00017-X.
[26] P. Honohan, "The retreat of deposit dollarization," International Finance, 2008, doi:10.1111/j.1468-2362.2008.01223.x.
[27] International Monetary Institute, "The internationalization of the Renminbi," The Internationalization of the Renminbi, **9781315884**, 1–221, 2013, doi:10.4324/9781315884967.
[28] J. a Frankel, "NO SINGLE CURRENCY REGIME IS RIGHT FOR ALL COUNTRIES OR AT ALL TIMES," Time, 1999.
[29] G. Michalopoulos, Global Divergence in Trade, Money and Policy-the internationalization of the euro: trend, challenges and risks, 2006.
[30] S. Hyun, S. Lee, "On the Determinants of Currency Internationalization: Implications for Korean Won Internationalization (in Korean)," Economic Analysis, **19**(1), 43–75, 2013.
[31] C. won Park, Essays on International Monetary System Issues, Hanyang University, 2012.
[32] D.W.Y. Leung, J. Fu, "Interactions between CNY and CNH Money and Forward Exchange Markets," SSRN Electronic Journal, 2014, doi:10.2139/ssrn.2448569.
[33] F.X. Diebold, K. Yilmaz, "Measuring financial asset return and volatility spillovers, with application to global equity markets," Economic Journal, 2009, doi:10.1111/j.1468-0297.2008.02208.x.
[34] Y.W. Cheung, D. Rime, "The offshore renminbi exchange rate: Microstructure and links to the onshore market," Journal of International Money and Finance, 2014, doi:10.1016/j.jimonfin.2014.05.012.
[35] M. Funke, C. Shu, X. Cheng, S. Eraslan, "Assessing the CNH-CNY pricing differential: Role of fundamentals, contagion and policy," Journal of International Money and Finance, 2015, doi:10.1016/j.jimonfin.2015.07.008.
[36] M.D. Bordo, C.M. Meissner, M.D. Weidenmier, "Identifying the effects of an exchange rate depreciation on country risk: Evidence from a natural experiment," Journal of International Money and Finance, 2009, doi:10.1016/j.jimonfin.2008.10.004.

[37] M. Schneider, A. Tornell, "Balance sheet effects, bailout guarantees and financial crises," Review of Economic Studies, 2004, doi:10.1111/j.1467-937X.2004.00308.x.
[38] I. Petrova, M.G. Papaioannou, D. Bellas, "Determinants of Emerging Market Sovereign Bond Spreads: Fundamentals Vs Financial Stress," IMF Working Papers, 2010, doi:10.5089/9781455210886.001.
[39] IMF, Recent Experiences in Managing Capital Inflows—Cross-Cutting Themes and Possible Policy Framework, 2011, doi:10.5089/9781498339353.007.
[40] M.A. Segoviano Basurto, C. Caceres, V. Guzzo, "Sovereign Spreads: Global Risk Aversion, Contagion or Fundamentals?," IMF Working Papers, 2010, doi:10.5089/9781455200795.001.
[41] J. Hilscher, Y. Nosbusch, "Determinants of sovereign risk: Macroeconomic fundamentals and the pricing of sovereign debt," Review of Finance, 2010, doi:10.1093/rof/rfq005.
[42] C. Moser, "The Impact of Political Risk on Sovereign Bond Spreads - Evidence from Latin America," Proceedings of the German Development Economics Conference, Göttingen 2007 / Verein Für Socialpolitik, Research Committee Development Economics, 2007.
[43] L.E. Kodres, K. Hartelius, K. Kashiwase, "Emerging Market Spread Compression: Is it Real or is it Liquidity?," IMF Working Papers, 2008, doi:10.5089/9781451868722.001.
[44] J. Pan, K.J. Singleton, "Default and recovery implicit in the term structure of sovereign CDS spreads," Journal of Finance, 2008, doi:10.1111/j.1540-6261.2008.01399.x.
[45] M. Teo, "The liquidity risk of liquid hedge funds," Journal of Financial Economics, 2011, doi:10.1016/j.jfineco.2010.11.003.
[46] E. Warshaw, "Parallel Currency Markets and the Monetary Exchange Rate Model: A VECM Application to Turkey Over 1987–1998," Eastern European Economics, 2016, doi:10.1080/00128775.2016.1194215.
[47] L.S. Lee Jaeha, "Causal relationship between the KRW onshore spot market and the offshore NDF KRW spot market," Korea Financial Management Association, **17**(2), 211–227, 2000.
[48] S. Johansen, "Statistical analysis of cointegration vectors," Journal of Economic Dynamics and Control, 1988, doi:10.1016/0165-1889(88)90041-3.
[49] L.F. and I. Yu, A STUDY ON THE TRANSMISSION OF MONEY MARKET TENSIONS IN EMEAP ECONOMIES DURING THE CREDIT CRISIS OF 2007 – 2008, 2009.
[50] HKMA, Introduction of Intraday Repo under the Renminbi Liquidity Facility, HKMA, 2014.
[51] G. Ferrucci, Empirical Determinants of Emerging Market Economies' Sovereign Bond Spreads, 2003, doi:10.2139/ssrn.597422.
[52] Y.W. Cheung, C.H. Hui, A. Tsang, "The RMB central parity formation mechanism: August 2015 to December 2016," Journal of International Money and Finance, 2018, doi:10.1016/j.jimonfin.2018.04.011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 746-753 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Framework to Align Business Processes: Identification of the Main Features

Joaquina Marchão[1,*], Leonilde Reis[2], Paula Ventura Martins[1]

[1]*Faculty of Sciences and Technology, Faro, University of Algarve, 8005-139, Portugal*

[2]*School of Business Administration, Polytechnic Institute of Setubal, 2014-503, Portugal*

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 04 March, 2021
Online: 27 March, 2021

Keywords:
ITIL
COBIT
Business Process

ABSTRACT

Information and Communications Technologies are developing faster today than ever before, giving an important contribution to the global economy. Organizations in developed and developing economies explore new technologies to gain advantage and add value. That evolution also brings an increasing complexity to the organizations' management. The alignment of organizational practices with international standards and best practices worldwide accepted in this domain is a relevant topic. To identify gaps in Information and Communications Technologies management area, a brief analysis of international standards will be considered in the state-of-the-art. Considering that Information Technology Infrastructure Library and Control Objectives for Information and related Technology are the most used in the literature review, this paper will propose an Information and Communications Technologies management framework based on those two standards. The approach pretends to solve some gaps found in process alignment, continuing improvement of Information and Communications Technologies services in the context of the organization, driving stakeholder satisfaction and cost optimization. Concluding, the final goal of this paper is to present the framework features analysed, to allow an integrative and multidisciplinary vision, leading to cost optimization, increasing communication, and stakeholder satisfaction.

1. Introduction

This work is an extended version of the paper [1] originally presented in 2020 at the 15th Iberian Conference on Information Systems and Technologies (CISTI).

Information and Communications Technologies (ICT) is one of the main motors of economic development in the XXI century and a success factor of strategic business evolution. In this context, the theme of ICT management and governance processes in organizations is considered relevant, especially on the alignment of organizational objectives and ICT operational efficiency and management of services based on stakeholders needs. A company must implement information technology that is oriented to customers and stakeholders to be competitive [2]. Applications have become an integral part of business strategies while creating new opportunities for alliances and collaborations [3]. Core assets of organizations, business processes shape the functioning and efficiency of organizations [4]. Process modelling guidelines are a valuable instrument for increasing the quality of process models [5], and the human factor is of great importance in Business Process Management (BPM) implementation [6]. Early involvement of stakeholders from top management to the operational level is essential for successful implementation of BPM [7]. The way processes are designed and performed affects both the quality of service that customers perceive and the efficiency with which services are delivered [8].

Managing and using IT services is considered very important for modern businesses to improve their performance [9]. The management of IT services is increasingly driven by the necessity of cost reduction and quality improvement. As a result, new organizational models have been created and implemented aiming at achieving economies of scale, while improving customer satisfaction and experience has been one of the main focuses of recent research in Service Science [10]. Information Technology (IT) grows very rapidly and affects the success of the business lifecycle of an organization [11]. The constant evolution and diversity of technologies and components used in hardware, software and communications networks [12], results on complex management and governance processes in ICT departments

*Corresponding Author: Joaquina Marchão, Email: jamarchao@gmail.com

involving different knowledge and different teams. ICT services are always changing, seeking to respond to business' demands. The growing complexity of customer needs is one of the prevailing problems faced by IT enterprises at present. At the same time, quick response to unexpected problems and externally imposed requirements are testing the IT change management [13]. Additionally, the high availability of information technology (IT) applications and infrastructure components is a significant factor for the success of organizations because more and more business processes depend entirely on IT services [14].

The alignment of organizational practices with international standards and best practices worldwide accepted on this matter is a relevant topic. Among other frameworks and international standards, it was selected the Information Technology Infrastructure Library (ITIL) [15] and Control Objectives for Information and related Technology (COBIT) [16]. Based on those best practices, some models or frameworks are used to manage the complexity of organizations [17]. IT service delivery should be done in a cost-efficient manner. Combining ITIL and COBIT can be valuable for organizational goals. ITIL should be used to define strategies, plans, and processes, on the other hand, COBIT will be used for metrics, benchmarks, and audits [18].

ICT management and governance thematic involve multidisciplinary knowledge. This work is based on a real context analysis of stakeholders from IT support area, and related operational processes and flows. The proposed framework pretends to solve some gap alignment between ITIL and COBIT. The main goal of this paper is to identify the framework features to allow the alignment between several process areas. Operational efficiency and optimization of ITC cost of service, and to stakeholder's satisfaction are other benefits.

The paper presents 5 sections. The first section is a brief introduction to the problem under study. Section 2 presents the state-of-the-art related to this theme. Section 3 shows an overview of the initial processes. In section 4 the framework features are identified based on the processes under analysis and respective interactions. Section 5 includes conclusions and perspectives for future work.

2. State-of-the-Art

ITIL and COBIT are the international most implemented IT frameworks in organizations around the world. ITIL has a focus on IT departments and management services, and COBIT has a focus on organizations' objectives and business control. It was considered to analyse both ITIL and COBIT because these are the most used international best practices. It is expected the framework under construction could solve gaps found in the alignment processes of ICT services for continuous improvement, leading to stakeholders' satisfaction and cost optimization. A brief description of the international standards considered relevant to the framework is presented below.

2.1. ITIL

ITIL's best practices are oriented towards ICT Service Management through the stages defined in ITILv3, the life cycle management of ICT services are Strategy, Design, Transition, Operation, and Continual Service Improvement. Figure 1 shows the stages of ITIL related to the scope.

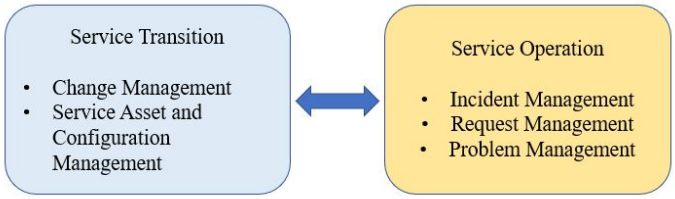


Figure 1: ITIL v3, stages selected within the scope of this paper, adapted from [15]

Stages of Service Operation (components: incident management, request management, and problem management) and Service Transition (components: change management and service asset and configuration management) are used by ICT service support. Stakeholders perceive the service value in the Service Operations stage. Data collection and analysis began with reported incidents and requests for new service: which services were more affected by incidents and why. Consequently, the analysis covered the processes, technology, and people involved. After reflecting, we concluded that the scope should not only be based on this stage but also include others, such as the Service Transition stage, to understand the changes that occur in the production environment and manage the impact in related service assets and configurations, on a day-to-day work basis with stakeholders. ITIL presents good practice guidelines for ICT service management, with scope in the life cycle of services. The implementation of this framework is complex because the literature provides only some general guidance [19]. Implementing ITIL is usually long, expensive, and risky [20], and should require an organizational change affecting procedures, functions, and common attitudes within the organization [21]. The alignment between IT business processes and business objectives is important to achieve stakeholder's satisfaction.

2.2. COBIT

COBIT [16] is a reference model of Governance and Management Objectives, allowing the alignment of ICT management with organizational and business objectives. COBIT defines the components underlying the decisions to be made, how and by whom. This model does not guide what strategies to follow, what technologies to adopt, or what the best architecture to choose. It can address results in three different roles: managers, users, and auditors. Managers seek risk assessment and control of investments; users seek the quality of services; and auditors seek to assess ICT management work. COBIT defines factors and components that should be considered by the organization to build and maintain an appropriate governance system: processes, organizational structures, policies and procedures, information flows, culture, ethics and behaviour, competencies and capabilities, services, applications and infrastructure of the organization, and not just ICT. This framework treats information and related technologies as assets, which must be managed. It presupposes a mental change in the organization, both at the level of ICT managers and business managers, and the latter is responsible for ICT related assets.

Figure 2 shows the COBIT objectives related to the scope.

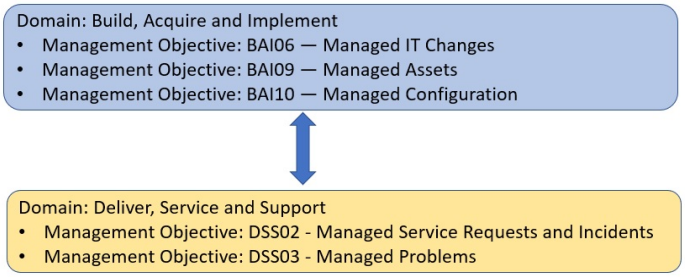


Figure 2: COBIT, Objectives selected within the scope of this paper, adapted from [16]

COBIT presents numerous objectives towards the alignment between the organization and IT. It has a transversal vision for the organization. The interceptions between ICT management objectives and organizational objectives allow the definition of metrics to achieve strategic objectives with standards. For its implementation, it requires cooperation between the different responsibility levels. This framework allows you to define a strategy to generate maximum value creation through IT investments. One of the basic elements in management is the creation of value added [22].

2.3. ITIL vs COBIT

ITIL is the framework most used for IT service management in IT departments. COBIT is the framework most used for governance and management concerning organizations' objectives. It was designed to support the audit and evaluation of IT activities. The sponsorship of the board of directors is important. IT governance has become critical to successful business operations. Many organizations implement IT governance to ensure that they align their IT strategy with the overall business goal [23]. Both can be used by different organizations' types and dimensions, and organizations usually begin implementing ITIL before implementing COBIT. ITIL and COBIT can be applied to improve the quality of service [24]. By planning a mature IT governance implementation of IT services is expected to do well and the embodiment of good IT Governance [25]. Table 1 presents ITIL and COBIT analysis.

Table 1: ITIL and COBIT analysis

Parameter	ITILv3	COBIT19
Scope	Best practices for IT service management, implemented by any organization providing internal and external IT services. Focus on the service lifecycle, value, and customers.	Best practices for strategic teams and people responsible for audit and compliance. Focus on governance and management objectives, enterprise goals and stakeholders' needs.
Advantages	Improve management and utilization of IT resources.	Integration with the ITIL framework and other models.
	Increase efficiency by eliminating redundant tasks and standardizing concepts.	Alignment between governance and management business objectives with IT governance and management
	Alignment of services with business needs.	
	Optimize IT costs.	
	Improve the quality of service delivery.	
	Increase stakeholder confidence.	
Gaps	Does not contain detailed process maps.	Does not guide how to assess processes.
	Does not define measures for process improvement.	Does not provide a "road map" for continuous improvement.
	Difficult implementation.	Difficult implementation.
	Does not provide working instructions.	
	Does not align IT processes and business objectives.	

The objective of the work is to develop a framework to solve organizational problems, and gaps identified on ITIL and COBIT, in particular: describing some process maps, how to assess them and how to align IT with business objectives.

Table 2 presents the alignment between ITIL processes and COBIT objectives, for the scope of this paper.

Table 2 : Alignment between ITIL and COBIT

ITIL v3	COBIT19
Incident and Request Management	DSS02 – Managed Service Requests and Incidents
Problem Management	DSS03 – Managed Problems
Change management	BAI06 – Managed IT Changes
Service Asset and Configuration Management	BAI09 – Managed Assets BAI10 – Managed Configuration

Table 2 shows that only Service Asset and Configuration process of ITIL has two different objectives in COBIT, the other processes have one corresponding in COBIT objectives.

The analysed related work has focus on: tool implementation and assignment of responsibilities [26]; tool implementation, indicators, and management metrics [27]; business process integration, patterns [28]; business process integration, an approach quality-oriented [29]; and a survey presenting the state-of-the-art about business process integration and exploring its trends [4]. The paper with the survey demonstrated that most integrated stakeholders' views are obtain by a notation. In conclusion, it can be argued that there are no business process approaches integrating the stakeholders' perspective. The proposed framework presented in this paper is oriented on the

alignment of business process definition and modelling, with continuous improvement.

3. Initial Processes

Figure 3 presents the original processes and flows. Incident and request management process; change management process and service asset and configuration process.

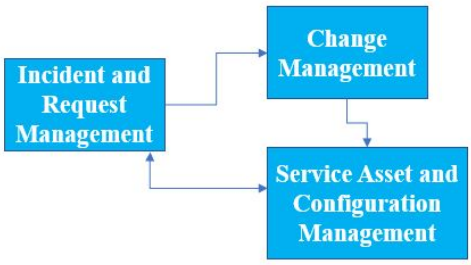


Figure 3: Original processes and flows

Incident/request management process and the service asset and configuration management process are integrated but change management process is managed in a different technological platform. The change management process is oriented to software development and managed in the enterprise project management and planning system. Without technological integration, other mechanisms were applied between the change management process and the incident/request management process, and the change management process and the service asset/ configuration management process to allow the information flow.

4. Identifying framework features

The objective of improving stakeholder satisfaction and optimizing costs for frontend area of service support was the basis to start the research of operational and standard literature. A project was developed to structure the information, organize resources, and control the achievement of the goals. The final phase will propose a framework for alignment and continuous management improvement within the context of ICT departments, to be validated by other organizations. In fact, knowing international best practices is not a guarantee of success because these models present what should be implemented, but not how to do it. The implementation exercise is a challenge in which many organizations fail [30].

The proposed approach is based on ITIL and COBIT frameworks yet adjusted to organizational specificities, considering literature review, and based on professional experience in the area.

The basic artifacts were directly observed for the collection of data. Third-party contracts for service support were analysed in detail, identifying resources used to perform the support function, skills and technical competencies, time-schedule allocation; technical system implemented and functionalities used to support the frontend area, number of incidents, problems and request services recorded per period, time used to solve a ticket and other data recorded that can be used to resource management; service catalogue and services available or under support; normative literature available in the organization and customers relationship; procedures and technical documentation used by technicians, and knowledge used to share relevant information.

Daily activity monitoring allows verifying procedural compliance, including attendance skills, technical knowledge, and efficiency in support; receipt of requests and records in the support system; workflow and work organization; priority management; relationships between teams of different levels of support and knowledge management. Customer complaints analysis provides an external perspective and perception of the service provided.

Literature review related with business processes management and best practices models ITIL and COBIT was carried out, informal interviews with key users, technicians, and external specialists in IT management area. Bottom-up analysis was important to find out what could be done best, which processes could be reviewed, and at what point. With the analysis of frequent points of failure, related services and the associated responsibilities matrix, adjustments could be defined and planned in the internal organizational work model. The software system used in support activities, with workflow associated with resolution of each category, the procedures in use to solve the issues and the human resources involved, were analysed to contribute to processes' review and to improve alignment for efficiency.

The "adopt and adapt" procedure is widely accepted as the best approach to ITIL [21]. With focus in the service operation' stage after analysis and validation, the operational support processes were redefined, such as incident and request management, and problem management, considering these are the support function's main processes.

As the support function is based on assets and configurations underlying the supported services, effective service management is considered essential, knowing changes are being implemented and anticipating impacts in the production environment - where the end-user support function is developed. Asset configurations and change management processes were also redefined. These are associated with the service transition stage in the ITIL framework. The approach proposes an integrative and multidisciplinary vision to standardize concepts, alignment and process improvement, procedures, operational relationships, and other communications difficulties between department areas and stakeholders.

Figure 4 presents the features defined in this paper. It was identified with three main operational areas and four main processes and related flows.

The ICT Infrastructure Management area has, among others, the responsibility of specialized support of second intervention level – expert team, which includes maintenance and technological evolution of ICT communications' infrastructures - voice and data, monitoring and administration of servers that support corporate systems, implementation and management mechanisms to control infrastructures' security - logical and physical including data centres and maintenance of ICT recovery plan in case of disaster.

Application and Systems Management area is responsible for corrective maintenance and evolution of production systems or new solutions development that respond to business or regulatory needs, managing the impact of the changes in the existing systems architecture and related interfaces.

The responsibilities of ICT Support Management area focus on the Helpdesk function, frontend team, basic resolution support the first intervention level in data network access, roll-out of workstations, and support for peripherals, software configurations, installations, and upgrades.

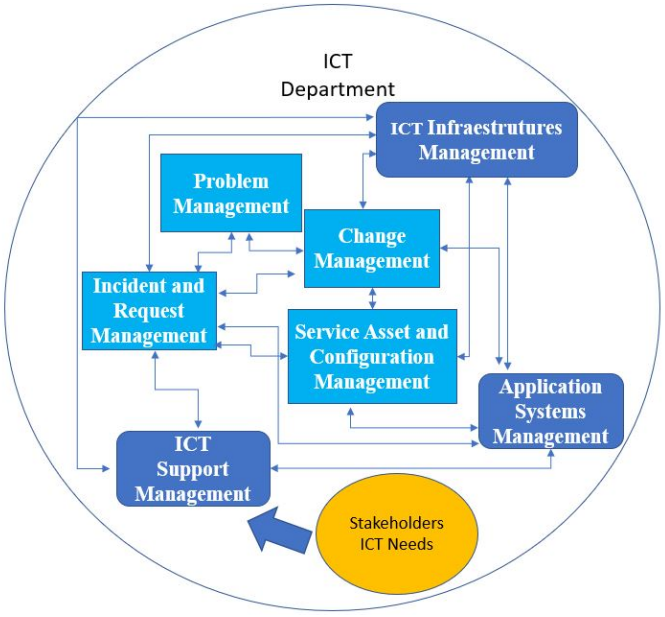


Figure 4 - Operational areas and related processes

Should be noted in Figure 4 that infrastructure management and application management share processes of change, asset and configuration management, and both interact with support management through the incident and request management process.

Problem management is a distinct process, which aims to solve complex situations. The concept is to activate and manage the problem as a project, bringing together the necessary technical teams with the appropriate skills and expertise, to find a satisfactory answer or resolution. From experience, it is expected that it will be impossible to solve all the problems, sometimes it could end up unsolved, with a temporary resolution or a workaround solution.

The bidirectional flow presented in Figure 4 is based on the necessary and constant feedback to optimize and adjust the processes, in a continual improvement cycle.

4.1. ICT Support Management

ICT Support Management starts on helpdesk, the frontend area of the IT department. ICT service requests and incident management require hard work for the team in terms of human relationship and flexibility in new situations adjustment. This support can be done on-site or remotely. The performance of the team is closely related to the level of technical knowledge but also the experience in the existing technology platforms in the organization. As support begins in the service, it is relevant to sensitize these resources to the theme of customer relationships, by the potential impact on the image of the department and the organization. The incident and request management process is what best translates daily operations of supporting areas. Reported incidents are classified by priority levels according to the impact and urgency of each situation. The resolution underlies the procedures and responsibilities at the different levels of intervention.

Repeated errors or incidents for which there is no known solution are managed in the problem management process lack detailed investigation and analysis, and technical experts with different skills can intervene. Control of the infrastructure components and systems that make up ICT services and their relationships is carried out in the service asset and configuration management process. In this process, changes to configurations and assets installed in production environments are managed, following solutions found for problems or newly installed assets.

Evolutionary or corrective maintenance interventions, which lead to the implementation of changes in the production environments that support the services provided, should be managed in the change management process, with the inherently associated risk management, by the disruptive impact they can cause on the services and consequent operation of the organizations.

4.2. ICT Technological Management of Infrastructures and Systems

Technological Management of Infrastructures and Systems plays a central role in the continuity of the organization itself. These teams ensure the availability of the most appropriate technological resources at every moment, helping the development and speeding up business processes.

The approach presented, having a multidisciplinary vision, powers the aggregation of several areas of knowledge, namely in ICT infrastructure management, application management, and support management areas. Security management is also incorporated, given its cross-cutting nature across all areas, and is increasingly relevant. In this context, it is considered essential an integrative approach to support the organization's business strategy, providing integrity, availability, confidentiality, and ubiquity to one of the organization's most valuable assets – information.

For the approach design, the ITIL and COBIT standards were analysed. The contribution of ITIL to the design of the approach will enhance the introduction of good practice in the operational management of ICT services when ITIL is well implemented, and ITIL enables organizations to provide services with great efficiency, quality, and cost reduction [31]. Framework COBIT contributes to the organization's good practices in governance, information management, and ICT management.

4.3. Improved processes

Business Process Management (BPM) is the art and science of overseeing how work is performed in an organization and it is not about improving the way individual activities are performed but about managing entire chains of events, activities, and decisions that ultimately add value to the organization, and its customers. These chains of events, activities, and decisions are called processes [8].

Processes underlying this approach aim to represent the various iterations in the context of ICT assistance. It should be

noted that the processes interact with each other, depending on the situation under analysis.

The following processes were analysed: Incident and Request Management, Problem Management, Change Management, and Service Asset and Configuration Management, as shown in Figure 5. It presents the flows between the different processes involved in the operational management of IT services support. The communications between the processes are bidirectional to allow a continuous improvement.

Figure 5 shows the processes within the scope of this paper.

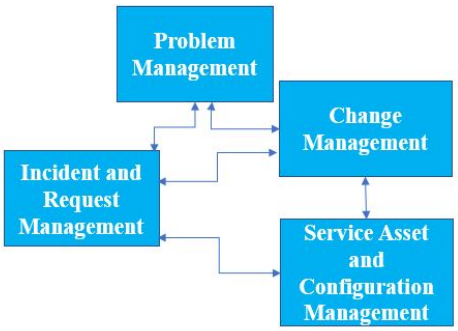


Figure 5: Processes and related flows

Incident and Request Management

The Incident and Request Management Process could be the first process to be implemented [32]. This process can begin with inputs from monitoring tools or stakeholders contacted through self-service portals, emails, or phone calls. Frontend teams are working to resolve anomalies and restore the normal function of services as quickly as they can. This is the main process for IT frontend support, with direct contact with stakeholders, asking for assistance or asking for a new service. This process allows, e.g., assistance optimization and time resolution control. Through this process, the stakeholders interact with the technical team and form their opinion on the IT services performance in general [33]. Incident and request management interacts with the change management process to send feedback when some changes are needed. Incident and request management process also communicates with problem management process whenever a resolution is unknown or cannot be processed by frontend team.

Figure 6 presents the primary process in the frontend support area, incident, and request management process.

Using a prioritization matrix, some incidents can be classified with high urgency and severe impact on business processes. These cases were treated as major incidents and a specific sub-process was defined.

Major Incident Management

Major Incident Management process, presented in Figure 7, is a sub-process of Incident and Request Management process, specifically defined to manage incidents with highest priority. A major incident demands a quick response from the technical team. A senior supervisor could lead the situation, control the status of resolution and be aware of risks and challenges the situation demands, managing the communication plan to the board of managers and sometimes also the communication outside of the organization. After the conscient decision, decide if or when to apply the disaster recovery plan to restore the services and put the company working again with basic resources.

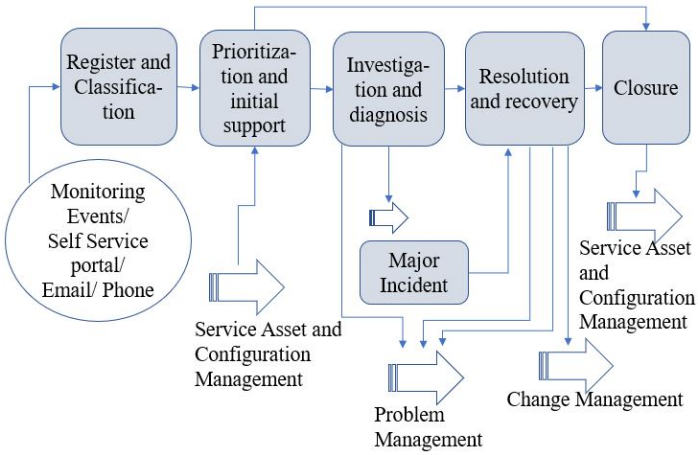


Figure 6: Incident and Request Management Process

Figure 7 shows major incident management process.

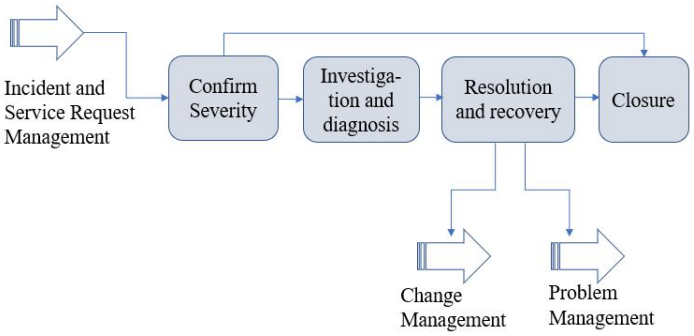


Figure 7: Major Incident Management Process

The major incident process interacts with the change management process to send feedback after resolution, to avoid similar incidents in the future. The major incident process interacts with the problem management when the resolution needs deeper analysis and resolution with expert teams.

Problem Management

Through the problem management process an expert team with different skills analyse and solve complex incidents. This process interacts with incident and request management, when the frontend team cannot resolve an incident based on existing procedures or known errors, the issue is transferred to an expert team. Some inputs from the change management process can be useful to solve problems and after conclusions some data could be sent to the change management process, e.g., to document and correct errors. Every problem analysed is managed as a project, with specific tasks and schedules for each expert team. Periodic meetings are planned to discuss solutions and to align following strategies. The maturity of the teams and their knowledge of IT architecture implemented are important factors for the success of the operationalization on this process.

Figure 8 presents the scheme of the problem management process.

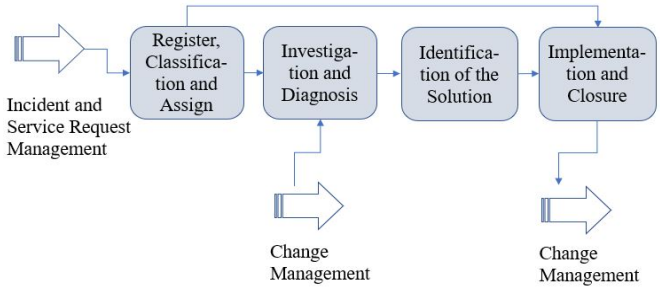


Figure 8: Problem Management Process

Change Management

The change management process ensures documentation and control of changes in software development and systems implemented in the production environment. This includes changes in software applications and other ITC configuration items such as network infrastructure and hardware that supports the systems used.

Figure 9 presents the scheme for the change management process.

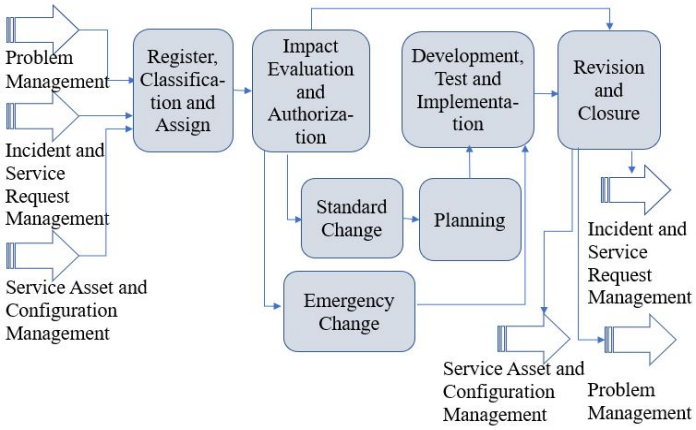


Figure 9: Change Management Process

The change management process can be triggered by the problem management process, incident and service request management process, and service assets and configurations management process.

Service Asset and Configuration Management

Service asset and configuration management is a process to manage changes of infrastructure assets and other controlled assets in production environment. Multiple data are structured in a database of configuration items and the information is used to support other processes. Figure 10 presents the scheme of the Service Asset and Configuration Management process.

The service asset and configuration management process communicate with incident and request management process, when the technical support team needs to check some data to make a diagnosis e.g., equipment characteristics or software versions. Communicates with the change management process when the team update the database configuration items, e.g., manage software licenses, upgrade software versions, or manage equipment lifecycles in the production environment.

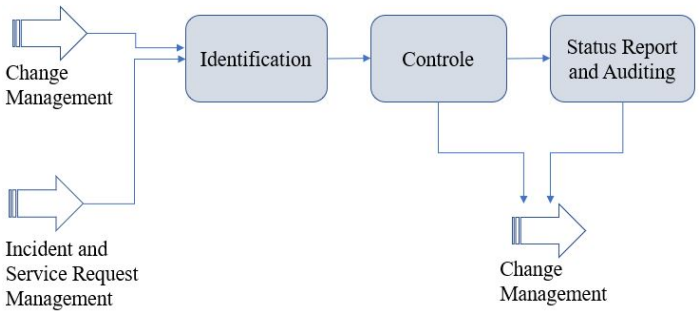


Figure 10: Service Asset and Configuration Management Process

4.4. Framework features

Table 3 summarizes the framework's steps developed.

Table 3: Phases and steps included in framework

Phase	Steps
Initiation	Identify the problem
	Communication with stakeholders (managers)
	Research resources (AS IS processes analysis)
	Problem description, detailed (involve strategic business objectives and alignment metrics)
	Preparation of processes proposal TO BE, applying ITIL and aligning with business objectives (strategy) through COBIT
Planning	Communication with stakeholders (key-users) to detail functional needs
	Communication with operational teams (technicians) to identify specific needs and constraints regarding technology
	Business processes definition TO BE
	Business processes validation (key-users)
	Alignment validation between business processes defined with strategy (managers)
	Define alignment metrics and indicators between strategy and processes

The framework is under development. Table 3 presents the main steps of the initiation and planning phases. The execution, monitoring and closure phases will complete the framework.

At this moment the project lifecycle is not concluded, nevertheless it is important to refer that in the following phases, not detailed in this proposal, it could be necessary adjust the plan to reach stakeholders objectives.

5. Conclusions and Future Work perspectives

Currently, ICT governance and management are inseparable from Corporate Governance. Constant alignment and focus on stakeholders' needs constitutes the basis for organizations' optimization costs. ICT departments can provide solutions that enhance the business or create new business opportunities.

Based on the analysis of international best practices ITIL and COBIT we intend to contribute to aligning IT goals with business goals, a departmental but also organizational view. The framework under construction will present the mechanisms to solve some GAPs found in ITIL and COBIT, in particular: describing some process maps, how to assess them and how to align some business processes. Using ITIL to organize work and manage IT services will improve operational efficiency. Using COBIT from its top-down perspective, the alignment of business goals to IT goals can be achieved and controlled evidencing in that way IT value to business, thus enabling a multidisciplinary and integrative strategic vision.

In future work we will carry out a systematic review of the literature that analyses other approaches. Based on the contributions of this paper and the future Systematic Literature Review, it is intended to present the final proposal of framework implementation and validate it in an organizational context to achieve stakeholder satisfaction and cost optimization.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] J. Marchão, L. Reis, P.V. Martins, "Framework de Alinhamento de Áreas e Processos de Negócio das TIC," in 15ª Conferencia Iberica de Sistemas y tecnologias de Informacion (CISTI), 2020.

[2] I.K. Raharjana, S. Susmiandri, A. Justitia, "Applying IT Services Business Relationship Management on Security Outsource Company," Proceeding of the Electrical Engineering Computer Science and Informatics, **5**(5), 2018, doi:10.11591/eecsi.v5i5.1622.

[3] A. Sahid, Y. Maleh, M. Belaissaoui, Information System Evolution, Emerald Publishing Limited: 29–66, 2020, doi:10.1108/978-1-80043-810-120211004.

[4] R. Belchior, S. Guerreiro, A. Vasconcelos, M. Correia, "A survey on business process view integration," ArXiv, 2020.

[5] D.T. Avila, R.I. dos Santos, J. Mendling, L.H. Thom, "A systematic literature review of process modeling guidelines and their empirical support," Business Process Management Journal, **27**(1), 1–23, 2020, doi:10.1108/BPMJ-10-2019-0407.

[6] M. Hrabal, D. Tuček, V. Molnár, G. Fedorko, "Human factor in business process management: modeling competencies of BPM roles," Business Process Management Journal, **27**(1), 275–305, 2020, doi:10.1108/BPMJ-04-2020-0161.

[7] C. Czarnecki, Business Process Management Cases, Springer International Publishing, Cham, 2018, doi:10.1007/978-3-319-58307-5.

[8] M. Dumas, M. La Rosa, J. Mendling, H.A. Reijers, Fundamentals of Business Process Management, Springer Berlin Heidelberg, Berlin, Heidelberg, 2018, doi:10.1007/978-3-662-56509-4.

[9] M. Gërvalla, N. Preniqi, P. Kopacek, "IT infrastructure library (ITIL) framework approach to IT governance," in IFAC-PapersOnLine, 181–185, 2018, doi:10.1016/j.ifacol.2018.11.283.

[10] V. Cavalcante, S. Bianchi, A. Braz, F. Amorin, N. Nauata, "Investigating business needs fluctuations on IT delivery operations," in Annual SRII Global Conference, SRII, 19–26, 2014, doi:10.1109/SRII.2014.13.

[11] A. Limanto, A.F. Khwarizma, Imelda, R.Y. Rumagit, V.P. Pietono, Y. Halim, S. Liawatimena, "A study of Information Technology Infrastructure Library (ITIL) framework implementation at the various business field in Indonesia," in 2017 5th International Conference on Cyber and IT Service Management, CITSM 2017, 2017, doi:10.1109/CITSM.2017.8089244.

[12] A.D. Suryawan, Veronica, "Information Technology Service Performance Management Using COBIT and ITIL Frameworks : A Case Study," in Proceedings of 2018 International Conference on Information Management and Technology, ICIMTech 2018, 223–228, 2018, doi:10.1109/ICIMTech.2018.8528197.

[13] M. Thirumaran, D. Ponnurangam, K. Rajakumari, G. Nandhini, "Evaluation model for Web service change management based on business policy enforcement," in Proceedings - 2012 International Symposium on Cloud and Services Computing, ISCOS 2012, 63–69, 2012, doi:10.1109/ISCOS.2012.13.

[14] P. Kubiak, S. Rass, "An Overview of Data-Driven Techniques for IT-Service-Management," IEEE Access, **6**, 63664–63688, 2018, doi:10.1109/ACCESS.2018.2875975.

[15] OCG, The official Introdution to ITIL Service Lifecycle, Stationery Office, United Kingdom, 2007.

[16] ISACA, Information Systems Audit and Control Association 2018. COBIT 2019 Framework: Governance And Management Objectives, ISACA, Schaumburg, 2018.

[17] M.M.F. Roquete, "Modelo de maturidade para apoio à implementação de uma filosofia de gestão orientada a processos numa organização," 2018.

[18] S. Sahibudin, M. Sharifi, M. Ayat, "Combining ITIL, COBIT and ISO/IEC 27002 in order to design a comprehensive IT framework in organizations," in Proceedings - 2nd Asia International Conference on Modelling and Simulation, AMS 2008, 749–753, 2008, doi:10.1109/AMS.2008.145.

[19] A. El Yamami, K. Mansouri, M. Qbadou, E.H. Illousamen, "Toward a new multi-agents architecture for the adoption of ITIL framework by small and medium-sized enterprises," in Colloquium in Information Science and Technology, CIST, 40–45, 2016, doi:10.1109/CIST.2016.7805091.

[20] R. Pereira, M.M. Da Silva, "A maturity model for implementing ITIL V3 in practice," in Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC, 259–268, 2011, doi:10.1109/EDOCW.2011.30.

[21] R.A.S. Esteves, Implementação do Processo Gestão da Configuração da framework ITIL – um Estudo de Caso, Instituto Politécnico de Bragança, 2012.

[22] D. Ivanov, A. Tsipoulanidis, Global Supply Chain and Operations Management, 2019.

[23] M. Leketi, M. Raborife, "IT Governance Frameworks and their Impact on Strategic Alignment in the South African Banking Industry," in 2019 IST-Africa Week Conference, IST-Africa 2019, 2019, doi:10.23919/ISTAFRICA.2019.8764872.

[24] A. Ekanata, A.S. Girsang, "Assessment of capability level and IT governance improvement based on COBIT and ITIL framework at communication center ministry of foreign affairs," in 2017 International Conference on ICT for Smart Society, ICISS 2017, 1–7, 2018, doi:10.1109/ICTSS.2017.8288871.

[25] F. Hartawan, J.S. Suroso, "Information technology services evaluation based ITIL V3 2011 and COBIT 5 in center for data and information," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 44–51, 2017, doi:10.1007/978-3-319-54430-4_5.

[26] M. Kloppenburg, J. Kettenbohrer, D. Beimborn, M. Boegle, Business Process Management Cases, Springer International Publishing, Cham, 2018, doi:10.1007/978-3-319-58307-5.

[27] F. Bustamante, W. Fuertes, P. Diaz, T. Toulkeridis, "Integration of IT frameworks for the management of information security within industrial control systems providing metrics and indicators," Proceedings of the 2017 IEEE 24th International Congress on Electronics, Electrical Engineering and Computing, INTERCON 2017, 15–18, 2017, doi:10.1109/INTERCON.2017.8079672.

[28] M. Aouachria, A. Leshob, J. Gonzalez-Huerta, A.R. Ghomari, P. Hadaya, "Business Process Integration: How to Achieve Interoperability through Process Patterns," Proceedings - 14th IEEE International Conference on E-Business Engineering, ICEBE 2017 - Including 13th Workshop on Service-Oriented Applications, Integration and Collaboration, SOAIC 207, 109–117, 2017, doi:10.1109/ICEBE.2017.26.

[29] A. Qusef, A. Hamdan, S. Murad, "Business process integration with quality characteristics: Quality-oriented process," Proceedings - 2017 International Conference on Engineering and MIS, ICEMIS 2017, 2018-Janua, 1–5, 2018, doi:10.1109/ICEMIS.2017.8273030.

[30] Y. Mahy, M. Ouzzif, K. Bouragba, "Supporting ITIL processes implementation using business process management systems," in Proceedings - 2016 3rd International Conference on Systems of Collaboration, SysCo 2016, 2017, doi:10.1109/SYSCO.2016.7831338.

[31] J. Nabais, A.M. Pinto, A. Cruz, J. Cardoso, "Interface design for IT service management practice," in Iberian Conference on Information Systems and Technologies, CISTI, 2014, doi:10.1109/CISTI.2014.6876929.

[32] L. Lema-Moreta, J. Calvo-Manzano, "A proposal for implementation of ITIL incident management process in SMEs," in 2017 IEEE 2nd Ecuador Technical Chapters Meeting, ETCM 2017, 1–5, 2018, doi:10.1109/ETCM.2017.8247494.

[33] J. Persse, "The ITIL Process Manual - Key Processes and their Application," The ITIL Process Manual, 55, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 754-762 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Development of an EEG Controlled Wheelchair Using Color Stimuli: A Machine Learning Based Approach

Md Mahmudul Hasan[1, *], Nafiul Hasan[2], Mohammed Saud A Alsubaie[3]

[1]*Centre for Accident Research and Road Safety- Queensland, Queensland University of Technology, Brisbane, QLD 4059, Australia*

[2]*Department of Electrical and Electronic Engineering, Bangladesh Army University of Engineering and Technology, Qadirabad, 6431, Bangladesh*

[3]*Department of Mathematics, Taif University, Taif, 26513, Saudi Arabia*

ARTICLE INFO

Article history:
Received: 16 December, 2020
Accepted: 26 February, 2021
Online: 27 March, 2021

Keywords:
Brain-computer interface
Electroencephalogram
Principal component analysis
Color stimuli
Machine Learning

ABSTRACT

Brain-computer interface (BCI) has extensively been used for rehabilitation purposes. Being in the research phase, the brainwave-based wheelchair controlled systems suffer from several limitations, e.g., lack of focus on mental activity, complexity in neural behavior in different conditions, and lower accuracy. Being sensitive to the color stimuli, the EEG signal changes promises a better detection. Utilizing the Electroencephalogram (EEG changes due to different color stimuli, a methodology of wheelchair controlled by brainwaves has been presented in this study. Red, Green, Blue (primary colors) and Yellow (secondary color) were chosen as the color stimuli and utilized in a 2 × 2 color window for four-direction command, namely left and right, forward and stop. Alpha, beta, delta and theta EEG rhythms were analyzed, time and frequency domain features were extracted to find the most influential rhythm and accurate classification model. Four classifiers, namely, K- Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest Classifier (RFC) and Artificial Neural Networks (ANN) were trained and tested for assessing the performance of each of the EEG rhythm, with a five-fold cross-validation. Four different performance measures, i.e. sensitivity, specificity, accuracy and area under the receiver operating characteristic curve were utilized to examine the wholescale performance. The results suggested that Beta EEG rhythm performs the best apart from all the rhythms for the color stimuli based wheelchair control. While comparing the performance of the classifiers, ANN-based classifier shows the best accuracy of 82.5%, which is higher than the performance of the three other classifiers.

1. Introduction

This paper is an extension of work originally presented in 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 [1]. The presented paper [1] utilized electroencephalogram (EEG) for wheelchair control using color stimuli where the current article is expanded further to validate the EEG based wheelchair control system using multiple machine learning models. Also, this paper examines the utility of the different color stimulus on the EEG based wheelchair control system.

EEG is a reflection of our neurons activity which is associated with all kind of human behaviours- thoughts, emotional state, eye vision etc. Due to its nature, EEG changes its value of features with respect to different influencer like eye vision. The EEG rhythms are defined by their frequency range, named delta, theta, alpha, beta and gamma corresponds to 1-4 Hz, 4-8 Hz, 8-13 Hz, 13-30 Hz and 36-44 Hz respectively.

To make life easier of the person with quadriplegia and paralyzed patients, many initiatives have been taken, but most of the latest works are eye blink based control systems [2]. But eye blink can not be a good command due to its uneasiness to control. In [2], a system is introduced by means of a different range of attention level and double eye blink (both identified by EEG) to

*Corresponding Author: Md Mahmudul Hasan, 2 Rochester Terrace, Brisbane, QLD 4059, Australia; Email: mahmudul.hasan.eee.kuet@gmail.com

make a brain-controlled wheelchair. Four range of attention level as four direction command and double eye blink as on/off command were used at the preliminary level. Nonetheless, retain the attention level at an indefinite range requires sophisticated long term training, thus making these alike systems less feasible in real-world application. By using an indefinite number of eye blink as logic and cognitive level, a brain-computer interface (BCI) was developed with the wireless operation and interfaced with RS-232C in another study [3]. In order to make more advanced EEG based wheelchair, it was tried to be wireless. Depending on modification and attention; eye blinking signal the system was tried to be constructed in the study.

In some cases, steady-state visually evoked potentials (SSVEP) showed promising results [4]. There are some eye blink pattern-based works, which are also providing very optimistic results [5,6]. Neurosky mindwave headset was utilized in [6], which detected the eye blinks, counted and rendered to Arduino UNO by Bluetooth to control a DC motor and thus the wheelchair. In [6], EEG signal was extracted by MindWave mobile application and sent to the Rasberry pi microprocessor. The output of Rasberry pi and joystick are taken as input for Arduino Mega to control the DC motor through the motor driver. They introduced a virtual map by analyzing familiar wheelchair routes as well. In advance, eye blinking with glancing a model has been proposed in few studies [7,8]. Specifically in [7] autoregressive neural network was built to classify EEG eye signals such as eye blinking, eye glancing in left and right. In [8] Steady-State Visual Evoked Potential (SSVEP) based work and implementation of artificial neural network and support vector machine is introduced to classify flickering frequency lights. Here multiclass support vector machine worked better than ANN while input features was obtained from Fast Fourier Transform.

Most of the previous research on wheelchair control have worked with the eye blink, SSVEP or Internet of Things (IoT). But sill there are very few works on the EEG based wheelchair control. The use of Brain-computer interface (BCI) is currently in the research phase. The existing systems suffer from several limitations, e.g., lack of focus on mental activity, complexity in neural behaviour in different conditions, and lower accuracy. Being sensitive to the color stimuli, the EEG signal changes promises a better detection. A previous study worked with EEG signal for wheelchair control, where time-frequency domain features were extracted from the signals and used for the wheelchair control with a color stimuli pattern. By applying artificial neural network, the study found the beta EEG band as the most influential frequency band where alpha as the least influential band [1]. This study showed promising results but utilized only one classifier (ANN without hyperparameter tuning) for the classification approach and only one performance measure (mean square error) for the assessment of the system [1]. However, considering the sensitivity and specificity metrics are most important for an EEG based control system. As higher sensitivity with lower specificity leads to the higher false decision and the opposite trend causes the missing of a lot of negative states, a compromise between the two metrics is crucial.

The proposed study examines the EEG based wheelchair control using a 2 × 2 matrix shape color screen with the combination of four different colors, being Red, Green, Blue and Yellow. These four different colors have been proven to be sensitive to the EEG signal changes [1], which were designed to indicate the left/right/forward/stop functioning, and were assigned as four numerical values (0,1,2,3) to be considered as ground truth. The principal component analysis (PCA) was conducted to separate the background effect of color stimuli on the EEG signal. After the rhythm separation from EEG, the data analysis and classification was done. The data was recorded using the BIOPAC® data acquisition unit, and the pre-processing and feature extraction was done using the Acqknowledge-4.1® software [9,10].

So far, there is no specified classification model to be used for a specific dataset. Considering the computational complexity and the learning mechanisms, four supervised classification models, namely, K- Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest Classifier (RFC) and Artificial Neural Networks (ANN) were trained, tested, and compared for evaluating the performance of each of the EEG rhythm. Being a smaller datest with respect to the number of features, five-fold cross-validation uas used to validate the performance of the classifiers. Four different performance measures (sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve) were utilized to examine the system performance.

The following part of this paper is organized as follows- a brief methodology, including experimental design and tools, then the result section with the findings. Last, the paper was concluded followed by a short discussion on the outcomes.

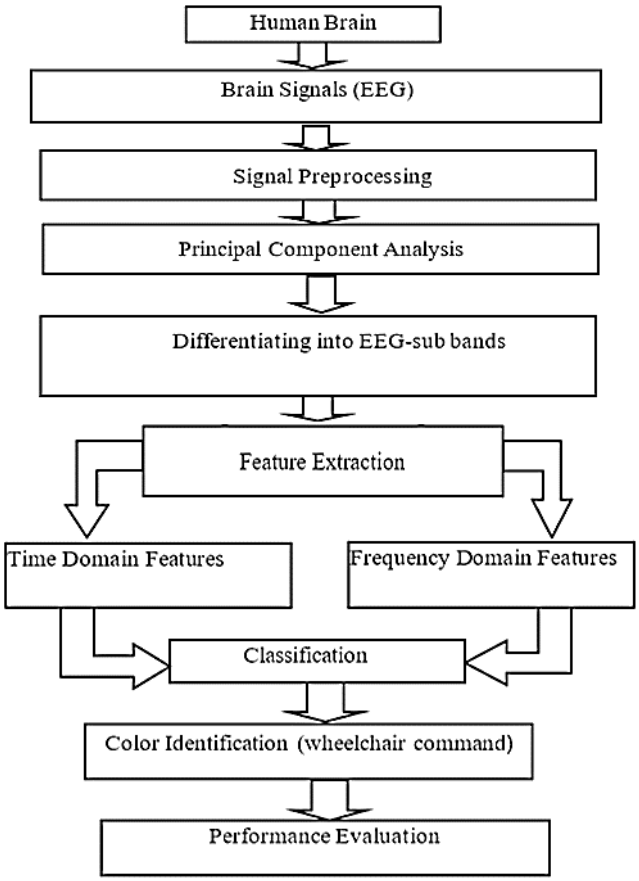


Figure 1: Block diagram for the proposed EEG based wheelchair control system

2. Methodology

2.1. Experimental Design

The roadmap for building an electroencephalogram based wheelchair by color stimuli is shown in figure 1. From the selected three participants, EEG were obtained by the BIOPAC® signal acquisition system. The eye blink artefacts were removed using the EOG blink removal techniques, with the help of Acqknowledge 4.1® software. Then FIR bandpass filter was used to separate bands while PCA for demolishing background effect. Afterwards, total twenty features were extracted for each of the EEG band, which are sensitive to color stimuli and the selected features were supplied towards the machine learning tools as the independent variable. The four different colors were coded to 0-3 as the dependent variable and were labeled for four different color detection. Four different classification models, namely K-nearest neighbours (KNN), support vector machines (SVM), artificial neural network (ANN) and random forest classifier (RFC) classifier models were developed in python 3.6.9 platform in Google Colab platform. Best classifier and rhythm were evaluated by their performance.

2.2. Experimental Equipment

2.2.1. Hardware tool

In this experimentation, BIOPAC MP 36 was used at Biomedical Engineering (BME) lab, KUET, which is shown in figure 2 [1].

2.2.2. Pre-processing and feature extraction software tool

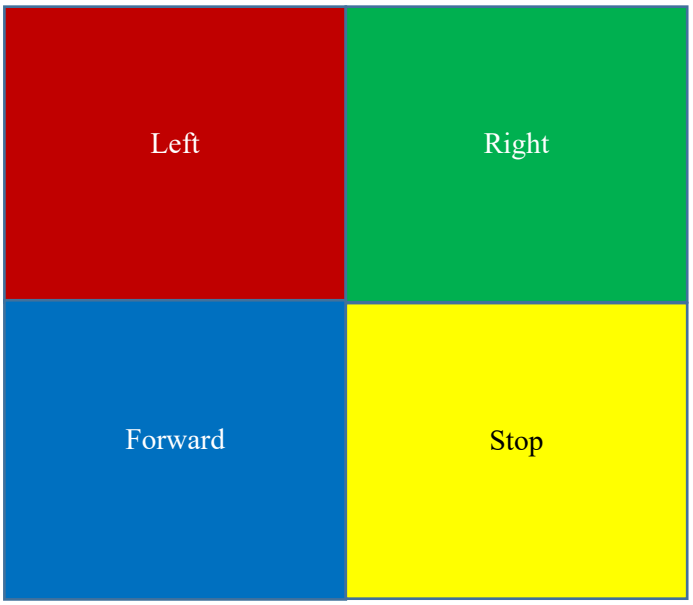
BIOPAC® student Lab Pro and Acqknowledge 4.1® software was used for PCA and feature extraction. Machine learning based classification models were used in the python 3.6.9 version in Google Colab platform.

2.3. Participants

In this experiment, three subjects participated who were male, healthy and not suffering from any color blindness or psychological illness. and they were The color blindness was tested using the Ishihara 38 Plates CVD Test [11,12], to check their vision and to ensure they are not suffering from difficulties in choosing colors, especially the deuteranopic vision (red-green color blindness). Then, the Color stimulus matrix was shown in a computer monitor (21.5" with a 1920 ×1080 resolution). The color stimuli was consisting of Red, Green, Blue (primary colors) and yellow (one of the secondary colors), which were utilized in a 2 × 2 dimension color window for four-direction command, namely left and right, forward and stop. The participants were instructed to focus on each of the color for 15 seconds long, with their normal blinking. It took 60 seconds (1 minute) in total to complete a full visualization of the four colors. The color matrix sub windows for the specific color were programmed to give a pulse after every 15 seconds, such that the participants can automatically focus on the specified sub-window. In total, 20 trials of 20 minutes were taken for visualization of the four colors sequentially, i.e., red, green, yellow and blue. The electrodes were placed on the right central (C4), and the right occipital (O2) position, with a reference electrode in A2 position.



(a)



(b)

Figure 2. Representational view while conducting an experiment in BME lab, KUET (a) EEG electrode positions (b) Color arrangement in the screen for the wheelchair control system

2.4. Experimental Procedure

2.4.1. Signal Preprocessing

Because of muscle movement, eye blinking, hand movement and the background effect behind the colour stimulus, the primarily obtained EEG signals could have contained noise. Along with line frequency was 50 Hz. To pre-process raw EEG was gone through bandpass finite impulse response (FIR) filter with a range of 0.5 to 44 Hz, as it removes the non-linear trends of the signals. Later the signal was further smoothed, taking a moving average over a small period of the signal. Figure 3 shows the EEG signal representation with in Acqknowledge software.

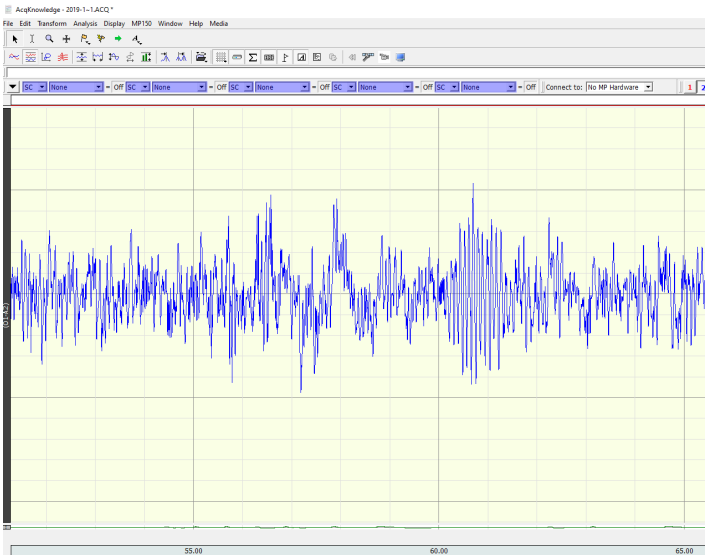


Figure 3. EEG signal representation in Acqknowledge-4.1® software

2.4.2. Principal Component Analysis

As a data reduction tool PCA decomposes input signal to a new signal, removing the components which are not significant. After the signal pre-processing and feature extraction, the PCA was used to select the most significant features, which included standard deviation (STDDEV), skew (sk), kurtosis (k), power spectrum density (PSD) mean, PSD max, PSD skew, Fast Fourier Transform (FFT) mean, FFT max, FFT skew. Total of nine features were chosen out of 20 features based on the PCA analysis. The features were selected based on the threshold that was put on 'k' principal components. It is worth noting that most of the selected features were frequency domain features.

2.4.3. Separation of the frequency sub-bands

The build-in bandpass filter was used in Acknowledge 4.1® software to find alpha, beta, delta and theta rhythm. The frequency ranges used for delta, theta, alpha and beta EEG bands are 1-4 Hz, 4-8 Hz, 8-13 Hz, 13-30 Hz, respectively.

2.4.4. Feature Extraction

Total twenty features were extracted in time and frequency domain, including maximum value (E_{max}), standard deviation (STDDEV), skew (sk), kurtosis (k), moment of order from one to five, Lyapunov exponent, mutual information (m), correlation coefficient (c), power spectrum density (PSD) mean, PSD max, PSD skew, Fast Fourier Transform (FFT) mean, FFT max, FFT skew were extracted for each subject, color, rhythm using the Acknowledge 4.1® software. An epoch size of 1 second was utilized to extract all the EEG features.

2.4.5. Feature Scaling

The features extracted from time and frequency domains have a different range in their magnitude. Different machine learning models work with various features putting them in the same matrix; it is necessary to put all the features in a same range, which is referred to as feature scaling. There are two common types of feature scaling that is done in preliminary data: standardization and normalization [13].

As a part of the normalization process, MinMaxScaling was done in this study in python platform using the MinMaxscaler() function from sklearn library. Here the data is shrunk within a range between [-1,1].

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

The min-max scaling can be defined by the equation (1), where, x_{new} is the normalized value of a feature point x, within a range x_{min} and x_{max} [13].

2.5. Classification

To control a wheelchair on the decision of the appropriate movement, classification is the final step which is done with the help of machine learning models. Machine learning is the application of artificial intelligence, which provides a system capable of learning nature from a given data. There are three categories of Machine Learning models and applications, supervised learning, unsupervised learning and reinforcement learning [14]. Supervised learning is extensively used for the classification and regression problem [14]. Previous studies worked with EEG have used supervised learnings, especially K-Nearest Neighbour [15], Support Vector Machine [16], Random Forest Classifier [17] and Artificial Neural Network [18]. Based on the previous studies, these four classifiers were chosen for the data classification in this research.

2.5.1. K-nearest neighbours (KNN)

KNN is a simple supervised learning algorithm which is very popular and widely used for classification and regression problems. At the very starting point, KNN read the value of K, type of distance D and test data; then it finds the K nearest neighbours D to the test data and thus sets the maximum label class of K to test data. The same process is gone through as an iterative process named looping. In details, it's algorithm initializes the value of K from 1 (setting as initial iteration value). After loading data, iteration from initial K =1 (generally) to total number of training data point. Then, distances specifically Euclidean distance between test data and each row of training data is measured and sorted in ascending order to get topmost k rows from the sorted array and the most frequent class is returned as the predicted class [19]. The value of K was tuned, and the K for best efficiency was chosen in the classifier model in this research to reduce overfitting.

2.5.2. Support Vector Machines (SVM)

SVM aims to obtain a hyperplane which classifies the data point (data points can be at any side of hyperplane) in feature dimensional space while depending on both linear and non-linear regression. Data points distance across to hyperplane are called support vector whose detection can exchange hyper plane's location [19]. The model used a Gaussian kernel for SVM classifier in this research due to the non-linear trend of the dataset. Two parameters- 'C' and 'gamma' was adjusted within a set of values using the grid search algorithm to reduce overfitting.

2.5.3. Random Forest Classifier (RFC)

Random forests are made of individual decision trees with a logic of group of weak learners to finally make a strong learner

while the decision trees operate as divided or conquer. A class is predicted from every decision tree and a final class is predicted by model depending on their vote [19]. Two parameters were tuned in the RFC models, namely, 'n_estimate', which implies the number of trees in the forest and 'max-depth' which signifies the depth of each tree.

2.5.4. Artificial Neural Network (ANN)

An ANN consists of neurons which assess the weighted sum of the inputs contemplating that there is a bias and passes the sum applying activation function such as sigmoid, RBF etc. The research used a feed-forward neural network model (with backpropagation algorithm) which relates input with appropriate output to obtain low squared error (output v/s expected output) by applying gradient descent. The number of hidden layers and the number of neurons in each hidden layer was tuned here in the ANN model. Regularisation was done using the dropout layer with a dropout rate of 0.5 to reduce overfitting.

2.6. Performance Measures

2.6.1. Sensitivity or True Positive Rate (TPR)

Sensitivity is the proportion of the true positives (desired factor), which is correctly identified from the given test set [20]. The definition of sensitivity can be provided by equation (2), where, TP = True Positive and FN = False Negative.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+ FN}} \quad (2)$$

2.6.2. Specificity or True Negative Rate (TNR)

Specificity is the proportion of true negative (undesired factor) in which was correctly excluded from the given test sets [20]. The definition of specificity can be provided by equation (3), where TN = True Negative and FP = False Positive.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+ FP}} \quad (3)$$

2.6.3. Accuracy

Accuracy is defined as the proportion of true results (either true positive or true negative) in an experiment [21,22]. The definition of accuracy can be provided by equation (3), given that TP = True positive, TN= True Negative, FP= False Positive and FN = False Negative

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (4)$$

2.6.4. Area under the receiver operating characteristic (ROC) curve (AUC)

To find out the best compromise between sensitivity and specificity, ROC is plotted. It is a plot of the sensitivity (true positive rate) against the (1- specificity) or false positive rate, where all the possible combination of TPR and FPR are plotted, showing the trade-off between them [16,20,23]. As sensitivity and specificity are two major parameters of performance measures, AUC under ROC always provide a compromise between them.

Five-fold cross-validation was done while evaluating the performance measures. For each of the validation, the dataset was divided in a ratio of 4:1 for training and test data, respectively. The mean value and the standard deviation (SD) were noted, considering the five experimental validations. As the classification is a four-class problem, one versus all method was used in all the classification approach, splitting the four-class problem in binary class. Thus, the mean sensitivity, specificity and AUC was calculated from the obtained confusion matrix.

3. Results

3.1. Data Visualization

The data points found from the selected features were plotted in violin plots to observe the range of each of the features. The following Figure 4 shows that the time and frequency domain features are having a versatile variation in the range in the horizontal box plot. The plots were done for the observation of the entire feature sets. The Maximum value/magnitude (of EEG signal) and the power spectral density maximum value show a greater range than the other features while the PSD mean and FFT (max) are showing the lowest range. Range of the difference features varies among themselves either in the time domain or in the frequency domain. Thus, feature scaling was done on the given dataset.

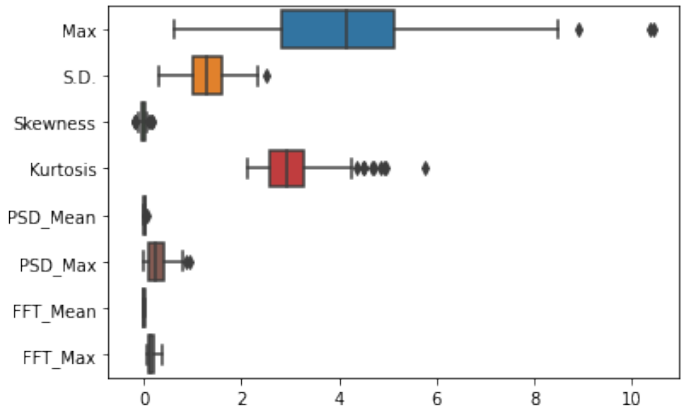


Figure 4: Box plot of part of the EEG beta rhythm features showing the varying magnitude of different features

3.2. Classification Performance

After necessary feature extraction and scaling, all the features were scaled and were supplied towards the machine learning models with necessary parameter tuning. Four different performance measures were evaluated, namely, sensitivity or true positive rate (TPR), specificity or true negative rate (TNR), accuracy and area under the receiver operating characteristic curve (AUC). The obtained results are listed below in Table 1.

3.2.1. Scenario 1: Delta Rhythm

The plots of the performance measures (mean ± SD) obtained from the classification of colours using the delta rhythm features from four different classifiers, KNN, SVM, RFC and ANN, are shown in the following Figure 5. The plots show that the gap between sensitivity and specificity is higher in KNN (gap = 15.3%) and RFC (gap = 7.9%). The gap is less in SVM (2.2%) and ANN (0.3%). Overall, ANN gives an accuracy of 62.8%, which performs the best.

Table 1: Performance measures (mean value) for EEG based wheelchair control using four different classifiers, with a five-fold cross-validation

EEG Rhythm	Performance Metrics	KNN	SVM	RFC	ANN
Delta	Sensitivity	60.8	55.4	55.1	63.3
	Specificity	45.5	57.6	47.2	63
	Accuracy	53.5	56.2	54.6	62.8
	AUC	53.9	56.7	57	69.6
Theta	Sensitivity	63.6	66	60.2	80.8
	Specificity	55.6	62.8	60.6	62.1
	Accuracy	59.6	63.9	60	71.5
	AUC	59	64.5	70.6	73.2
Alpha	Sensitivity	92.3	82.1	55.7	62.1
	Specificity	25	62.5	60.7	52.9
	Accuracy	59.2	72	58.2	72
	AUC	55.1	71.8	68.1	63.4
Beta	Sensitivity	82.7	78.2	79	88.5
	Specificity	68.4	67.2	77.5	75.3
	Accuracy	**75.4**	**72.7**	**77.5**	**82.5**
	AUC	75	76.9	88.3	89.1

3.2.2. Scenario 2: Theta Rhythm

The plots of the performance measures (mean ± SD) obtained from the classification of colours using the Theta rhythm features are shown in the following Figure 6. The plots show that the gap between sensitivity and specificity is higher in ANN (18.8%), KNN (8.05%), and SVM (3.16%). Less gap is observed in the case of RFC (0.4%). Overall, ANN gives an accuracy of 71.5%, which performs the best.

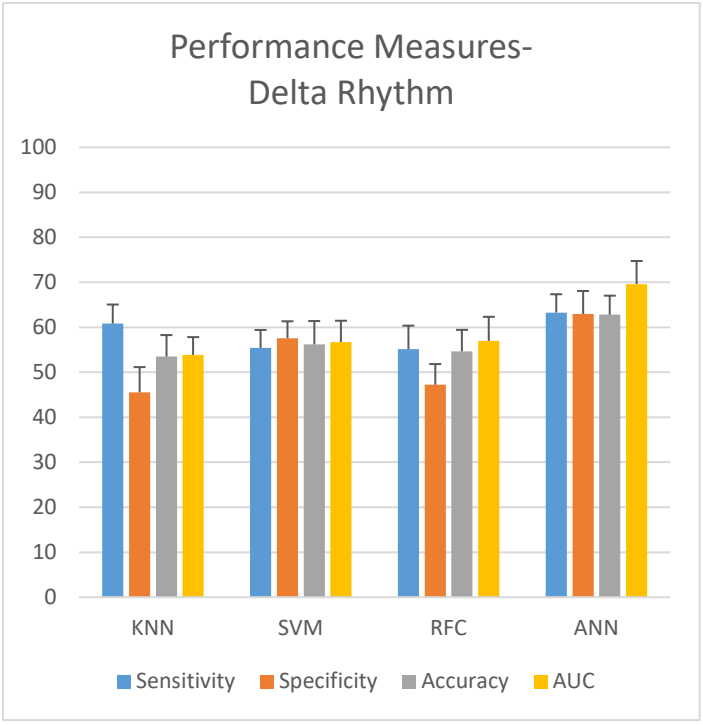


Figure 5: Performance measurement of delta rhythm

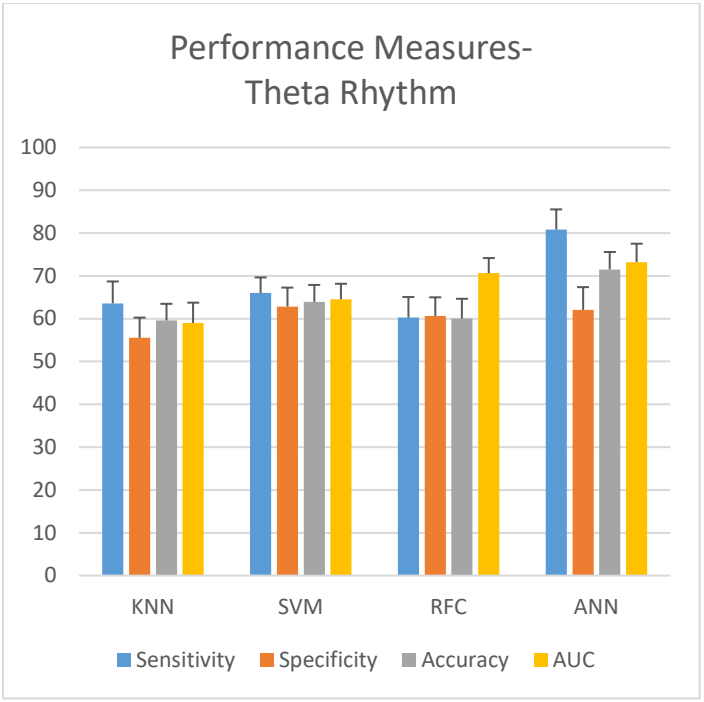


Figure 6: performance measurement of theta rhythm

3.2.3. Scenario 3: Alpha Rhythm

The plots of the performance measures (mean ± SD) obtained from the classification of colours using the Alpha rhythm features are shown in the following Figure 7. The plots show that the gap between sensitivity and specificity is higher in KNN (67.3%), SVM (19.7%) and ANN (9.19%). The gap is lees in RFC (5%). Overall, ANN gives an accuracy of 72%, which performs the best.

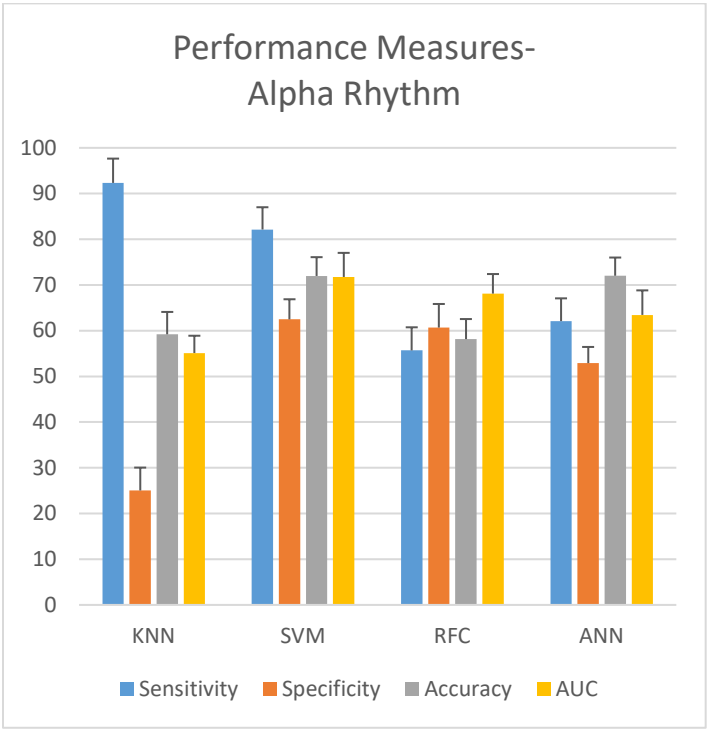


Figure 7: performance measurement of alpha rhythm

3.2.4. Scenario 4: Beta Rhythm

The plots of the performance measures (mean ± SD) obtained from the classification of colours using the Beta rhythm features are shown in the following Figure 8. The plots show that the gap between sensitivity and specificity is higher in KNN (14.3%), ANN (13.2%), and SVM (11%). The gap is lees in RFC (1.45%) and Overall, ANN gives an accuracy of 82.5%, which performs the best.

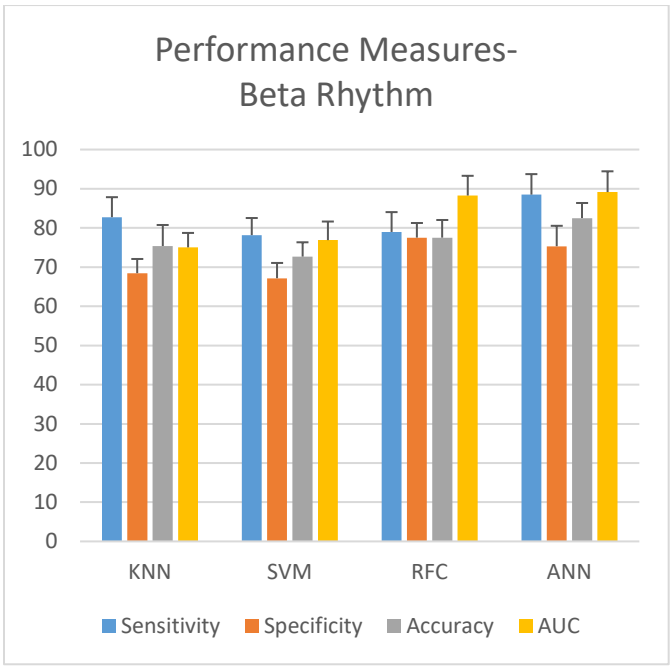


Figure 8: performance measurement of beta rhythm

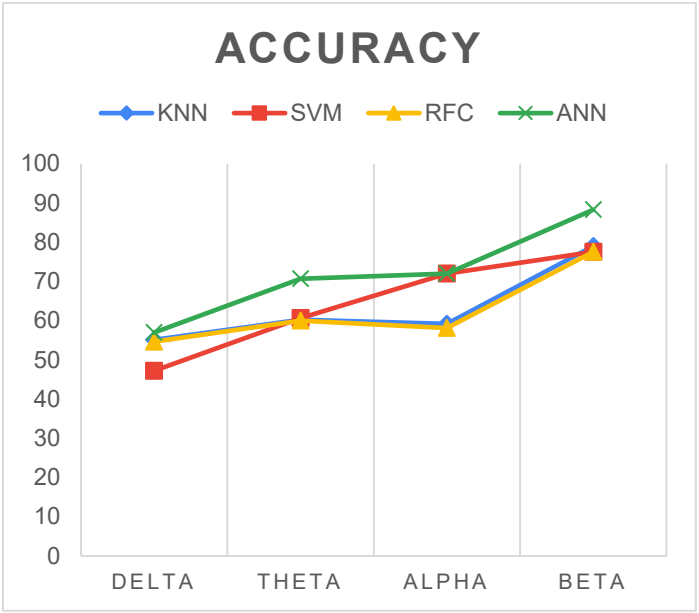


Figure 9: Comparison of accuracy matrics for four different classifiers

3.3. Choosing the best EEG Rhythm

To choose the best EEG frequency band, the overall accuracy measure was considered as the reference metrics as it is difficult to compare different classifiers using several factors. The plots of the accuracy for four different classifiers corresponding to the four different frequency band are shown in Figure 9. From the given figure, it is evident that the accuracy for beta rhythm is better than any other frequency bands for all the four classifiers. So, in the rest of the paper, the Beta EEG rhythm will be considered for further analysis.

3.4. Choosing the best classifier

The plots for the area under the ROC curve for the classifiers built using the Beta EEG features are shown in Figure 10 below. The figure illustrates that the ANN classifiers show the best compromise between sensitivity and specificity, with covering the highest area under the ROC curve (AUC = 0.89). Thus, the next part of the paper will compare the performance of the frequency bands considering ANN classifier.

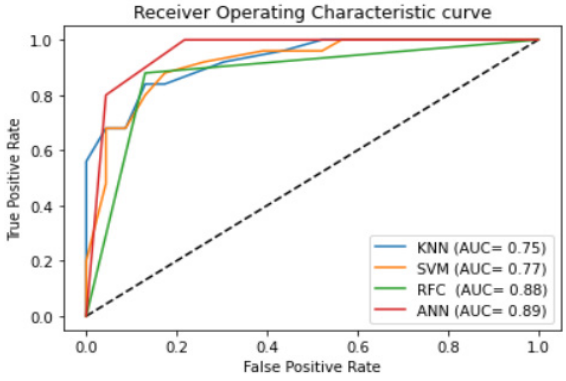


Figure 10: Comparison of AUC for four different classifiers built using the Beta EEG Features.

The plot of the area under the AUC curve for 5 different experiments in 5-fold cross-validation with the Beta rhythm using ANN classifier (see Figure 11). The AUC for beta rhythm-based models ranged from (0.75-0.92), with a mean of 0.89 and standard deviation of 0.07. This signifies that, the beta rhythm shows an excellent performance than the other frequency bands of EEG signal in EEG based wheelchair control using the color stimulus.

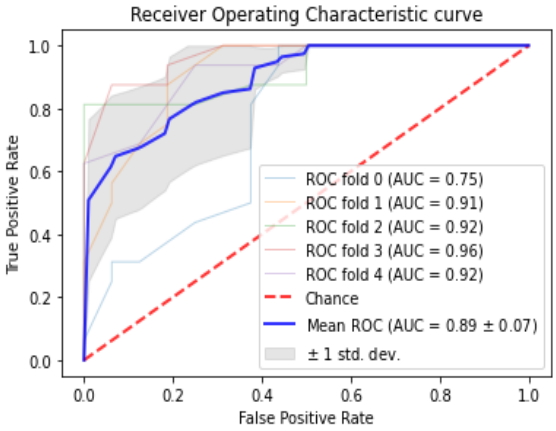


Figure 11: ROC Plots for Beta rhythm-based ANN model with 5-fold CV

4. Discussion

This study used four different classifiers for assessing the performance of the EEG rhythms for EEG based wheelchair

control using colour stimuli. The results revealed that Beta rhythm performs the best among the other rhythms while the Alpha rhythm performs the worst. The delta and theta rhythm showed an average performance. The reason behind the better performance of the beta rhythm could be it is more prominent in the concentration while the theta and alpha are responsible for drowsiness and relaxation, respectively. It also revealed that the maximum performance was obtained using the ANN Classifier, with a sensitivity, specificity and accuracy of 88.5%, 75.3% and 82.5%. Moreover, ANN based model with beta EEG based dataset shows promising AUC (0.89), which is a good compromise between sensitivity and specificity. The finding of this study is consistent with the previous study [1], where the authors found the Beta band as the best performing rhythm, though they have used only one classifier (ANN), and one performance metrics (Accuracy). The possible reason behind the best performance of ANN could be the backpropagation algorithms, which is strong enough to learn the inherent features and complex structure of the data. This is likely due to the simultaneous update of the weights while running every iteration for each forward pass and backward pass. The study [1] also achieved a higher accuracy than the current study, which is probably due to not using the cross-validation techniques. Rather, a hold-out validation was used in the previous study [1], which most probably created overfitting problem. The given problems were overcome in the current study.

On the other hand, the Delta rhythm performs the worst (ANN accuracy = 62.8%). However, some difference is found in the obtained results due to different parameter setting in the neural network (such as hidden layers, and the number of neurons). For future implications, three major issues are recommended. First, using the four-colour stimuli that we have used and validated using EEG sub-bands. Second, using the Beta EEG-band power while utilizing the brain-computer interface system for wheelchair control, as it shows comparatively higher performance in identification of the different color stimuli. Third, it is recommended to use the color codes in a small monitor attached with the wheelchair handle, which would help the user to focus on the screen and thus enhance the detection accuracy of the control system. However, the scope of the paper is not out of limitations. The correlation of the EEG signal with different color stimuli was out of the scope of the study; however it could be considered for future research. The study consisted of only three healthy subjects, who are not paralysed. The lower sample size may lack the generalization for proper implication; however future research could be done on large number of sample and involving peoples suffered by paralysis. The background effect is one of the main challenges while using the color stimulus. Inter-individual difference among participants is another factor, which is needed to be considered. As the paper represented a novel methodology of EEG based wheelchair control using four different colours (for left-right/ forward/backward action), research is required to find out the feasibility in real-world condition as well.

5. Conclusion

In order to develop EEG controlled user-friendly wheelchair, using this proposed model, an analysis was done in this study to find out the feasibility of the time and frequency domain features. Here the designed system was obtained after applying several steps- feature scaling, tuning of classifiers and finally with five-fold cross-validation of the developed models. The analysis from the results of the study shows that beta rhythm shows the best accuracy with KNN (75.4%), SVM (72.7%), RFC (77.5%) and ANN (82.5 %). On the other hand, the Beta EEG rhythm offers the lowest accuracy with all classifiers. However, the experiment could be done on a greater number of participants to validate the model based on leave one participant out approach. Also, efficiency will increase with the addition of more EEG channels which can be considered for future work.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors want to acknowledgment the staffs of the Biomedical Engineering Laboratory, KUET, for providing with the signal recording facilities.

References

[1] N. Hasan, M.M. Hasan, M.A. Alim, "Design of EEG based wheel chair by using color stimuli and rhythm analysis," in 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE: 1–4, 2019, doi:10.1109/ICASERT.2019.8934493.

[2] A. Dev, M.A. Rahman, N. Mamun, "Design of an EEG-Based Brain Controlled Wheelchair for Quadriplegic Patients," in 2018 3rd International Conference for Convergence in Technology (I2CT), IEEE: 1–5, 2018, doi:10.1109/I2CT.2018.8529751.

[3] J.-S. Lin, K.-C. Chen, W.-C. Yang, "EEG and eye-blinking signals through a Brain-Computer Interface based control for electric wheelchairs with wireless scheme," in 4th International Conference on New Trends in Information Science and Service Science, IEEE: 731–734, 2010.

[4] K.-T. Kim, S.-W. Lee, "Towards an EEG-basedIntelligent Wheelchair Driving System with Vibro-tactile Stimuli," in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE: 2382–2385, 2016, doi:10.1109/SMC.2016.7844595.

[5] P. Lahane, S.P. Adavadkar, S.V. Tendulkar, B.V. Shah, S. Singhal, "Innovative Approach to Control Wheelchair for Disabled People Using BCI," in 2018 3rd International Conference for Convergence in Technology (I2CT), IEEE: 1–5, 2018, doi:10.1109/I2CT.2018.8529473.

[6] A. Maksud, R.I. Chowdhury, T.T. Chowdhury, S.A. Fattah, C. Shahanaz, S.S. Chowdhury, "Low-cost EEG based electric wheelchair with advanced control features," in TENCON 2017-2017 IEEE Region 10 Conference, IEEE: 2648–2653, 2017, doi:10.1109/TENCON.2017.8228309.

[7] H.T. Nguyen, N. Trung, V. Toi, V.-S. Tran, "An autoregressive neural network for recognition of eye commands in an EEG-controlled wheelchair," in 2013 International Conference on Advanced Technologies for Communications (ATC 2013), IEEE: 333–338, 2013, doi:10.1109/ATC.2013.6698132.

[8] R. Singla, A. Khosla, R. Jha, "Influence of stimuli colour in SSVEP-based BCI wheelchair control using support vector machines," Journal of Medical Engineering & Technology, **38**(3), 125–134, 2014, doi:https://doi.org/10.3109/03091902.2014.884179.

[9] D. Acquisition, Analysis with BIOPAC MP Systems, AcqKnowledge 4 Software Guide. pdf, Biopac Systems, Inc. ISO 9001: 2008.

[10] B.B. flow Monitor, Biopac Systems, Inc. ISO 9001: 2000.

[11] D. Flück, "Ishihara 38 Plates CVD Test| Colblindor," Colblindor. Wordpress. Nd Web, 7–11, 2016.

[12] H.M. Marey, N.A. Semary, S.S. Mandour, "Ishihara electronic color blindness test: An evaluation study," Ophthalmology Research: An International Journal, 67–75, 2015.

[13] C. Colantuoni, G. Henry, S. Zeger, J. Pevsner, "SNOMAD (Standardization and NOrmalization of MicroArray Data): web-accessible gene expression data analysis," Bioinformatics, **18**(11), 1540–1541, 2002, doi:https://doi.org/10.1093/bioinformatics/18.11.1540.

[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, "Scikit-learn: Machine learning in Python," The Journal of Machine Learning Research, **12**, 2825–2830, 2011.

[15] F. Riaz, A. Hassan, S. Rehman, I.K. Niazi, K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," IEEE Transactions on Neural Systems and Rehabilitation Engineering, **24**(1), 28–35, 2015, doi:10.1109/TNSRE.2015.2441835.

[16] S. Bhattacharyya, A. Khasnobish, S. Chatterjee, A. Konar, D.N. Tibarewala, "Performance analysis of LDA, QDA and KNN algorithms in left-right limb movement classification from EEG data," in 2010 International conference on systems in medicine and biology, IEEE: 126–131, 2010.

[17] H. Lee, S. Choi, "Pca+ hmm+ svm for eeg pattern classification," in Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings., IEEE: 541–544, 2003.

[18] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," Computer Methods and Programs in Biomedicine, **108**(1), 10–19, 2012, doi:https://doi.org/10.1016/j.cmpb.2011.11.005.

[19] K. AlSharabi, S. Ibrahim, R. Djemal, A. Alsuwailem, "A DWT-entropy-ANN based architecture for epilepsy diagnosis using EEG signals," in 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), IEEE: 288–291, 2016, doi:10.1109/ATSIP.2016.7523093.

[20] W. Zhu, N. Zeng, N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," NESUG Proceedings: Health Care and Life Sciences, Baltimore, Maryland, **19**, 67, 2010.

[21] A.-M. Šimundić, "Measures of diagnostic accuracy: basic definitions," Ejifcc, **19**(4), 203, 2009.

[22] A.-M. Šimundić, "Measures of diagnostic accuracy: basic definitions," Medical and Biological Sciences, **22**(4), 61, 2008.

[23] D. Justin, R.S. Concepcion, A.A. Bandala, E.P. Dadios, "Performance Comparison of Classification Algorithms for Diagnosing Chronic Kidney Disease," in 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE: 1–7, doi:10.1109/HNICEM48295.2019.9073568.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 4, 762-769 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

xASTESJ

ISSN: 2415-6698

An Explanatory Review on Cybersecurity Capability Maturity Models

Adamu Abdullahi Garba[1,*], Maheyzah Muhamad Siraj[2], Siti Hajar Othman[2]

[1]*Yobe State University Damaturu, Computer Science, Yobe State University Damaturu, 1144, Nigeria*

[2]*School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310, Malaysia*

ARTICLE INFO

Article history:
Received:26 June, 2020
Accepted:14 August, 2020
Online: 28 August, 2020

Keywords:
Cybersecurity Maturity Model
Capability Maturity Model
Security Models
Confidentiality

ABSTRACT

Cybersecurity is growing exponentially day by day in both the public and private sectors. This growth also comes with a new and dynamic cyber-threats risk that causes both sectors' performance to halt. These sectors must update their cybersecurity measures and must understand the capability and maturity of their organization's cybersecurity preparedness. Cybersecurity maturity models are widely used to measure how ready an organization is when it comes to cybersecurity. The main aim of this article is to conduct a comprehensive review of the current cybersecurity capability maturity models using a systematic review of published articles from 2011 to 2019. A comparative study was conducted based on Halvorsen and Conradi's taxonomy. The review indicated almost all the cybersecurity maturity model consists of similar elements like maturity levels and processes but significantly lacks the validation process, it was observed each of the models were predominantly designed for a specific purpose and also for different organization size and application domain.

1. Introduction

Cybersecurity is a method of protecting organization assets, through the identification of threats that can compromise the critical information stored in the organization systems, it also involves the protection, identification, and responding to threats. However, cybersecurity evolves from computer security which means securing the physical components of a computer system from any damage, to information security which means securing the stored information in a computer system from unauthorized assess by maintaining it Confidentiality, integrity and availability (CIA) then to cybersecurity which includes both computer security and information security and also adding the security of information being transferred across a different medium (wired & wireless) and also access from anywhere. The advancement of the cybersecurity domain is dated back to the 1950s. The field of cybersecurity emerged as a result of Robert Morris testing the worlds' network vulnerability in 1980 when he uses a virus he created to test the size of the internet, to protect organization assets, an organization needs to improve their cybersecurity practices. The knowledge of cybersecurity has also been used negatively, the Russian in the 1980s attacked around 400 military computers in the US which include the pentagon computers[1–3]. Therefore, knowledge of cybersecurity capability maturity models is essential as the research area is new and growing exponentially, critical review in the existing models and their applications is important to know, to fill this gap this paper intends to answer the following objective:

- To identify currently available cybersecurity capability maturity models available for this study from 2011 to 2019 using systematic review (SR).
- To identify the main difference between the cybersecurity capability maturity model and their levels
- To understand the application of the cybersecurity capability maturity models.

This paper consists of eight sections, section 1 is the introduction, the second section 2 discusses on the evolution of cybersecurity capability maturity models, section 3 discusses on the method used in conducting the research, section 4 explains the Review on cybersecurity capability maturity models, section 5 explains the results and discussion of the comparative analysis of the identified models, it further discusses the importance of the research and explained how the objective was achieved, section 6 explains the future direction of the research and from where other authors can continue to explain the research direction and lastly, section 7 is the conclusion.

*Corresponding Author: Adamu Abdullahi Garba, adamugaidam@gmail.com

1.1. Review Method and Protocol

The systematic literature review is defined as " a well-defined study or methodology for identifying, analyzing, and also interpreting all available evidence related to a specific research question [4,5]. This method was mostly used in medicine [6], but not now it has been adopted by many fields of studies like social science, information system, and computer science, software engineering [5]. In this study systematic review, "which aims for exhaustive searching, quality assessment, inclusion, and exclusion criteria which are typically narrative with tabular form was adopted "[7]. The aim or analysis of this method is to explain what is known for practices, what remains unknown, and recommendation for future research.

This systematic search started with a well-developed review protocol based on the procedures of the SR review. The protocol includes: background study as evolution, review method, research objectives, and data extraction criteria, and for this study. This section helps to increase the accuracy of the review and also reduces bias in conducting the research. Table 1 describes the review criteria.

Table 1: Inclusion and exclusion criteria

Included article	Excluded article
Full text and available	Full text but not available
Year of publication from 2011 – 2019	Outside range of the year
Published in English	Non-English
Only focus on the domain (cybersecurity)	Were outside domain
Must be a model	Not related to objectives
Found in the selected database.	Duplicated studies

1.2. Inclusion and exclusion criteria

This section is mainly to set up the criteria of inclusion and exclusion for the researchers to follow when doing the study. This research considers the following articles (emerald, IEEE explore, Scopus, the web of science and science direct) published in English, also published from 20011 to 2019 in the digital database. Table 1 shows the steps used in conducting the research.

As part of Step1, we searched articles that have the phrase "cybersecurity", "security model" AND "maturity", from different databases. After following the protocol mention in table 1, we collected 220 articles relevant to your objective

Step 2 is where we used the inclusion and exclusion criteria to determine article very close to your objective by reading abstract [5], we removed all the papers that do not have the word "cybersecurity model" and "cybersecurity capability maturity model". At the end of this phase, we only obtained 30 articles.

Step 3, the article obtained in step 2 was critically analyzed and read fully with more depth analysis. Based on the full- text, the previous criteria were applied to identify the actual articles that are related to our objective. The articles used are fully cybersecurity oriented. At the end of this phase, only 7 articles were obtained. These seven articles are selected based on table 1 criteria.

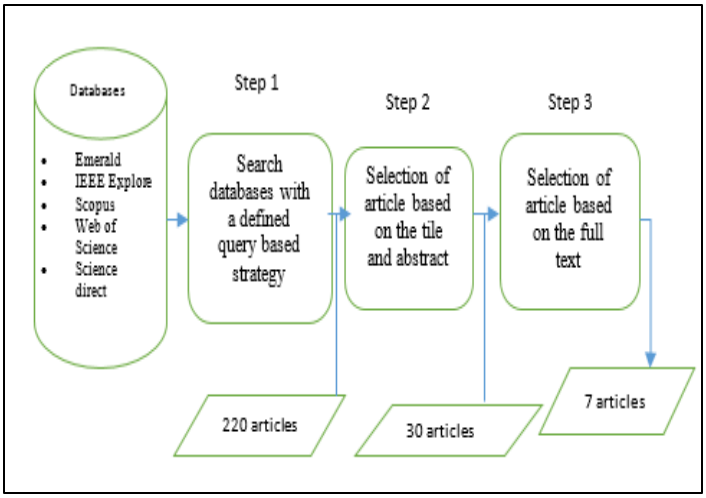


Figure 1. Search and selection of articles considered for this study.

2. Cybersecurity Capability Maturity Models

The cybersecurity Capability maturity model has emerged from the capability maturity model been design from the quality management field in the 1930s. it becomes popular in the 1990s when it was first developed by software engineering institutions [8]. Today all these models are based on this basic model, the model has a set of a structured set of operations and activities that improve over time [9]. the model is later being adapted into many fields of studies to identify or measure the maturity level of an organization or process or even product quality as they are widely known. The capability maturity model (CMM) which was for software industries has some key elements for an effective software development process[10] the model has 5 basic process maturity levels called, initial, repeatable, defined, managed, and optimizing [11]. A best- practices and process efficiency is provided for every five levels of each process for evaluating the maturity [12]. Also [13] conducted a study regarding the capability maturity models, in his research he identified and compares many maturity models for software domain and product quality, while [14] surveyed maturity models specifically for knowledge management to find out how far it contributed to the measurement of knowledge management, but the study only emphases on one special type of maturity model, therefore it is not suitable for general mapping of the maturity model research, the model was designed for software products as guidance as well as for management excellence in producing quality software"[15].

The cybersecurity maturity model offers a framework for assessing the maturity of a security program and guidance on how to reach the next level [16]. The cybersecurity maturity model provides a pathway that enables the organization to measure where they are along that path. This can be a valuable tool not only for improving Cybersecurity efforts but also for collaborating with upper management and getting the support needed to enhance Cybersecurity culture in organizations. There are various Cybersecurity Maturity Models from which to choose, Based on the systematic review performed regarding the currently available Cybersecurity models published to the knowledge of the author from 2011 to 2019 are; Cybersecurity Capability Maturity Model (C2M2), Electricity Subsector Cybersecurity Capability Maturity Model (ES-C2M2), Oil and Natural Gas Subsector

Cybersecurity Capability Maturity Model (ONG-C2M2), National Initiative for Cybersecurity Education-Cybersecurity Capability Maturity Model (NICE-C2M2), Community Cyber Security Maturity Model (CCSMM), African union maturity model for cybersecurity (AUMMCS) and Federal Financial Institute of Examination Council Capability Maturity Model (FFIEC- CMM). These identified models are selected because they focus on cybersecurity, other models were found during the SR but were not fully focusing on cybersecurity, like the National Institute of Standards and Technology Cybersecurity Framework (NIST CSF), The NIST CSF differs from the C2M2, as NIST doesn't consider the CSF a maturity model, Instead of 10 domains, NIST represents five Cybersecurity functions: identity, protect, detect, respond and recover Models identified section 5 shows Comparisons Cybersecurity Capability Maturity Models results then followed by a discussion on the Comparisons Cybersecurity Capability Maturity Models.

3. Methodology

Based on the systematic review performed regarding the currently available cybersecurity capability maturity models published, the author has identified the following: C2M2, ES-C2M2, ONG-C2M2, NICE-C2M2, CCSMM, FFIEC- CMM, and AUMMCS as explained in the previous section. The author adapted Halverson and Conradi taxonomy of software process improvement (2001), this taxonomy consists of 21 features peculiar to software process and is grouped into 5 categories: **general, process, organization, quality**, **and result**. Each category refers to:

- **General**: the features that describe the overall attribute of improvement.
- **Process:** the feature that explains the way the organization uses the features.
- **Organization**: this explains the relationship between the features and organization and how they work simultaneously.
- **Quality:** this explains the feature related to the quality dimension.
- **Result:** this explains the feature of the results as the result of using the environment, the cost of achieving the result. In this paper, only **general, process, organization, and result**s are adapted as the other one has no relation to Cybersecurity Capability Maturity Models. The feature that falls under each category is modified to suit Cybersecurity terms as shown in Table 2 below.

Table 1: Halverson and Conradi taxonomy

Category	Feature
General	Cybersecurity oriented
	Origin
	Purpose
	Prescriptive/ descriptive
	Maturity level
Process	Field Applicable
	Define role
	Depth of assessment
	Assessment
	Assessor
Organization	Actors
	Organization size
	Level of documentation
	Organization Environment
Result	Implementation cost

The features related to ***General*** group are defined below:

- **Cybersecurity Oriented:** this feature depicts which model was purposely designed for Cybersecurity maturity.
- **Origin:** this criterion is used to know the country, lab, organization that created or design the model e.g. the US.
- **Prescriptive/Descriptive**: the criteria used here is either Prescriptive: if the model is enforcing a rule to be used if the model is adapted, while descriptive: if a model is describing or classifying its objectives and how to follow it, not enforcing rules.
- **Maturity Level**: the criteria are used to understand the level of maturity for each model number 1- 5 are used, the more level a mode is the more level of the maturity increases.
- **Field Applicable**: the criteria is used to know the area where the model is applicable criteria include: organization, paper lab. University.
- **Define Role**: this feature explains how well the roles and functions are evaluated using " Yes" if a role is defined and else " No" is used.
- **Depth of Assessment:** the criteria are used here is either "General" if the assessment is not in-depth and "specific" if the assessment is in-depth that is more than one level.
- **Actors**: the criteria used here are " management, staff, communities or states " to know who is using the model directly.
- **Organization Size**: this criterion is used to know the size of the organization for appropriate adaption, criteria used here are: large, medium, small, or all.
- **Level of Documentation**: criteria use are either "high" when a model has implementation guide and other supporting documents that will help adaptor to implement the model, "medium "is when no more details are available on the implementation guide but there are white papers and other supporting documents, "low" in both implementation and white paper are not available but other introductory documents are available.
- **Organization Environment**: criteria "Overall" is used if the model focuses on the entire organization while " Explicit" if the model focuses on a specific unit or department in the organization.
- **Assessment:** the feature is explained by the name of a process to be assessed in the organization e.g. risk, maturity, customer, employee, organization.
- **Assessor:** the criteria use in this feature are " internal" if the assessor is from the environment the model is implemented, " external" if the assessor comes from outside the workplace, and "internal and external" where the assessor can be both.

- **Validation Method**: this criteria is use to know the method of validation includes: survey, case study. Experiment.
- **Implementation Cost**: this criteria is use to know how much to spend when implementing the model.

4. Review

This section explains the main structure and domain found in the identified maturity models based on their focus on cybersecurity. Based on the SR the identified from scientific articles are C2M2, ES-C2M2, ONG-C2M2, NICE-C2M2, CCSMM, FFIEC- CMM, and AUMMCS models.

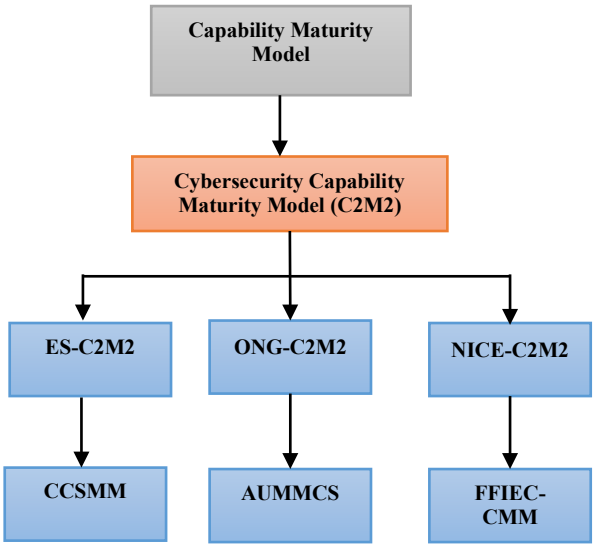


Figure 2: Descriptive diagram of maturity models

Figure 2 above elaborates more on the identified maturity models, it indicates C2M2 is the origin or the first model to be designed in the cybersecurity domain, The difference from the models is either from the number of levels like C2M2 has 4 levels while CCSMM has 5 levels, application area, also C2M2 can be assessed both internally and externally while CCMM can only be asses externally. However, the most important concept of all is that they are only design for cybersecurity specifications, i.e cybersecurity orientation.

4.1. Cybersecurity Capability Maturity Model (C2M2)

The Cybersecurity Capability Maturity Model was designed by Carnegie Mellon University in collaboration with the US Department of Energy in 2014 [17]. The model has ten domains and each domain is a grouping of cybersecurity practices. Also, many objectives are grouped to be in one domain which represents achievements the model contains ten domains with grouped objectives and Maturity levels. (Appendix *A shows the domain and also the maturity level*)

4.2. Electricity Subsector Cybersecurity Capability Maturity Model (ES-C2M2)

The Electricity Subsector Cybersecurity Capability Maturity Model was designed by the department of energy USA to protect the electricity subsector from any form of cybersecurity attacks. [18]. This model was designed as a subsector of the C2M2 i.e. Independent guidance. Both models' general purpose is almost the same, which is to improve cybersecurity capabilities by allowing continuous benchmarking. This model also has ten domain and are the same as the C2M2. The model was developed with the main four primary sector functions as listed below.

- Generation
- Transmission
- Distribution
- Markets

There is a difference between the models in reporting of incidents, C2M2 mentioned ISACs in general in the DOE form while ES-C2M2 threat and vulnerability incident are reported to electricity sector information sharing and analysis center specifically [18]. The model is purposely for electricity sector organizations.

4.3. Oil and Natural Gas Subsector Cybersecurity Capability Maturity Model (ONG-C2M2)

The Oil and Natural Gas Subsector Cybersecurity Capability Maturity Model was developed or derived from ES-C2M2 first version, the main reason for this model was to address the threats and vulnerabilities uniquely characterized by the Oil and Gas subsector. This model can be used to support cybersecurity capabilities in the ONG subsector, it enables organizations to evaluate and benchmark their capabilities. The ONG-C2M2 was designed to address problems in the oil and gas sector only. The development process was extensively cantered with public and private sector experts through pilot facilitation, working sessions. This model uniquely includes the exploration, gathering, production, processing storage, and transportation of petroleum liquids and natural gas. The critical areas where threats can occur, from exploration to storage as technology is used to do all the processes, where security has been not tightened well, an attacker from nowhere can hinder the process or even stops the organizational activities. The ONG-C2M2 and the ES-C2M2 are derived from the C2M2, therefore they share the same domain and maturity levels only place of applications, and the purpose of design differs. This is the reason the author did not repeat the same tables to avoid repetitions, see Appendix *A shows the domain and also the maturity level*)

4.4. National Initiative for Cybersecurity Education Capability Maturity Model (NICE)

The NICE model was designed by then US President George Bush under the directive of national security in 2008, [19]. The model was designed purposely to select the staff with cybersecurity background and skills. The model comprises three key components focusing on staff security structure at the management and the role of staff, the model was officially published in 2014 [19]. The model consists of three domains: Process and Analytics, Integrated Governance, and Trained Professionals and Enabling Technology, which also three maturity levels (*limited level, progressing level, and optimized level*).

4.5. Federal Financial Institute of Examination Council Capability Maturity Model (FFIEC- CMM)

The Federal Financial Institute of Examination Council Capability Maturity Model was published to guide or navigate the increasing complexity of the cyber risk landscape. This assessment tool was designed to help managers assess their institution's cybersecurity preparedness, evaluate its risk, and determine what risk management practices and controls are needed to attain the desired state. This tool has two-part as shown below:

- **The Inherent Risk Profile**: these are risks posed to the organization by technologies and connection types, delivery channels, online and mobile products, and other external threats.
- **Cybersecurity Maturity**: this helps the organization to measure the level of risk and corresponding controls. The level starts from baseline to innovation. The model contains five domains and some assessment factors

4.6. African Union Maturity Model for Cybersecurity (AUMMCS)

The African union maturity model for cybersecurity was made available in 2014 by the center for the cybersecurity at the University of Johannesburg on security and protection of personal data at the convention of African member states, this model covers three sections: *electronic transactions, personal data protection and promoting cybersecurity and combating cybercrime* [20]. The model was intended to help member states of the African Union to evaluate their cybersecurity status against a specific part of the convention. This model can be utilized in two ways: one as a self-assessment by a specific country against the specification of the convention, two as a comparison by the AU between different member states in other to see how they can compare themselves as requirements are concerned. The model only covers the promotion of cybersecurity and combating cybercrime[20]. The model has four objectives:

- A national culture of cybersecurity does exist.
- A national Cybersecurity policy does exist.
- Public-private partnerships, initiated by the government, do exist.
- Cybersecurity capacity building on all levels, driven by the government, does exist.

The model also has four MLS maturity levels:

- ML0: Nothing Exists At All.
- Ml1: Very Basic Position.
- Ml2: Progressed Position.
- Ml 3: Stable Position.

4.7. The Community Cybersecurity Maturity Model (CCM2)

The community cybersecurity maturity model was developed in San Antonio Taxes by the Centre for Infrastructure Assurance and Security (CIAS) [21]. this model was designed to the needs of state and community to the development of a practical and sustainable program of cybersecurity in taxes United States. This model identified the character of community and state as their cyber-security program mature, such aspect includes knowledge, security policies, procedures, information exchange, and cybersecurity training and education. The main importance of this model is to respond to the linkage that exists among states since more communities made up a state. Also, the model is shown in a three-dimensional way [21]. This model is made up of five maturity levels with the lowest initial level showing characteristics for communities that do not share a minimal level of cyber preparedness in the four key areas knows as *Awareness, information sharing, processes and procedures, and integration.* [21]. An example, like top managers at level one, would have little or even no awareness of any cybersecurity threat and its damage, also have little or no information sharing on the cyber event between entities within organization cities or states. Also, few processes or procedures would be in place in the community to handle any cyber threat and lastly, lack of or no mechanism in place (security exercise) to evaluate the level of preparedness of the community or its capacity to respond to any threats. Initially, the model focus on designing the roadmap for the community to follow than later was identified that it is not yet robust enough to adequately represented what is needed for a community to be secure individual community must have a certain level of security as well program necessary to address prevention and detection of cyber threats. The model shows how important information sharing is with other communities so that current threats picture can be obtained and to be able to alert other communities that might be affected and share measures are taken. The model was then expended to include three-dimension to include a third axis that will indicate characteristics and activities at an individual organizational level as well as at a state level.

4.8. Comparative Evaluation of Cybersecurity Capability

The section shows a well and detailed explanation of the identified maturity models, a descriptive diagram of the maturity models, also tables 3 shows a summary of the comparisons among the models using the adopted taxonomy features from Halverson and Conradi's taxonomy of software process improvement.

4.9. Maturity Model

Note: 1 yet to be determined

Table 3 indicated how Halverson and Conradi's taxonomy features were used in explaining the identified cybersecurity capability maturity models. The features give a full description of all the models, such as their origin, reasons for creating the model, number of maturity levels, where it is applicable, who can use the model in the organization, how depth the implementation guidelines, etc. This description will help other organizations to see the features of each model and where it can be applied. Furthermore, the table can be an inside for top management of an organization that has less knowledge of cybersecurity to decide or decide on what type of model would suit their organization. The research aims to identify models from 2011 to 2019, but mostly relevant material from 2011 to 2019 was used, this is because only a few models are specifically cybersecurity oriented, those what were identified but did not fit the inclusion criteria includes like Control Objectives for Information and Related Technology Organization (COBIT), Project Management Maturity Model (OPMM) and Siemens Knowledge Management Maturity Model (KMMM) were not used in this research.

Table 3: Comparative review on cybersecurity capability maturity model

Model / Features	C2M2	ES-C2M2	ONG-C2M2	NICE-C2M2	CCSMM	FFIEC-CMM	AUMMCS
Cybersecurity Oriented	Yes	Yes (derived fromvC2M2)	Yes (derived from ES-C2M2)	Yes	Yes	Yes	Yes
Origin	The US.Dept of Energy	The US.Dept of Energy	The US.Dept of Energy	The US.Dept of Energy	CIAS	US Federal Financial Institute Of Examination Council	Centre For Cyber Security At The University of Johannesburg
Maturity level	4	3	4	3	5	5	4
Purpose	Assessment of cybersecurity capabilities for any organization comprises of a maturity model evaluating a tool	Tailored to energy subsector	Tailored to the oil and natural gas subsector	Tailored to three areas: process and analytics, integrated governance, skilled practitioners and technology for work development	Tailored to communities yardstick to know the security posture	Tailored to as assessment tools to identify organizational risk and determine their cybersecurity maturity	Tailored to ensuring citizens and government and business are protected African member states
Actors	Management	Management	Management	Staff	Communicates	Management/ Employees	States
Organization Size	large	large	large	large	Medium	large	All
Level of Documentation	Medium	Medium	Medium	Medium	Low	High	Medium
Organization Environment	Overall	Overall	Overall	Explicit	Explicit	Explicit	Overall
Define role	Yes	Yes	Yes	Yes	No	No	No
Depth of Assessment	Specific	Specific	Specific	General	Specific	Specific	General
Field Applicable	Organization	Electricity	Oil and Natural Gas	Workforce	Communities	Financial Organization	African states
Prescriptive/ descriptive	Both	Both	Both	Both	Descriptive	Both	Both
Assessment	Organization maturity	Electricity grid protection	Oil and gas protection	Organization maturity	Community protection	Organization maturity	Data protection
Assessor	Internal and external	Internal and external	Internal and external	Internal and external	external	external	external
Implementation Cost	1	1	1	1	1	1	1i

5. Result and Discussion

Cybersecurity has been growing exponentially day by day both in private and public sectors, so also cyber threats. These threats are dynamic and organizations need to be updated on the measures they should adopt to secure the critical assets. In this article, a review was carried out to identify the most recent mechanism used in protecting and also identifying the maturity of cybersecurity in an organization. The research is limited to models from 2011 to 2019 and also those designed specifically for cybersecurity as specified in the first objective. Table 3 elaborated on the comparisons of the identified models, the table shows most of the models have basic similarities, such as domains and levels, but also differ in some areas which include the level of implementation and guidelines, the actor's role, the field of application, and also assessment. Objective two of this article was to identify the maturity and the level of the models, the models have almost similar maturity description. Some models use levels like C2M2, while others use the baseline to innovation and others use initial to vanguard to describe how maturity increase from one level to the next. Overall they have basic similarities.

However, some models were derived from other models like in the case of the ES-C2M2, and the ONG-C2M2 models are derived from C2M2. Most models are more specific than generic. The last objective was to understand the application domain of the identified models. Most of the models' application domain includes organization, oil and gas section, communities, banking sectors, and even continent as a form of guidelines as shown in table 3. Certain models are designed to be used for the entire organization like C2M2 while other are not like NICE-C2M2. The complete adoption of a model seems to be impossible as the most model is designed for a specific purpose as in NICE-C2M2 which was designed for skilled staff. This discussion further shows that organization can assess their needs before selecting an appropriate model to measure their cybersecurity maturity level.

6. Research Direction

This paper explains the cybersecurity maturity models properties and their similarities and Applications domain, based on the reviews of all the available models, no any author explains the validation process of the proposed model before implementation, therefore, a future research can focus on how cybersecurity capability maturity models are evaluated and also cost of implementation of the model in an organization as no model explains the financial standpoint.

7. Conclusion

In conclusion, cybersecurity measures is an essential entity to be known by all organizations, identifying organizational maturity level and knowledge on cybersecurity is a most, also knowing what model to be used in identifying the maturity level is important. There is limited research on cybersecurity capability maturity models and their application as the research area is new and growing exponentially. This research will serve as the first step in knowing the relevant cybersecurity capability maturity models available and also areas of application. However, all the identified models are fully based on cybersecurity but adopting can be impossible, however, the models can be adapted and customized. Tables 3 give a clear view of all the models and how to choose a suitable model for any organization based on the features used. Furthermore, only C2M2 focuses on the entire organization while others focused on cybersecurity. Lastly, all models found after the SR lacks cost implementation, therefore, to know how much to spend for implementing any model depends highly on the size of the organization and the number of critical assets to be protected.

Conflict of Interest

The author declares no conflict of interest

Acknowledgment

This journal would not have been possible without the exceptional support and guidance of my supervisors Dr. Maheyzah Muhamad and Dr. Siraj Siti Hajar Othman. Their enthusiasm, knowledge, and exacting attention to details have been an inspiration and kept this journal work on track. I would also like to thank all the reviewers that gave their comments to make this paper acceptable to the community of knowledge.

References

[1]. M. Dunn Cavelty, "Breaking the Cyber-Security Dilemma: Aligning Security Needs and Removing Vulnerabilities," Science and Engineering Ethics, **20**(3), 701–715, 2014, doi:10.1007/s11948-014-9551-y.
[2]. Nicholas R., The Cybersecurity Dilemma, Duke University, 2011.
[3]. R.W. Taylor, Cyber Crime and Cyber Terrorism, 4th ed., Pearson Education, Inc., United State, 2019.
[4]. Z. Soltani, N.J. Navimipour, "Customer relationship management mechanisms: A systematic review of the state of the art literature and recommendations for future research," Computers in Human Behavior, **61**, 667–688, 2016, doi:10.1016/j.chb.2016.03.008.
[5]. Kitchenham, Guidelines for performing Systematic Literature Reviews in Software Engineering, Durham Durham, UK, 2007, doi:10.1145/1134285.1134500.
[6]. A.D. Oxman, "Systematic Reviews: Checklists for review articles," BMJ, **309**(6955), 648–651, 1994, doi:10.1136/bmj.309.6955.648.
[7]. M.J. Grant, A. Booth, "A typology of reviews: An analysis of 14 review types and associated methodologies," Health Information and Libraries Journal, **26**(2), 91–108, 2009, doi:10.1111/j.1471-1842.2009.00848.x.
[8]. M.C. Paulk, "A History of the Capability Maturity Model for Software," The Software Quality Profile, **1**(1), 5–19, 2009.
[9]. C. V Weber, S.M. Garcia, M. Bush, "Key Practices of the Capability Maturity Model," 1993.
[10]. Y. Goksen, E. Cevik, H. Avunduk, "A Case Analysis on the Focus on the Maturity Models and Information Technologies," Procedia Economics and Finance, **19**(15), 208–216, 2015, doi:10.1016/s2212-5671(15)00022-2.
[11]. R.M. Adler, "A dynamic capability maturity model for improving cyber security," 2013 IEEE International Conference on Technologies for Homeland Security, HST 2013, 230–235, 2013, doi:10.1109/THS.2013.6699005.
[12]. P. Byrnes, M. Phillips, "-- ~ 47L ~ Software Engineering Method Description ESC-TR-96-002," (April), 1996.
[13]. D. Budgen, M. Turner, P. Brereton, B. Kitchenham, "Using Mapping Studies in Software Engineering," Ppig, **2**, 195–204, 2008.
[14]. N. Khatibian, T. Hasan gholoi pour, H. Abedi Jafari, "Measurement of knowledge management maturity level within organizations," Business Strategy Series, **11**(1), 54–70, 2010, doi:10.1108/17515631011013113.
[15]. G.B. White, "The Community Cyber Security Maturity Model The Center for Infrastructure Assurance and Security," Proceedings of the 40th Hawaii International Conference on System Sciences, (June), 1–8, 2007, doi:10.1109/HICSS.2007.522.
[16]. L. Johnson, Cybersecurity framework, 2020, doi:10.1016/b978-0-12-818427-1.00012-4.
[17]. W. Miron, K. Muita, "Technology Innovation Management Review

Cybersecurity Capability Maturity Models for Providers of Critical Infrastructure," Technology Innovation Management Review, **4**(October), 33–39, 2014, doi:10.22215/timreview/837.

[18]. A. Sorini, E. Staroswiecki, 8. Cybersecurity for the Smart Grid, Elsevier Ltd, 2017, doi:10.1016/B978-0-12-805321-8.00008-2.

[19]. J.A.C.-M. and I.D.S.-G. Angel Marcelo Rea-Guaman, Tomás San Feliu, "Comparative Study of Cybersecurity Capability Maturity Models Angel," Computer Standards and Interfaces Software Process Improvement and Capability Determination Conference 2017, **60**, 1–2, 2018, doi:10.1016/j.csi.2018.05.002.

[20]. S.H.B. Von Solms, "A maturity model for part of the African Union Convention on Cyber Security," Proceedings of the 2015 Science and Information Conference, SAI 2015, 1316–1320, 2015, doi:10.1109/SAI.2015.7237313.

[21]. G.B. White, "The community cyber security maturity model," 2011 IEEE International Conference on Technologies for Homeland Security, HST 2011, 173–178, 2011, doi:10.1109/THS.2011.6107866.

Appendix A

Table 4: C2M2 Domain description

Domains	Grouped Objectives
Asset, Change and Configuration Management	Manage Asset inventory Manage Asset configuration Manage changes to Asset Management Activities
Cybersecurity Program Management	Established Cybersecurity Program Strategy Sponsor Cybersecurity Program Established And Maintain Cybersecurity Architecture Perform Secure Software Development Management Activities
Event and Incident Response, Continuity of Operation	Detect Cybersecurity Events Escalate Cybersecurity Events And Declare Incidents Respond To Incident And Escalated Cybersecurity Events Plan Continuity Management Activities
Identify and Access Management	Established And Maintain Identities Control Assess Management Activities
Information Sharing and Communications	Share Cybersecurity Information Management Activities
Risk Management	Established Cybersecurity Risk Management Strategy Manage Cybersecurity Risk Manage Activities
Situational Awareness	Perform Logging Perform Monitoring Established And Maintain A Common Operating Picture Management Activities
Supply Chain and External Dependencies Management	Identify Dependencies Manage Dependency Management Activities
Threat And Vulnerability Management	Identify And Respond To Threats Reduce Cybersecurity Vulnerabilities Management Activities
Workforce Management	Assign Cybersecurity Responsibilities Control The Workforce Life Cycle Develop a Cybersecurity Workforce Increase Cybersecurity Awareness Management Activities

Table 5: C2M2 Maturity level description

Maturity indicator level MIL	Level description
Level 0	This level has no practices or processes defined. There is no stable environment for activities. MIL 0 is given as a result of the domain objective not achieved.
Level 1	This level contains a set of initial practices. This level activities are usually ad hoc and chaotic. MIL 1 is scored if there is an initial practice performed
Level 2	This level has more stable practice compared to MIL, more confidence is achieved at this level as the result of the performance and is sustained over time.
Level 3	At MIL 3 policy is applied to the practices to further stabilize the operations in the organization and is guided by top-management directives. Also, the staff s' are fully trained and fully funded.

Table 6: ES-C2M2 and ONG-C2M2 domain description

Domain	Practices
Risk	Risk Assessment
Assets	Asset, Change, and Configuration Management
Access	Identity and Access Management
Threat	Threat and Vulnerability Management
Situation	Situational Awareness
Sharing	Information Sharing And Communication
Response	Event And Incident Response, Continuity Of Operations
Dependences	Supply Chain And External Dependencies Response Management
Workforce	Workforce Management
Cyber	Cybersecurity Program Management

Table 7: ES-C2M2 and ONG-C2M2 maturity level

Maturity Level	Description
MIL 0 " Not Performed"	This level describes the domain has achieved nothing.
MIL 1 " Initial"	This level shows only initial practices are performed
MIL 2 "Performed"	The level is characterized by having well-documented practices, stakeholders' involvement, and provision of standards or guidelines for practice implementation.
Mil 3 "Managed "	This level shows all practices and activities are fully guided by policy, also practice is only assigned to adequate skills personal. The formed policy are periodically evaluated for improvement

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 763-775 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Frequency Scaling for High Performance of Low-End Pipelined Processors

Athanasios Tziouvaras[*,1], Georgios Dimitriou[2], Michael Dossis[3], Georgios Stamoulis[1]

[1]*Department of Computer Engineering, University of Thessaly, Volos, 38221, Greece*

[2]*Department of Computer Science, University of Thessaly, Lamia, 35100, Greece*

[3]*Department of Computer Science, University of Western Macedonia, Kastoria, 52100, Greece*

ARTICLE INFO

Article history:
Received: 30 November, 2020
Accepted: 02 March, 2021
Online: 27 March, 2021

Keywords:
Better-than-worst-case
Timing analysis
Adaptive clock scaling
IoT processors
RiscV design implementation
Timing speculation

ABSTRACT

In the Internet of Things era it is expected that low-end processor domination of the embedded market will be further reaffirmed. Then, a question will arise, on whether it is possible to enhance performance of such processors without the cost of high-end architectures. In this work we propose a better-than-worst-case (BTWC) methodology which enables the processor pipeline to operate at higher clock frequencies compared to the worst-case design approach. We employ a novel timing analysis technique, which calculates the timing requirements of individual processor instructions statically, while also considering the dynamic instruction flow in the processor pipeline. Therefore, using an appropriate circuit that we designed within this work, we are able to selectively increase clock frequency, according to the timing needs of the instructions currently occupying the processor pipeline. In this way, the error-free instruction execution is preserved without requiring any error-correction hardware. We evaluate the proposed methodology on two different RiscV Rocket core implementations. Results with the SPEC 2017 CPU benchmark suite demonstrate a 12% to 76% increase on the BTWC design performance compared to the baseline architectures, depending on the appearance rate of instructions with strict timing requirements. We also observe a 4% to 37% increase on power consumption due to the operation of the pipeline at higher clock frequencies. Nevertheless, the performance increase is up to nine times larger than the power consumption increase for each case.

1 Introduction

Traditional microprocessor design ensures an error free instruction execution on general purpose processors. According to the established model, the designer designates the clock frequency and the voltage values of the processor, so that no timing violation of the critical path occurs. Thus, the design revolves around the timing analysis of the worst-case scenario, and the critical path acts as a strict timing threshold, constraining the processor performance.

In contrast to the traditional model, the better-than-worse-case (BTWC) paradigm attempts to relax any critical path restrictions through timing speculation (TS), by scaling up and down the processor voltage or clock frequency, allowing timing errors to occur. The resulting errors can then be resolved by an integrated rollback error correction mechanism. Such a paradigm presents many design opportunities for performance enhancement and power reduction.

This paper is an extension of work originally presented in MOCAST conference in [1] and is loosely based on the BTWC design paradigm, primarily focusing on the performance increase of the processor pipeline for low-cost processors. In this work, we present instruction-based clock-scaling, a methodology that improves performance by executing instructions in varying clock frequencies, according to the opcode of the executing instructions. A proposed timing analysis methodology detects instruction opcodes that may run at high speed, as well as instruction opcodes that must run at low speed. The instruction pipeline is then fed with multiple clock signals, multiplexed in a way that when a critical instruction is decoded, a slower clock is selected for the cycle that exhibits the maximum timing delay, reverting to the previous clock in the following cycle. Thus, the typical clock selected can be faster than the one designated via traditional timing analysis.

Such an approach diverges from classic BTWC design techniques, in that we utilize the knowledge about each individual instruction timing requirements, obtained from our timing anal-

[*]Corresponding Author: Athanasios Tziouvaras, attziouv@inf.uth.gr

ysis methodology. With that knowledge, our model foresees any upcoming timing errors and corrects them a priori. In this way, we eliminate any execution error probability, thus no error recovery mechanism is deployed, as each instruction is certain to meet its timing requirements. Since performance penalty induced by error correction is non-trivial, our more deterministic error detection mechanism avoids the performance implications of traditional BTWC techniques. Furthermore, the developed methodology is architecture independent making it applicable to any given single-issue in-order-execute design, without requiring intrusive changes to its microarchitecture.

The above technique has been implemented and tested on two different post-layout RiscV Verilog processor implementations, using the SPEC 2017 CPU benchmark suite. The same tests have been applied on the implementations without clock-scaling, and the results obtained show a clear improvement in processor performance, between 12% and 76%, and an average 3.7-fold improvement in performance-to-power ratio, despite the expected increase in power consumption due to high frequency operation. Further comparison with state of the art TS methodologies highlights the effectiveness of our technique when considering throughput improvement.

The rest of this paper is organized as follows. Section 2 contains a research review on the topics of our research, whereas Section 3 discusses the proposed processor timing analysis technique. Our opcode-based clock-scaling technique is presented in detail in Section 4. The processor model used and the experiments conducted are described in Section 5. Finally, Section 6 gives the conclusion of our work.

2 Previous research

Following decades of standard worst-case processor design methodologies, a considerable amount of research is being conducted in BTWC designs in the last years. The BTWC is a paradigm that encompasses various techniques which approach the critical path timing requirements more flexibly than traditional designs [2]. In particular TS, which is a type of BTWC design, violates critical path restrictions, allowing and then correcting any resulting execution errors. Such technique enables researchers to experiment with energy-performance tradeoffs and manages to increase the performance and lower the power consumption of the circuit [3].

The application of TS on processor designs has led to the development of Razor [4], [5]. Razor employs TS to improve the performance-to-power ratio of the design, and utilizes “shadow latches”, which identify and correct the timing errors made by the alteration of the voltage. This correction mechanism operates in real time and ensures the error free instruction execution. Another research work focuses on the dynamic frequency scaling of a superscalar processor, supported by error recovery mechanisms to compensate for resulting timing errors [6]. In that work, researchers deploy both local and global error correction mechanisms, which ensure a correct instruction execution when the processor is overclocked at higher frequencies.

BTWC design methodologies have also proven able to address the ever increasing process variation effects or the uncertainty caused by the environment and the fabrication process of the integrated circuits [7]. As a result, a significant amount of research compensates for process, voltage and temperature (PVT) variation, while exploring possible TS benefits [8]. In such cases, probabilistic methods can be deployed to model the PVT fluctuations [9], while guardbanding has been proposed in one work to safeguard the circuit against timing violations [10]. Another work shows that process variation effects result in pipeline imbalances as long as timing delay is concerned [11]. In order to overcome such problem, a framework has been developed to tighten the timing of the circuit using a time stealing technique that equalizes the timing delays of each pipeline stage. A third work exhibits that the mitigation of PVT effects may be achieved by a properly developed framework [12]. In that work, the researchers manage to model the PVT effects and create a framework that enables error-power and error-frequency tradeoffs. Finally, another work demonstrates novel techniques which may be used to design PVT resilient microprocessors [13]. Such techniques include the monitoring of critical paths by sequential circuits that detect timing errors, or the monitoring of each pipeline stage for worst-case delays. In both cases, the designers also propose error recovery mechanisms and exhibit a significant performance increase by utilizing clock frequency scaling.

Previous work has also shown that designers may obtain efficient energy-performance tradeoffs at a higher architecture level [14] when applying BTWC paradigms. A marginal cost analysis demonstrates the potential of circuit voltage scaling on architectural level, highlighting the optimal operational point of a target processor. Furthermore, the design process of a microprocessor could also be aligned to facilitate TS friendly microarchitecture adjustments. Specifically the optimization of the most frequently exercised critical paths may result in clock frequency scaling and lower power dissipation on existing TS architectures [15]. Along the same lines, the slack redistribution of the most frequently occurring timing paths of a processor may lead to architectures with lower power consumption and minimum error rate [16]. Another work in [17] proposes a TS cache design which manages to lower the energy consumption of the system while maintaining high cache hit ratios within various cache organizations.

As the TS paradigm revolves around scaling the clock frequency in real time, research is also focused on clock adaptation techniques. Specifically, previous work in [18] manages to adapt the clock frequency of a POWER 7 processor core by adjusting the voltage level in firmware level. Combined with a critical path monitoring mechanism, researchers achieve voltage scaling when the critical path is not excited while using the available timing margin as a guardband mechanism. Another work utilizes a unary coding scheme to enable the PLL to quickly adapt to the required clock changes in real time [19]. This approach can be applied on a single core clock to enable its dynamic frequency change without imposing significant delays. A research that also underlines the importance of a robust real time clock adaptation scheme is [20]. In this work, researchers manage to deploy a clock adaptation scheme which can reduce the clock frequency in an AMD 28nm microprocessor core improving the power efficiency of the system. This approach utilizes a phase generator which can modify the clock’s phase in order to stretch its period. Similarly, in [21] authors employ a dynamic clock adjustment technique on a simple processor pipeline to adjust the clock frequency according to the application type that is being executed

on the processor pipeline.

TS designs are often prone to exhibit metastable behavior, resulting in non-deterministic timing phenomena which should be taken into account when designing a TS processor [5], [22]. This issue usually appears when the input data arrives close to a rising clock edge, resulting in the possibility of undetected errors. In the course of minimizing the metastable behavior of timing speculative circuits, design methods have been proposed, utilizing a time borrowing technique alongside of a careful examination of the data path timing, which simplifies the issue of metastability by moving such behavior to the circuit error path only [23]. Despite the progress being made on this issue, a more recent work claims that circuit metastable behavior in TS designs is not yet efficiently addressed [22]. In this regard the mean time between failures of such designs discourages any possible industrial applications.

Another thought provoking aspect in BTWC techniques is the error recovery mechanism and the performance penalty it imposes on the design [24]. Due to that penalty significant effort has been made on the improvement of the deployed error prediction mechanisms [25]. While some designs employ statistical methods to successfully detect specific instruction sequences which have pre-analyzed timing requirements [10], others tend to focus on monitoring the critical path excitation by individual instructions [26], [27]. Another approach revolves around the identification of timing critical instructions during runtime, using that knowledge to improve the energy-efficiency ratio of the processor [28]. Finally a study on the cmos recovery mechanism reveals the impact of the technology on such techniques, as researchers develop a hardware model sufficient to simulate timing speculation designs [29]. The same research also underlines the importance of a fine-grained error recovery mechanism in BTWC designs. Although the penalty imposed by the execution or by the unsuccessful prediction of a critical instruction is relatively low, results indicate that the performance loss due to error recovery is non-trivial. Moreover, in some extreme cases the design's performance deteriorates to the point that the TS design displays lower throughput than the baseline processor [10].

From all the above reviewed work we have concentrated our interest on the issues of error recovery in the TS design paradigm, as well as on the issues of metastability observed in that paradigm. Our motivation has been to study such issues and come out with a novel technique that exploits TS in a way that any possible speculation errors are detected dynamically and recovered before they appear, thus avoiding the costly error recovery mechanisms, and at the same time eliminating metastability phenomena altogether. In the following sections we introduce our opcode-based timing analysis and clock scaling technique for error-free timing speculation in pipelined microprocessors.

3 Timing Analysis in Processor Datapaths

3.1 Static and Dynamic Timing Analysis

Timing analysis is a technique traditionally used in order to analyze the timing requirements and timing delays of a digital circuit. By employing the traditional timing analysis approach, designers designate the optimal clock frequency for a design which is usually derived by the timing requirements of the critical path.

Standard static timing analysis (STA) is performed either at flop to flop or at input to output basis. STA is used to calculate the worst-case delay of the circuit and to detect any timing violations that may occur under certain design constraints. STA can be used to identify the critical path of the circuit, which is a major factor for the clock period selection of the design. However, STA is overly pessimistic, as it calculates the worst-case slack for timing paths and thus, it is considered inefficient for the BTWC paradigm.

On the contrary, dynamic timing analysis (DTA) is used to extract more accurate timing information from a digital circuit. DTA utilizes a large set of vectors which are used as circuit inputs that excite the circuit paths for various input combinations. An exhaustive DTA will calculate every possible timing path for a given circuit, at a very high time cost though. DTA eliminates the path pessimism as it analyzes the timing requirements of every possible input of the circuit. Although DTA would be more appropriate for BTWC design timing analysis, its time cost renders its usage impossible on large designs, as it trades accuracy for completion time.

3.2 The Pseudo Dynamic Timing Analysis Concept

The technique proposed in this work revolves around the BTWC design paradigm. In this sense, we develop a timing analysis methodology that extracts timing information from a circuit to enable us to take advantage of the timing differences of the circuit paths. To this end, we analyze each individual instruction supported by the processor, with respect to the unique timing requirements it presents. In order to analyze each instruction independently, we isolate from the circuit all the possible paths an instruction may take, while declaring the rest of the paths as false. We iterate this process until we exhaust all the available supported instructions. In the sequel, we perform STA on each separate path group related to each individual instruction. As a result, the timing results we obtain depict the worst-case delay of each instruction, instead of depicting the worst-case delay of the processor critical path.

In order to present our technique, we refer to a standard timing analysis tool, with which we conduct timing analysis on the post-layout netlist of a recent open-source 64-bit six-stage pipelined processor implementation of the Risc-V ISA. Figure 1 depicts our initial approach to the problem. First, we pick an instruction supported by the processor. Then, we examine the processor ISA to identify the instruction opcode, while ignoring any register or data fields facilitated within the instruction word. For that opcode we commence a "case_analysis" operation, in which the designer may set any of the circuit inputs to constants, and let the tool being utilized perform a flop to flop timing analysis given the fact that some of the circuit inputs are set to a fixed value. In our case these values represent the current instruction opcode field. The tool propagates any generated signals through the processor pipeline to analyze the timing of each pipeline stage separately while performing STA. As a result we calculate the timing requirements for each pipeline stage separately with respect to the fixed instruction opcode and thus, for the corresponding instruction. That report gives the worst-case scenario of the analyzed instruction with respect to timing.

```
1   Start
2   instruction_number = number_of_supported_instructions
3   While (instruction_number > 0)
4      instruction =  next_instruction
5      opcode =  instruction [opcode]
6      set_case_analysis [opcode]
7      stage_delays = STA_timing_report [netlist, opcode]
8      worst_case[instruction_number] = max[stage_delays]
9   instruction_number --
10  End
```

Figure 1: The proposed timing analysis methodology which studies each instruction as an individual entity.

```
1   Start
2   instruction_number = number_of_supported_instructions
3   While (instruction_number > 0)
4      instruction =  next_instruction
5      for (all possible opcode transitions)
6            opcode =  instruction [opcode]
7            set_case_analysis [opcode, opcode_transition]
8            stage_delays = STA_timing_report
             [netlist, opcode, opcode_transition]
9            worst_case[instruction_number] = max[stage_delays]
10     instruction_number --
11  End
```

Figure 2: The timing analysis methodology which compensates for the dynamic voltage change of the opcode field.

In the sequel, we keep record of the slowest pipeline stage timing, before moving on to the next instruction. Finally we pick another instruction and reiterate this process, until all the available instructions are exhausted.

It becomes clear that the proposed timing analysis methodology is a hybrid between STA and DTA. In this sense, it performs STA for each instruction of the processor ISA, but it further analyzes all possible instructions, giving a DTA flavor to the result. However, it is not a full DTA, since it only varies the opcode field of the instruction word, thus not considering input values neither for any other field of the instruction word nor for any other part of the circuit. In particular, the implementation of the architecture we work on supports 215 instructions, which means that we only need 215 analyses in our method, instead of 264 which would be needed for a full DTA on the variations of the instruction word alone, or many more if we were to consider other circuit inputs as well. On the other hand, our approach is not as pessimistic as classic STA, and with only little higher complexity it can produce designs that exhibit significantly better performance than designs produced by STA. Our second approach presented next can produce even more efficient designs.

3.3 Dynamic Opcode Value Changes

Using the proposed technique we manage to analyze the timing requirements of each processor instruction individually. To this end, we are iteratively affixing certain circuit inputs at constant voltage values. Specifically, we are bounding the opcode field bits to static binary values in order to analyze each instruction behavior.

In real time digital circuits, inputs change in a dynamic way as new values are stored into the pipeline registers at the rising edge of the clock signal, immediately before they are needed and used. As a result, new instructions are fetched for execution on each clock cycle and thus, inputs representing the opcode bits of each instruction are not constrained to fixed voltage values; instead they dynamically change, resulting in an unpredictable transient timing behavior in every pipeline stage after the fetch stage. Such behavior appears at the decode stage due to the opcode bits per se, as well as at all subsequent stages due to the control bits produced by the opcode bits and is propagated to such stages. For this reason, the discrepancy between our initial concept and a real time system behavior in timing deviations should be addressed.

In order to compensate for the dynamic voltage change in the processor pipeline inputs, we employ a modification of the aforementioned timing analysis methodology. Since our focus is now around the timing variance created by the dynamic behavior of the instruction opcode field, we consider the opcode as a bit sequence whose length is defined by the ISA. Consequently we have to take into account every possible value transition which leads to the currently analyzed bit sequence. Normally the amount of all possible combinations grows exponentially with the length of the sequence. We consider this approach unsuitable for our needs as its high time cost makes it impossible for practical application. Furthermore, we aim at the development of a methodology, which can be employed to analyze any ISA, without depending on the instruction opcode length of the design.

The solution we propose to resolve this is based on the observation that in processor architectures not all possible bit transitions lead to valid bit sequences of the opcode field. More specifically, as the instructions succeed one another during the instruction fetch stage, the number of valid opcode bit combinations is constrained by the number of the instructions supported by the ISA. So instead of analyzing the timing delay of each possible opcode bit sequence transition, we focus on the analysis of each possible instruction succession.

Figure 2 depicts the proposed solution, which relies on the initial concept as described earlier, augmented with the dynamic value change compensation approach we discussed. In this solution, we analyze each instruction's timing requirements individually as before, but instead of using fixed voltage values to describe the currently examined instruction, we analyze each possible opcode transition that could lead to the opcode bits of the current instruction. Such transitions represent any rising and falling voltage values that could result in that particular bit sequence. The timing analysis of such cases is studied individually, while the timing analysis tool propagates all generated signals through the processor pipeline. We still use the "case_analysis" function, as it can be expanded to include rising and falling voltage change. When we complete the timing analysis of an instruction, we save the worst-case delay. Afterwards we proceed with the analysis of the next instruction, until all supported instructions are exhausted.

The method we proposed in this section uses STA to find the worst-case delay path of each individual instruction. But in order to achieve an accurate timing result we utilize an exhaustive iterative analysis resembling more now that of a DTA method. However,

even if we restrict value variations within the instruction opcode, with an opcode field of x bits, a standard DTA approach would require 2^x iterative timing analyses to effectively analyze the timing of the opcode length, as each possible x-bit combination may lead to the required sequence. Instead, using our methodology, the number of iterations needed for each opcode analysis is only equal to the total number of supported instructions, in our case 215 iterations.

The proposed timing analysis technique lies somewhere between STA and DTA, closer to STA with respect to complexity, but closer to DTA with respect to output quality. We call this technique Pseudo Dynamic Timing Analysis (PDTA). In the following section we discuss the application of PDTA to adaptively scale the clock frequency of the core pipeline of a processor.

4 Clock scaling using opcode-based pseudo dynamic timing analysis

4.1 Adaptive clock scaling in pipelined processors

Adaptive clock scaling is often used for power control in modern high-performance processor architectures. Processor cores can be slowed down when not in full use, in an attempt to reduce power consumption and avoid overheating. In some cases, cores can be sped up for a limited time, in order to boost performance for cycle-hungry applications. On the other hand, low-cost processors that are preferred for embedded and low-performance systems can also use clock scaling in order to increase performance, especially if this is achieved in a fairly cheap manner.

Another way to effectively scale up the clock frequency is to deepen the processor pipeline. In this way each stage's latency is reduced and the system may operate lower clock periods. Previous works in [30] and [31] demonstrate that deepening the processor pipeline results in an increase in circuit area and power consumption due to the implementation of additional pipeline registers, Furthermore, deep pipelined processors require more complicated forwarding, control and stalling mechanisms thus, further impairing the design's area and power requirements. As a result, authors in [30] and [31] conclude that the increase of the pipeline stage amount does not necessarily result in performance increase as the costs of wrong branch predictions and pipeline flushing become greater.

Our work focuses on low-cost IoT processors which present significantly low area and power requirements as stated in [32], [33]. Under this premise we do not opt in deepening the pipeline width or enhancing the complexity of the system, instead we focus on increasing the processor throughput while preserving the system microarchitecture as it is. In this section we will present how PDTA can be employed for such a purpose.

4.2 Scaling clock by opcodes

PDTA can be used to acquire timing information for each instruction separately and thus, we aim to use such information for adaptive clock scaling based on the instruction opcode. In order to validate our clock-scaling technique, we will continue using the same 64-bit six-stage pipelined processor that we referred to in Section 3, which is a single-issue in-order-execute RISC-V processor implementation. We will refer to this implementation as "baseline processor". In order to tighten the timing of the processor's functional units we also deploy a second implementation that utilizes pipelined functional units. As a result time consuming operations require more clock cycle to complete but they display lower latency. We will refer to this implementation as "pipelined execution". We classify the obtained results into 11 instruction classes as shown in Table 1. Each instruction class contains a group of individual instructions with similar timing requirements. We also pinpoint the slowest pipeline stage in terms of delay, for every class and we refer to such stage as "critical stage". Finally, we assign a "worst-case delay" value to each class, which is the highest instruction delay in the corresponding group. The classes go as follows:

- The *Logical* instruction class which includes logical operations such as *and*, *ori* and *xor*.
- The *Shift* instruction class which includes shift operations such as *shift left logical* or *shift arithmetic*.
- The *Comparison* instruction class which includes *bit comparison* operations.
- The *Jump* instruction class which includes jump operations such as *jump and link*, *jump register* or *jump*.
- The *Multiplication* instruction class which includes integer *multiplication* operations.
- The *Division* instruction class which includes any *integer division* operations.
- The *Other arithmetic* instruction class which includes all other *integer arithmetic* operations except for multiplication and division, such as addition or subtraction.
- The *Memory access* instruction class which includes any *memory access* operation such as load word or store byte.
- The *FP Multiplication* instruction class which includes *floating point multiplication* operations.
- The *FP Division* instruction class which includes *floating point division* operations.
- The *Other FP arithmetic* instruction class which includes all other *floating point arithmetic* operations except for FP multiplication and division, such as FP addition or subtraction.

Table 1: Analysis of the instruction classes of the RiscV Rocket core architecture.

Instruction class	Slowest pipeline stage (critical stage)	Baseline worst case delay	Pipelined execution worst case delay
Logical	Execute stage	1.2 ns	1.2 ns
Shift	Execute stage	1.5 ns	1.5 ns
Comparison	Execute stage	1.5 ns	1.5 ns
Jump	Execute stage	1.1 ns	1.1 ns
Multiplication	Execute stage	2.9 ns	1.5 ns
Division	Execute stage	3.3 ns	1.1 ns
Other arithmetic	Execute stage	1.9 ns	1 ns
Memory access	Memory stage	3.9 ns	1.3 ns
FP Division	Execute stage	3.7 ns	1.3 ns
FP multiplication	Execute stage	3.2 ns	1.1 ns
Other FP arithmetic	Execute stage	3.0 ns	1 ns

After studying the aforementioned instruction classes we observe the following:

- Each pipeline stage presents unique timing requirements depending on the instruction being executed.
- Some pipeline stages may produce error free results while utilizing higher clock frequencies than the others.
- The error free instruction execution is preserved if we satisfy the timing requirements of each individual pipeline stage for the executing instruction.

We deduce that we can dynamically adapt the clock period of the processor during the run time in order to achieve higher throughput, while also guarantying the error free instruction execution. To this end, we isolate the critical instructions, i.e. the instructions that constrain the clock frequency and we display the results we obtained for each implementation in Table 2. We track down all critical instruction classes for our architecture and we assign a minimum operational clock period to each one of them. Due to the prior timing analysis, we are guaranteed that each critical instruction will execute without errors at the designated clock period. We also consider a typical clock period for each design which is suitable for the error free execution of non-critical instructions.

Our design focuses on letting the pipeline operate at high clock frequencies when critical instructions are absent. When a critical instruction is detected, we downscale the clock frequency, as soon as the critical instruction enters the pipeline stage which would otherwise cause a timing error. We refer to such stage as critical stage. Figure 3 shows an execution instance of a small instruction sequence on the Rocket core Baseline implementation. We track the minimum clock period with respect to the pipeline stages involved and we mark the critical stages that contribute to frequency downscaling. Under this premise, the critical stage is the slowest pipeline stage in which the clock frequency needs to be adjusted so that no timing errors to occur. In this example, the pipeline under examination may operate at higher clock frequencies during the 1st to 5th and 8th to 10th clock cycle, while the clock frequency must be lower at the 6th and 7th cycle.

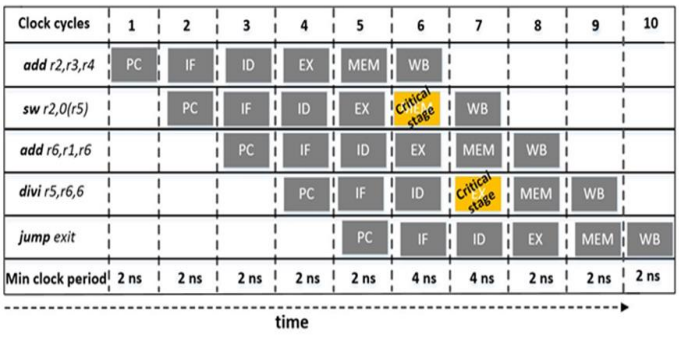


Figure 3: An instruction execution instance of the Rocket core implementation displaying the minimum operational clock period during each stage.

Table 2: The clock periods for critical instructions along with the typical clock period for the Rocket core implementation.

	RiscV baseline	RiscV pipelined execution
Critical instruction class	Multiplication, Division, Memory access, FP division, FP multiplication	Shift, Comparison, Multiplication
Critical instruction clock	4 ns	1.5ns
Typical clock	2 ns	1.3 ns

4.3 Dynamic clock scaling mechanism

We will now present in detail the circuit we designed to enforce the adaptive selection of the clock frequency. The circuit uses the information extracted from the timing analysis proposed in Section 3, to designate whether the clock frequency should be adapted. As the decision making will be occurring in real time, our design needs to employ reliability and speed. Figure 4 displays the designed clock control unit, which is charged with such task, and is implemented on the Rocket core implementation. It consists of an instruction snooping module and a clock selection module. In order to make the design nonintrusive for the processor architecture, we attach the two-module circuit at the side of the fetch and the decode stages of the processor pipeline.

Instruction snooping module: To be able to change the clock frequency dynamically, we need information about the class of the instructions that are headed for execution. To this end, we implement an instruction snooping circuit that receives a copy of the instruction word coming out of the instruction cache. This circuit monitors the instructions fetched and tracks down their progress in the pipeline. Moreover, it utilizes lookup tables which contain both the critical instruction opcodes and the critical pipeline stage for each corresponding instruction as calculated in Section 4.2. Using this information, the circuit produces a logical output on whether the clock frequency must be changed, driving with that output the clock selection module. **Clock selection module:** This module propagates the appropriate clock pulse selected by the frequency selector mechanism of the instruction snooping module. The clock selection module inputs are the frequency selection signal generated by the instruction snooping module and two PLL signals, one of high and one of low frequency. The frequency selection signal determines which pulse will be selected for the pipeline clock, when the instruction arrives at the critical stage. If frequency is indeed switched to low, the module must revert to the high frequency in the following cycle. The selection circuit is a simple multiplexor with insignificant contribution to the total delay of the module.

In the implementation of the second module, we observed that reverting to the original frequency may result in an unstable pipeline clock behavior, as shown in Figure 5. Such a phenomenon exists due to the frequency difference between the clocks and may prove catastrophic for the instruction execution. We address this problem in the way shown in Figure 6, by generating another low frequency pulse signal with a 180 degrees phase shift of the original. We also design a phase selector circuit which is responsible for selecting the appropriate phase when necessary. The phase selector is signaled by the frequency selector when a frequency scaling event is about to happen. It then designates the selected PLL phase so that no unstable behavior is exhibited. The complexity of the phase selection mechanism depends on the number of PLLs required.

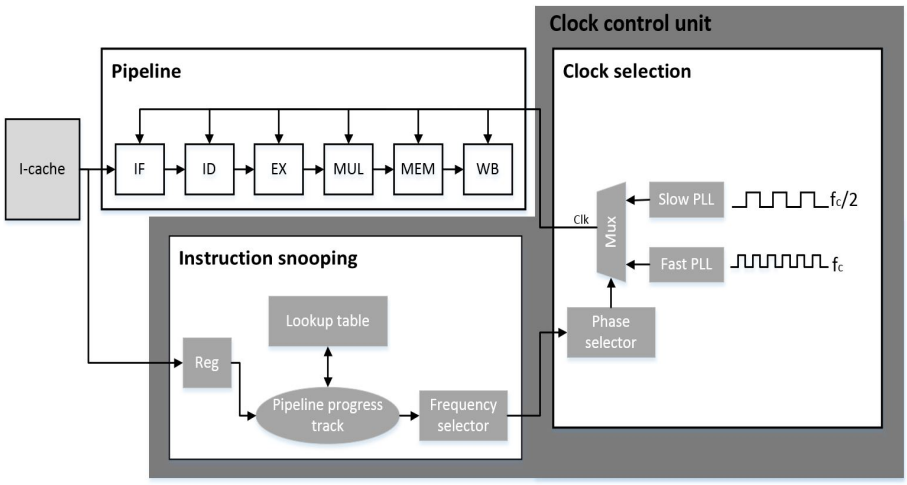


Figure 4: The clock control unit integrated in the rocket core.

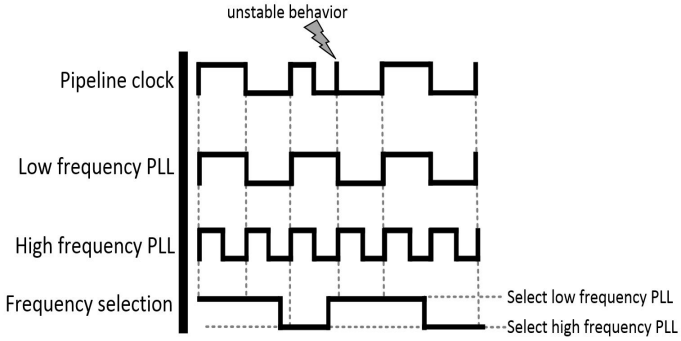


Figure 5: Unstable clock behavior due to subsequent clock selections.

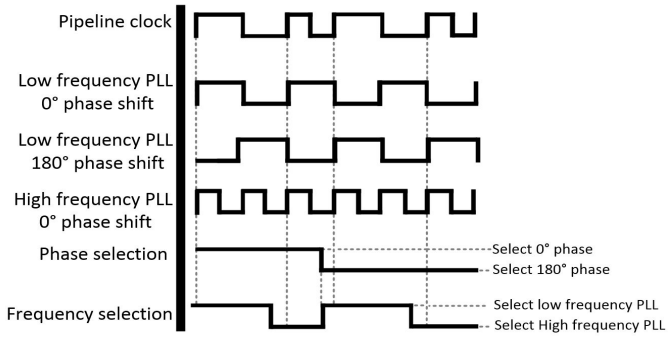


Figure 6: The clock instability compensation technique.

In the case of the implemented Baseline RiscV pipeline, where one period is an integer multiple of the other (2ns and 4ns), the generation of one additional shifted PLL resolves the problem. In other processor implementations though, additional PLLs may be needed, each with a specific phase shift, in order to enable the phase selector to compensate for all possible unstable behaviors of the pipeline clock. We adopt this approach as we acknowledge the need for a robust real time clock scaling mechanism. In contrast with [19] and [20] which manage to change the clock frequency for up to 7.5% of the core clock speed, we require much higher adaptation values. For that reason we do not change the clock frequency directly, instead we pre-generate the number of PLLs required and we proceed in selecting the appropriate candidate each time.

In general, when we have two clocks with a period ratio m:n, m and n being mutually primes, we need m PLLs for the phase shifts of the first and n PLLs for the phase shifts of the second clock, giving a total of m+n PLLs. Such a solution to the possible instability problem serves as the most efficient in terms of performance. As we saw earlier, clock frequencies used are the highest possible, with longer clock periods just enough to cover the critical pipeline stage delay. But this choice may result in a large number of PLLs. A cheaper solution would be to always use a slow clock period that is a multiple of the typical period. In this way we would not need that many PLLs, sacrificing performance for simpler implementation. In some cases, like the one examined above, it occurs that the optimal performance solution coincides with the cheapest solution, but this is definitely not the general case though.

4.4 Clock tree synthesis and distribution network

The proposed methodology requires the adoption of multiple PLLs to formulate the processor core clock frequency. The processor clock is dynamically selected according to the information obtained by the clock control unit and any change should be enforced within a very limited amount of time. Previous works in [21] and [15] have shown that cycle-to-cycle clock selection is possible and thus, the feasibility of the proposed methodology is ensured. To this end we opt to synthesize a single clock tree that reaches every register of the design, as we consider the propagation of all the available PLLs very costly. Under this premise, we perform the clock tree synthesis operation as if the design was operating under a single core clock with a constant clock frequency. The resulting clock distribution network is implemented by using timing constrains for the lower possible period that is obtainable by the clock in our designs, i.e. 2 ns and 1.3 ns for the corresponding “baseline” and “pipelined execution“ implementations.

The clock control logic is implemented in a global level; on the clock tree root where the available PLLs are driven in order to select the most appropriate to propagate through the constructed clock tree.

In this way we avoid the expensive area, routing and energy costs of multiple PLL distribution, but we impose timing skew in the clock tree network. Also local, cell level clock gating for heavily gated clock networks such as ours results comes with various challenges, as a previous works in [34] and [35] suggest, that are not within the scope of this work.

In order to compensate for such a delay we design the clock selection unit to generate the outputs that control the clock selection process within the available timing margins. In this sense, we implement a low complexity and low latency control circuit capable of generating outputs with low delay, before the imminent clock pulse. As a result each decision on the dynamic clock frequency change is made within the timing margin available in order to properly distribute the clock pulse throughout the clock network in time. Also, the clock tree synthesis process we employ does not require the routing of all the PLLs of the design; instead it manages to propagate the dynamically selected PLL through a single synthesized clock tree.

5 Evaluation

In this section we discuss the evaluation process we employ in order to evaluate the PDTA methodology. To this end, we elaborate on the CAD toolflow and simulation environment we utilize and we present the results we obtain in terms of speedup, power consumption and overhead of the PDTA.

5.1 CAD toolflow and simulation

After rigorous consideration of many open-source simple processor cores that have been used in architecture-oriented research in the last decade, we have opted for the RiscV Rocket Core [36] processor implemented in Verilog for evaluating our methodology. We have implemented two processor versions as mentioned in Section 4, both of which include our clock control design; we then compiled and tested the circuits using a number of benchmarks, and produced a final evaluation of our ideas. For the front-end design flow we used the Synopsys Design Compiler [37] in conjunction with the NanGate 45nm Open Cell Library [38], whereas for the back-end place and route process we used the Synopys IC Compiler [39]. Afterwards, we used Synopsys PrimeTime [40] to apply the PDTA methodology we described in Section 3 on the generated post-placement netlists. We subsequently performed post-layout simulations, back annotating the design with Mentor Graphics Modelsim [41]. In order to analyze the performance and power consumption of the system, we selected the SPEC CPU2017 benchmark suite and we utilized the RiscV toolchain to compile each benchmark to generate the required binaries in accordance with the RiscV architecture. Finally, we employed Synopsys Power Compiler [42] to generate power reports for each benchmark. The parameters of both the baseline processor and the pipelined execution implementations are displayed in Figure 7. The processor supports in-order instruction issue and execution with 64-bit instruction length. It also employs a BTB of 512 entries using the g-share prediction mechanism. We have also incorporated an L1 cache to the processor; in particular a 4-way associative 16KB i-cache and a 4-way set associative 16KB

Processor Parameters shared on both implementations	
ISA : Risc V Implementation : Rocket core Instruction width : 64 bits Instruction issue : In order, single issue Pipeline depth : 6 stages	
I-cache	**BTB**
Associativity : 4-way Size: 16 KB Round trip : 1 cycle TLB size : 512 entries	Prediction mechanism : gshare History length: 9 bits Prediction levels : 2 Table entries : 512 entries
Baseline Processor Parameters	**Pipelined Execution Processor Parameters**
D-cache Associativity : 4-way Size: 16 KB Round trip : 4 cycles TLB size : 512 entries **Execute stage:** All operations : 1 cycle **Worst-case clock period :** 4ns **PDTA PLL count:** #4	**D-cache** Associativity : 4-way Size: 16 KB Round trip : 7 cycles TLB size : 512 entries **Execute stage:** Multiplication : 2 cycles Division : 3 cycles Other arithmetic : 2 cycles FP Division : 3 cycles FP multiplication : 3 cycles Other FP arithmetic : 3 cycles **Worst-case clock period :** 1.5 ns **PDTA PLL count:** #18

Figure 7: The configuration parameters of both processor implementations.

d-cache with LRU replacement policy. The access time of the data cache is 4 and 7 clock cycles for the baseline and the pipelined execution implementations correspondingly, while the access time for the instruction cache is 1 clock cycle. The clock period for each implementation is defined by the slowest pipeline stage as described in section 4. Finally, the amount of PLLs required for the implementation of the PDTA methodology is 4 for the baseline and 18 for the pipelined execution RiscV implementations.

5.2 Speedup

We have run the Spec2017 CPU benchmarks on the baseline processor implementation, on the pipelined execution implementation and on their corresponding BTWC versions. We present the normalized instruction throughput improvements we obtained according to our experiments in Figure 8 where results indicate an average performance increase in instruction throughput of 1.6 and 1.3 correspondingly. In the same figure we also present the appearance rate of critical instructions for each processor implementation. Further result analysis discloses the following information:

Firstly, designs with relaxed timing constrains benefit more from the PDTA methodology when compared with designs that display tighter timing requirements. This behavior is expected as the PDTA methodology exploits timing differences between individual processor operations. As a result, the more relaxed the system timing, the higher performance increase is achieved. Secondly, frequent appearance of critical instructions throttles the system's performance as the design is forced to operate under the worst-case clock period. As a result benchmarks that display low critical instruction appearance rates, also display higher throughput increase.

In order to further assess the effectiveness of PDTA methodology we compare the obtained throughput results with other state of the art timing speculation techniques. Tables 3 and 4 below demonstrate

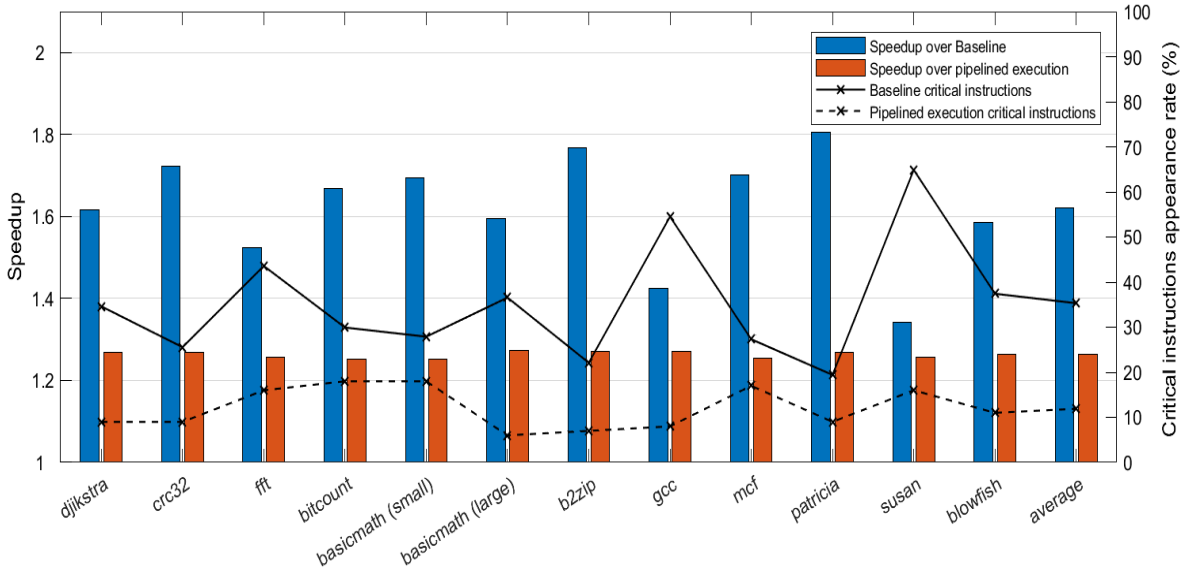


Figure 8: Normalized throughput improvement and critical instruction appearance rate of the proposed design methodology compared to the corresponding baseline processors.

the outcomes of such comparison. Table 3 depicts the normalized throughput improvement of the application-adaptive guardbanding technique proposed by A. Rahimi et. Al. in [10]. We compare the PDTA methodology of both RiscV implementations with the best and worst performance-wise design corners explored by [10]. The application-adaptive guardbanding technique outperforms our methodology when it comes to the best-case design corners, but PDTA proves to be more efferent in terms of performance in worst-case design corners. Table 4 below displays the PDTA results in conjunction with Blueshift optimization as described in [15]. In this work the proposed TS methodology is applied on both Razor [4] and OpenSPARC T1 processors displaying significant performance improvements. According to table 4 PDTA design approach achieves better throughput improvements if compared with Razor or OpenSPARC T1 processor when the baseline RiscV is considered. In contrast, the pipelined execution RiscV design is slightly behind the Razor processor in terms of performance, while it still surpasses the OpenSPARC T1 with the Blueshift design paradigm. PDTA comparison with state of the art TS methodologies highlights the competitive edge of our methodology as its performance is measured on average the same level if not above, compared to other TS approaches.

Table 3: Throughput improvement comparison between Application-adaptive guardbanding and PDTA.

Benchmark	**Adaptive Guardbanding Best/Worst design** corner [10]	**PDTA Baseline /PDTA Pipelined**
djikstra	1.87 / 1.36	1.61 / 1.28
patricia	1.89 / 1.38	1.8 / 1.26
susan	1.81 / 1.58	1.4 / 1.25
blowfish	1.84 / 1.35	1.6 / 1.25
Average	1.88 / 1.25	1.6 / 1.26

Table 4: Throughput improvement comparison between Blueshift OpenSPARC, Razor and PDTA methodology.

Benchmark	**Blueshift OpenSPARC [15]**	**Blueshift Razor [4]**	**PDTA Baseline /PDTA Pipelined**
b2zip	1.18	1.37	1.8 / 1.28
gcc	1.25	1.39	1.42 / 1.29
mcf	1.04	1.05	1.7 / 1.25
Average	1.15	1.27	1.64 / 1.273

5.3 Power consumption compared to the baseline processor

Due to clock frequency scaling, our design often tends to operate at higher frequencies. As higher frequencies are more power hungry, we expect a higher power usage compared to the baseline processor. To verify that assumption, we measure the power consumption of the BTWC design and we compare it to its relevant baseline processors in Figure 9. Results show that the power consumption increase is higher for the benchmarks that present more opportunities for aggressive frequency scaling. Specifically, an increase of 4% to 36% in power consumption is observed, depending on each benchmark's capacity for frequency scaling. Nevertheless, by dividing the performance improvement over the power increase for each benchmark, we get an average of 3.7 improvement rate for the performance-to-power ratio, which is a quite significant overall improvement that we observe with our technique.

5.4 Overhead of the PDTA methodology

In order to properly evaluate the overhead of the PDTA methodology we measure both the PDTA design costs in terms of area and power complexity and the PDTA analysis cost in terms of time requirements.

Regarding the design costs of the PDTA, we quantify the overhead in power and area of the post-layout implementation of the

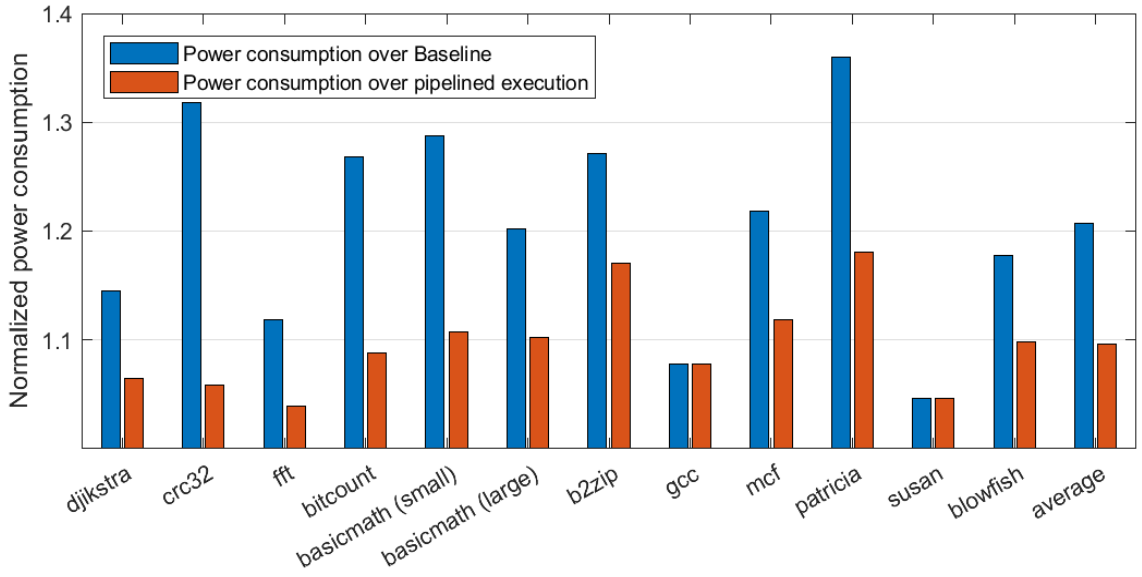


Figure 9: Normalized power consumption increase of the proposed methodology compared to the baseline processors.

clock control circuit and instruction snooping modules. Such modules are the essential components of the PDTA design and are used as described in section 4. Table 5 lists the power and area requirements of the aforementioned modules, along with the power and area requirements of the baseline processors. We observe that the power overhead of the control unit is less than 2% of the total average power consumption of the baseline pipelines, while the area overhead almost 0.001%. Further, Table 6 depicts the area overhead comparison between the state of the art methodologies and the PDTA approach which is proposed in this work. We observe that the PDTA methodology achieves the least area overhead compared with the rest and thus, we conclude that the PDTA is well suited for low-end, low power pipelined processors.

Regarding the time requirements of the PDTA, we measure the amount of iterations and the amount of time required for the PDTA analysis to complete. The results we obtain are depicted in Table 7 in comparison with the standard STA and DTA methodologies. We define the iteration count as the amount of times the corresponding timing methodology in invoked in order to sufficiently cover the timing paths of the pipeline under examination. To this end, the DTA methodology examines every possible bit-transition and thus, it requires $2^{instruction_length}$ iterations (2^{64} for the 64-bit RiscV Rocket core implementations). On the other hand, the STA methodology examines the worst case scenario only, for each instruction path and thus, the required timing iterations are analogous to the complexity of the design. The PDTA methodology resides in between the DTA and STA approaches as the amount of the required iterations is depended on the bit length of the opcode field of the ISA, as discussed in section 3. As a result, the PDTA requires 2^7 iterations in order to properly analyze the timing paths of the RiscV Rocket core pipeline, as the opcode field of the rocket core ISA is 7 bits. In order to evaluate the time requirements of each methodology, we run the DTA, STA and PDTA approaches on the same RiscV rocket core pipeline using an Intel i7 coffee lake processor with 6 cores and 16 GB of DDR4 DRAM. Results indicate that the STA analysis finishes in a 20 second period of time while the PDTA methodology requires 5 minutes. In contrast, the DTA requires over 100 hours to finish and thus, it is considered time costly for timing analysis in processor pipelines. We conclude that the PDTA methodology manages to efficiently manage the tradeoffs between STA and DTA as it provides detailed timing reports for each ISA-supported instruction while also requiring a reasonable time to finish.

Table 5: The power and area overhead of the clock control and instruction snooping circuits in comparison to RiscV .

Implementation	Average power	Area
RiscV Baseline	65.67mW	0.24mm^2
RiscV Pipelined execution	149.04mW	0.55mm^2
Clock control and instruction snooping	98.21uW	321um^2

Table 6: Area overhead comparison between PDTA methodology and the state of the art.

Benchmark	Area overhead
iRazor [5]	13.6%
Application-adaptive guardbanding [10]	0.022%
TS Cache [17]	1.8%
Active management [18]	0.12%
DynOR [21]	5 – 13%
Optimal In Situ Monitoring [26]	3.15%
Razor-Lite [43]	4.42%
Bubble Razor [44]	21%
PDTA (this work)	0.001%

Table 7: Time requirements of DTA, STA and PDTA to complete the timing analysis of RiscV pipeline.

Implementation	Iterations Theoretical	Iterations (RiscV)	Time (RiscV)
DTA	$2^{instruction_length}$	2^{64}	Over 100 hours
STA	Supported instructions	215	20 seconds
PDTA (this work)	2^{opcode_length}	2^{7}	5 minutes

5.5 PVT tolerance considerations

In order to evaluate our design we utilize a single design corner that operates in 0°and 0.72 V. We select the aforementioned corner as the 0.72 V is considered low power pipeline operation and thus, it stays within the scope of this work. We also set a clock uncertainty of 10% to compensate for the process variation effect which may induce clock jitter and uncertainty to the integrated circuit. Evaluating the proposed methodology with a full range of dynamic variations as well as static process parameters variations is possible but the PDTA analysis should be conducted independently for each individual design corner. A higher voltage than 0.72 volts would result in shorter delay instruction paths, while lower operating temperatures would lead to higher delays in the low-voltage region of 0.72 volts as previous work in [45] demonstrates. To this end, the PDTA analysis should be conducted for each design corner in order to extract the exact timing information for the corresponding operating Voltage and Temperature values. On the other hand, the process variation effect can be emulated by setting clock jitter and clock delay values, similarly to our approach. The methodology of the PDTA does not require any modifications in order to function properly within different PVT effects and thus, it can produce accurate timing results given the exact operating condition of the integrated circuit.

6 Conclusion

The BTWC design paradigm promises to alleviate critical path constraints which have negative effects on processor timing. In this paper we present PDTA, a timing analysis methodology, which shifts the focus from a general critical path analysis to the less constrained analysis of paths that are actually followed by individual instructions. To this end we design and implement a circuit capable of identifying the timing requirements of any incoming instruction and selecting the appropriate pipeline clock out of a number of deployed PLLs. Thus we are able to scale up the clock frequency beyond its worst-case operational limit. We evaluate our methodology using a RiscV processor architecture which presents differences in pipeline stage timing. Results demonstrate an average performance increase of 1.62x, as well as a 3 to 4-fold improvement in performance-to-power ratio, compared to the baseline processor. The main contributions of this work to the current state of the art include:

- We propose a novel methodology implemented on the circuit level which considers instruction opcodes for performance increase. Previous consideration of opcodes for performance increase has been compiler-only consideration, i.e. instructions with expensive opcodes are identified at compile time and avoided in code generation.
- The timing analysis we present is focused on individual instructions rather than on whole instruction sequences. In this way we collect information that is more applicable in real systems, as certain instruction sequences do not appear in a steady and predictable rate.
- Our methodology identifies the timing requirements of individual incoming instructions. Since we are a priori aware of such constraints, we do not deploy any error detection or error correction mechanism, while also avoiding any metastable behavior which commonly manifests in such cases.
- We explore an architecture-oriented approach for the BTWC design paradigm. Our work studies the processor architecture to extract timing information based on the ISA of the processor. As a result, we consider this approach to have greater applicability as it can be used on any processor architecture without requiring any further adjustments.
- The proposed methodology does not make any intrusive changes on the processor architecture. This means that it can be applied without any additional design costs, and it maintains the original pipeline intact.

Overall, the BTWC approach presents many opportunities for designers to experiment with novel methodologies and innovative techniques. We believe that research will produce various studies related to this field in the future, as it promises to redefine the design paradigm of modern integrated circuits. Finally, given the dominance of low-cost processors in embedded and many domain-specific designs, ideas that boost performance of processors within those fields in a simple and cheap way will appear, some of which successfully establishing new directions in microprocessor design.

Conflict of Interest The authors declare no conflict of interest.

References

[1] A. Tziouvaras, G. Dimitriou, M. Dossis, G. Stamoulis, "Adaptive Operation-Based ALU and FPU Clocking," in 2020 9th International Conference on Modern Circuits and Systems Technologies (MOCAST), 1–4, 2020, doi: 10.1109/MOCAST49295.2020.9200282.

[2] T. Austin, V. Bertacco, "Deployment of better than worst-case design: solutions and needs," in 2005 International Conference on Computer Design, 550–555, 2005, doi:10.1109/ICCD.2005.43.

[3] R. Ye, F. Yuan, J. Zhang, Q. Xu, "On the premises and prospects of timing speculation," in 2015 Design, Automation Test in Europe Conference Exhibition (DATE), 605–608, 2015, doi:10.5555/2755753.2755890.

[4] D. Ernst, Nam Sung Kim, S. Das, S. Pant, R. Rao, Toan Pham, C. Ziesler, D. Blaauw, T. Austin, K. Flautner, T. Mudge, "Razor: a low-power pipeline based on circuit-level timing speculation," in Proceedings. 36th Annual IEEE/ACM International Symposium on Microarchitecture, 2003. MICRO-36., 7–18, 2003, doi:10.1109/MICRO.2003.1253179.

[5] Y. Zhang, M. Khayatzadeh, K. Yang, M. Saligane, N. Pinckney, M. Alioto, D. Blaauw, D. Sylvester, "iRazor: Current-Based Error Detection and Correction Scheme for PVT Variation in 40-nm ARM Cortex-R4 Processor," IEEE Journal of Solid-State Circuits, **53**(2), 619–631, 2018, doi:10.1109/JSSC.2017.2749423.

[6] V. Subramanian, M. Bezdek, N. D. Avirneni, A. Somani, "Superscalar Processor Performance Enhancement through Reliable Dynamic Clock Frequency Tuning," in 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'07), 196–205, 2007, doi:10.1109/DSN.2007.90.

[7] T. Austin, V. Bertacco, D. Blaauw, T. Mudge, "Opportunities and challenges for better than worst-case design," in Proceedings of the ASP-DAC 2005. Asia and South Pacific Design Automation Conference, 2005., I/2–I/7 Vol. 1, 2005, doi:10.1109/ASPDAC.2005.1466113.

[8] C. C. Wang, K. Y. Chao, S. Sampath, P. Suresh, "Anti-PVT-Variation Low-Power Time-to-Digital Converter Design Using 90-nm CMOS Process," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, **28**(9), 2069–2073, 2020, doi:10.1109/TVLSI.2020.3008424.

[9] B. Poudel, A. Munir, "Design and Evaluation of a PVT Variation-Resistant TRNG Circuit," in 2018 IEEE 36th International Conference on Computer Design (ICCD), 514–521, 2018, doi:10.1109/ICCD.2018.00083.

[10] A. Rahimi, L. Benini, R. K. Gupta, "Application-Adaptive Guardbanding to Mitigate Static and Dynamic Variability," IEEE Transactions on Computers, **63**(9), 2160–2173, 2014, doi:10.1109/TC.2013.72.

[11] T. Abhishek, S. Smruti R., T. Josep, "ReCycle: Pipeline Adaptation to Tolerate Process Variation," in Proceedings of the 34th Annual International Symposium on Computer Architecture, ISCA '07, 323–334, Association for Computing Machinery, New York, NY, USA, 2007, doi:10.1145/1250662.1250703.

[12] S. Sarangi, B. Greskamp, A. Tiwari, J. Torrellas, "EVAL: Utilizing processors with variation-induced timing errors," in 2008 41st IEEE/ACM International Symposium on Microarchitecture, 423–434, 2008, doi:10.1109/MICRO.2008.4771810.

[13] K. A. Bowman, J. W. Tschanz, S. L. Lu, P. A. Aseron, M. M. Khellah, A. Raychowdhury, B. M. Geuskens, C. Tokunaga, C. B. Wilkerson, T. Karnik, V. K. De, "A 45 nm Resilient Microprocessor Core for Dynamic Variation Tolerance," IEEE Journal of Solid-State Circuits, **46**(1), 194–208, 2011, doi:10.1109/JSSC.2010.2089657.

[14] A. Omid, M. Aqeel, L. Benjamin C., P. Sanjay J., H. Mark, "Energy-Performance Tradeoffs in Processor Architecture and Circuit Design: A Marginal Cost Analysis," in Proceedings of the 37th Annual International Symposium on Computer Architecture, ISCA '10, 26–36, Association for Computing Machinery, New York, NY, USA, 2010, doi:10.1145/1815961.1815967.

[15] B. Greskamp, L. Wan, U. R. Karpuzcu, J. J. Cook, J. Torrellas, D. Chen, C. Zilles, "Blueshift: Designing processors for timing speculation from the ground up." in 2009 IEEE 15th International Symposium on High Performance Computer Architecture, 213–224, 2009, doi:10.1109/HPCA.2009.4798256.

[16] A. B. Kahng, S. Kang, R. Kumar, J. Sartori, "Designing a processor from the ground up to allow voltage/reliability tradeoffs," in HPCA - 16 2010 The Sixteenth International Symposium on High-Performance Computer Architecture, 1–11, 2010, doi:10.1109/HPCA.2010.5416652.

[17] S. Shen, T. Shao, X. Shang, Y. Guo, M. Ling, J. Yang, L. Shi, "TS Cache: A Fast Cache With Timing-Speculation Mechanism Under Low Supply Voltages," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, **28**(1), 252–262, 2020, doi:10.1109/TVLSI.2019.2935227.

[18] C. R. Lefurgy, A. J. Drake, M. S. Floyd, M. S. Allen-Ware, B. Brock, J. A. Tierno, J. B. Carter, "Active management of timing guardband to save energy in POWER7," in 2011 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), 1–11, 2011, doi:10.1145/2155620.2155622.

[19] T. Hashimoto, Y. Kawabe, M. Kara, Y. Kakimura, K. Tajiri, S. Shirota, R. Nishiyama, H. Sakurai, H. Okano, Y. Tomita, S. Satoh, H. Yamashita, "An adaptive clocking control circuit with 7.5% frequency gain for SPARC processors," in 2017 Symposium on VLSI Technology, C112–C113, 2017, doi:10.23919/VLSIT.2017.7998133.

[20] A. Grenat, S. Pant, R. Rachala, S. Naffziger, "5.6 Adaptive clocking system for improved power efficiency in a 28nm x86-64 microprocessor," in 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 106–107, 2014, doi:10.1109/ISSCC.2014.6757358.

[21] J. Constantin, A. Bonetti, A. Teman, C. Müller, L. Schmid, A. Burg, "DynOR: A 32-bit microprocessor in 28 nm FD-SOI with cycle-by-cycle dynamic clock adjustment," in ESSCIRC Conference 2016: 42nd European Solid-State Circuits Conference, 261–264, 2016, doi:10.1109/ESSCIRC.2016.7598292.

[22] S. Beer, M. Cannizzaro, J. Cortadella, R. Ginosar, L. Lavagno, "Metastability in Better-Than-Worst-Case Designs," in 2014 20th IEEE International Symposium on Asynchronous Circuits and Systems, 101–102, 2014, doi:10.1109/ASYNC.2014.21.

[23] K. A. Bowman, J. W. Tschanz, N. S. Kim, J. C. Lee, C. B. Wilkerson, S. L. Lu, T. Karnik, V. K. De, "Energy-Efficient and Metastability-Immune Resilient Circuits for Dynamic Variation Tolerance," IEEE Journal of Solid-State Circuits, **44**(1), 49–63, 2009, doi:10.1109/JSSC.2008.2007148.

[24] X. Wang, W. H. Robinson, "Error Estimation and Error Reduction With Input-Vector Profiling for Timing Speculation in Digital Circuits," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, **38**(2), 385–389, 2019, doi:10.1109/TCAD.2018.2808240.

[25] Z. Li, T. Zhu, Z. Chen, J. Meng, X. Xiang, X. Yan, "Eliminating Timing Errors Through Collaborative Design to Maximize the Throughput," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, **25**(2), 670–682, 2017, doi:10.1109/TVLSI.2016.2587810.

[26] H. Ahmadi Balef, H. Fatemi, K. Goossens, J. Pineda De Gyvez, "Timing Speculation With Optimal In Situ Monitoring Placement and Within-Cycle Error Prevention," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, **27**(5), 1206–1217, 2019, doi:10.1109/TVLSI.2019.2895972.

[27] E. Tune, Dongning Liang, D. M. Tullsen, B. Calder, "Dynamic prediction of critical path instructions," in Proceedings HPCA Seventh International Symposium on High-Performance Computer Architecture, 185–195, 2001, doi:10.1109/HPCA.2001.903262.

[28] J. Xin, R. Joseph, "Identifying and predicting timing-critical instructions to boost timing speculation," in 2011 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), 128–139, 2011, doi:10.1145/2155620.2155636.

[29] M. de Kruijf, S. Nomura, K. Sankaralingam, "A unified model for timing speculation: Evaluating the impact of technology scaling, CMOS design style, and fault recovery mechanism," in 2010 IEEE/IFIP International Conference on Dependable Systems Networks (DSN), 487–496, 2010, doi:10.1109/DSN.2010.5544278.

[30] H. Y. Cheah, S. A. Fahmy, N. Kapre, "Analysis and optimization of a deeply pipelined FPGA soft processor," in 2014 International Conference on Field-Programmable Technology (FPT), 235–238, 2014, doi:10.1109/FPT.2014.7082783.

[31] A. Hartstein, T. R. Puzak, "The optimum pipeline depth for a microprocessor," in Proceedings 29th Annual International Symposium on Computer Architecture, 7–13, 2002, doi:10.1109/ISCA.2002.1003557.

[32] V. Agarwal, R. A. Patil, A. B. Patki, "Architectural Considerations for Next Generation IoT Processors," IEEE Systems Journal, **13**(3), 2906–2917, 2019, doi:10.1109/JSYST.2018.2890571.

[33] B. M. Tariq, "A Study of Current Trends in the Design of Processors for the Internet of Things," in Proceedings of the 2nd International Conference on Future Networks and Distributed Systems, ICFNDS '18, Association for Computing Machinery, New York, NY, USA, 2018, doi:10.1145/3231053.3231074.

[34] L. Weicheng, S. Emre, S. Can, T. Baris, "Clock Skew Scheduling in the Presence of Heavily Gated Clock Networks," in Proceedings of the 25th Edition on Great Lakes Symposium on VLSI, GLSVLSI '15, 283–288, Association for Computing Machinery, New York, NY, USA, 2015, doi:10.1145/2742060.2742092.

[35] C. Chang, S. Huang, Y. Ho, J. Lin, H. Wang, Y. Lu, "Type-matching clock tree for zero skew clock gating," in 2008 45th ACM/IEEE Design Automation Conference, 714–719, 2008, doi:10.1145/1391469.1391653.

[36] K. Asanovic, R. Avizienis, J. Bachrach, S. Beamer, D. Biancolin, C. Celio, H. Cook, P. Dabbelt, J. Hauser, A. Izraelevitz, S. Karandikar, B. Keller, D. Kim, J. Koenig, Y. Lee, E. Love, M. Maas, A. Magyar, H. Mao, M. Moreto, A. Ou, D. Patterson, B. Richards, C. Schmidt, S. Twigg, H. Vo, A. Waterman, "The Rocket Chip Generator," Technical report, EECS Department, University of California, Berkeley, April 2016.

[37] "Synopsys Design Compiler," https://www.synopsys.com/implementation-and-signoff/rtl-synthesis-test/design-compiler-graphical.html.

[38] C. H. M. Oliveira, M. T. Moreira, R. A. Guazzelli, N. L. V. Calazans, "ASCEnD-FreePDK45: An open source standard cell library for asynchronous design," in 2016 IEEE International Conference on Electronics, Circuits and Systems (ICECS), 652–655, 2016, doi:10.1109/ICECS.2016.7841286.

[39] "Synopsys IC Compiler," https://www.synopsys.com/implementation-and-signoff/physical-implementation/ic-compiler.html.

[40] "Synopsys PrimeTime," https://www.synopsys.com/implementation-and-signoff/signoff/primetime.html.

[41] "Mentor Graphics Modelsim," https://www.mentor.com/products/fv/modelsim/.

[42] "Synopsys Power Compiler," https://www.synopsys.com/implementation-and-signoff/rtl-synthesis-test/power-compiler.html.

[43] I. Kwon, S. Kim, D. Fick, M. Kim, Y. Chen, D. Sylvester, "Razor-Lite: A Light-Weight Register for Error Detection by Observing Virtual Supply Rails," IEEE Journal of Solid-State Circuits, **49**(9), 2054–2066, 2014, doi: 10.1109/JSSC.2014.2328658.

[44] M. Fojtik, D. Fick, Y. Kim, N. Pinckney, D. Harris, D. Blaauw, D. Sylvester, "Bubble Razor: An architecture-independent approach to timing-error detection and correction," in 2012 IEEE International Solid-State Circuits Conference, 488–490, 2012, doi:10.1109/ISSCC.2012.6177103.

[45] R. Kumar, V. Kursun, "Reversed Temperature-Dependent Propagation Delay Characteristics in Nanometer CMOS Circuits," IEEE Transactions on Circuits and Systems II: Express Briefs, **53**(10), 1078–1082, 2006, doi: 10.1109/TCSII.2006.882218.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 765-769 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Brain Tumor Classification Using Deep Neural Network

Gökalp Çınarer[1,*], Bülent Gürsel Emiroğlu[2], Recep Sinan Arslan[1], Ahmet Haşim Yurttakal[1]

[1]*Bozok University, Computer Technologies Department, Yozgat-66100, Turkey*

[2]*Kırıkkale University, Computer Engineering Department, Kırıkkale-71451, Turkey*

ARTICLE INFO

Article history:
Received: 13 July, 2020
Accepted: 15 September, 2020
Online: 12 October, 2020

Keywords:
Brain Tumor
Classification
Deep Neural Network

ABSTRACT

Brain tumors are a type of tumor with a high mortality rate. Since multifocal-looking tumors in the brain can resemble multicentric gliomas or gliomatosis, accurate detection of the tumor is required during the treatment process. The similarity of neurological and radiological findings also complicates the classification of these tumors. Fast and accurate classification is important for brain tumors. Computer aided diagnostic systems and deep neural network architectures can be used in the diagnosis of multicentric gliomas and multiple lesions. In this study, the Deep Neural Network classification model with Synthetic Minority Over-sampling Technique pre-processing was used on the Visually Accessible Rembrandt Images dataset. The proposed model for the classification of brain tumors consists of 1319 trainable parameters and the proposed method has achieved 95.0% accuracy rate. Precision, Recall, F1-measure values are 95.4%, 95.0% and 94.9% respectively. The proposed decision support system can be used to give an idea to doctors in the detection of glioma type tumors.

1. Introduction

Brain tumors are tumor types located in the brain. The mortality rate is high. There are two different types: benign and malignant. Benign brain tumors, which have a slow growth rate, are easily separable from brain tissue. Malignant brain tumors grow faster and can damage nearby brain tissue[1]. There have been many successful studies in which computer were used in medical applications in recent years. Skin cancer [2], breast cancer [3], thyroid nodules [4] are some of these studies.

In addition, tumors originating from brain tissue are called primary brain tumors, while tumors formed by the splashing of cancerous cells that appear elsewhere in the body are called secondary brain tumors. The most common (65%) and most malignant brain tumor is glioblastoma [5]. Glioblastoma multiforme (GBM) is one of the deadliest cancer types and is a common type of malignant glioma in the central nervous system. With the development of computed tomography and magnetic resonance imaging (MRI), it has become increasingly clear that gliomas may have a multifocal or multicentric [6]. Although the diagnosis of multiple lesions becomes easier with the introduction of computed tomography into the field, the differential diagnosis of multiple cerebral lesions still remains a problem. In multifocal lesions, the lesions are localized in more than one part of the brain and there is a macroscopic or microscopic level relationship between them. In multicentric gliomas, there is neither macroscopic nor microscopic level relationship between lesions. The propagation path of multicentric lesions is unclear, but this situation is slightly different in multiple lesions. Multiple lesions create multiple cerebral areas using a systematic propagation pathway [7]. The correct detection of this condition is very important in the treatment of the disease.

Early diagnosis of brain tumors can change the course of the disease and save lives. On the other hand, computers perform complex operations in a very short time. Computers are used in many areas of life, especially in health, as it produces automatic, fast and accurate results. For this reason, computer-aided diagnostic (CAD) systems are frequently used today. CAD systems analyze the complex relationships in medical data and assist the doctor in decision making [8]. Salvati divided multifocal gliomas into four different categories. These are diffuse, multicentric, multiple and multi-organ. In addition, in his study, he divided his patients into multicentric and multifocal tumor groups and found that there was no significant difference between patients [9]. In a spike case, as malignant glial tumors damage nerve cells, it seems possible that the multicentricity of gliomas may also accompany multiple sclerosis [10]. DNN can learn quickly,

* Corresponding Author: Gökalp Çınarer, gokalp.cinarer@bozok.edu.tr

automatically identify features, and can be successfully applied in computer learning systems. Besides that, DNNs are also used in some medical image analysis. It can be preferred especially in brain tumor segmentation and classification. It is designed the DNN model to segment brain separate gray and white matter from cerebrospinal fluid and analyzed the segmentation accuracy with DNN [11]. Similarly DNN model for the segmentation of brain tumors is proposed [12].

In this study, Deep Neural Network (DNN) architecture was proposed to classify the MR brain image features as n/a, multifocal, multicentric and gliomatosis. Synthetic Minority Over-sampling Technique (SMOTE) was used as a pre-processing. It has been determined that the pre-processing has a significant effect on the classification success.

2. Material and Methods

2.1. Dataset

In the study, The Repository of Molecular Brain Neoplasia Data (REMBRANDT) dataset [13] was used. The REMBRANDT dataset is a dataset of The Cancer Imaging Archive (TCIA) database shared by the National Cancer Institute. TCIA is an accessible archive containing medical images [14]. This data set contains images of 33 patients. Each patient has an average of 20 MR images, on separate axial, sagittal and coronal planes. For each patient, there are feature key values provided by 3 different neuroradiologists from Thomas Jefferson University (TJU) Hospital. This feature set is called VASARI (Visually Accessible Rembrandt Images). VASARI dataset was used in the classification process.

2.2. SMOTE

SMOTE is a method used in unbalanced data. Every minority class sample is taken and synthetic samples are created by looking at any or the sample's entire k neighbour. Thus, the minority class is over-sampled [15]. The main difference from other sampling methods is the production of synthetic samples by looking at their neighbours, instead of copying and reproducing the minority class samples. Thus, it realizes class distribution without providing additional information.

2.3. Deep Neural Networks (DNN)

Machine learning technologies potentiate many technologies in the modern world. People use these technologies in many areas such as webcams, e-commerce activities, and mobile devices. In addition, it offers good results in complex problems. These are identifying objects from images, converting speech to text and matching web contents with the interests of the users. Traditional ML techniques have some limitations in processing raw data. Extracting features representing data in the raw form and converting them into feature vectors used as input for classifier is a difficult process that requires expertise [16].

Deep learning allows abstract representation of data by creating simple but nonlinear modules. In this way, it is possible to learn very complex functions. This structure reveals the distinguishing features between classes for classification problems. Deep learning produces successful results for many problems that cause problems in the advancement of artificial intelligence methods. It is used in many areas such as image recognition [17-19], speech processing [20, 21], and gene mutation prediction [22]. It is thought that DNN will make faster progress thanks to the new neural network structures proposed with an increase in computation capacity and amount of data [16].

Deep neural network structure is shown in Figure 1.

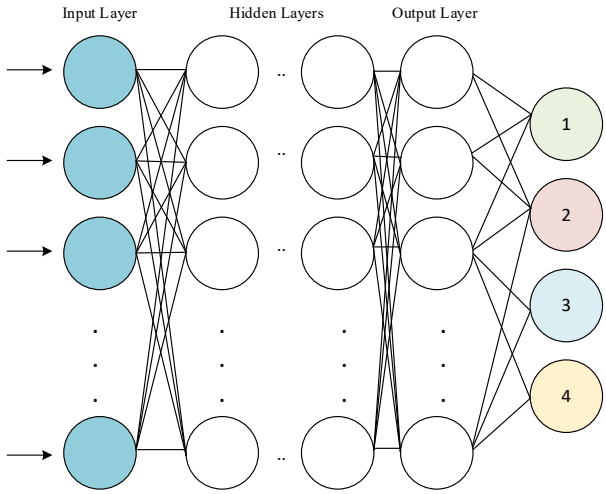


Figure 1: Deep Neural Network Schema [16]

The development of computer hardware structure and the widespread use of GPU architecture facilitates and speed up the training of the complex and multi-layered neural networks as shown in Figure 1. In this way, more outstanding [23].

The mathematical model used in the neural network training stage is shown in Eq-1.

$$f_{output} = af(m_n af(m_{n-1} \dots af(m_1 x))) \quad (1)$$

According to Equation, m_n represents the weight matrix in the nth layer and af is activation function. This structure gives better results than single-layer networks [24].

In this study, a deep neural network design with 1 input, 1 output and 2 hidden layer was designed.

The increasing of software technologies shaped under three specialized topics such as artificial intelligence, machine learning and deep learning has caused an increase in interest in these three subjects. Deep Neural Networks finds the relationship between a range of inputs and outputs. Successful results in many areas such as image processing, natural language processing, voice recognition [25].

3. Results

The application was carried out on the open source Python program. Tensorflow and Keras libraries were used. The count plot of the data to be classified into classes is given in Figure 2.

All images included in the Rembrandt data set consist of axial, sagittal and coronal image planes with an average of 20 images per plane for a total of 33 patients. In the data set, 30 different features were obtained from each tumor brain MR image. Rembrandt images (VASARI) were analyzed by three TJU radiologists and features of the images were extracted. In this way, a data set

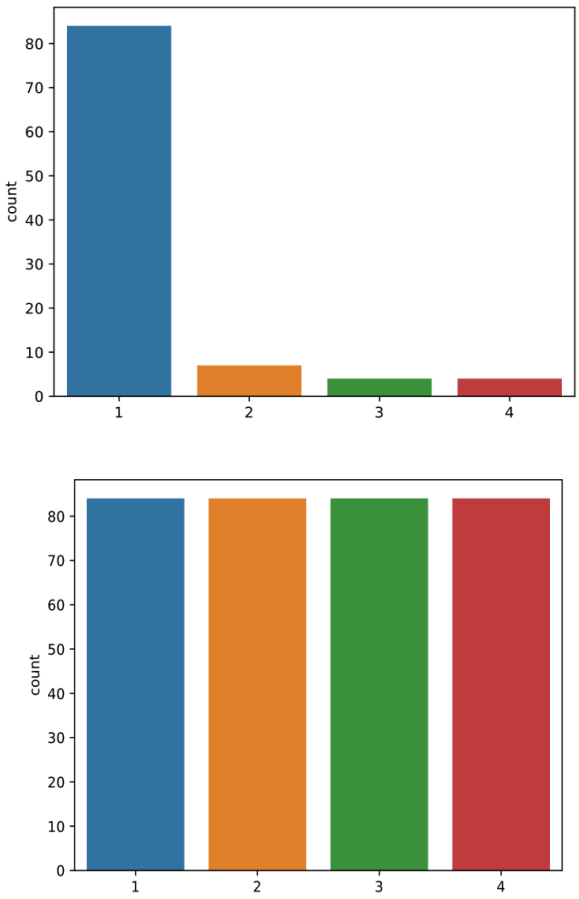


Figure 3: Count plot of data after SMOTE

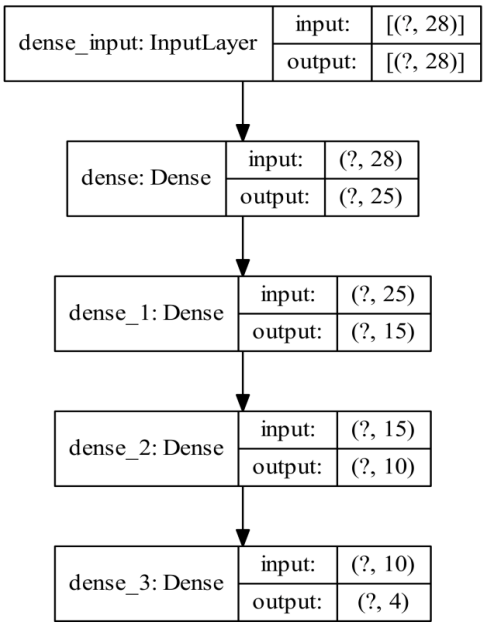


Figure 4: Proposed DNN Model

consisting of 99 patients and 30 different characteristic features used in the study. The data were classified according to n/a, multifocal, multicentric and gliomatosis. The data appears to be unbalanced. (84 n/a patient, 7 multifocal patient, 4 multicentric patient and 4 gliomatosis patient). For this reason, SMOTE pre-treatment was applied. Figure 3 shows the count plot after the SMOTE pre-processing.

The proposed DNN model is given in Figure 4. There are 3 hidden layers in the model, having 25, 15, and 10 filters, respectively. Model includes 1319 trainable parameters.

The activation function used in hidden layers is ReLu. The activation function used in the output layer is SoftMax. The optimizer who used the error back propagation phase is Adam. In this study, the data set was formed as 70% training and 30% test group. Figure 5 shows the change in accuracy on training and test data over 100 epochs.

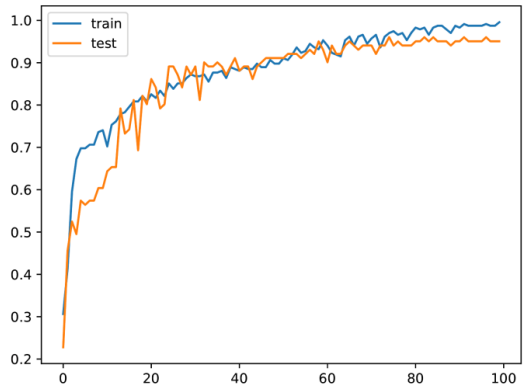


Figure 5: Accuracy graph

As can be seen from the graphic, the proposed method has overcome the over-fitting problem. Figure 6 shows confusion matrix.

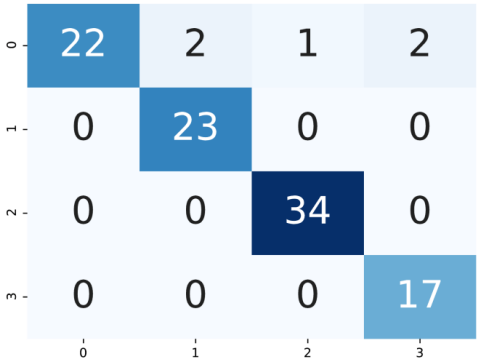


Figure 6: Confusion matrix

DNN model was used to categorize brain tumors into four classes (n / a, multifocal, multicentric and gliomatosis). Four basic measurement indicators were used when evaluating the classification results. The proposed DNN classifier knew the opinion of radiologists with 95% accuracy. The success rate of the model was also examined with other performance metrics such as precision recall and F1. Respectively Precision, Recall, F1-measure values are 0.954, 0.950 and 0.949.

4. Discussion

Deep learning methods are used in many research fields of medical image analysis, from image recording, radiomic feature

selection, segmentation and classification. Estimating the type of brain tumors is quite difficult and troublesome in the medical field.The aim of this study is to classify MR images of brain tumors as n/a, multifocal, multicentric and gliomatosis. In this study, unlike other studies, multiple cerebral lesions were classified. These lesions can have similar features and are difficult to classify. Studies in the literature determine tumor levels or predict whether the tumor is benign or malignant. Kumar and Vijay Kumar classified brain images as tumors and non-tumors using the Feed Forwarded Artificial neural network-based classifier and reached 91.17% accuracy with the ensemble classifier [26]. Glioblastoma and low-grade gliomas using logistic regression algorithm is classified with 0.88 accuracy [27]. Mahajani classified the MRI brain image as normal or abnormal with Back Propagation Neural Network (BPNN) and K Nearest Neighbor (KNN) algorithms. 70% accuracy using the KNN classifier and 72.5% accuracy using the BPNN classifier [28]. In the literature this dataset was tested with KNN, Random Forest (RF), Support Vector Machines (SVM), Linear Discriminant Analysis (LDA) machine learning algorithms. SVM algorithm with 90% accuracy rate was found to be better compared to other algorithms [29]. As seen in the examples, machine learning algorithms are commonly used approaches in classification. However, since deep learning models can analyze complex relationships effectively, it has gained an important place in machine learning in recent years. In another study, VGG-19 architecture used to classify brain tumors grades (I, II, III and IV) and achieve 0.90 accuracy and 0.96 average precision [30]. Also, in the other study it is classified the grades of glioma tumor using Convolutional Neural Networks (CNN) with an accuracy of 91.16% [31]. Compared to other algorithms, it has been seen that the applied method has advantages over traditional classifier algorithms. The complexity of the brain structure in the analysis of the brain image and the lack of imaging standards prevent the correct application of computer-aided learning systems in classifying brain tumors. Therefore, radiologist views come to the fore in determining these criteria. Table 1 shows some deep learning studies in the literature.

Table 1: Deep learning approaches for brain tumor classification

Aim and Method	Accuracy
Tumor Segmentation-DNN[32]	0.88
Brain Tumour, CNN[33]	91.43
Brain Tumour, Capsule Network[34]	86.56
Brain Tumoru, DNN[35]	90.66

In addition, proposed classification model can be tested with different parameters and radiomic features. Obtaining the data by experienced radiologists and determining the radiologist views with high accuracy is an indicator of the success of the algorithm. The Deep Learning approach facilitates researchers and trains the model using its own hyper parameters without specifying many features beforehand. But it is very important to apply the correct parameters to the model. In this study, the deep learning model was used and high accuracy was obtained with the correct parameters applied.

5. Conclusion

In the article, multifocal, multicentric and gliomatosis classification of brain tumors with SMOTE pre-processing of the DNN model were examined. The results show that the proposed method is very successful. The proposed method gives automatic, fast and accurate results. It can give an idea as a second look when doctors are unstable.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

No funding,

References

[1] J.C. Buckner, P.D. Brown, B.P. O'Neill, F.B. Meyer, C.J. Wetmore, J.H. Uhm, "Central nervous system tumors," in Mayo Clinic Proceedings, 2007, doi:10.4065/82.10.1271.

[2] N. Razmjooy, M. Ashourian, M. Karimifard, V. V. Estrela, H.J. Loschi, D. do Nascimento, R.P. França, M. Vishnevski, "Computer-Aided Diagnosis of Skin Cancer: A Review," Current Medical Imaging Formerly Current Medical Imaging Reviews, 2020, doi:10.2174/1573405616666200129095242.

[3] J.Y. Kim, J.J. Kim, L. Hwangbo, H.B. Suh, S. Kim, K.S. Choo, K.J. Nam, T. Kang, "Kinetic heterogeneity of breast cancer determined using computer-aided diagnosis of preoperative MRI scans: Relationship to distant metastasis-free survival," Radiology, 2020, doi:10.1148/radiol.2020192039.

[4] D. Fresilli, G. Grani, M.L. De Pascali, G. Alagna, E. Tassone, V. Ramundo, V. Ascoli, D. Bosco, M. Biffoni, M. Bononi, V. D'Andrea, F. Frattaroli, L. Giacomelli, Y. Solskaya, G. Polti, P. Pacini, O. Guiban, R. Gallo Curcio, M. Caratozzolo, V. Cantisani, "Computer-aided diagnostic system for thyroid nodule sonographic evaluation outperforms the specificity of less experienced examiners," Journal of Ultrasound, 2020, doi:10.1007/s40477-020-00453-y.

[5] H. Ohgaki, P. Kleihues, Epidemiology and etiology of gliomas, Acta Neuropathologica, 2005, doi:10.1007/s00401-005-0991-y.

[6] C.G. Patil, P. Eboli, J. Hu, Management of Multifocal and Multicentric Gliomas, Neurosurgery Clinics of North America, 2012, doi:10.1016/j.nec.2012.01.012.

[7] R.P. Thomas, L.W. Xu, R.M. Lober, G. Li, S. Nagpal, "The incidence and significance of multiple lesions in glioblastoma," Journal of Neuro-Oncology, 2013, doi:10.1007/s11060-012-1030-1.

[8] T.R. Jensen, K.M. Schmainda, "Computer-aided detection of brain tumor invasion using multiparametric MRI," Journal of Magnetic Resonance Imaging, 2009, doi:10.1002/jmri.21878.

[9] M. Salvati, L. Cervoni, P. Celli, R. Caruso, F.M. Gagliardi, "Multicentric and multifocal primary cerebral tumours. Methods of diagnosis and treatment," Italian Journal of Neurological Sciences, 1997, doi:10.1007/BF02106225.

[10] L.C. Werneck, R.H. Scola, W.O. Arruda, L.F.B. Torres, "Glioma and multiple sclerosis: case report.," Arquivos de Neuro-Psiquiatria, 2002, doi:10.1590/S0004-282X2002000300024.

[11] P. Moeskops, M.A. Viergever, A.M. Mendrik, L.S. De Vries, M.J.N.L. Benders, I. Isgum, "Automatic Segmentation of MR Brain Images with a Convolutional Neural Network," IEEE Transactions on Medical Imaging, 2016, doi:10.1109/TMI.2016.2548501.

[12] P. Dvořák, B. Menze, "Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, doi:10.1007/978-3-319-42016-5_6.

[13] L. Scarpace, A.E. Flanders, R. Jain, T. Mikkelsen, D.W. Andrews, "Data From REMBRANDT," The Cancer Imaging Archive, 2015, doi:10.7937/K9/TCIA.2015.588OZUZB.

[14] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, F. Prior, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," Journal of Digital Imaging, 2013, doi:10.1007/s10278-013-9622-7.

[15] N. V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, 2002, doi:10.1613/jair.953.

[16] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F.E. Alsaadi, "A survey of deep neural network architectures and their applications," Neurocomputing, 2017, doi:10.1016/j.neucom.2016.12.038.
[17] H. Fujiyoshi, T. Hirakawa, T. Yamashita, Deep learning-based image recognition for autonomous driving, IATSS Research, 2019, doi:10.1016/j.iatssr.2019.11.008.
[18] G. Hu, C. Yin, M. Wan, Y. Zhang, Y. Fang, "Recognition of diseased Pinus trees in UAV images using deep learning and AdaBoost classifier," Biosystems Engineering, 2020, doi:10.1016/j.biosystemseng.2020.03.021.
[19] M. Trokielewicz, A. Czajka, P. Maciejewicz, "Post-mortem iris recognition with deep-learning-based image segmentation," Image and Vision Computing, 2020, doi:10.1016/j.imavis.2019.103866.
[20] A.L. Maas, P. Qi, Z. Xie, A.Y. Hannun, C.T. Lengerich, D. Jurafsky, A.Y. Ng, "Building DNN acoustic models for large vocabulary speech recognition," Computer Speech and Language, 2017, doi:10.1016/j.csl.2016.06.007.
[21] J. Novoa, J. Fredes, V. Poblete, N.B. Yoma, "Uncertainty weighting and propagation in DNN–HMM-based speech recognition," Computer Speech and Language, 2018, doi:10.1016/j.csl.2017.06.005.
[22] M.K.K. Leung, H.Y. Xiong, L.J. Lee, B.J. Frey, "Deep learning of the tissue-regulated splicing code," Bioinformatics, 2014, doi:10.1093/bioinformatics/btu277.
[23] A. Yiğit, "İŞ SÜREÇLERİNDE İNSAN GÖRÜSÜNÜ DERİN ÖĞRENME İLE DESTEKLEME," Trakya Üniversitesi, Fen Bilimleri Enstitüsü, Yayınlanmış Yüksek Lisans Tezi, 2017.
[24] G.E. Dahl, T.N. Sainath, G.E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2013, doi:10.1109/ICASSP.2013.6639346.
[25] Y. Zhang, M. Pezeshki, P. Brakel, S. Zhang, C. Laurent, Y. Bengio, A. Courville, "Towards end-to-end speech recognition with deep convolutional neural networks," in Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2016, doi:10.21437/Interspeech.2016-1446.
[26] P. Kumar, B. Vijay Kumar, "Brain tumor MRI segmentation and classification using ensemble classifier," International Journal of Recent Technology and Engineering, 2019.
[27] K.L.C. Hsieh, C.M. Lo, C.J. Hsiao, "Computer-aided grading of gliomas based on local and global MRI features," Computer Methods and Programs in Biomedicine, 2017, doi:10.1016/j.cmpb.2016.10.021.
[28] P.P.S. Mahajani, "Detection and Classification of Brain Tumor in MRI Images," International Journal of Emerging Trends in Electrical and Electronics, 2013.
[29] G. Cinarer, B.G. Emiroglu, "Classificatin of Brain Tumors by Machine Learning Algorithms," in 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2019 - Proceedings, 2019, doi:10.1109/ISMSIT.2019.8932878.
[30] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, S.W. Baik, "Multi-grade brain tumor classification using deep CNN with extensive data augmentation," Journal of Computational Science, 2019, doi:10.1016/j.jocs.2018.12.003.
[31] S. Khawaldeh, U. Pervaiz, A. Rafiq, R.S. Alkhawaldeh, "Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks," Applied Sciences (Switzerland), 2017, doi:10.3390/app8010027.
[32] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.M. Jodoin, H. Larochelle, "Brain tumor segmentation with Deep Neural Networks," Medical Image Analysis, 2017, doi:10.1016/j.media.2016.05.004.
[33] J.S. Paul, A.J. Plassard, B.A. Landman, D. Fabbri, "Deep learning for brain tumor classification," in Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging, 2017, doi:10.1117/12.2254195.
[34] P. Afshar, A. Mohammadi, K.N. Plataniotis, "Brain Tumor Type Classification via Capsule Networks," in Proceedings - International Conference on Image Processing, ICIP, 2018, doi:10.1109/ICIP.2018.8451379.
[35] D. Nie, H. Zhang, E. Adeli, L. Liu, D. Shen, "3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, doi:10.1007/978-3-319-46723-8_25.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 784-789 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

BLDC Motor Vibration Identification by Finite Element Method and Measurements

Jerzy Podhajecki *, Stanisław Rawicki

Faculty of Technology, The Jacob of Paradies University, Gorzów Wielkopolski, 66-400, Poland

ARTICLE INFO

Article history:
Received: 30 December, 2020
Accepted: 11 March, 2021
Online: 31 March, 2021

Keywords:
BLDC motor
Vibrations
Finite element method

ABSTRACT

Within the results of scientific research, vibrations and their reduction have been described for the brushless direct current motor with permanent magnets (BLDC motor). In this paper, calculations (the finite element method using commercial Finite Element Software Ansys) and measurements were performed to identifying the sources of vibrations in BLDC motor. The article presents numerical and experimental research on the resonant frequencies of the stator and the rotor; transient vibrations of the stator due to Maxwell forces in the motor have been analysed. It was shown that the natural frequencies were the main source of vibrations. The vibration sources indentification made possible formulation of better principles of choice of constructional motor parameters with the aim of attaining minimalization of vibrations.

1. Introduction

Due to the high efficiency, electrical BLDC motors are becoming more widely used within the framework of practical applications. In this field, development of new permanent magnets with the very big energy density is additionally of great importance [1]. The large and impressive technical progress within rotating electrical machines is connected with dynamical extension of power electronics converters. For motors there is possible the modern speed control on a large range. Present-day instantaneous torque/flux control makes possible considerable improving the electrical machines operation. For practical applications, disadvantage of PWM inverters consists in enlargement of the overall motor noise. Here it can be noticed that the special variant is substantial when the phenomenon of resonance appears within the framework of supply harmonics and resonant frequencies of machine's parts. Designers of electric motors ought to undertake an obligation to create less noisy devices [2].

Magnetic sources of vibrations in BLDC motor are caused by Maxwell forces, cogging torque ripple and torque ripple [3, 4]. If the natural frequencies of the stator or the rotor coincide with frequencies of magnetic forces or torque ripple resonance may occur. Also, as a result of imperfections of construction or assembly, rotor eccentricity can occur and distortion of the magnetic field may appear. Owing to this disadvantageous situation, the magnetic force distribution changes and the unbalanced magnetic force exists [5]. Vibrations can also be due to friction of different parts of machine, e.g. bearings. Aerodynamic noise is generated by the circulation of the cooling medium of the machine.

In literature analytical methods have been also used to describe mathematically Maxwell magnetic forces harmonics taking into consideration both the spatial and time frequency order [3]. For correct designing, the best numbers of slots were found and given in order to attain avoidance of low harmonic radial forces and considerable reduction of vibrations.

The numerical investigation to predict vibrations are based on creation of magnetical and mechanical finite element models [6]. According to the general strategy within the scientific methodology, authenticity of research work may be realized by comparing numerical results with experimental ones. In the magnetic motor model the flux distribution in air-gap and the cogging torque–obtained by calculations - are compared with measurements. In the case of the mechanical model natural vibrations frequencies are compared for theoretical and practical activities. Among other things the accuracy of calculated results depends on proper modeling of magnetic, mechanical and acoustic parameters of the motor structure and the quality of the model of the numerical mesh. The magnetic forces can be minimized by optimizing changes within the geometric parameters of the magnetic circuit.

In literature concerning modern power electronics it was shown that the selecting appropriate strategies of Pulse Width Modulation (PWM) could ensure avoidance of the mechanical resonance [2]. For the other example it was illustrated that geometrical modification of the rotor was the effective method to

*Corresponding Author: Jerzy Podhajecki, Email: jerzypodh@o2.pl

increase natural frequencies and by this constructional way the phenomenon of resonance has been eliminated [7].

2. BLDC motor and magnetic calculations

2.1. The BLDC motor

Taken to theoretical and practical investigations electronically controlled BLDC motor, with the stator having concentrated windings and with the interior rotor possessing permanent magnets, has been presented as three separate parts in Figure 1. The specification of characteristic parameters and different data of the motor are presented in paper [8].



Figure 1: The analysed BLDC motor (stator, rotor) with PWM converter for power supply

Within the fundamental governing equations there is the Maxwell's equation [9]:

$$rot\ H = J \tag{1}$$

where H is the magnetic intensity, J – the current density.

In accordance with Gauss's Law, taking into account divergence theorem we have for the magnetic flux density B the following relation:

$$div\ B = 0 \tag{2}$$

The relationship between inducing field strength and magnetization intensity is given by:

$$B = \mu H \tag{3}$$

where μ is the magnetic permeability.

The current density distribution of the coil is here described as follows:

$$J = 1 \cdot J_z \tag{4}$$

Numerical solution of 2-D magnetostatic problem is based on the equation:

$$\Delta^2 A_z = \mu J_z \tag{5}$$

The outer boundary of the stator yoke has been applied for all outer nodes of the stator yoke:

$$A_z\,(x, y) = A_{z1}\,(x, y) = 0 \tag{6}$$

The stress tensor method is used for calculation of the magnetic force.

2.2. Magnetic field results

Within the framework of computations concerning the finite element methodics, several assumptions are here taken into consideration. The magnetic field reckoning is performed using the principles connected with the magnetostatic analysis.

For the 2-D FEM model, the BLDC motor is fed with rectangular wave current The rotor is rotating at one mechanical degree of freedom.

The two-dimensional model for magnetic field calculations is shown in Figure 2 (so-called geometrical model) and in Figure 3 (here it is the numerical mesh – the quarter of the model).

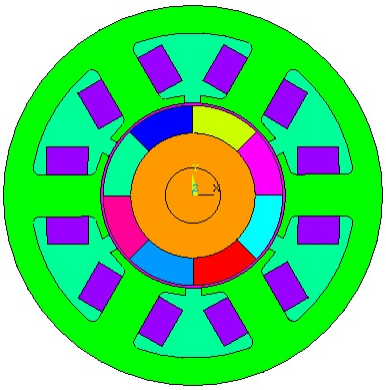


Figure 2: Geometrical model of BLDC 6/8 motor (6 - stator teeth, windings and rotor with 8 - magnets applied)

Figure 4 presents the distribution of the tangential and radial flux densities in motor air-gap. By other means and for the selected rotor position, Figure 5 illustrates the calculated magnetic flux distribution.

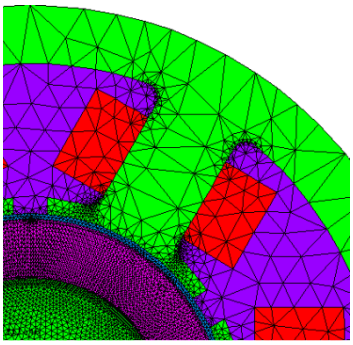


Figure 3: Numerical mesh of the magnetic model: stator (the part of stator iron, windings) and rotor with magnets (1/4 of the BLDC motor model)

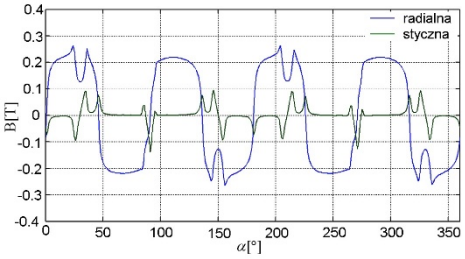


Figure 4: Radial (rad) and tangential (st) flux densities distribution in the middle of the air-gap in BLDC motor

The Maxwell forces variations identify the mode orders in the electrical machine. The dominant mode shapes are associated with Maxwell force distribution in the air-gap shown in Figure 6. On the ground of detailed analysis of diagrammatic presentation in Figure 6, it can be noticed that the general pattern is repeated twice; here the lowest order of radial force and dominant mode order are equal two [10]. It means that the most important circumferential modes of vibrations of the stator are the second and fourth mode of natural frequencies.

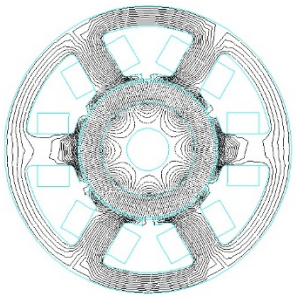


Figure 5: Magnetic flux distribution in BLDC motor

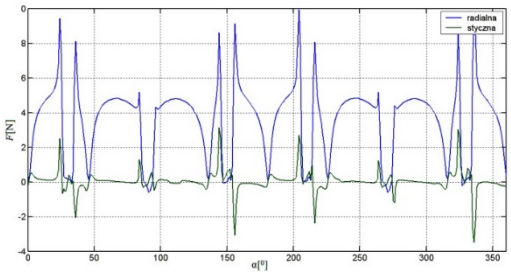


Figure 6: The magnetic forces radial and tangential distribution acting on stator poles in the air-gap in BLDC model

3. Modal analysis

3.1. Modal analysis of stator

Two important reasons: determination of the natural frequency and recognition of the mode shape of structure are the sought purpose of modal analysis. We are interested in characteristics of a structure's dynamics connected with varying loads.

The equation for the free vibrations of the undamped system can be expressed in matrix notation as follows [11]:

$$\{[K] - \omega^2 [M]\} \{u\} = 0 \qquad (7)$$

where: *[K]* - stiffness matrix, *[M]* - mass matrix, ω – angular frequency, u – displacement matrix. Frequencies: $f = \omega / 2\pi$ and displacement matrix: u are the solutions of the equation (7).

The natural frequencies of vibrations can be excited by radial Maxwell forces acting on the stator poles. The maximum amplitude of vibrations occurs for resonance when the excitation frequency matches with the natural frequency of the motor and the spatial distribution of Maxwell force is similar to one of the mode of the natural frequencies of the stator.

The numerical calculations using Lanchos algorithm of the natural frequencies were performed for two various stator's models: only alone core in the first variant and the core with the winding in the second case. Within such methodics the stator core is modeled as steel. However in literaturę it was shown that the influence of windings on vibrations could be also important [12]. Because of the complex structure, windings can be taken into account in simplifying manner by adding mass or by increasing density of stator teeth [13].

To realize calculations with the using finite element method (FEM) and allowing for windings, density of stator teeth has been increased by such manner to accomplish finally the smallest possible error for the dominant second and fourth mode. The compromise value of the density was determined as the following value: $\rho = 1000$ kg / m^3.The selected result is given in Figure 7.

Here it ought to be underlined that – despite of efforts – taking windings in a simplified way in the physical model into consideration causes some errors for other modes of natural frequencies of stator (Table 1). Experimental studies using hammer impact method were conducted to verify the results of the numerical calculations and the radial vibrations were recorded using accelerometers Brüel & Kjær. The results have been given in Figure 8. The comparison between numerical results and outcomes of measurements can be realized on the ground of data given in Table 1.

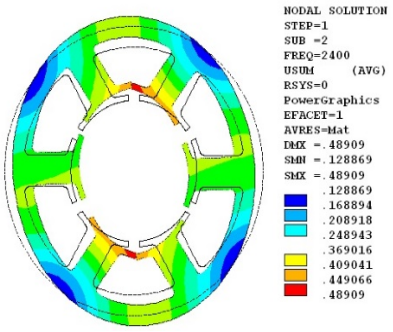


Figure 7: Modal analysis results of the for stator core with windings modelled by increasing stator density (r=2, f=2400Hz)

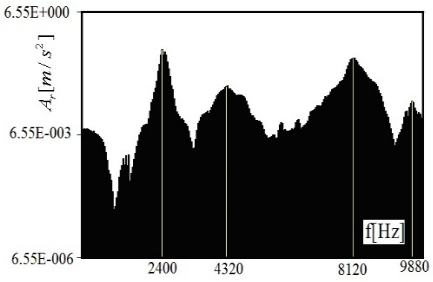


Figure 8: The natural frequencies of the stator - radial acceleration Ar [m / s2] of stator structure obtained by hammer impact method

Table 1: Natural frequencies of the stator [Hz]

Mode order	FEM - stator yoke	FEM - stator yoke with windings	Measurements (stator yoke with windings)
1	3026	2400	2400
2	4024	5366	4320
3	6753	7101	8120
4	12600	10223	9880

3.2. Modal analysis of rotor

The numerical calculations of the natural frequencies using Lanchos algorithm were performed for two cases: the first variant of the separate rotor and the second case of the rotor with bearings modeled as spring elastic - damping element. The selected result is given in Figure 9.

The shaft is made of steel and ferrite magnets. Calculations were performed for different stiffness to elastic - damping element to fit the measurements best of all [7]. Experimental studies using shaker method were conducted to verify the results of the numerical results and the radial vibrations were recorded using accelerometers Brüel & Kjær by [11]. The obtained exemplary results of calculations and measurements using shaker methods are presented in Table 2.

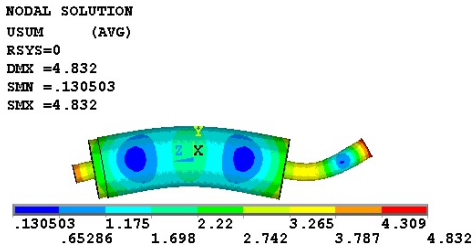


Figure 9: Modal analysis of rotor model without bearings – numerical results bending mode, f = 9766 Hz

Table 2: Calculation bending natural frequencies of the rotor [Hz] for different stiffness of bearings and measurements

Mode of deformation	FEM - 10^6 [N/m]	FEM -10^7 [N/m]	Measurements
first bending mode	4019	4136	3950
second bending mode	9770	9797	10550

4. Magnetic forces and transient analysis

The magnetic forces depend on the number of teeth, the number of poles and the speed of rotation. The frequencies of magnetic forces in 3-phase BLDC motor can be calculated as follows [4]:

$$f_{lf} = n \cdot p \cdot \frac{N}{60} \tag{8}$$

$$f_{tr} = 6n \cdot \frac{p}{2} \cdot \frac{N}{60} \tag{9}$$

$$f_{ct} = n \cdot l_{cm} \cdot \frac{N}{60} \tag{10}$$

where: n - arbitrary integer, p - number of pairs of poles, l_{cm} – number of stator slots, N – rotor velocity [rpm]. The rotation speed was chosen to be 4000 rpm. The first frequencies of Maxwell forces, torque ripple and cogging torque are equal: f_{lf} = 533 Hz, f_{tr} = 800 Hz, f_{ct} = 711 Hz.

The transient analysis is governed by the following equation [9]:

$$[M]\ddot{[u]}+[C]\dot{[u]}+[K][u]=F(t) \tag{11}$$

where: $[M]$ - mass matrix, $[K]$ - stiffness matrix, u – vector of displacement; $[C]$ - damping matrix. Newmark method has been employed for differential equation solution. The calculations were carried out for the rotor velocity N = 4000 rpm. Magnetic forces calculated in magnetic analysis are loads in the mechanical analysis. In keeping with the energy absorbing in the constructional structure of the BLDC motor, mainly due to the windings, structural damping has been presented [9]. However due to to lack of the experimental data damping are assumed to zero. The maximum amplitude of vibration occurs near the resonant frequencies the second and fourth natural frequency of the stator. The results are given in Figure 10 and Figure11. The stator deformation due to magnetic forces during no-load for selected time-steps are presented (Figure 12 and Figure 13).

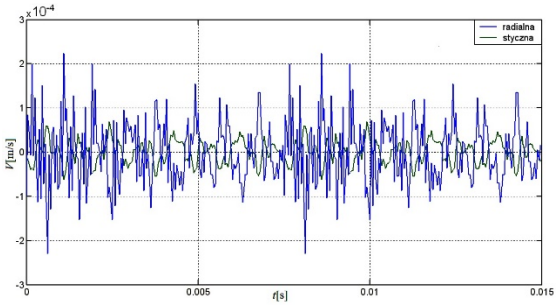


Figure 10: Numerical transient analysis of stator vibrations due to magnetic forces - velocity versus time V [m/s] - radial (blue) and tangential velocity (green) for selected outer node

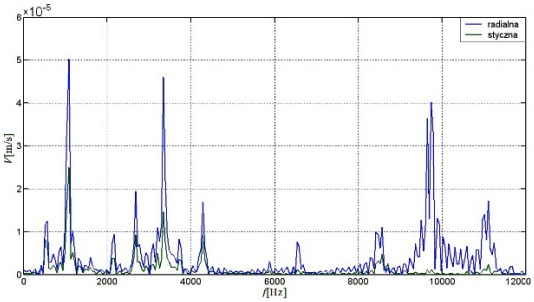


Figure 11: Numerical transient analysis - Fourier analysis of radial velocity for selected outer stator node of stator V [m/s] - radial (blue) and tangential velocity (green)

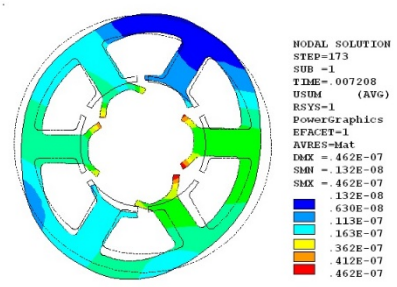


Figure 12: Numerical transient analysis - the stator deformation and displacement U [m] for selected time-step

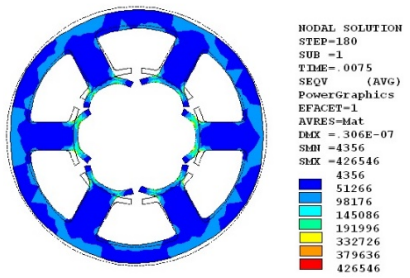


Figure 13: Numerical transient analysis of stator deformation and Von Mises stress σ_{eqv} [Pa] for selected time-step

5. Measurements of BLDC in no-load mode

Experimental studies were conducted to verify the numerical results. The rotation speed in no-load mode was equal 4000 rpm. The maximum oscillation amplitude occurs near the natural frequencies of the stator (second mode vibration: f = 2960 Hz and fourth mode vibration: f = 10020 Hz) and rotor (second bending mode of rotor: f = 10550 Hz measured for the rotor placed within the motor structure by shaker method) – Figure 14. Here it is advisable to make a note of additional information that the investigated motor was suspended by elastic springs. The current was recorded in no-load mode and its Fourier transformation is presented in Figure 15. It can be noticed that the vibrations spectrum was similar for different values of velocity of the rotor. The main sources of vibrations are resonant frequencies excited by magnetic forces and in keeping with different electric machine imperfections creating harmonic forces exciting the bending mode of the rotor [3]. It may be concluded from comparison of Fourier analyses (FFT). In Figure 14 this analysis is realized within the radial velocity of vibrations however in Figure 15 the current of the motor stator can be investigated.

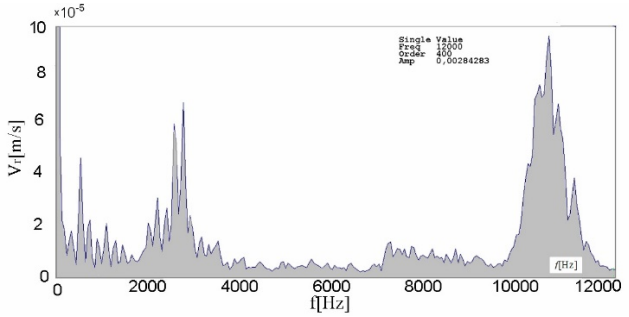


Figure 14: Measurement of stator vibrations - Fourier analysis of radial velocity Vr [mm/s] for selected point

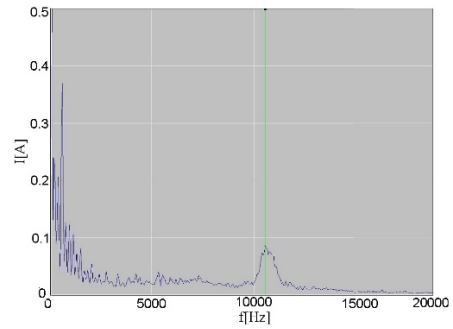


Figure 15: Measurement of current supply I [A] of BLDC motor during no-load mode - Fourier analysis

6. Conclusions

Due to the high efficiency, electrical BLDC motors are becoming more widely used as torque motors in control. systems and instrumentation. One of the disadvantages of BLDC motor is the increase in noise due to the use of PWM converters, so the study of the possibilities of reducing motor noise is an important problem. The article investigated the natural frequencies of the stator and the rotor, as well as the possibility of resonance due to coincidence of the frequency of magnetic forces and the pulsation of the motor torque with natural vibrations of the stator or rotor.

The study of the magnetic forces in the air gap using the finite element method allowed us to determine the main modes of the magnetic forces, the influence of which can lead to the occurrence of resonance and an increase in motor noise. The analysis of the natural frequencies of the stator was carried out on the basis of the equation of free oscillations without taking into account damping. The influence of the windings was taken into account in a simplified way, by adding mass to the stator by increasing density of stator teeth, which led to a small error for most important circumferential modes of vibration of the stator (the second and fourth mode natural frequencies) and significant error in comparison with the experimental results for the other natural frequencies.

The natural frequencies of the rotor were also determined by comparing experimental and numerical results calculated for different stiffness values of the bearings implemented as spring elastic - damping elements.

Transient stator vibrations due to magnetic forces were depending on the number of teeth, the number of poles, and the speed of rotation. The rotation speed was equal 4000 rpm. Experimental studies were conducted to verify the results. The maximum oscillation amplitude occurs near the natural frequencies of the stator and rotor. Hence, it is concluded that the main sources of vibration are the natural frequencies of the motor.

During motor operation and rotation of the rotor, the form of the stator vibrations may change the dominant order of mode. These scientific results should be taken into account when designing the BLDC motor.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgment

The computations were performed PLGrid infrastructure.

References

[1] J. Gieras, Permanent magnet motor technology: design and applications", 3rd ed. Taylor and Francis Group, ISBN 978-1-4200-6440-7, 2010.

[2] J. Besnerais et al., "Prediction of Audible Magnetic Noise Radiated by Adjustable-Speed Drive Induction Machines", IEEE Transactions on Industry Applications, **46**(4), 1367-1373, 2010, doi: 10.1109/TIA.2010.2049624.

[3] J. Besnerais et al., "Analysis of the electromagnetic acoustic noise and vibrations of a high-speed brushless DC motor," 8th IET International Conference on Power Electronics, Machines and Drives (PEMD 2016), Glasgow, UK, 1-10, 2016, doi: 10.1049/cp.2016.0284.

[4] H.J. Lee et al., Noise source identification of a BLDC motor", Journal of Mechanical Science Technology", **22**(708), 2008, doi: 10.1007/s12206-008-0110-9.

[5] U. Kim et al., "Effects of magnetically induced vibration force in brushless permanent-magnet motors", IEEE Transactions on Magnetics, **41**(6), 2164-2172, 2005, doi: 10.1109/TMAG.2005.847628..

[6] M. Furlan et al., "The Magnetic Noise of a DC Electric Motor - Modeling of Three-Times-Coupled Electromagnetic, Mechanical and Acoustic Phenomena". In: Wiak S., Krawczyk A., Trlep M. (eds) Computer Engineering in Applied Electromagnetism. Springer, Dordrecht, 2005, doi: 10.1007/1-4020-3169-6_48.

[7] J.D. Ede et al., "Rotor resonances of high-speed permanent-magnet brushless machines", IEEE Transactions on Industry Applications, **38**(6), 1542-1548, 2002, doi: 10.1109/TIA.2002.804765.

[8] J. Podhajecki, "Numerical and experimental vibration analysis of BLDC motor", 19th International Symposium on Electromagnetic Fields in Mechatronics, Electrical and Electronic Engineering (ISEF), Nancy, France, 1-2, 2019, doi: 10.1109/ISEF45929.2019.9096907.

[9] A. Belahcen, Magnetic Forces and Magnetostriction in Electrical Machines", 951-22-7183-4, 2004.

[10] M. Islam et al., "Noise and vibration characteristics of permanent magnet synchronous motors using electromagnetic and structural analyses", IEEE Energy Conversion Congress and Exposition, Phoenix, AZ, USA, 3399-3405, 2011, doi: 10.1109/ECCE.2011.6064228.

[11] J.P. Lecointe et al., Five methods of stator natural frequency determination: Case of induction and switched reluctance machines", Mechanical Systems and Signal Processing. **18**, 1133 – 1159, 2004, doi: 10.1016/j.ymssp.2003.08.002.

[12] W. Cai, "Impact of stator windings and end-bells on resonant frequencies and mode shapes of switched reluctance motors", IEEE Transactions on Industry Applications, **38**(4), 1027-1036, July-Aug. 2002, doi: 10.1109/TIA.2002.800594.

[13] C. Schlensok, Electromagnetically excited audible noise - Evaluation and optimization of electrical machines by numerical simulation", COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, **26**, 727-742, 2007, doi: 10.1108/03321640710751181.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 788-800 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Dependency Head Annotation for Myanmar Dependency Treebank

Hnin Thu Zar Aye*, Win Pa Pa

Natural Language Processing Lab, University of Computer Studies, Yangon, Yangon, 11411, Myanmar

ARTICLE INFO

Article history:
Received: 22 September, 2020
Accepted: 17 November, 2020
Online: 24 November, 2020

Keywords:
Dependency head
Universal Dependencies
Treebank
Annotation schemes

ABSTRACT

Complete manual annotation of dependency treebank needs resources like annotators and annotation tools and takes long time and has high possibility of inconsistent annotations for free word order languages such as Myanmar. This paper describes a dependency head annotation scheme with Universal part-of-speech and Universal Dependencies for Myanmar dependency treebank. Currently 22,810 sentences and 680,218 tokens were annotated from three corpora for Myanmar dependency treebank. Some language specific issues are also described with examples. Raw syntactic structures were annotated automatically by UDPipe according to the Universal Dependencies based on Universal-part-of-speech tag scheme. Then unsupervised annotated dependency head structures have been manually updated in post processing. To be reliable and speedy post process with reduced errors for manual updating, selected sentences were added to the training data after being updated. After that the model has been retrained and the remaining sentences were parsed by UDPipe. Post processing was repeated until all sentences were updated. Some specifications of dependency annotation schemes in sentences encountered in post processing are presented with examples. For parsing performance of annotated data, cross validation tests and parsing experiments were performed. Moreover, annotated treebank data have also been evaluated by CoNLL 2017 evaluation script for parsing performance. Results of parsing experiments and evaluation are also reported by unlabeled and labeled attachment scores and demonstrated that the proposed method is a suitable way for building Myanmar dependency trees. Moreover, syntax structures of treebank are also analyzed and syntax information is also presented. This dependency head annotation for dependency treebank is the first work for Myanmar language as far as we know.

1. Introduction

A treebank annotated with syntax or dependency structure is an essential and important resource for natural language processing systems in any language. Treebank annotated dependency structures is called dependency treebank. While phrasal constituents and syntax rules could not provide a direct role in sentences, syntactic dependency information of a sentence can describe directed grammatical relations between words. Moreover, dependency grammar is also able to deal with morphologically rich and relatively free word order languages. Dependency treebank is also critical resource in any language to develop natural language processing applications [1].

Many dependency treebanks have been constructed manually for many languages such as Czech, German, Danish, French, Portuguese, Estonian, Russian, Dutch, Danish, Turkish, Basque, Italian, English [2], Norwegian [3], Finnish [4], Romanian [5], Ancient Greek and Latin [6], Vietnamese [7],

Annotating dependency syntactic information in sentences is still a hard task for Myanmar having free word order nature. Moreover, currently there is still low resource for syntactic information for Myanmar language.

The Myanmar grammar is different from other languages of ASEAN countries such as Thailand, Vietnam, Malaysia and these languages already have treebank resources. Myanmar has been similar structures with the other SOV order languages such as Japanese, Chinese, and Korean and also a head final language. According to these properties, for Myanmar, a dependency-based head finalization has been proposed for statistical machine translation (SMT) in [8]. Although the proposed method was being able to improve a baseline SMT result without requiring parallel

*Corresponding Author: Hnin Thu Zar Aye, hninthuzaraye@ucsy.edu.mm

training information, it depends on dependency parser of source-side language to achieve higher performance than unsupervised baseline result [8]. Having similar syntactic structures of Japanese and Myanmar, a parsing approach has been proposed by applying SMT by using Japanese as pivot in [9]. The proposed method performed well with satisfying results due to the good performance of Japanese parser [9]. Although Myanmar dependency syntax structures were applied competently for SMT [8] and parsing [9] without annotated Myanmar corpus, there was dependence on intermediate language dependency parser like English and Japanese as limitation

Progress of unsupervised dependency parsing researches is increasing with the shared tasks of the tenth conference on computational natural language learning (CoNLL-X) in recent years [10]. A Universal part-of-speech tag (U-POS) set has been proposed as standard for research in unsupervised induction of syntactic structure [11]. Universal Dependencies (UD) is a project developing cross-linguistically consistent treebank annotation for many languages [12]. Currently, there are over 100 treebanks of more than 70 languages available in the UD inventory. Treebanks have derived UD format from existing formats in many languages like Korean [13]. UDPipe has been proposed to easily perform basic natural language processing tasks from tokenization to parsing in CoNLL-U format, the revised version of CoNLL-X format, for treebanks of UD without requiring any other external data [12]. UDPipe has been applied in dependency treebank building [14, 15].

Currently, Asian language treebank (ALT) project has developed a Myanmar syntax treebank with a parallel corpus by annotators using web-based tool [16]. Building treebank needs annotators and applied tools for processes of tree building. Therefore, it becomes hard for low resource language like Myanmar.

Having limitation of related works [8, 9], no resource for dependency parsing and information of Myanmar, improvement of unsupervised parsing researches [10], and simple trainable facility of UDPipe with CoNLL-U format, as our motivation, we annotated a corpus by applying U-POS tags and unsupervised dependency parsing by UDPipe of UD project to get raw syntax information for Myanmar [17]. Then, as future work of unsupervised annotation, manual post processing is carried out on the unsupervised parsed results with reference dependency structures to build dependency treebank in order to apply in parsing Myanmar sentences by deep learning approaches.

This paper presents dependency head annotation to build Myanmar dependency treebank by U-POS tag sets and UD 2.0 guidelines by updating in post processing on unsupervised annotated corpus. Building Myanmar dependency treebank contains two main parts. The first is automatic annotating by unsupervised dependency parsing by UDPipe to get raw universal dependency syntactic information [17]. The second is manual post processing by dependency structures for correct dependency heads. During post processing, UDPipe is applied in a bootstrapping manner to be consistent updating with reduced post processing time..

The organization structure of paper is as follow: Section 2 briefly describes nature of Myanmar language and sentences. Section 3 presents corpus information and overview of annotation scheme. Section 4 describes corpus pre-processing tasks before annotation. Section 5 presents Myanmar language specific tags and their related U-POS tags and mapping scheme between language POS and U-POS tags. Section 6 describes post checking and updating unsupervised dependency heads in sentences with examples. Section 7 reports parsing experiments with post processed data. Section 8 discusses about parsing and evaluation results of treebank. Section 9 concludes all sections and presents future work.

2. Myanmar Language

Myanmar (Burmese) is a member of the Lolo-Burmese grouping of the Sino-Tibetan language belonging to the Southern Burmish branch of the Tibeto-Burman languages. It is an official language in Myanmar. It is also the first language of the Bamar people, the principal ethnic group and related ethnic groups of the country, and a second language of ethnic minorities in Myanmar. Myanmar (Burmese) is a tonal, pitch-register, and syllable-timed language, largely monosyllabic and analytic, with a subject–object–verb word order. It is morphological rich and agglutinative language.

2.1. Nature of Myanmar Sentences

Myanmar sentences can be written with formal or colloquial style. Two types of Myanmar sentence construction are simple and complex sentence types. Simple sentence has only one nominal phrase, action maker or subject, and one verb phrase. Complex sentence has two or more clauses or simple sentences, joined with conjunction, or post positional markers, or particles to modify the followed part which might be phrase or clause of main sentence.

Sentence / (Type)	ကိုယ် တွင် ညို သော်လည်း ၊ မျက်နှာ နှင့် လက် များ မှာ မည်း နေ သည် ။ (Complex sentence)													
Clauses / (Type)	***ကိုယ်***တွင် ***ညို*** သော်လည်း ၊ (Dependent clause)					***မျက်နှာ*** နှင့် ***လက်*** များ မှာ ***မည်း***နေ သည် ။ (Independent root clause)								
Phrases/ (Type)	***ကိုယ်*** တွင် (Noun)		***ညို*** သော်လည်း ၊ **(Adjective)**			***မျက်နှာ*** နှင့် ***လက်*** များ မှာ (Noun)					***မည်း***နေ သည် ။ **(Adjective/Root)**			
Words	***ကိုယ်***	တွင်	***ညို***	သော်လည်း	၊	***မျက်နှာ***	နှင့်	***လက်*** များ		မှာ	***မည်း***	နေ	သည်	။
Morphenes	***ကိုယ်***	တွင်	***ညို***	သော်လည်း	၊	***မျက်နှာ***	နှင့်	***လက်***	များ	မှာ	***မည်း***	နေ	သည်	။
Morpheme translation	body	at	brown	although	,	face	and	leg	-s	-	black	-	-	.
Sentence translation	Although body is brown, face and legs are black .													

Table 1: Suffixes of example sentence from Figure 1

Types	Nominal marker (post positional marker)	Plural marker (particle)	Verb suffix (particle)	Verb marker (post positional marker)
Morpheme Words	“တွင်”,“များ”, “မှာ”	“များ”	“နေ”	“သည်”

Table 2: Composition of Myanmar Dependency Treebank

Corpus Source	Domain	Sentences	Tokens	Average Length	Remark
myPOS	Wikipedia	11,010	239,534	21.75	manually annotated
Web News	Websites’ News	1,800	66,710	37.06	manually annotated
ALT	Wiki news	10,000	373,974	37.39	Automatic annotated
	Total	22,810	680,218		

Myanmar noun phrases are usually ended with different types of nominal markers, postpositional markers (PPMs), to identify their roles in sentences such as subject, object but most colloquial style noun phrases are written without these markers. Myanmar sentences might have one or more phrases or clauses. Some subordinate clauses are usually used to represent more detail meaning for modified parts and also placed before modified ones. Having these nested parts, defining correct dependency between sub and main parts of the sentence takes more time. Figure 1 illustrates an example sentence structure.

There are four phrases in sentence of Figure 1. In Myanmar language, adjective ended with verb maker suffix is verb phrase. In the example sentence, sentence ended verb phrase is composed of adjective with post positional verb marker Four suffix types of morphemes of the example sentence are described in Table 1. Bold and italic words are main content words of each phrase. Main root phrase of a sentence or clause is final verb phrase of that sentence. Dependent sentence or clause modifies the independent root sentence. Therefore, main root of example sentence in Figure 1 is the last right most verb phrase.

Moreover, most formal sentences are usually ended with verb or verb phrase. However, some sentences may not be ended with verbs because of the Myanmar sentence writing style in which there are presence of hiding verbs hidden for actions of being or having or living or coming actions. Moreover, Myanmar sentences can be written by many forms according to nature of free word order language, and some special cases of sentence construction of Myanmar grammar. Moreover, emphasized noun phrases can be put at the beginning of the sentence according to writer’s idea.

These conditions can also be complicated and time-consuming issues to define correct dependency heads and relations for phrases in sentences. Besides above nested cases, there arises one issue in dependency tree building.

3. Corpus Annotation

This section presents corpus statistic and overall architecture of dependency head annotation.

3.1. Corpus Information

Currently two corpora have been annotated manually by dependency head information and one corpus has been annotated in unsupervised way for Myanmar treebank. The myPOS corpus consists of 11,010 sentences written by formal and colloquial format from the Myanmar Wikipedia including various areas such as economics, history, news, politics and philosophy [18]. The sentence writing style of myPOS corpus is very similar to current mostly used standard formal and colloquial style. Therefore, we selected first this corpus to annotate to get similar syntactic structures of current written styles of Myanmar sentences and annotated them as first standard sentences for other corpora. Web News corpus contains 1,800 sentences written by current modern writing styles in Myanmar news websites. ALT corpus contains 10,000 translated sentences from Wiki news. The statistical information of current Myanmar dependency treebank is presented in Table 2.

3.2. Overview Structure of Annotation

Using difference corpora in treebank, word segmentation and POS tagging style of each corpus is different. Being morphological rich and agglutinative language, most words are segmented to provide morphological level syntax information in this work. Therefore, pre-processing is needed to carry out to be the same word segmentation and POS-tagged scheme among different corpora. U-POS tags, CoNLL-U shared task format, shared Japanese model, and UDPipe of UD project were used to get raw universal dependencies in this work.

The corpora were transformed to CoNLL-U format by adding U-POS tags in pre-processing. After transformation, they were annotated by unsupervised dependency parsing by using shared Japanese model in UD project [17] because of similar conditions in grammar structure of Myanmar and Japanese and dependency structures and UD of Japanese [19,20]. Then unsupervised annotated dependency heads were manually post checked by human annotator to be more correct dependency head nodes. One annotator was done manual post checking on automatic annotated results by learning Myanmar language grammar books and books written by linguistic experts of Myanmar language. Therefore, dependency head annotation of Myanmar dependency treebank has two steps: initial corpus pre-processing and post processing on unsupervised annotated results of pre-processing.

In post processing, to be consistent and fast checking and updating, selected updated data were repeatedly trained with all other not updated data in corpus by UDPipe. After checking and updating manually selected 2,000 sentences, they were added to the training data, Then the training model was retrained to parse the remaining sentences in corpus by the updated model. Parsed

results were updated manually and trained again until all sentences were post processed in corpus. The overall architecture of annotation is illustrated in Figure 2. Detail explanation with examples will be presented in Section 4 and Section 6.

4. Pre-processing

This section presents overview of corpus pre-processing step. Before dependency structures annotation, it is needed to check and update word segmentation and POS tagging of used corpora because some corpora might have different segmentation and POS tagging styles based on their original created purposes. Using unique word segmentation and POS tagging can be easy to transform dependency corpus. It is also better and easier having unique word segmentation and POS tagging than different formats to transform dependency corpus.

To have a consistent Myanmar POS tag scheme, a new general POS tag scheme has been defined. It will be explained in following sub section. Moreover, U-POS tags were also added to the CoNLL-U format in order to get raw universal dependencies syntax structures.

In Myanmar sentences, some words might also have different POS tag forms for the same word. Tagging correctly for each content word in sentences is important for correct dependency head. One example of these issues is presented in Table 3. The quality of word segmentation and POS tagging can mainly impact providing correct dependency information of sentences in corpus. Therefore, rechecking word segmentation and POS tagging of used corpus before annotation is very important.

Then, the correct POS tagged corpus was transformed CoNLL-U format by adding related U-POS tags for most language POS tags by manual python script as UD 2.0 format. In adding U-POS tags by python script, it is needed to check again manually U-POS tags of Myanmar conjunctions because it is also needed to be correct form of two conjunction tags of U-POS as described in Table 5 according to the content words of sentence.

5. Part-of-speech Tagging Scheme

There are ten main POS tags in Myanmar language such as noun, pronoun, adjective, verb, adverb, post-positional marker, particles, conjunctions, interjection and punctuation [21]. For easy mapping to U-POS tags, 16 POS tags have been defined for language specific tags which are presented in Table 4 with examples.

Most tags of these were already defined in ALT [6] except proper noun (PRPN) and text number (TNUM). In the previous work of this, these two tags were defined as noun, “N” [17]. In most Myanmar corpora, proper nouns and text numbers are tagged as noun, but they are already defined as proper noun and number in U-POS tag set. Therefore, these two tags were also added in Myanmar language tag set in order to be fast and easy mapping between Myanmar POS tags and U-POS tags.

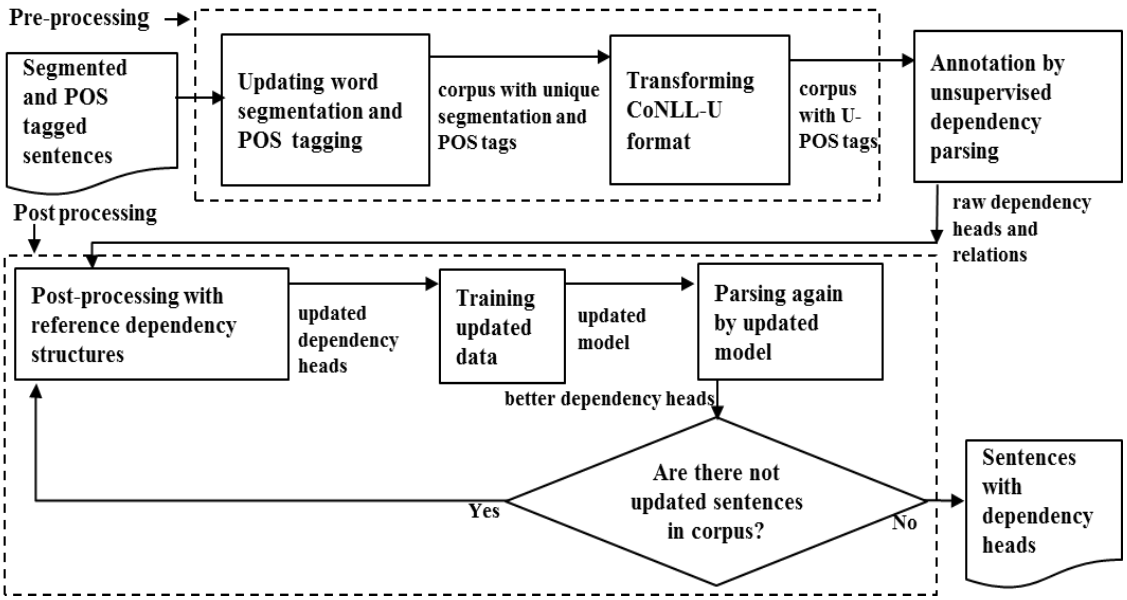


Figure 2: Overview architecture of dependency head annotation

Table 3: Example POS tagging for same word

Words	Correct Tags	Meanings	Examples in phrases
နေ	verb	live	Myanmar : ရန်ကုန် မှာ နေ Words in English: Yangon in live - Translation : Live in Yangon.
နေ	particle	describing continuous action	Myanmar : လေ့လာ နေ သည် Words in English: study - - Translation : is studying

Table 4: Myanmar POS tags for language specific tag set

POS Tag	Description	Example words
N	Noun	သတင်းစာ (Newspaper), ခန်းမ (Hall), ကုမ္ပဏီ (Company)
PRPN	Proper Noun	စက်တင်ဘာ (September), အာရှ (Asia)
NUM	Number	ဝ (0), ၁(1), ၂(2), ၃(3), ၄(4)
TNUM	Number text letter	သုည (zero) , တစ် (one) , နှစ် (two), သုံး (three), လေး(four)
FOR	Foreign word	Carsh, Federal, process, Bit
ABB	Abbreviation	ညွှန်/ချုပ် (Director-General)
PRON	Pronoun	ကျွန်တော်/ ကျွန်မ/ ကျွန်ုပ် (I), ထို (that), မည်သူ (who)
ADJ	Adjective	ကျယ်ဝန်း (wide), ယဉ်ကျေး (polite), မြင့်မား (high), လေးလံ (heavy)
ADV	Adverb	မကြာခဏ (frequently), အလွန် (very), လောလောဆယ် (currently)
V	Verb	စား (eat), သွား (go), ရေးသား (write), ဖြစ် (be), ရှိ (has/have), တည်ရှိ (exist)
CONJ	Conjunction	နှင့် (and), သို့မဟုတ် (or), ၍/သောကြောင့် (because)
PART	Particle	များ/တို့/တွေ (plural marker), ခန့် (about), ခဲ့ (past marker), နိုင်(can), ကြ (plural action marker)
PPM	Post positional marker	သည်/က/မှာ (nominal marker for subject), ၌ (at) , ၏ (of) , ဖြင့် (by)
PUNC	Punctuation	၊ , ။ , “ , (or -LRB-,) or -RRB- , “, ”
SB	Symbol	% , $
INT	Interjection	အိုး (Oh), အမလေး (Oh my god!)

Table 5: Mapping scheme between Universal POS and Myanmar language specific POS tags

U-POS Tag	Description [Examples]	Myanmar Language POS
NOUN	Noun	N
		FOR
PROPN	Proper Noun	PRPN
		ABB
NUM	Number	NUM
		TNUM
PRON	Pronoun	PRON
ADJ	Adjective	ADJ
ADV	Adverb	ADV
VERB	Verb	V
CCONJ	Coordinating Conjunction [နှင့် (and), သို့မဟုတ် (or)]	CONJ
SCONJ	Subordinating Conjunction [၍/ သောကြောင့် (because), နှင့်တစ်ပြိုင်နက် (as soon as), ပါက (if)]	
PART	Particle	PART
ADP	Adposition	PPM
PUNCT	Punctuation	PUNCT
SYM	Symbol	SB
INTJ	Interjection	INT

Table 6: Mapping scheme between Universal POS and Myanmar language specific POS tags

U-POS Tag	Description	myPOS	ALT	Web News	Total
NOUN	Noun	55,590	71,246	18,237	145,073
PRON	Pronoun	2,680	7,864	331	10,875
PROPN	Proper Noun	12,377	23,168	3,375	38,920
NUM	Number	6,261	11,240	2,259	19,760
ADJ	Adjective	7,108	7,776	1,393	16,277
ADV	Adverb	2,868	4,686	852	8,406
VERB	Verb	34,046	62,876	11,057	107,979
CCONJ	Coordinating Conjunction	4,624	6,368	1,584	12,576
SCONJ	Subordinating Conjunction	6,587	7,493	1,452	15,532
PART	Particle	52,413	78,268	13,294	143,975
ADP	Adposition	38,838	64,624	8,674	112,136
PUNCT	Punctuation	15,845	28,255	4,199	48,299
SYM	Symbol	199	116	3	318
INTJ	Interjection	98	0	0	98

All conjunctions in Myanmar sentences were tagged as CONJ which were mapped to one of two conjunctions of U-POS tags: CCONJ and SCONJ, coordinating and subordinating conjunction respectively. Mapping scheme between U-POS tags and Myanmar POS tags can be seen in Table 5. Total frequencies of U-POS tags in each corpus of treebank are listed in Table 6.

6. Post-processing

Annotation method and types of annotation schemes are important in dependency treebank construction. The annotation was based on Universal Dependencies. In the UD annotation scheme, dependency relations are expressed between words and main content words attached leaf words such as function words in sentences [22]. Main parts of syntactic structures are head nodes and relation links called as dependency relation labels between words.

Generally Myanmar nouns, adjectives, and verbs are formally written with suffixes such as post positional markers or particles. Currently, dependency relation labels are automatic unsupervised annotated results without post processing in treebank. Only dependency head information was post processed.

Overview of the linking structures of dependency head node words between dependent words in most occurred phrases in sentences had been presented in our previous work [17]. In that work, only construction of dependency link arc connections between heads and dependents words had been presented. Therefore, the arcs directed to the heads from dependents.

After the previous work, unsupervised dependency annotation results have been post processed to update dependency head links based on the dependency structures described in our previous work. Updating dependency relation labels needs more time because of few annotators and issues of sentence writing styles. Therefore, dependency relations labels are still automatic unsupervised annotated labels after post processing. To describe full referenced dependency information, in this paper, dependency relation labels between head node words and dependent words will be presented with unsupervised annotated labels in sample sentences according to the viewpoint of the language typology.

6.1. Proper Noun and Possessive Phrase

The specific unique names of common nouns are called as proper nouns as in other languages. Myanmar proper nouns are usually found in before or after common nouns [21]. In Figure 3, two proper nouns, following two common nouns can be seen in example sentence. In that sentence, example possessive case can also be seen. Possessive case of a noun or pronoun can be written by a post positional suffix marker, “ ၏”, to show possession of following right noun by left proper noun of it.

6.2. Compound Noun

Myanmar compound nouns might have two or more words and POS tags of these words can be nouns, verbs, adjectives or adverbs [21]. An example compound noun in sentence of Figure 4 contains two consecutive nouns and means “memorial stamp”. In this case, the left noun modifies the right one.

6.3. Numeral Phrase

Most numeral noun phrases can be written by three general formats which are described in Table 7 with their meanings. Counting amount can be written by digit number or text number [21]. Sample numeral phrases in sentence can be seen in Figure 5.

6.4. Adjective Phrase

Adjectives modify nouns and are usually placed before or after noun in Myanmar sentences. Myanmar adjectives can be written as simple or simple adjective with suffix particles or transformed adjective [21]. Suffixes of adjective or example adjective phrases are presented in Table 8. Sample dependency of adjective phrase in sentence can be seen in Figure 6.

6.5. Adverb Phrase and Verb Phrase

Adverbs can be written as simple or transformed adverb with the suffix particle, “ စွာ”, followed by verb, or adjective, or adverb.

Example adverb phrases can be seen in Table 9. Example dependency of adverb phrase in sentence can be seen in Figure 7.

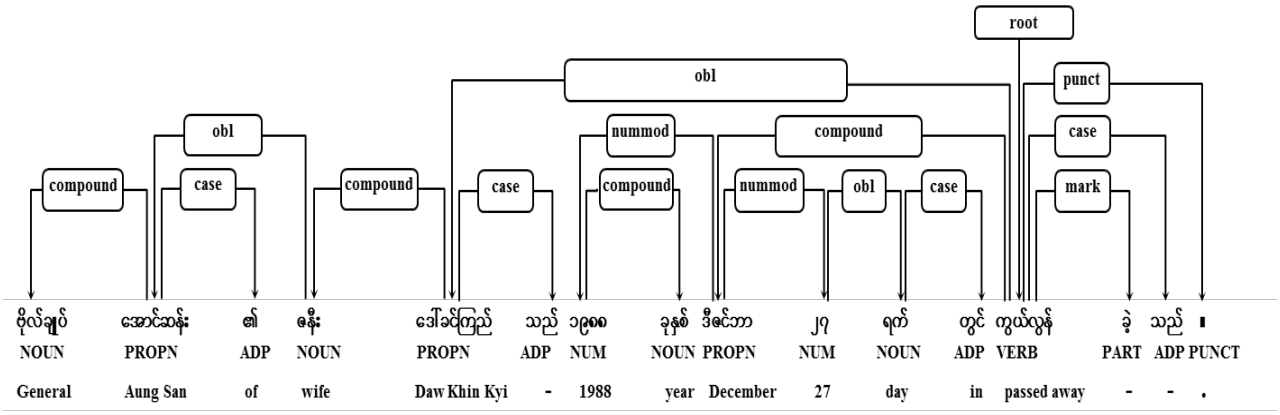


Figure 3: Proper noun and possessive case examples in sentence

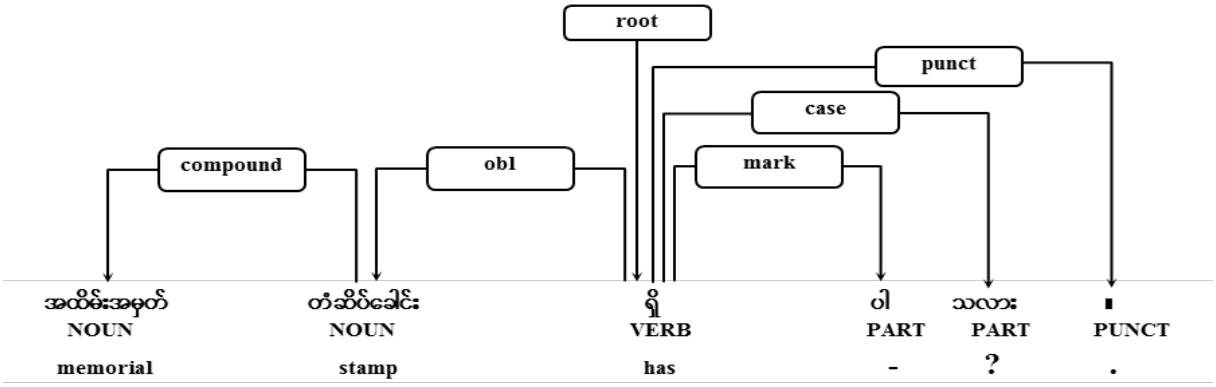


Figure 4: Compound noun example in sentence

Table 7: Numeral phrase formats

Numeral Phrase Forms	Example	Meaning	Remarks
N NUM/TNUM PART/N	မှတ်ချက် ၂/နှစ် ချက် comment 2/two - (Glossary)	2/two comments	
N N (noun affixed particle, "အ", to form common nouns) NUM/TNUM	မှတ်ချက် အချက် နှစ်ဆယ် comment fact twenty (Glossary)	twenty comments	if counted amount is exact numbers of multiple of ten, hundred, thousand, etc.
N N(transformed or common noun) NUM/TNUM PART/N	မှတ်ချက် အချက် နှစ်ဆယ် comment fact twenty two - (Glossary)	twenty two comments	

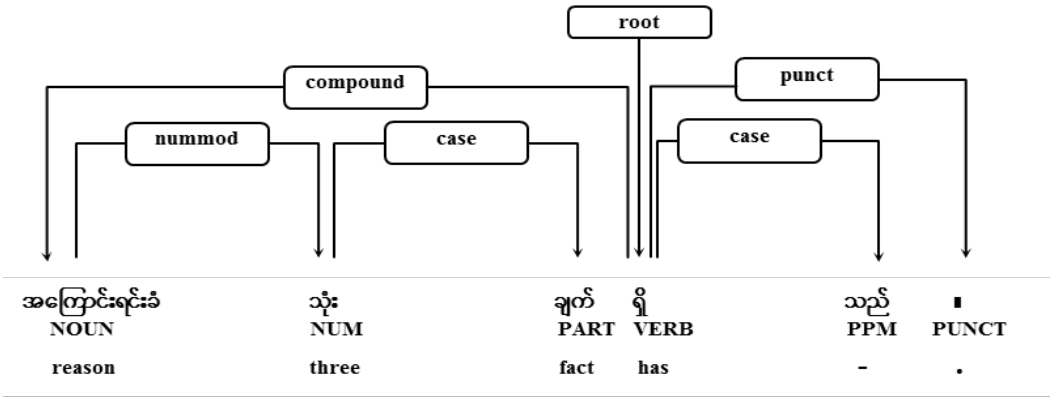


Figure 5: Numeral phrase example in sentence

Table 8: Suffixes of adjective or adjective phrases and examples

Particles used to suffix or transform adjective or adjective phrase		သော, သည့်, မည့်, တဲ့
Example Myanmar adjective	***Translations***	***Remarks***
ကြီး သော အိမ် **big** - house	big house	Simple adjective followed suffix
အိမ် **ကြီး** house **big**	big house	Simple adjective
အစိုးရ ပေး သည့် အိမ် government provide - house	house provided by government	Transformed adjective

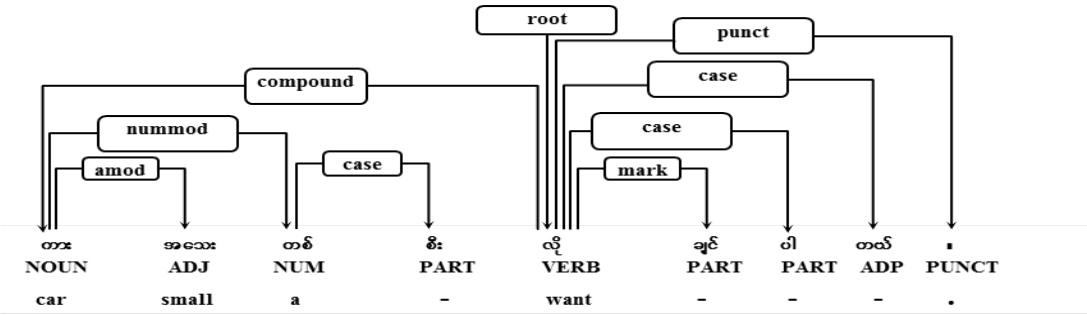


Figure 6: Adjective phrase example in sentence

Table 9: Example of adverb forms

Example Myanmar adverb phrases	**Meanings**	**Remarks**
မြန်မြန်(quickly)	quickly	Simple adverb
မြန်မြန်သွက်သွက် (quickly)	quickly	Simple adverb
လျင်မြန် စွာ quick -	quickly	Transformed adverb with suffix particle

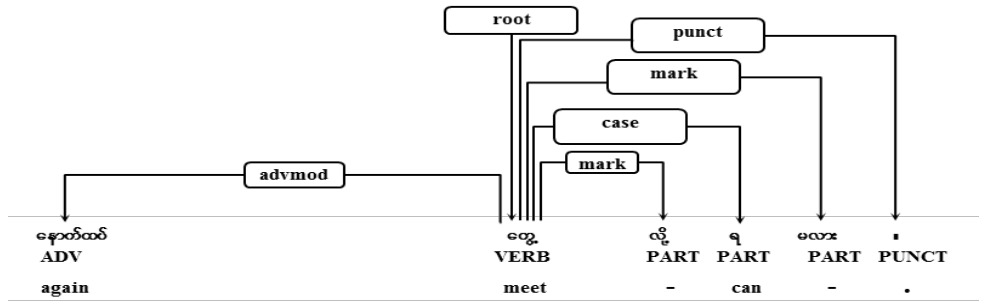


Figure 7: Adverb and verb phrase example in sentence

Table 10: Suffixes of verbs

Suffixes to verb	**Description of usage and example**	**Suffix Type**
သည်, ၏, ပြီ, မည်, မယ်, တယ်	to form verb by showing tense Examples: သွား သည် (go) သွား မယ် (will go)	post positional markers
မှာ, ပါ, စမ်း	to express giving order or answering action Examples: လုပ် ပါ (giving order or requesting “Do” action) မောင်း မှာ (answering or forecasting “drive” action based on the meaning of content words in sentence)	particles

Verb phrases in sentences are usually formed by one of the verb ended post positional markers or particles [21] as presented in Table 10 .

In addition, zero or more particles can be followed by verbs before verb ended markers to express the full state of action. In example sentence of Figure 7, the verb phrase, " တွေ့လို့ရမလား", means "can meet" and "တွေ့" is main verb and the suffix sequence with three particles, "လို့ ရ မလား", means asking politely for the requested action. That main verb phrase is also modified by left adverb phrase, " နောက်ထပ်". As a result, the verb, "တွေ့", is root of that sentence. The whole sentence means "Can meet again?",

6.6. Conjunctions

In Myanmar language, conjunctions are used to combine not only clauses or simple sentences but also related words or phrases. Moreover, simple sentences can be connected by suffixes: post positional markers, " က, မှာ, ကို", or particles "ဟု , သော , သည့် , မည့်, တဲ့" to form noun or adjective clause in sentence. If two or more simple sentences are connected by post positional markers or particles or conjunctions, combined sentence becomes complex sentence and clauses represent the roles as subjects or objects or adverbs. Most clauses ended by conjunctions are usually adverb modifiers in main sentence. Therefore, the role of dependent clauses can be divided into three types: noun clause, adjective clause, and adverb clause based on their roles in sentences. If combined clauses or sentences contains the same subject, subject can be omitted in dependent clause or independent clause or in both. Similarly, object noun can also be omitted in dependent clause or independent clause if it is placed in one [21, 23].

Some conjunctions described in Table 11 are used to connect words, phrases to connect two sentences to give coordinated extra meaning [21, 23].

Table 11: Example conjunctions

Conjunctions	**Usage**
နှင့် , လည်းကောင်း…လည်းကောင်း, ရော…ပါ, ရော… ရော, သို့မဟုတ်, ဖြစ်စေ…ဖြစ်စေ, ဖြစ်ဖြစ်…ဖြစ်ဖြစ်, သော်လည်းကောင်း…သော်လည်းကောင်း, မှတစ်ပါး, ပြီး	to connect words, phrases in sentence for extra meaning
ထို့ပြင်, ထို့အပြင်, ၎င်းပြင်, သည့်ပြင်	to give connection between prior sentence and next by giving coordinated extra meaning.

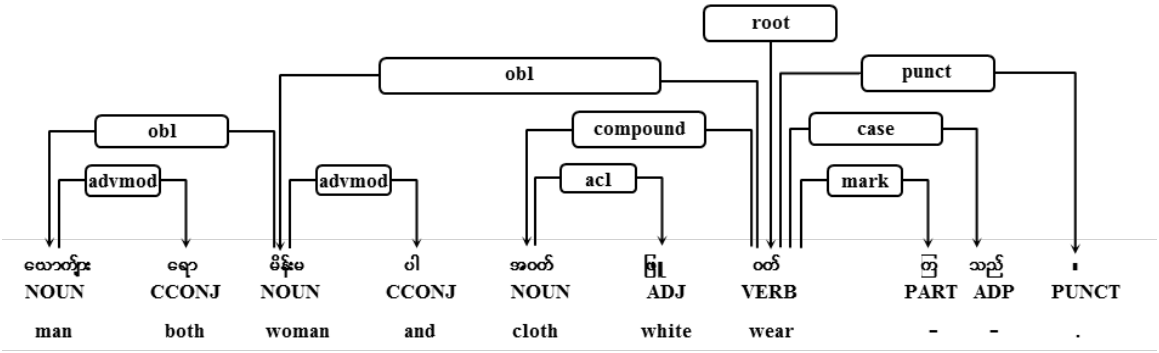


Figure 8: Coordinated phrase example in sentence

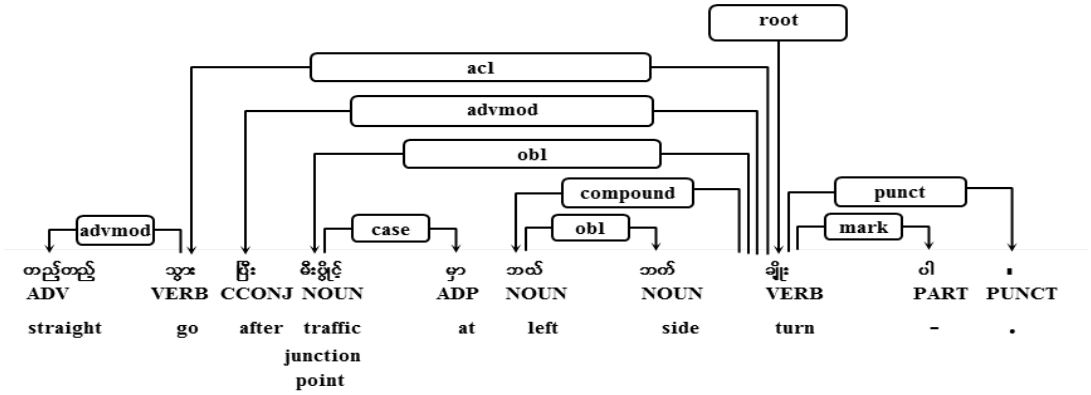


Figure 9: Coordinated clauses example in sentence

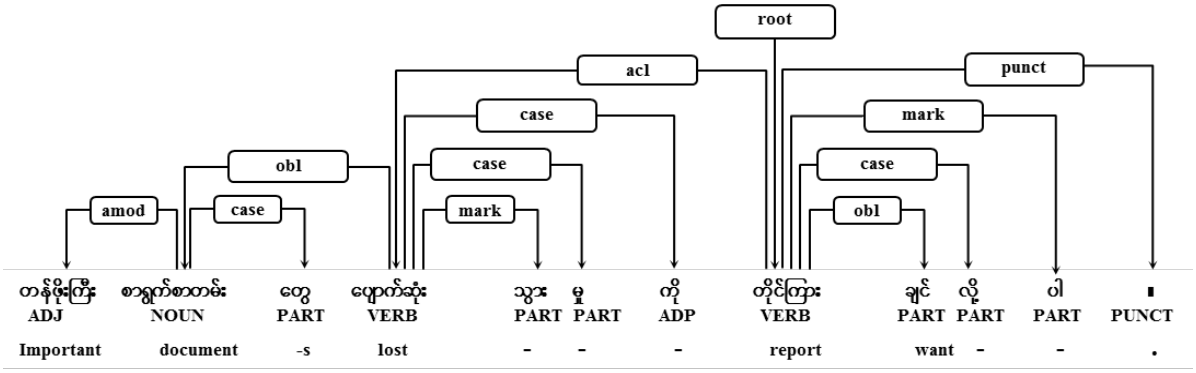


Figure 10: Complex sentence with noun clause

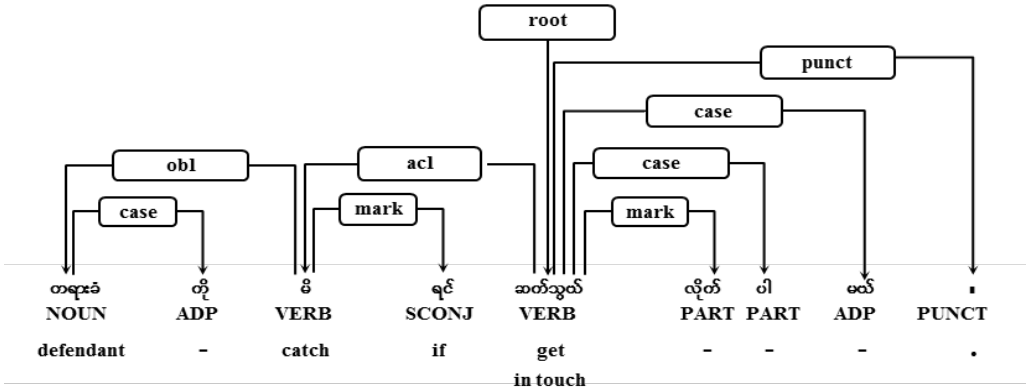


Figure 11: Complex sentence joined with subordinating conjunction

Table 12: Average parsing precision scores of each corpus

Corpus	Sentences		Average Scores	
	Train	Test	UAS	LAS
myPOS	10,010	1,000	89.06	86.67
Web News	1,620	180	87.03	84.83
ALT	9,000	1,000	91.40	90.24

Sample coordinated phrase with conjunctions in sentence can be seen in Figure 8. Moreover, complex sentence combined two simple sentences by coordinating conjunction can also be seen in Figure 9. The left side of conjunction is dependent sentence or clause to give first action and the right side of conjunction is independent clause to give final action as shown in Figure 9. Subjects of clauses are omitted and main sentence means "After go straight, turn left at the traffic junction point.". Therefore, the conjunction is tagged as "CCONJ", coordinating conjunction, for U-POS tag.

Some Myanmar conjunctions such as "လျှင်/ရင်/ပါက, သောကြောင့်, နှင့်တစ်ပြိုင်နက်, သကဲ့သို့, စေရန်, သောအခါ, သော်လည်း,.." , "if, because, as soon as, as/like, in order to, when, although,.." , are frequently used to connect clauses to provide the required meaning for main clauses [21, 23]. Therefore, these types of conjunctions are tagged as "SCONJ", subordinating conjunction, for U-POS tag.

Example complex sentence including noun clause as an object role in sentence can be seen in Figure 10. Main complex sentence means "Lost of important documents is wanted to be reported." The PPM, " ကို", is used to connect the left clause to represent as object noun for the action of root verb, "တိုင်ကြား", of main sentence.

In addition, example complex sentence joined with subordinating conjunction type, " ရင်" (if) , can be seen in Figure 11. Subjects are omitted in both dependent clause and main independent clause. It means that "If defendant is caught, I will get in touch".

7. Parsing Experiment

The main purpose of building treebank is to use in dependency parsing. This section presents parsing experiments executed for performance of treebank. UDPipe is an open-source trainable pipeline processing tool to perform segmentation, POS tagging, lemmatization and dependency parsing without any other external

data for multiple languages. And it is also available under open-source Mozilla Public Licence (MPL) and provides bindings for C++, Python, Perl, Java and C#. UDPipe 1.2 has been used in our parsing experiments.

We divided 90% and 10% of total sentences from each corpus of treebank for training and test data to evaluate each corpus by 10-fold cross validation test. We arranged alternate order of test sentences range and all the rest for train data in each corpus and split them into equally-sized parts for each validation test by python script. We calculated average score of all validation tests to report as total average score of cross validation test for each corpus.

8. Results and Evaluation

In this section, the statistical results of parsing experiments to evaluate parsing performance of treebank and evaluation results will be described. For each corpus performance of treebank, average parsing precision scores are calculated from the total results of cross validation tests in each corpus and they are listed in Table 12. Average precision scores measured by unlabeled attachment score (UAS) and label attachment score (LAS) of each corpus are over 89% and 86%, over 87% and 84 %, and 91% and 90% in myPOS, Web News, and ALT respectively .

Each corpus of treebank was evaluated by the CoNLL 2017 UD parsing evaluation script. The evaluation results will be described in following sub section. In addition, syntax structures of the referenced dependency types were also analyzed by manual python script which counts types of dependency structures being countable types in each sentence. The analyzed results will be described in next sub sections.

8.1. Evaluation

Corpora contained in treebank were evaluated by the CoNLL 2017 UD parsing evaluation script, "conll17_ud_eval.py" [24], to know their parsing accuracy. Gold standard file and system output file are input to the evaluation script to evaluate the data. The parsing model was trained with all current data of treebank to generate system output of each corpus. The current accuracies comparing standard annotated data with system output parsing result of each corpus can be seen in Figure 12.

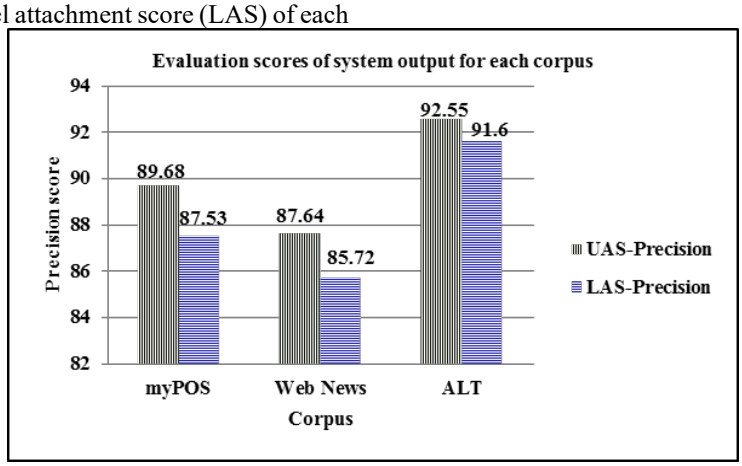


Figure 12: Precision scores of treebank data

Table 13: Information of sentence lengths in treebank

Sentence Type	myPOS	ALT	Web News	Total sentences of each range	Remark
Simple short	2,253	13	74	2,340	<=10 words
Short	6,463	3,858	732	11,053	>=11 and <=30 words
Normal range	1,929	4,398	568	6,895	>=31 and <=50 words
Long	349	1,687	401	2,437	>=51 and <=100 words
Very long	16	44	25	85	>100 words

Table 14: Phrases and clauses structures in treebank

Phrase Types	myPOS	ALT	Web News	Total
Noun	27,083	48,466	6,316	81,865
Proper Noun	10,514	17,750	3,013	31,277
Numeral Noun	5,790	10,852	2,125	18,767
Compound Noun	21,096	25,706	9,537	56,339
Adjective	9,535	11,993	1,581	23,109
Adverb	3,027	4,580	837	8,444
Verb	16,593	19,136	3,584	39,313
Phrases with Conjunctions	2,830	3,534	921	7,285
Clauses	7,027	9,837	1,950	18,814

Precision scores of UAS and LAS in system outputs of myPOS and Web News data are over 89% and 87%, and 87% and 85% respectively in comparing manually updated standard data. System output precision scores of UAS and LAS for ALT data is over 92% and 91% in comparing with automatic unsupervised annotated standard data.

As sentences of ALT corpus are longer than myPOS and Web News corpus, precision scores of unsupervised annotated ALT is more than manual updated standard data of myPOS and Web News corpus. However, system outputs of myPOS and Web News corpus are better and more similar to reference dependency structures than system outputs of ALT in manual random checking on system output trees.

Although currently Myanmar ALT has been developed manually, there is still no dependency resource for Myanmar. Constructing treebank manually is an error-prone and slow process. The high precision scores of cross validation tests and evaluations by post processed model illustrate that the proposed method can produce efficient precision scores for dependency parsing. With consistent and faster annotation, the proposed method will provide fast dependency tree building for Myanmar which has free word order and issues mentioned in Section 2

The length of sentences is also one main part of treebank characteristics for having various types of syntax and syntactic structures. Therefore, the ranges of sentences in treebank were also analyzed by classifying five levels by python script and resulted sentence ranges are listed in Table 13.

8.2. Syntax Analysis

The syntax information is one of the most important characteristics of treebank to know syntax status. It is difficult to count all syntax information of natural language sentences by program script exactly because of the issues of phrases and many sentence construction types discussed in Section 2. However, overview syntax structures of formal sentences from each corpus of treebank could be extracted by python script based on sequence order of words and POS tags information of phrases and clauses written by formal literature written style. The program counts the related phrase and clause types from word and POS tag sequences of sentences annotated as example dependency structure types described in Section 6. However, some informal phrases and clauses not ended with the formal related post positional markers or particles, could not be counted by the program. Overall countable syntax structures are listed in Table 14.

9. Conclusion

This paper has presented annotating dependency heads based on Universal Dependencies framework for Myanmar dependency treebank. Myanmar POS tags and U-POS tags of treebank, issues of tagging, and mapping scheme between two tag sets have also been presented. In addition, dependency head annotation schemes have also been described with sample sentences in line with grammatical point of views. This work is first for Myanmar to the best of our knowledge. Parsing experiments have also been executed for performance of treebank and results have also been presented. To conclude, contribution is first dependency head annotation for building Myanmar dependency treebank and can provide useful information for direct dependency parsing for Myanmar sentences by Myanmar model in future.

As future work, firstly we would intend to add more annotated sentences from same and different domains such as News articles data and Myanmar grammar books data to treebank because correct syntactic forms are able to provide faster annotation and better useful syntax information for Myanmar dependency treebank. The next work is to update unsupervised dependency labels according to the standard of Universal Dependencies based on Myanmar grammatical point of views.

References

[1] N. Green, “Dependency Parsing”, In WDS 2011 Proceedings of Contributed Papers, WDS 2011, 137–142, 2011, doi:10.1007/1-4020-4889-0_3..

[2] T. Kakkonen, “Dependency treebanks: methods, annotation schemes and tools”, In Proceedings of the 15th NODALIDA conference, 94–104, 2005, doi: arXiv:cs/0610124

[3] P. E. Solberg, “Building gold-standard treebanks for Norwegian”, In Proceedings of the 19th Nordic Conference of Computational Linguistics, 2013 (NODALIDA 2013), Linköping University Electronic Press, 459-464, 2013.

[4] K. Haverinen, J. Nyblom, T. Viljanen, V. Laippala, S. Kohonen, A. Missilä, S. Ojala, T. Salakoski, and F. Ginter, “Building the essential resources for Finnish: the Turku Dependency Treebank”, Language Resources & Evaluation, **48**(3), 493-531, 2014, doi: 10.1007/s10579-013-9244-1.

[5] C. MĂRĂNDUC, and C. A. PEREZ “A Romanian Dependency Treebank”, In International Journal of Computational Linguistics and Applications, **6**(2), 83-103,2015.

[6] D. Bamman, and G. Crane, "The Ancient Greek and Latin Dependency Treebanks," In Language Technology for Cultural Heritage 2011, Springer, 79-98, 2011, doi:10.1007/978-3-642-20227-8_5.

[7] K. Nguyen, “BKTreebank: Building a Vietnamese Dependency Treebank”, In Proceedings of the 11th International Conference on Language Resources and Evaluation 2018 (LREC-2018),European Languages Resources Association (ELRA), 2164-2168, 2018.

[8] C. Ding, Y. K. Thu, M. Utiyama, A. M. Finch, and E. Sumita, “Empirical Dependency-Based Head Finalization for Statistical Chinese-, English-, and French-to-Myanmar (Burmese) Machine Translation”, in Proceedings of the 11th International Workshop on Spoken Language Translation, 184-191, 2014.

[9] C. Ding, Y. K. Thu, M. Utiyama, and E. Sumita, “Parsing Myanmar (Burmese) by Using Japanese as a Pivot”, Proceedings of 14th International Conference on Computer Applications, 158-162, 2016

[10] D. Mareˇcek, “Twelve Years of Unsupervised Dependency Parsing”, ITAT 2016 Proceedings, CEUR Workshop Proceedings , **1649**, 56–62, 2016..

[11] S. Petrov, D. Das, and R. McDonald, “A Universal Part-of-Speech Tagset”, In Proceedings of the 8th International Conference on Language Resources and Evaluation, 2089-2096, 2012.

[12] M. Straka, J. Hajic, and J. Strakov ˇ a´, “UDPipe: Trainable Pipeline for Processing CoNLL-U Files Performing Tokenization, Morphological Analysis, POS Tagging and Parsing,” In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), 4290–4297, 2016.

[13] J. Chun, N. Han, J. D. Hwang, and J. D. Choi, “Building Universal Dependency Treebanks in Korean”, In Proceedings of the 11th International Conference on Language Resources and Evaluation 2018 (LREC-2018). European Languages Resources Association (ELRA), 2194-2202, 2018.

[14] E. Badmaeva, and F. M. Tyers, “A Dependency Treebank for Buryat”, In 15th Internationtal Workshop on Treebanks and Linguistic Theories (TLT15), 1-12, 2017.

[15] T. Rama, and S. Vajjala, “A Dependency Treebank for Telgu”, In Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories (TLT16), 119–128, 2018.

[16] Y. Kyaw Thu,, W. Pa Pa, M. Utiyama, A. Finch, and E. Sumita, “Introducing the Asian Language Treebank (ALT)”, In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16) 2016 May , 1574-1578, 2016. DOI: 10.1109/ICSDA.2016.7918974

[17] H. T. Z. Aye, W. P. Pa, and Y. K. Thu, “Unsupervised Dependency Corpus Annotation For Myanmar Language”, In Proceedings of the 2018 Oriental COCOSDA-International Conference on Speech Database and Assessments, IEEE, 78-83, 2018, doi:10.1109/ICSDA.2018.8693009

[18] K. W. W. Htike, Y. K. Thu, Z. Zhang, W. P. Pa, Y. Sagisaka, and N. Iwahashi, "Comparison of Six POS Tagging Methods on 10K Sentences Myanmar Language (Burmese) POS Tagged Corpus", In Proceedings of the CICLING 2017, April 2017..

[19] S. Mori, H. Ogura, and T. Sasada, "A Japanese word dependency corpus", Proceedings of the 9th International Conference on Language Resources and Evaluation 2014 May 26, 753–758, 2014.

[20] T. Tanaka, Y. Miyao, M. Asahara, S. Uematsu, H. Kanayama, S. Mori, and Y. Matsumoto, "Universal Dependencies for Japanese", In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16) , 1651-1658, May 2016.

[21] Department of the Myanmar Language Committee, Myanmar Grammar, Ministry of Education, The Republic of the Union of Myanmar, 2013.

[22] P. Osenova, and K. Simov, "Treebanks, Linguistic Theories and Applications Multilingual Treebanks: the Universal Dependencies Case", In 30th European Summer School in Logic, Language and Information, 2018.

[23] U Thaung Lwin, Advanced New Method Myanmar Grammar) Second Edition, New Method Book Store (နည်းသစ်စာအုပ်တိုက်), 2015.

[24] A. Kutuzov, "Dependency Parsing (Project A) ", In INF5830, Fall 2017,University of Oslo, 2017

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 790-800 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Advanced Multiple Linear Regression Based Dark Channel Prior Applied on Dehazing Image and Generating Synthetic Haze

Binghan Li[*,1], Yindong Hua[2], Mi Lu[1]

[1]*Department of Electrical & Computer Engineering, Texas A & M University, College Station, TX, 77840, USA*

[2]*Department of Electrical & Computer Engineering, Stony Brook University, Stony Brook, NY, 11794, USA*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 02 March, 2021
Online: 31 March, 2021

Keywords:
Haze Removal
Dark Channel Prior
Multiple Linear Regression
Convolutional Neural Network
Object Detection
Synthetic Haze

ABSTRACT

Haze removal is an extremely challenging task, and object detection in the hazy environment has recently gained much attention due to the popularity of autonomous driving and traffic surveillance. In this work, the authors propose a multiple linear regression haze removal model based on a widely adopted dehazing algorithm named Dark Channel Prior. Training this model with a synthetic hazy dataset, the proposed model can reduce the unanticipated deviations generated from the rough estimations of transmission map and atmospheric light in Dark Channel Prior. To increase object detection accuracy in the hazy environment, the authors further present an algorithm to build a synthetic hazy COCO training dataset by generating the artificial haze to the MS COCO training dataset. The experimental results demonstrate that the proposed model obtains higher image quality and shares more similarity with ground truth images than most conventional pixel-based dehazing algorithms and neural network based haze-removal models. The authors also evaluate the mean average precision of Mask R-CNN when training the network with synthetic hazy COCO training dataset and preprocessing test hazy dataset by removing the haze with the proposed dehazing model. It turns out that both approaches can increase the object detection accuracy significantly and outperform most existing object detection models over hazy images.

1 Introduction

This paper is an extension of work initially presented in conference name [1]. Computer vision has recently played a major role in broad applications on urban traffic, such as autonomous and assisted driving, traffic surveillance, and security maintenance. However, the existence of haze, mist, dust, and fumes can severely degrade the visibility of images captured outside. Haze generates reduced contrasts, fainted surfaces, and color distortion to outdoor scenes, which will inevitably complicate many advanced computer vision tasks, including object classification and segmentation. Since the depth information of haze is non-linear and dependent over a global scene, haze removal becomes a challenging task. Most computer vision algorithms are designed based upon haze-free input images. They benefit a lot from haze removal, making it a highly desired task in computational photography and computer vision applications.

Many algorithms have been proposed to restore clear images from hazy images. Polarization-based methods presented in [2]–[3] analyze the polarization effects of atmospheric scattering and remove haze through as few as two images with different degrees of polarization. Depth-based approaches introduced in [4], [5] evaluate depth information upon some assumptions or priors, then estimate transmission map $t(x)$ and atmospheric light A from it. Some recent CNN-based haze-removal models proposed in [6]–[7] are built upon various powerful CNNs to self-learn transmission map $t(x)$ directly from large-scale image datasets. Among effective conventional dehazing algorithms, Dark Channel Prior (DCP) [8] is generally accepted due to its novel prior and outstanding performance. In most non-sky patches of the haze-free image, at least one color channel contains dark pixels with extremely low intensity, which is primarily generated by the air light. However, the estimation on the medium transmission $t(x)$ and atmospheric light A is not precise, especially when the scene object is inherently similar to the air light over a large local region and no shadow is cast on it. And the restored image looks unnaturally dark when there is a sky region with sunlight.

In this paper, the authors propose a novel Multiple Linear Regression Dark Channel Prior based model (MLDCP). Trained with

[*]Corresponding Author: Binghan Li, Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, 77840, USA, Email: libinghan1994@outlook.com,

the training dataset in REalistic Single Image DEhazing (RESIDE) [9], the MLDCP model can optimize the rough estimation of transmission map $t(x)$ and atmospheric light A by self-learning. The authors show experimentally on RESIDE test dataset that their model achieves the highest SSIM and PSNR values (two important full-reference metrics) compared with DCP and some other well-known state-of-the-art dehazing algorithms and CNN-based architectures. The authors further evaluate the effect on object detection in the hazy environment when dehazing the test images by their MLDCP model. Besides, the experimental results demonstrate that MLDCP not only enhances the performance of object detection in the hazy environment but also outperforms most dehazing algorithms on this task with higher detection accuracy.

The authors also present a straightforward and flexible algorithm to generate synthetic haze to any existing image datasets, inspired by a reversed MLDCP model. The authors aim to enhance object detection performance in the hazy environment by utilizing synthetic hazy images as training datasets. In the experiment, this algorithm is applied to MS COCO training dataset [10] by adding synthetic haze to the images and build a new Hazy-COCO training dataset. The authors evaluate the mean average precision (mAP) of Mask R-CNN [11], a widely adopted object detection and segmentation model, by training the network with the Hazy-COCO training dataset. The experimental results indicate that it leads to an impressive improvement when preprocessing training datasets with the inverse MLDCP algorithm.

2 Related Work

2.1 Overview of Dehazing Algorithms

2.1.1 Background Knowledge

In computer vision, the widely used atmospheric scattering model to describe the generation of a hazy image is as follows:

$$\boldsymbol{I}(x) = \boldsymbol{J}(x)t(x) + \boldsymbol{A}(1 - t(x)) \tag{1}$$

where $I(x)$ is the observed intensity (hazy image), $J(x)$ is the scene radiance (haze-free image), $t(x)$ is the medium transmission map, and A is the atmospheric light. The first term $J(x)t(x)$ is called attenuation and the second term $A(1 - t(x))$ is called airlight [12].

The medium transmission map $t(x)$ describes the portion of the light that is not scattered and reaches the camera [8]. When the atmosphere is homogeneous, the transmission matrix $t(x)$ can be defined as:

$$t(x) = e^{-\beta d(x)} \tag{2}$$

where β is the scattering coefficient of the atmosphere, and $d(x)$ is the scene depth representing the distance between the object and camera.

Most state-of-the-art single image dehazing algorithms exploit the atmospheric scattering model (1) and estimate the transmission matrix $t(x)$ and the atmospheric light A in either physically grounded or data-driven ways. Then the haze-free images $J(x)$ can be recovered by computing the reformulation of (1):

$$\boldsymbol{J}(x) = \frac{1}{t(x)}\boldsymbol{I}(x) - \boldsymbol{A}\frac{1}{t(x)} + \boldsymbol{A} \tag{3}$$

2.1.2 Conventional Single Image Dehazing Algorithms

Haze removal is a challenging task due to the non-linear and dependent depth maps over a global scene in hazy images. Many efforts have been made to tackle this challenge by exploiting natural images priors and depth statistics. Most conventional dehazing algorithms focus on predicting two critical parameters, medium transmission matrix $t(x)$ and global atmospheric light A, which are necessary to recover haze-free images via computing (3). In [12], an automated method is proposed based on the observation that the contrast of a haze-free image is higher than that of a hazy image. Furthermore, a Markov Random Fields (MRFs) framework is implemented to estimate the atmospheric light A by maximizing the local contrast of a hazy image. The output results are visually impressive but may not be physically valid. Assuming that the transmission and surface shading is uncorrelated in local areas, the authors eliminate the scattered light by locally estimating the optical transmission map of the scene with constant constraints in [13]. Despite its compelling results, it may fail in the cases with heavy haze and lead to the inaccurate estimation of color and depth maps.

A widely recognized single image dehazing algorithm called Dark Channel Prior (DCP) is proposed in [8], which can estimate the transmission map $t(x)$ more reliably. A regular pattern is found that in most non-sky patches of haze-free images, at least one color channel (dark channel) has some pixels whose intensity is very low and even close to zero. Then this pixel-based observation can be formally described by defining the dark channel J^{dark} as:

$$\boldsymbol{J}^{dark}(x) = \min_{y\in\Omega(x)}(\min_{c}\boldsymbol{J}^{c}(y)) \approx 0 \tag{4}$$

where c indicates RGB color channels and y refers to the pixel in a local patch $\Omega(x)$ centered at x. Adding minimum operators to both sides of the transformation of (1):

$$\min_{y\in\Omega(x)}(\min_{c}\frac{\boldsymbol{I}^{c}(y)}{\boldsymbol{A}^{c}}) = \widetilde{t}\min_{y\in\Omega(x)}(\min_{c}\frac{\boldsymbol{J}^{c}(y)}{\boldsymbol{A}^{c}}) + 1 - \widetilde{t} \tag{5}$$

Transmission map $\widetilde{t}$ can be put outside of the minimum operators based on the fact that $\widetilde{t}$ is a constant in the patch.

Since the dark channel of a haze-free image can be approximately taken as 0, the multiplicative term in (5) can be eliminated by adding (4). Then transmission map $\widetilde{t}$ can be predicted by:

$$\widetilde{t} = 1 - \omega\min_{y\in\Omega(x)}(\min_{c}\frac{\boldsymbol{I}^{c}(y)}{\boldsymbol{A}^{c}}) \tag{6}$$

The additional parameter ω is a constant parameter that optionally controls the degree of haze removal. Even in a haze-free image, the haze still exists among distant objects. A small amount of haze will keep the vision perceptual natural with the sense of depth. The dehazing parameter determines how much haze will be removed.

In case that the recovered scene radiance $J(x)$ is prone to noise when transmission map $t(x)$ is extremely low, DCP restricts $t(x)$ by a lower bound t_0, which is set to 0.1 in [8]:

$$\widetilde{t} = \max(t(x), t_0) \tag{7}$$

As for the estimation of atmospheric light A, it is defined as the color of the most haze-opaque regions in [12], which refers to the brightest pixels in a hazy image. However, this assumption only

applies when there is no sunlight in local regions. This limitation is optimized in [8] by considering the sunlight and adopting the dark channel to detect the most haze-opaque. DCP picks the top 0.1 percent brightest pixels in the dark channel, among which the pixels with the highest intensity in an input image are selected as the atmospheric light.

2.1.3 Limitations of DCP

DCP has some limitations that it may fail to accurately estimate transmission map $t(x)$ and atmospheric light A, when object surfaces are essentially analogous to the air light over a local scene without any projected shadows. Although DCP in [8] takes sunlight into consideration, the influence of sunlight is still tremendous when there is strong sunlight in the sky region. It will underestimate the transmission map of these objects and overestimate the haze layer. Thus the brightness of the restored image is darker than the real-world haze-free image. The authors compare a group of hazy images and recovered images from DCP, the color distortion in the sky region can be observed obviously in Figure 1:



(a) Real-world Hazy image (b) Dehazed image via DCP

(c) Real-world Hazy image (d) Dehazed image via DCP

Figure 1: Limitations of DCP

Additionally, the constant parameter ω in (6) is fixed to 0.95 in [8] without any changes corresponding to different haze distributions. Excessive haze removal will create color distortion, and images with insufficient haze removal will remain blurred. In order to solve this problem, the authors propose a dehazing parameter adaptive method in [14] based on DCP that estimates dehazing parameter ω locally instead of globally. It can automatically adjust the value of ω according to the distribution of haze. DCP in [8] implements soft matting to optimize the estimation of transmission map $t(x)$, which is further enhanced to be more accurate and efficient by utilizing an explicit image filter called *guided filter* in [15]. In [16], the authors point out that traditional DCP has not fully exploited its potential and will generate undesirable artifacts due to inappropriate assumptions or operations. Then it introduces a novel method that estimates transmission map $t(x)$ by energy minimization. The energy function combines DCP with piecewise smoothness and obtains an outstanding performance compared to conventional pixel-based dehazing algorithms. Several attempts have been made in [17], [18] to process the color distortion and optimize the restoration in some bright regions, such as the sky and reflective surfaces.

2.1.4 Overview of CNN-based Dehazing Algorithms

In recent years, neural network has made significant progress in numerous computer vision tasks [19], [20] and natural language processing tasks [21], [22]. Various Convolutional Neural Networks (CNNs) are designed to obtain a more accurate estimation of transmission matrix $t(x)$ by self-learning the mapping between hazy images and corresponding transmission maps, which outperform most conventional dehazing algorithms. DehazeNet in [23] is an end-to-end system built upon a deep convolutional neural network whose layers are specially designed to embody established priors in haze removal. Furthermore, a novel non-linear activation function is executed to improve the quality of output recovered images. A multi-scale deep neural network is proposed in [7] to dehaze a single image, which consists of a coarse-scale stage that roughly predicts the transmission matrix $t(x)$ over a global view, and a fine-scale stage that refines the rough estimation locally. In [24], the DCP energy function is defined as the loss function in a fully-convolutional dilated residual network. Feeding the network with real-world outdoor images, it minimizes the loss function completely unsupervised during the training process. A light-weight CNN model called All-in-One Dehazing Network (AOD-Net) in [6] is designed based on a reformulated atmospheric scattering model. It generates the recovered images directly and can be widely embedded into other deep CNN models to enhance the performances of some high-level tasks over hazy images.

2.1.5 Dehazing Benchmark Datasets and Metrics

Traditional haze-removal algorithms used to evaluate and compare dehazing performances by merely presenting a group of hazy images and dehazed images restored from various dehazing algorithms. The enhancement of dehazing performance is expected to be observed from the visual comparison of images. However, it is not convincing to prove that a new dehazing algorithm outperforms other algorithms only from human eyes perception. Two widely adopted image metrics to evaluate and compare single image dehazing algorithms are PSNR and SSIM [25]. PSNR refers to the peak-signal-to-noise ratio, which is generally applied to evaluate the image quality. SSIM refers to the structural similarity index measure, a well-known metric to measure the similarity between two images. Since it is generally impossible to capture the same visual scene with and without haze, while all other environmental conditions stay identical, it is incredibly challenging to measure the SSIM value between a hazy image and its haze-free ground truth

image. Therefore, recent efforts have been made to create synthetic hazy images from haze-free images based on the depth information.

In [26], the authors build two sets of images without haze and with synthetic haze from both real-world camera captured scenes and synthetic scenes to evaluate the performance of the proposed visibility enhancement algorithm. They utilize the software to generate 66 synthetic images built upon a physically-based road environment. By obtaining the depth information, four different fog types are added to 10 camera images and finally create a dataset with over 400 synthetic images in total. A fog simulation is proposed in [27] by simulating the underlying mechanism of hazy image formation (1) and utilizing the standard optical model for daytime haze. The fog simulation pipeline is leveraged to add synthetic fog to urban scenes images in Cityscapes dataset [28] and generate a Foggy Cityscapes dataset. Foggy Cityscapes dataset consists of 550 refined high-quality synthetic foggy images with detailed semantic annotations and additional 20000 synthetic foggy images without sufficient annotations.

The REalistic Single Image DEhazing (RESIDE) dataset [9] is the first large-scale dataset for benchmarking single image dehazing algorithms, and it includes both indoor and outdoor hazy images. RESIDE dataset also contains a large-scale synthetic training set and two sets designed respectively for objective and subjective quality evaluations. Moreover, in the supplementary RESIDE-β set, they add annotations and object bounding boxes to an evaluation set consisting of 4322 real-world hazy images, which can be utilized to test the performance of object detection in the hazy environment. A rich variety of criteria beyond PSNR and SSIM is also provided in [9] to evaluate the performance of dehazing algorithms, including full-reference metrics, no-reference metrics, subjective evaluation, and task-driven evaluation. However, the criteria are only practically applicable to the global performance of haze-removal. It cannot embody the difference locally between two images, failing to judge if our MLDCP model outperforms traditional DCP specifically in the bright region. Therefore, the authors compare dehazing performances experimentally with both recovered images' visual quality and two pivotal metrics PSNR and SSIM.

2.2 Overview of Object Detection Models

Object detection is the combination of object classification and localization, which can both recognize and localize all object instances of specific categories in an image. Due to its close relationship with image and video analysis, object detection has been widely applied in various computer vision tasks, especially in autonomous driving, traffic surveillance, and some other smart city fields.

Fast R-CNN refers to Fast Region-based Convolutional Network [29]. The network is fed with an image and a set of object proposals and outputs a convolutional feature map. A region of interest (RoI) pooling layer is proposed to extract a fixed-length feature vector and feed it into a sequence of fully connected layers. The output layers contain a softmax layer that estimates the softmax probability over K object classes plus a background class, and another layer with offset values that refine the bounding box positions of an object. Fast R-CNN overcomes the disadvantages of R-CNN and SPPnet [30] with a higher detection accuracy as well as faster training and test speed. Later in [31], a Region Proposal Network (RPN) is added into Fast R-CNN generating a Faster R-CNN. The RPN aims to simultaneously propose candidate object bounding boxes and corresponding scores at each position. It then implements RoIPool in Fast R-CNN to extract features from each candidate box and perform classification as well as bounding-box regression. Both RPN and Fast R-CNN are trained independently but share the same convolutional layers. While achieving state-of-the-art object detection accuracy, the unified network can further increase the speed significantly.

Mask R-CNN [11] extends the Faster R-CNN by an output branch with a binary mask for each RoI, in parallel to a branch for bounding-box recognition and classification. However, the quantization of RoIPool in Faster R-CNN has a negative effect on predicting binary masks. To remove this harsh quantization, Mask R-CNN replaces RoIPool with a RoIAligh layer that aligns the extracted features with the input properly. It can efficiently detect object instances in an image and simultaneously generate a pixel-accurate segmentation mask for each instance. Due to its compelling performance and influential architecture, Mask R-CNN is widely used as a solid baseline to exploit more object detection tasks.

2.3 Domain Adaption Methods

Domain adaptation is a novel strategy that can be utilized to advance object detection models [32]–[33]. And it has been proved effective, especially in some extreme weather [34], [35], such as hazy, rainy, and snowy. A Domain Adaptive Faster R-CNN is proposed in [34] to enhance the cross-domain robustness of object detection. Based on H-divergence theory, two domain adaptation components on image level and instance level are integrated into Faster R-CNN architecture, aiming to reduce the domain discrepancy at both levels. Training data with images and full supervision is used as the source domain, and only unlabeled images in test data are available for the target domain. For both components, it adapts the classifier trained on a source domain and implements the adversarial training strategy to learn domain-invariant features. The method further incorporates a consistency regularization into the Faster R-CNN model to obtain a domain-invariant region proposal network (RPN).

Inspired by Domain Adaptive Faster R-CNN in [34], the authors in [36] adopt a similar idea and designs a Domain-Adaptive Mask-RCNN (DMask-RCNN). The source domain takes the clean images in the MS COCO dataset, and the target domain takes unannotated real-world hazy images in RESIDE dataset [9], and their dehazed output images by MSCNN [7] respectively. And DMask-RCNN adds a domain-adaptive branch after the base feature extraction layers in Mask R-CNN architecture, aiming to mask the generated features to be domain-invariant between the source domain and target domain. The experimental results in [34], [36] demonstrate that the domain adaptation method can enhance the performance of both Faster R-CNN and Mask R-CNN models when tackling the object detection task in the hazy environment. Moreover, this enhancement can be more effective when feeding the target domain with images restored by a robust dehazing algorithm.

2.4 Object Detection Datasets

Object recognition is a core task to understand a visual scene in computer vision, which involves several sub-tasks, including image classification, object detection, and semantic segmentation. All three tasks have high demands for image datasets. Object classification requires each image to be labeled with a sequence of binary numbers, which indicate if object instances exist in the image or not. Object detection is more challenging, which combines object classification and localization. It not only identifies which specific class an object belongs to but also locates it in an image. The object localization requires collecting bounding boxes that locate one or more objects in an image, which is a considerable workload in a large-scale image dataset. The PASCAL VOC challenge [37] is a widely recognized benchmark in visual object recognition and detection. The VOC2007 dataset contains 20 object categories spread over 11000 images. The annotation procedure is designed to be consistent, accurate, and exhaustive, and annotations are made available for training, validation, and test data. The ImageNet dataset [38] involves millions of cleanly labeled and full-resolution images in the hierarchical structure of WordNet [39]. The dataset provides an average of 500-1000 images to illustrate each synset in a majority of 80000 synsets of WordNet. Many object recognition algorithms have made a significant breakthrough by using the training resource of the ImageNet dataset. This benchmark dataset also plays an essential role in advancing object recognition tasks.

Semantic segmentation requires labeling each pixel in an image to a category, and it is an extremely time-consuming task to build a large-scale dataset with detailed semantic scene labeling. The Microsoft COCO dataset [10] consists of about 3.3 million images, and over 2 million of them are labeled. It collects 91 common object categories, which are fewer than the ImageNet dataset. However, in contrast to both the VOC2007 dataset and ImageNet dataset, MS COCO dataset contains significantly more instances per category and considerably more object instances per image. And every instance of each object category is fully labeled and segmented by a novel instance-level segmentation mask. MS COCO dataset has been widely utilized for training some more complex CNN architectures that aim to make further progress on object recognition as well as semantic segmentation tasks.

3 Proposed Methods

3.1 Multiple Linear Regression DCP Model

DCP estimates transmission map $t(x)$ from (6) and selects the pixels in the dark channel with the highest intensity among the top 0.1 percent brightest pixels as the atmospheric light A. Then hazy images can be recovered from (3) motivated by the estimations of $t(x)$ and A. However, the rough estimations in DCP introduced in Section 2.1.3 can generate some unpredicted deviations, which are theoretically impossible to be eliminated during the estimation process. Now the authors implement a multiple linear regression model to optimize the haze-removal algorithm in DCP.

Multiple linear regression is a statistical technique using several explanatory variables to predict the output response variable. As a predictive analysis, multiple linear regression aims to model the linear relationship between two or more independent variables and a continuous dependent variable. The authors still adopt the rough estimations on transmission map $t(x)$ and atmospheric light A in DCP. Three components $\frac{\boldsymbol{I}(x)}{t(x)}$, $\frac{A}{t(x)}$ and A in (3) can be regarded as explanatory variables, since $\boldsymbol{I}(x)$ refers to the pixel of input hazy image while $t(x)$ and A can be estimated from DCP. The scene radiance $\boldsymbol{J}(x)$, which refers to the pixel of output recovered image, can be regarded as the response variable. The authors add regression coefficient weights to each explanatory variable and a constant term (bias) to the atmospheric scattering model. Then (3) can be reformulated as a multiple linear regression model (8), which describes how the mean response $\boldsymbol{J}(x)$ changes with explanatory variables:

$$\boldsymbol{J}(x) = \omega_0 \frac{\boldsymbol{I}(x)}{t(x)} + \omega_1 \frac{\boldsymbol{A}}{t(x)} + \omega_2 \boldsymbol{A} + b \tag{8}$$

As the authors introduced, DCP is already an effective dehazing algorithm even with the rough estimations on $t(x)$ and A. The authors implement a multiple linear regression model to optimize DCP by learning the relationships between hazy images with haze-free images, which is extremely challenging to find out in traditional pixel-based dehazing algorithms. Since both $\boldsymbol{J}(x)$ and A are defined and estimated on RGB color channels, the dimensions of three weights and a bias should be (3,1), intending to refine the parameters at all three color channels. The Outdoor Training Set (OTS) in RESIDE dataset [9] provides 8970 outdoor haze-free images, each of which also contains 30 synthetic hazy images with haze intensity ranging from low to high. During the training process, the network learns the relationships between output recovered images and their relevant haze-free ground truth images. However, it is theoretically impossible to obtain the ground truth image of a real-world hazy image. The most efficient approach is referring to a real-world haze-free image as ground truth image and its synthetic hazy image as input image respectively. The various intensity of synthetic haze guarantees that the MLDCP model can be applied to more than a fixed haze intensity in real-world scenes.

When MLDCP model is trained on OTS dataset, the authors refer to haze-free images as target ground truth images $\boldsymbol{J}$, refer to synthetic hazy images as input images $\boldsymbol{I}$, and refer to the recovered images as output images $\boldsymbol{J}_\omega$. In order to simplify the formula (8) during the training process, the authors assign $\boldsymbol{x}_0 = \frac{\boldsymbol{I}(x)}{t(x)}$, $\boldsymbol{x}_1 = \frac{A}{t(x)}$ and $\boldsymbol{x}_1 = A$. Then (8) can be reformulated as:

$$\boldsymbol{J}_\omega(x) = \omega_0 \boldsymbol{x}_0 + \omega_1 \boldsymbol{x}_1 + \omega_2 \boldsymbol{x}_2 + b \tag{9}$$

The deviations between the output images $\boldsymbol{J}_\omega$ and the target ground truth images $\boldsymbol{J}$ are estimated by mean-squared error (MSE):

$$\boldsymbol{MSE} = (\boldsymbol{J} - \boldsymbol{J}_\omega)^2 \tag{10}$$

MSE measures the average squared difference between an observation's actual and predicted values, which is commonly adopted in linear regression models. The loss function of MLDCP is evaluated based upon MSE:

$$\boldsymbol{L}(\omega) = \frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{J}^{(i)} - \boldsymbol{J}_\omega^{(i)})^2 \tag{11}$$

During the training process, the authors calculate the gradient of loss function by Stochastic Gradient Descent (SGD), a widespread

and effective technique to address optimization tasks. Then the optimal weights and bias can be obtained by minimizing MSE through the following steps:

- Implement SGD to compute the derivative of loss function to find the gradient of the error generated from current weights and bias, and adjusts the weights by moving in the direction opposite of the gradient to decrease MSE. Then authors iteratively update three weights via:

$$\omega_k = \omega_k - \alpha \frac{\partial}{\partial \omega} \boldsymbol{L}(\omega) \tag{12}$$

  where α is the learning rate, and $k \in (0, 1, 2)$.

- Expand the derivative of loss function (11), then (12) can be reformulated as:

$$\omega_k = \omega_k - \alpha \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{J}^{(i)} - \boldsymbol{J}_{\omega}^{(i)}) x_k \tag{13}$$

- Update the bias with the same process repeatedly by training each image in OTS dataset via:

$$b = b - \alpha \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{J}^{(i)} - \boldsymbol{J}_{\omega}^{(i)}) \tag{14}$$

When the optimal weights and bias are obtained from the training process, they can be added to the haze-removal model (8) directly, and real-world hazy images can be dehazed in the same way as that in conventional DCP. The improved performance is determined by the quality of synthetic hazy images. If the synthetic haze is more close to the real-world haze, the MLDCP model can learn more realistic haze information during the training process, and the dehazing performance can be further enhanced.

3.2 Synthetic Haze Generated Model

There have been some large-scale image datasets released for the object detection task. However, few image datasets are built just for object detection in extreme weather like hazy and rainy. The Real-world Task-driven Testing Set (RTTS) in RESIDE [9] contains 4322 real-world hazy images annotated with both object categories and bounding boxes, which is an essential contribution taken as a test dataset for object detection task in the hazy environment. However, even with data augmentation techniques, its amount of training images is far less efficient in training an object detection model in comparison with the larger scale of the MS COCO training dataset [10]. It means a considerable workload to build a large-scale image dataset with detailed annotations and segmentation instances like MS COCO or PASCAL VOC. And the rarity of real-world hazy images, especially in urban cities, makes it more challenging to collect the same amount of high-quality hazy images as those in the MS COCO training dataset.

The authors propose an algorithm that generates synthetic haze to any existing large-scale haze-free image datasets without much computation. Inspired by the multiple linear regression haze-removal model, the authors propose a new algorithm that implements a similar inverse MLDCP model to add synthetic haze to an image. If an object detection model is trained on the synthetic hazy dataset generated by this algorithm, its performance of detecting an object in hazy weather can be significantly enhanced.

Since DCP is an effective dehazing algorithm, and the multiple linear regression model can further exploit its potential, the authors suppose that similar ideas can also be applied to synthesize higher quality hazy images based on the regular dark channel pattern. Coefficient weights and bias are added to (1), which is expected to maximally restore the deviations between haze-free images and their relevant synthetic hazy images. Then the reformulated atmospheric scattering model is obtained as follows:

$$\boldsymbol{I}(x) = \beta_0 \boldsymbol{J}(x) t(x) + \beta_1 \boldsymbol{A} t(x) + \beta_2 \boldsymbol{A} + b \tag{15}$$

The synthetic haze generated model is still trained on OTS in RESIDE dataset [9]. In contrast to MLDCP model, the authors refer to haze-free images as input images $\boldsymbol{J}$ and refer to synthetic hazy images in OTS as target images $\boldsymbol{I}$. And the output synthetic hazy images are taken as $\boldsymbol{I}_\beta$. The cost function is defined based on MSE that evaluates the deviations between the output images $\boldsymbol{I}_\beta$ and target images $\boldsymbol{I}$:

$$\boldsymbol{L}(\beta) = \frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{I}^{(i)} - \boldsymbol{I}_{\beta}^{(i)})^2 \tag{16}$$

The authors implement SGD during the training process to travel down the slope of cost function until it reaches the bottom lowest value, and three weights and a bias are updated iteratively by calculating the derivative of cost function as follows:

$$\beta_k = \beta_k - \alpha \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{I}^{(i)} - \boldsymbol{I}_{\beta}^{(i)}) x_k \tag{17}$$

$$b = b - \alpha \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{I}^{(i)} - \boldsymbol{I}_{\beta}^{(i)}) \tag{18}$$

where $k \in (0, 1, 2)$ and x_k equals to $\boldsymbol{J}(\mathbf{x})t(x)$, $At(x)$, and A in (15) respectively.

In fact, traditional DCP does not support the inverse functionality of generating synthetic haze to a haze-free image. Self-learning the haze information from input synthetic hazy images, the MLDCP model fully utilizes the superiority of DCP to generate a much higher quality synthetic haze with a more similar visual and inherent effect to real-world haze. Additionally, the inverse MLDCP model can add various intensities of haze to an image dataset, making it possible to train an object detection model on training datasets with different synthetic haze intensities corresponding to the density of real-world haze in test images. Since the method focuses on the enhancement by preprocessing the training dataset, it can be applied simultaneously with any improvements on the object detection models or preprocessing images in test datasets with more effective dehazing algorithms.

4 Experiment Results and Analysis

4.1 MLDCP Dehazing Performance

4.1.1 Experiment Setup

In the experiment, both training and test datasets are obtained from RESIDE dataset [9]. In the training process, 8970 haze-free images

and their corresponding synthetic hazy images in the OTS dataset are utilized for training the MLDCP model. In the test phase, the Synthesis Object Testing Set (SOTS) is used to test the dehazing performance. Since MLDCP focuses on recovering outdoor hazy scenes, the authors do not evaluate its dehazing performance over the 500 synthetic indoor images in SOTS. The dehazing performance of MLDCP is compared with conventional DCP on both SSIM and PSNR metrics over 500 outdoor synthetic hazy images and their haze-free ground truth images.

4.1.2 *Dehazing Performance on SOTS Dataset*

In Figure 2, the authors present the comparison of recovered images by conventional DCP and our MLDCP model. It is obviously observed that the recovered image by DCP (c) suffers from the color distortion in the sky region. In contrast, the recovered image by MLDCP (d) performs competently even in the sky region with strong sunlight. And the non-sky region in (c) is much darker than that in (d), and it looks unreal. In comparison with the input synthetic hazy image (a), MLDCP almost removes the haze completely. Besides, its recovered image (d) looks natural and shares a high similarity with the ground truth haze-free image (b).



(a) Synthetic hazy image



(b) Haze-free image



(c) Recovered image by DCP



(d) Recovered image by our model

Figure 2: Comparison on synthetic hazy image. (a) is a synthetic hazy image in SOTS, (b) is the haze-free ground truth image, (c) is recovered by conventional DCP algorithm, (d) is recovered by our MLDCP model.

The authors compare the dehazing performance of MLDCP with several conventional dehazing algorithms as well as some CNN-based haze removal models over two major image benchmark metrics PSNR and SSIM. PSNR evaluates the quality of recovered images, while SSIM measures the similarity between recovered images and haze-free ground truth images. In Table 1, MLDCP model increases PSNR and SSIM by 5.3 and 0.23 respectively in comparison with the conventional DCP. Compared with the performances of other widely adopted dehazing algorithms obtained from [9], the proposed MLDCP model achieves a reasonably acceptable PSNR value and the highest SSIM value. Considering that SSIM is essentially more dominant than PSNR when indicating the pixel-wise effect of haze removal, MLDCP outperforms all other dehazing algorithms in Table 1.

Table 1: Average SSIM and PSNR comparison among different dehazing algorithms over 500 outdoor synthetic hazy images in SOTS.

500 outdoor images		
Dehazing method name	**PSNR**	**SSIM**
Improved DCP model	23.84	**0.9411**
DCP	18.54	0.7100
FVR	16.61	0.7236
CAP	23.95	0.8692
NLD	19.52	0.7328
BCCR	17.71	0.7409
GRM	20.77	0.7617
DehazeNet	**26.84**	0.8264
MSCNN	21.73	0.8313
AOD-Net	24.08	0.8726

4.1.3 *Real-world Hazy Images Dehazing*

Even though synthetic benchmark datasets and metrics have achieved great success in comparing dehazing performances, the visual effect of synthetic haze is still different from the real-world haze. In Figure 3, the authors also present the recovered images by conventional DPC and MLDCP from real-world nature hazy scenes (first three rows) and hazy scenes in urban cities (last three rows). It is obviously observed that the MLDCP model performs better than conventional DCP in both sky and non-sky regions. The output images restored by MLDCP prevent the color distortion in sky regions, and the brightness over a global view seems more natural and closer to realistic haze-free scenes.

4.2 *Object Detection with Pre-dehazed Test Dataset*

4.2.1 *Experiment Setup*

Some object detection tasks on RESIDE RTTS dataset have been tested in [6], [9], [36]. The Domain-Adaptive Mask R-CNN model proposed in [36] achieves the highest detection accuracy on RTTS test dataset among several well-known object detection models including Faster R-CNN [31], Mask R-CNN [11], SSD [40] and RetinaNet [41]. A concatenation of dehazing algorithm and object detection modules is proposed in [6] to detect objects in the hazy environment. In [36], the authors further try more combinations of effective dehazing algorithms and object detectors in the cascade and evaluate their performances on the mean average precision (mAP) values. MLDCP model is utilized to dehaze the RTTS test dataset, and Mask R-CNN and DMask-RCNN modules are fed with the recovered images. Then the authors compare the detection accuracy with the results of dehazing-detection cascades in [36].

In Table 2, the proposed cascade of MLDCP dehazing model and Mask R-CNN or Mask R-CNN increases the detection accuracy by about 2% or 1.7% respectively, which is the highest in comparison with other dehazing approaches, including AOD-Net [6], MSCNN [7] and conventional DCP [8]. And pre-dehazing RTTS dataset with MSCNN and DCP can both enhance the performance of object

detection in the hazy environment. In [6], AOD-Net outperforms several dehazing algorithms on PSNR and SSIM values. However, it decreases the mAP result of DMask R-CNN from 61.72% to 60.47%, which means that a better dehazing performance on PSNR and SSIM values does not align with a higher accuracy on object detection in the hazy environment.

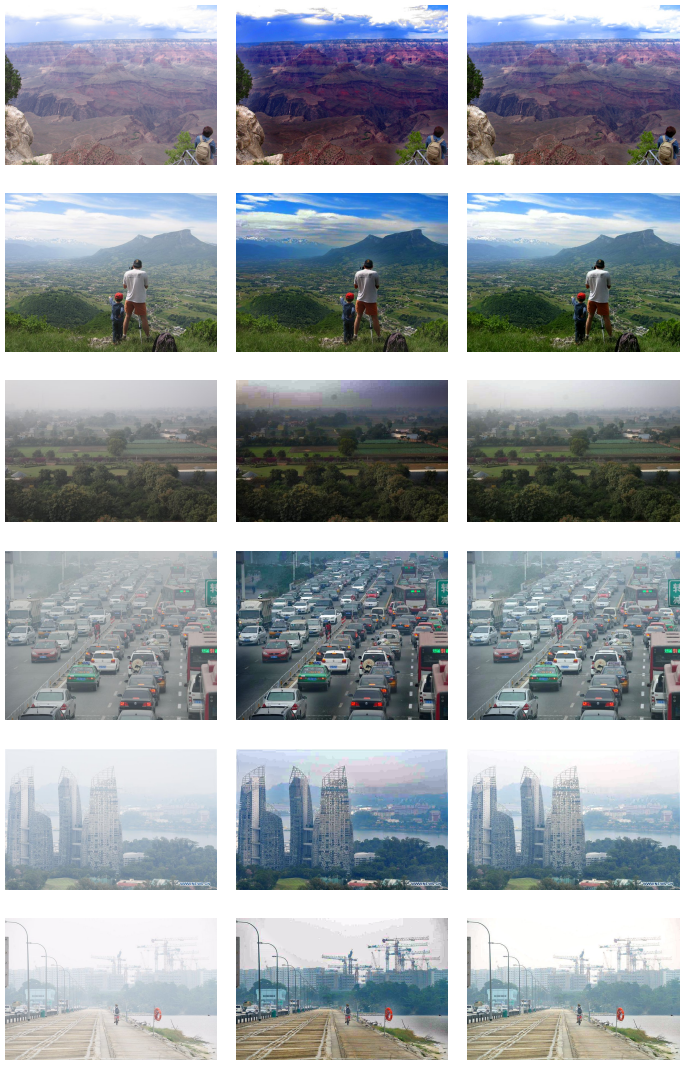


Figure 3: Visual comparison between real-world hazy images (first column), recovered images from DCP (second column) and recovered images from MLDCP (third column).

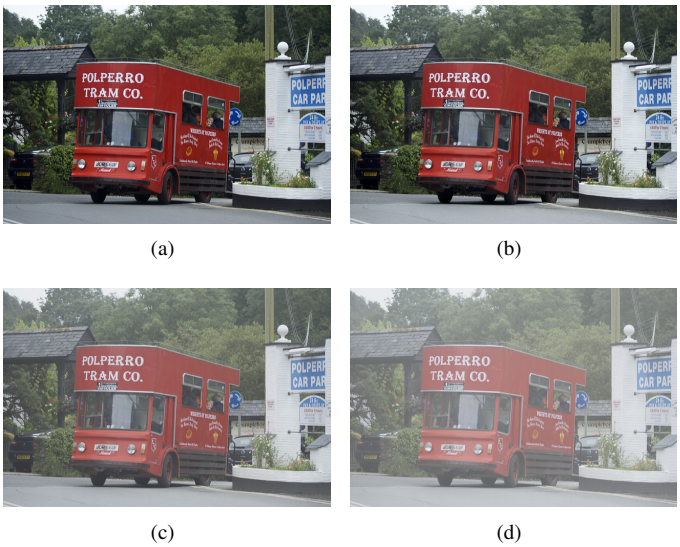
Table 2: Object detection accuracy (mAP) comparison among different dehazing-detection cascades

Framework	mAP(%)
Mask R-CNN	61.01
DMask R-CNN2	61.72
MLDCP + Mask R-CNN	**63.06**
AOD-Net + DMask R-CNN2	60.47
MSCNN + DMask R-CNN2	63.36
DCP + DMask R-CNN2	**62.78**
MLDCP + DMask R-CNN2	**63.42**

4.3 Object Detection with Hazy-COCO Dataset

4.3.1 Experiment Setup

The authors train Mask R-CNN on the new Hazy-COCO training dataset and evaluate its performance in comparison with the original Mask R-CNN trained on the MS COCO training dataset. During the training process, the backbone network is set as resnet101, and the training starts with the pre-trained COCO weights. The authors also tried various combinations of training stages to check the effect on its detection performance. It turns out that when training the head layers by ten epochs with a learning rate of 0.001 and fine-tuned layers from Resnet stage 4 and up by 60 epochs with a learning rate of 0.0001, the authors obtained the weights and bias with the highest detection accuracy on both RTTS and UG2 test datasets.



(a) (b)

(c) (d)

Figure 4: Comparison on synthetic hazy image. (a) is a haze-free image in MS COCO, (b) is the synthetic hazy image generated by inverse DCP, (c) is the synthetic hazy image with 0.1 haze density generated by our algorithm, (d) is the synthetic hazy image with 0.2 haze density generated by our algorithm.

Table 3: Object detection accuracy (mAP) comparison between MS COCO and Hazy-COCO training datasets

Test Dataset	Training Dataset	mAP(%)
RTTS	**COCO**	61.01
	Hazy COCO	**66.08**
RTTS+MLDCP	**COCO**	63.06
	Hazy COCO	**66.15**
UG2	**COCO**	32.07
	Hazy COCO	**38.53**

4.3.2 Experiment Results with Hazy-COCO Training Dataset

In Table.3, the authors evaluate the difference in detection performances when Mask R-CNN is pre-trained on MS COCO and Hazy-COCO datasets. The mAP results increase significantly by about 5% and 6% on RTTS and UG2 test datasets respectively, which presents how effective the Hazy-COCO training dataset is. Additionally, when training Mask R-CNN on the Hazy-COCO dataset, dehazing



Figure 5: Visual comparison between haze-free outdoor images from MS COCO dataset (left column) and the synthetic hazy images with 0.2 haze density generated by our algorithm (right column).

the RTTS test dataset by the MLDCP model can simultaneously increase the mAP result by 0.07%. However, this increment is much lower than the increment 2% when we preprocess the RTTS dataset individually. This is caused by the overlaps of enhancement effects between MLDCP and the advanced inverse DCP since they share some similarities over the dark channel regular pattern and multiple linear regression model.

4.3.3 Visual Effect of Synthetic Haze

Generally, traditional DCP does not support generating synthetic haze to a haze-free image by simply implementing the dark channel regular pattern inversely. This limitation becomes possible after the authors apply the multiple linear regression techniques to the atmospheric scattering model.

In Figure 4, the effect of haze generated by the inverse DCP algorithm can be barely observed in (b), in comparison with the haze-free image (a). Synthetic hazy images created by our advanced inverse DCP algorithm with haze density 0.1 and 0.2 are presented as (c) and (d) respectively. Both of the two synthetic hazy images are covered with seemingly realistic haze. And it is obviously observed that the synthetic haze in (d) is denser than the haze in (c). The authors also present four more groups of haze-free images from the

MS COCO dataset and their corresponding synthetic hazy images in Figure 5 to demonstrate the effectiveness of our algorithm.

5 Conclusion

As autonomous driving and traffic surveillance become widespread in smart cities, object detection in extreme environments like hazy has been paid special attention to. This paper proposed the approaches that increase the detection accuracy in the hazy environment from both dehazing test dataset and preprocessing training dataset.

First, an advanced Multiple Linear Regression Haze-removal Model was proposed, aiming to overcome the deficiencies of Dark Channel Prior. The authors implemented Stochastic Gradient Descent to update and find the optimal weights and bias to refine the rough estimation of two essential parameters transmission matrix $t(x)$ and atmospheric light A. The experimental results showed that MLDCP not only achieved higher PSNR and SSIM values than other state-of-the-art dehazing algorithms and CNN-based dehazing models, but also increased the detection precision by a higher rate when concatenated with object detection models. It demonstrated that sometimes exploiting practical conventional dehazing algorithms by machine learning techniques was superior to building more complicated neural networks.

Second, the authors proposed an inverse DCP algorithm formulated on the multiple linear regression model that could generate synthetic haze to any existing image datasets. The authors expected this technique to prevent from spending excessive time and workload on building a large-scale image dataset with synthetic or real-world haze, as well as detailed annotations for object detection and segmentation. Synthetic haze was added to the MS COCO dataset and trained Mask R-CNN on this Hazy-COCO training set. The experimental results presented a significant increase in detection accuracy in the hazy environment. However, this approach could only generate the average synthetic haze with the same density over all pixels in an image. On the contrary, real-world haze density normally varies from pixel to pixel. In future work, it is expected to optimize this limitation and generate more realistic synthetic haze.

References

[1] B. Li, W. Zhang, M. Lu, "Multiple Linear Regression Haze-removal Model Based on Dark Channel Prior," in 2018 International Conference on Computational Science and Computational Intelligence (CSCI), 307–312, IEEE, 2018, doi:10.1109/CSCI46756.2018.00066.

[2] Y. Y. Schechner, S. G. Narasimhan, S. K. Nayar, "Instant dehazing of images using polarization," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, volume 1, I–I, IEEE, 2001, doi:10.1109/CVPR.2001.990493.

[3] M. Lu, et al., Arithmetic and logic in computer systems, volume 169, Wiley Online Library, 2004, doi:10.1002/0471728519.

[4] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," ACM transactions on graphics (TOG), **27**(5), 1–10, 2008, doi:10.1145/1409060.1409069.

[5] S. G. Narasimhan, S. K. Nayar, "Interactive (de) weathering of an image using physical models," in IEEE Workshop on color and photometric Methods in computer Vision, volume 6, 1, France, 2003, doi:10.1109/ICIECS.2009.5365298.

[6] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, "An all-in-one network for dehazing and beyond," arXiv preprint arXiv:1707.06543, 2017.

[7] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in European conference on computer vision, 154–169, Springer, 2016, doi:10.1007/s11263-019-01235-8.

[8] K. He, J. Sun, X. Tang, "Single image haze removal using dark channel prior," IEEE transactions on pattern analysis and machine intelligence, **33**(12), 2341–2353, 2010, doi:10.1109/TPAMI.2010.168.

[9] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, Z. Wang, "Benchmarking single-image dehazing and beyond," IEEE Transactions on Image Processing, **28**(1), 492–505, 2018, doi:10.1109/TIP.2018.2867951.

[10] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, "Microsoft coco: Common objects in context," in European conference on computer vision, 740–755, Springer, 2014, doi:10.1007/978-3-319-10602-1_48.

[11] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision, 2961–2969, 2017, doi:10.1109/ICCV.2017.322.

[12] R. T. Tan, "Visibility in bad weather from a single image," in 2008 IEEE Conference on Computer Vision and Pattern Recognition, 1–8, IEEE, 2008, doi:10.1109/CVPR.2008.4587643.

[13] R. Fattal, "Single image dehazing," ACM transactions on graphics (TOG), **27**(3), 1–9, 2008, doi:10.1145/1360612.1360671.

[14] C. Chen, J. Li, S. Deng, F. Li, Q. Ling, "An adaptive image dehazing algorithm based on dark channel prior," in 2017 29th Chinese Control And Decision Conference (CCDC), 7472–7477, IEEE, 2017, doi:10.1109/CCDC.2017.7978537.

[15] K. He, J. Sun, X. Tang, "Guided image filtering," IEEE transactions on pattern analysis and machine intelligence, **35**(6), 1397–1409, 2012, doi:10.1007/978-3-642-15549-9_1.

[16] M. Zhu, B. He, Q. Wu, "Single image dehazing based on dark channel prior and energy minimization," IEEE Signal Processing Letters, **25**(2), 174–178, 2017, doi:10.1109/LSP.2017.2780886.

[17] J.-H. Kim, W.-D. Jang, J.-Y. Sim, C.-S. Kim, "Optimized contrast enhancement for real-time image and video dehazing," Journal of Visual Communication and Image Representation, **24**(3), 410–425, 2013, doi:10.1016/j.jvcir.2013.02.004.

[18] L. Shi, X. Cui, L. Yang, Z. Gai, S. Chu, J. Shi, "Image Haze Removal Using Dark Channel Prior and Inverse Image," in MATEC Web of Conferences, volume 75, 03008, EDP Sciences, 2016, doi:10.1051/matecconf/20167503008.

[19] Y. Hua, Y. Liu, B. Li, M. Lu, "Dilated Fully Convolutional Neural Network for Depth Estimation from a Single Image," in 2019 International Conference on Computational Science and Computational Intelligence (CSCI), 612–616, IEEE, 2019, doi:10.1109/CSCI49370.2019.00115.

[20] X. Wei, Y. Zhang, Z. Li, Y. Fu, X. Xue, "DeepSFM: Structure from motion via deep bundle adjustment," in European conference on computer vision, 230–247, Springer, 2020, doi:10.1007/978-3-030-58452-8_14.

[21] Y. Wang, X. Zhang, M. Lu, H. Wang, Y. Choe, "Attention augmentation with multi-residual in bidirectional LSTM," Neurocomputing, **385**, 340–347, 2020, doi:10.1016/j.neucom.2019.10.068.

[22] Y. Wang, Y. Huang, W. Zheng, Z. Zhou, D. Liu, M. Lu, "Combining convolutional neural network and self-adaptive algorithm to defeat synthetic multi-digit text-based CAPTCHA," in 2017 IEEE International Conference on Industrial Technology (ICIT), 980–985, IEEE, 2017, doi:10.1109/ICIT.2017.7915494.

[23] B. Cai, X. Xu, K. Jia, C. Qing, D. Tao, "Dehazenet: An end-to-end system for single image haze removal," IEEE Transactions on Image Processing, **25**(11), 5187–5198, 2016, doi:10.1109/TIP.2016.2598681.

[24] A. Golts, D. Freedman, M. Elad, "Unsupervised single image dehazing using dark channel prior loss," IEEE Transactions on Image Processing, **29**, 2692–2701, 2019, doi:10.1109/TIP.2019.2952032.

[25] A. Hore, D. Ziou, "Image quality metrics: PSNR vs. SSIM," in 2010 20th international conference on pattern recognition, 2366–2369, IEEE, 2010, doi: 10.1109/ICPR.2010.579.

[26] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, D. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," IEEE Intelligent Transportation Systems Magazine, **4**(2), 6–20, 2012, doi:10.1109/MITS.2012.2189969.

[27] C. Sakaridis, D. Dai, L. Van Gool, "Semantic foggy scene understanding with synthetic data," International Journal of Computer Vision, **126**(9), 973–992, 2018, doi:10.1007/s11263-018-1072-8.

[28] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in Proceedings of the IEEE conference on computer vision and pattern recognition, 3213–3223, 2016, doi:10.1109/CVPR.2016.350.

[29] R. Girshick, "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision, 1440–1448, 2015, doi:10.1109/ICCV.2015.169.

[30] K. He, X. Zhang, S. Ren, J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE transactions on pattern analysis and machine intelligence, **37**(9), 1904–1916, 2015, doi:10.1109/TPAMI.2015.2389824.

[31] S. Ren, K. He, R. Girshick, J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 91–99, 2015, doi:10.1109/TPAMI.2016.2577031.

[32] V. M. Patel, R. Gopalan, R. Li, R. Chellappa, "Visual domain adaptation: A survey of recent advances," IEEE signal processing magazine, **32**(3), 53–69, 2015, doi:10.1109/MSP.2014.2347059.

[33] R. Gopalan, R. Li, R. Chellappa, "Domain adaptation for object recognition: An unsupervised approach," in 2011 international conference on computer vision, 999–1006, IEEE, 2011, doi:10.1109/ICCV.2011.6126344.

[34] Y. Chen, W. Li, C. Sakaridis, D. Dai, L. Van Gool, "Domain adaptive faster r-cnn for object detection in the wild," in Proceedings of the IEEE conference on computer vision and pattern recognition, 3339–3348, 2018, doi: 10.1109/CVPR.2018.00352.

[35] Z. Chen, X. Li, H. Zheng, H. Gao, H. Wang, "Domain adaptation and adaptive information fusion for object detection on foggy days," Sensors, **18**(10), 3286, 2018, doi:10.3390/s18103286.

[36] Y. Liu, G. Zhao, B. Gong, Y. Li, R. Raj, N. Goel, S. Kesav, S. Gottimukkala, Z. Wang, W. Ren, et al., "Improved techniques for learning to dehaze and beyond: A collective study," arXiv preprint arXiv:1807.00202, 2018, doi: arXiv:1807.00202v2.

[37] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, A. Zisserman, "The pascal visual object classes (voc) challenge," International journal of computer vision, **88**(2), 303–338, 2010, doi:10.1007/s11263-009-0275-4.

[38] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition, 248–255, Ieee, 2009, doi:10.1109/CVPR.2009.5206848.

[39] G. A. Miller, WordNet: An electronic lexical database, MIT press, 1998, doi:10.7551/mitpress/7287.001.0001.

[40] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, "Ssd: Single shot multibox detector," in European conference on computer vision, 21–37, Springer, 2016, doi:10.1007/978-3-319-46448-0_2.

[41] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, "Focal loss for dense object detection," in Proceedings of the IEEE international conference on computer vision, 2980–2988, 2017, doi:10.1109/ICCV.2017.324.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 801-807 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Dilated Fully Convolutional Neural Network for Depth Estimation from a Single Image

Binghan Li[*,1], Yindong Hua[2], Yifeng Liu[3], Mi Lu[1]

[1]*Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, 77840, USA*

[2]*Department of Electrical & Computer Engineering, Stony Brook University, Stony Brook, NY, 11794, USA*

[3]*School of Business, Stevens Institute of Technology, Hoboken, NJ, 07030, USA*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 04 March, 2021
Online: 31 March, 2021

Keywords:
Depth Prediction
CNN
Dilated Convolutions

ABSTRACT

Depth prediction plays a key role in understanding a 3D scene. Several techniques have been developed throughout the years, among which Convolutional Neural Network has recently achieved state-of-the-art performance on estimating depth from a single image. However, traditional CNNs suffer from the lower resolution and information loss caused by the pooling layers. And oversized parameters generated from fully connected layers often lead to a exploded memory usage problem. In this paper, we present an advanced Dilated Fully Convolutional Neural Network to address the deficiencies. Taking advantages of the exponential expansion of the receptive field in dilated convolutions, our model can minimize the loss of resolution. It also reduces the amount of parameters significantly by replacing the fully connected layers with the fully convolutional layers. We show experimentally on NYU Depth V2 datasets that the depth prediction obtained from our model is considerably closer to ground truth than that from traditional CNNs techniques.

1 Introduction

This paper is an extension of work originally presented in conference name [1]. Depth prediction has always been a core task to understand the geometric relations within a 3D scene. It provides rich information about the distance of the objects in the image from the viewpoint of camera. This technique is necessary for many applications in computer vision including smoothing blurred parts of an image [2], [3], rendering of 3D scenes [4], virtual reality, self-driving cars [5], grasping in robotics [6] and autopilot [7]. However, predicting depth from images is a quite complex and challenging task. In absence of the environmental assumptions, the inherent ambiguity of mapping an intensity or color measurement into a depth value makes depth prediction an ill-posed problem. Many unique techniques have been proposed to tackle this problem, such as superpixels based algorithms [8], Structure-from-Motion (SfM) [9], data-driven methods [10] and CNN based approaches [11].

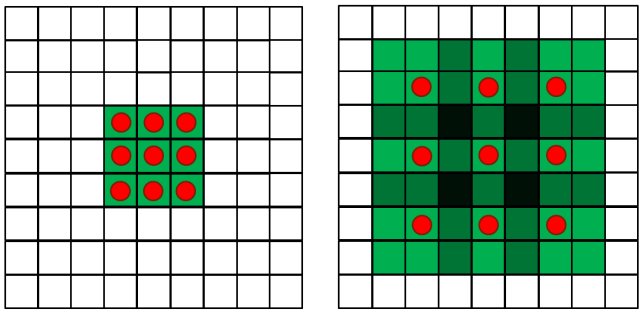
Neural Network has been widely applied on computer vision tasks and natural language processing tasks [12]–[14]. And Convolutional Neural Network (CNN) has achieved a great success on outperforming many state-of-the-art algorithms over object classification and detection [15], semantic segmentation [16], scene reconstruction, and face recognition [17]. Recently, depth prediction can also be addressed by CNNs due to the power that the ambiguous mapping between a single image and depth maps can be modeled via learning in the neural network. [11] increases the output resolution by efficiently learning the feature map up-sampling within the fully convolutional residual networks. The Multiple-Scale Deep Neural Network in [18] improves the depth prediction by estimating the global scene structure and refining it using local information. Three different computer vision tasks including depth prediction are addressed in [19] with a single multi-scale convolutional network architecture. [20] proposes a novel training loss in the CNN architecture, which enforces consistency between left and right depth maps, to perform end-to-end unsupervised single image depth estimation.

Despite Convolutional Neural Networks performs well on depth estimation, there are still some deficiencies remained to be improved. Some depth predicting CNN models still use pooling layers to extend receptive field. Pooling layers provide an effective approach to reduce the dimensionality of the network by summarizing the presence of features in patches of the feature map. However, they inevitably lose a lot of valuable information during the down sam-

[*]Corresponding Author: Binghan Li, Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, 77840, USA, Email: libinghan1994@outlook.com,

pling process. In contrast, fully connected layers will obtain the global relationship between pixels and image, and inherit all the combinations of the features from the previous layer. This will generate too many parameters, making fully connected layers incredibly computationally expensive.

Dilated convolutional neural network has been proved to be more effective than traditional CNNs in [16], [21].The feature of dilated convolutions is illustrated in Figure 1. Instead of contiguous pooling filters in traditional CNNs, dilation imposes filters that have spaces between each node. The dilated convolutions support exponential expansion of the receptive field without loss of resolution or coverage, which reduces the computation and memory costs. While preserving the dimensions of data at the output layer, dilated convolutions also maintain the ordering of data.



(a) 1-dilated convolution (b) 2-dilated convolution

Figure 1: Examples of dilated convolutions with receptive field. (a) refers to an 1-dilated convolution with a kernel of a 3x3 receptive field, which is inherently a standard convolution. (b) refers to a 2-dilated convolutions with a kernel of a 7x7 receptive field. The amount of parameters corresponding to each layer stays identical.

In this paper, a fully dilated convolutional neural network is proposed to enhance the performance of depth prediction. Our model is designed based on the framework in [18]. It contains two components: the global coarse-scale network, which estimates the coarse depth, and the local fine-scale network, which refines the coarse depth estimation combined with local information. In the coarse-scale network, the convolutional layers are replaced with the dilated convolutional layers and the fully connected layers are replaced with the convolutional layers. We train and evaluate our network on NYU Depth V2 datasets and make comparison on the depth predicting performances between our model with traditional VGG-16 model. The experiment results demonstrate that our model reduces the computational and memory costs, and achieves state-of-the-art performance on benchmarks.

2 Related Work

2.1 Depth Estimation

Depth estimation refers to a set of techniques and algorithms aiming to obtain the spatial information from a scene, which has been an essential computer vision task with a long history. Normally a 2D image taking from the camera is not able to represent the local spatial relations of a 3D scene. Since only one point of each pixel is projected in the real scene, its depth information is mathematically eliminated when projected into plane in a image. Depth information is a key prerequisite to perform multiple computer vision tasks. With the depth information, we are able to back project images captured from multiple views and the 3D scenes can be perfectly restructured by matching all the points. In order to accurately move the actuators in robotics, the depth estimation is required to multiple tasks such as perception, navigation, and planning. In the single image dehazing task, scene depth is a key parameter that supports to remove the haze locally instead of globally.

Depth estimation is an inherently ambiguous and complex task. Geometrically, infinite points in the scene are not projected, thus the depth information may be generated from considerable possible world scenes. For some specific computer vision tasks benefited from depth prediction, it will face more challenges. In the application of autonomous driving in particular, depth estimation will be degraded by occlusion, dynamic object in the scene and imperfect stereo correspondence. Despite the loss of depth information in the 3D dimension, depth map prediction has been investigated inspired by the analogy to how human eyes perceive depth information from depth cues.

2.2 State-of-the-art Algorithms

Humans estimate depth by comparing the images obtained from left and right eyes. Our eyes perceive and recognize the depth cues of the scene, then our brains will subconsciously analyze and recover the depth information easily. There are basically 4 categories of depth cues: Static monocular, depth from motion, binocular and physiological cues [22], making it possible to construct the spatial arrangement of objects in the scene. In computer vision, depth information is extracted mainly from monocular images and stereo images by exploiting epipolar geometry. And many efforts have been made to tackle the challenge of predicting depth from stereo vision depth estimation and monocular depth estimation.

Stereo vision is inspired by how human eyes calculate the approximate depth map from the minor difference between both viewpoints. It compares two differing views on a scene and predicts the relative depth information from the displacement in horizontal coordinates of corresponding image points. The local disparity can be obtained effortlessly using local appearance features. By contrast, estimating depth information from a single image requires a global view of the scene, which is one of the reasons why the monocular depth estimation has not been solved to the same degree as the stereo vision approach [18]. In [23], Scharstein summarizes a taxonomy and evaluation of existing dense two-frame stereo correspondence algorithms, including some earliest stereo vision based depth estimation algorithms [24], [25] with impressive performances. [26] proposes a probabilistic deep learning approach to model disparity and generate binocular training data to estimate model parameters, which outperforms state-of-the-art algorithms with fewer requirements for global detailed information of the scene.

Structure-from-Motion (SfM) is one of the most successful state-of-the-art techniques of depth estimation [27]. SfM estimates the camera motion from the relative pairwise camera positions of the extracted features. Then it predicts the depth information via triangulation from pairs of consecutive views, and recovers the 3D structure through the spatial and geometric relationship of objects in

the scene [9]. Despite its advantages over the lower cost and the less restricted environment, SfM has not been widely adopted in commercial applications due to its complex theories and the difficulty to further enhance the accuracy and speed.

Several techniques address the depth estimation task based on superpixels. In [28], the author proposes an algorithm to find an oversegmentation of the image that breaks up the planar surfaces into many small patches, which are named as superpixels. For each homogeneous superpixel in the image. In [4], the author uses a Markov Random Field (MRF) to infer depth information through a set of plane coefficients, which extract both the 3D location and orientation in the image. The MRF is trained via supervised learning to learn how different depth cues are associated with different depths. In [4], the author further extends the model by combining triangulation cues and monocular images cues, which supports to restructure a full and photorealistic 3D model of a larger scene. There are also some superpixels based algorithms deploying the Conditional Random Fields (CRFs) for the regularization of depth information. Deriving from the CRFs in [8], the author formulates monocular depth estimation as a discrete-continuous optimization problem. The continuous variables encode the depth of the superpixels in the input image, and the discrete variables represent the relationships between neighboring superpixels. In [29], the author extracts multi-scale image patches around the superpixel center and learn to encode the correlations between input patches and corresponding depths regressively with a deep CNN. Then it refines the depth estimation from the superpixel level to pixel level by using CRFs.

2.3 CNN Based Algorithms

A CNN-based depth estimation from a single image is a challenging task if without the local correspondences. A CNN model needs to self-learn the pixel-wise local details as well as the correlations between a pixel and the global scene during the training process.

In [18], the author presents a novel network to regress dense depth maps by implementing a deep network with two stages: global coarse-scale network and local fine-scale network. The coarse-scale network is initialized based on AlexNet [30] with five feature extraction layers, including convolution layers and max-pooling layers, and two fully connected layers followed. The network is able to integrate the understanding of a global view by making effective use of depth cues to predict the coarse depth information. The fine-scale network consists of convolutional layers and one pooling stage for the first layer edge features. Aligning with local details in the scene, the fine network fine-tune the coarse prediction by concatenating an additional low-level feature map.

Building upon the two-scale CNN architecture in [18], [19] authors a paper about predicting depth, surface normals and semantic labels with an improved multi-scale convolutional network. It replaces the Alexnet network with a deeper VGG-16 network [31]. Generating pixel-maps directly from an input image, this network can also align to many image details by using a sequence of convolutional network stacks applied at increasing resolution, without the need for low-level superpixels or contours.

A novel fully convolutional network incorporated with efficient residual up-sampling blocks is proposed in [11] to model the ambiguous mapping between monocular images and depth maps. It can output the depth maps with higher resolution, at the same time reduce the amount of parameters and train on one order of magnitude fewer data. This paper further proposes a scheme for up-convolutions and combine it with the concept of residual learning to create up-projection blocks for the effective up-sampling of feature maps, which has been proved to be more applicable when addressing high-dimensional regression tasks.

Fully connected layers in some CNN architectures will generate considerable parameters, causing several problems like slower training time, much memory consumption, and chances of overfitting. Some networks in [32], [33] replace fully connected layers with convolutional layers, which can decrease the image matrix to a lower dimension and reduce the amount of parameters. However, convolutional layers are not capable to encode the position and orientation of objects and lack the ability to be spatially invariant to the input data. Thus the output accuracy is much lower than that of fully connected layers. Some further improvements have been proposed to compensate the loss like the convolutional residual networks [11] we introduced above and the dilated convolution we will discuss in next section.

2.4 Dilated Convolution

The dilated convolution is a type of convolution that expands the kernel by inserting holes between the kernel elements. The standard convolution operator is defined in [21] as:

$$(F * k)(\mathbf{p}) = \sum_{\mathbf{s}+\mathbf{t}=\mathbf{p}} F(\mathbf{s})k(\mathbf{t}) \tag{1}$$

The dilated convolution operator has been referred to as convolution with a dilated filter. We refer to l as a dilation factor and the l-dilated convolution operator $*_l$ can be defined as:

$$(F *_l k)(\mathbf{p}) = \sum_{\mathbf{s}+l\mathbf{t}=\mathbf{p}} F(\mathbf{s})k(\mathbf{t}) \tag{2}$$

When $l = 1$, the discrete convolution is simply the 1-dilated convolution. The dilation factor l should be increased exponentially at each layer when building a network with multiple dilated convolution layers. At the same time, the number of parameters associated with each dilated convolution layer is identical. That's the reason why dilated convolution can significantly reduce the amount of parameters.

Dilated convolution is widely adopted to enhance the performance in computer vision tasks. The multi-scale context architecture with dilated convolutions presented in [21] has increased the precision of the advanced semantic segmentation models effectively. In [16], the author adopts dilated convolution in deep convolutional neural networks, which increases the dense computation of neural net responses and achieves a higher accuracy than state-of-the-art algorithms at semantic segmentation task. A network for congested scene recognition (CSRNet) in [34] is easily trained by replacing pooling layers with dilated kernels. And the results demonstrate that CSRNet improves the output performance significantly with lower mean absolute error (MAE) than state-of-the-art algorithms.

The Alexnet and VGG-16 models apply the pooling layers to downsample the input images and simultaneously extend the receptive field. However, this process generally results in a loss of pixel-wise details. Dilated convolutions can avoid the similar resolution degradation issue but obtain the same computation as pooling layers in Alexnet and VGG-16 networks. Derived from the application of dilated convolutions on semantic segmentation, a similar technique is implemented to tackle the depth estimation task. Dilated convolutions hold the superiority with exponentially expanding of the receptive fields, which supports obtaining the global relationships among pixels in an image without the resolution reduction when decreasing the amount of parameters.

3 Dilated Fully CNN Architecture

3.1 Overview of the Method

Our dilated fully CNN architecture designed for depth estimation is presented in Fig. 2. It is built upon the multi-scale deep network in [18], which incorporates two stages: the global coarse-scale network and local fine-scale network.

The upper component (Stack 1) is the global coarse-scale network, which is similar to the VGG-16 network and designed to predict the coarse depth information. The convolutional layers and the fully connected layers in VGG-16 are replaced with the dilated convolutions and convolutional layers respectively. The coarse stage contains dilation layers and fully connected layers (FCN). The dilation layers apply 3x3 convolutions with different dilation factors. The dilations are 1, 2, 3, 2, 3 and 4. Rather than reducing the feature map sizes by convolutional layers and implementing fully connected layers to learn details over the local scene, our network can remain the same feature map sizes with much fewer parameters and simultaneously obtain an overall view of the scene from fully connected layers without the resolution loss.

The bottom component (Stack 2) is taken as the local fine-scale network. While the coarse stage captures the global scene, we also need to obtain local information in the refined stage. The inputs images will pass a 9x9 convolutional layer with pooling. Then its output and the low-level feature maps output from coarse stage will be concatenated.

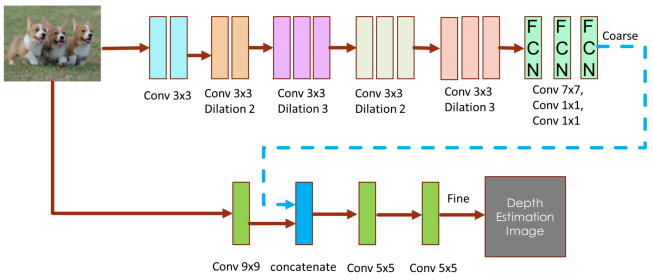


Figure 2: The architecture of two-stages dilated fully convolutional network

The architecture of the proposed dilated fully CNN with structural parameters is demonstrated in Table 1. In the coarse stage of VGG-16 network, the original frames of size 160x120 pixels are down-sampled by each convolutional layer. The front-end module that provides the input to the coarse stage of our network produces feature maps at 80x60 resolution. Our architecture remains the same resolution in the whole framework by replacing the convolutional layers with dilation convolutions. Table 1 demonstrates that the dilated convolution layers are particularly suited to coarse prediction due to its advantage of expanding the receptive field without the resolution reduction.

Table 1: Module architecture with the image size (size), the convolutional layer number (conv), the channel number (chan), the kernel size, and the dilation layer number in each layer.

	Layer	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	upsamp
Stack 1 (VGG)	size	160x120	80x60	40x30	20x15	10x7	1x1	1x1	1x1	80x60
	conv	2	2	3	3	3	-	-	-	-
	chan	64	128	256	512	512	4096	4096	4800	1
	ker.sz	3x3	3x3	3x3	3x3	3x3	-	-	-	-
Stack 1 (OUR)	size	80x60	80x60	80x60	80x60	80x60	80x60	80x60	80x60	-
	conv	2	2	3	3	3	1	1	1	-
	chan	64	128	256	512	512	512	512	1	-
	ker.sz	3x3	3x3	3x3	3x3	3x3	7x7	1x1	1x1	-
	dilation	1	2	3	2	3	4	-	-	-
	Layer	2.1	2.2	2.3	2.4					final
Stack 2	size	80x60	80x60	80x60	80x60					80x60
	conv	1	-	1	1					-
	chan	63	-	64	1					1
	ker.sz	9x9	-	5x5	5x5					-

3.2 Training Loss

Due to the ambiguity of the scale over a global scene in depth estimation, majority of the error is generated when obtaining the average scale of a scene. In order to compensate this training loss, we adopt the scale-invariant mean squared error in [18] as our loss function. Regardless of the absolute global scale, the scale-invariant error is applied to measure correlation among local pixels in the scene, which is defined in as:

$$\boldsymbol{D} = \frac{1}{2n}\sum_{i=1}^{n}(\log y_i - \log y_i^* + \alpha(y, y^*))^2 \tag{3}$$

where y and y^* represent the estimated depth map and the ground truth respectively. Each of them contains n pixels that are indexed by i. And $\alpha(y, y^*)$ aims to minimize the error between given y and y^*, which is defined as:

$$\alpha(y, y^*) = \frac{1}{n}\sum_{i}(\log y_i^* - \log y_i) \tag{4}$$

For any predicted depth map y, e^{α} is the scale that best aligns it to the ground truth y^*. The scale-invariant means that the error won't change with any scalar multiples of y.

We refer to this scale-invariant error as our training loss function that is formulated as:

$$\boldsymbol{L} = \frac{1}{2n^2}\sum_{i,j}((\log y_i - \log y_j) - (\log y_i^* - \log y_j^*))^2 \tag{5}$$

where $i, j \in \{1, 2 \ldots n-1\}$.

The difference between the estimated depth map y and the ground truth y^* at pixel i is defined as:

$$d_i = \log y_i - \log y_i^* \tag{6}$$

Then (5) can be re-formulated as:

$$\boldsymbol{L} = \frac{1}{n}\sum_{i}\boldsymbol{d}_i^2 - \frac{1}{n^2}\sum_{i,j}\boldsymbol{d}_i\boldsymbol{d}_j \tag{7}$$

(7) measures the error from the relationships between the output pixel *i* and pixel *j*. And each pair of predicted pixels and its corresponding ground truth pixels should share a similar amount of the difference between pixels, which can further reduce the errors.

4 Experiment Setup

4.1 Datasets and Implementation

We setup the experiment based upon NYU Depth V2 datasets [35], which is one of the largest RGB-D image datasets for indoor scene reconstruction. The raw dataset consists of 1449 RGB images with detailed object labels and annotattions with physical relations. Comprising 464 different indoor scenes and classified by 26 scene classes, those images are captured from a variety of buildings in modern cities. NYU Depth V2 dataset is significantly larger and more diverse than another similar Kinect scene dataset - NYU indoor scene dataset, which has limited diversity with only 67 scenes. In order to reduce the chances of overfitting, we shuffle the entire dataset. During the training process, only 800 images of the raw distribution are required. Then we take 200 images to execute the validation test and test our network with 449 images.

4.2 Training Procedure

In our network, we are committed to the universality of the structure. Since the image dataset will not generate any additional processing, our proposed network can be simply applied on other datasets.

Our network is implemented on Keras library running on top of Tensorflow. We train the network in two phases and use the GPU acceleration to speed up the training. First, dilated convolution layers and fully connected layers in the coarse-scale network are pretrained on the NYU Depth V2 training dataset. During this process, the parameters stay identical. Second, the coarse stage outputs concatenated with the edge features of input images are referred as the input images of the fine-scale stage. The convolutional layers in the fine-scale network refine the coarse prediction by aligning it with the detailed information in a local scene.

To make a fair comparison with the performance of traditional VGG-16 framework, two frameworks are both trained on NYU Depth V2 dataset. And they generate output images with the same size, 80x60 for each. The size of input image in VGG-16 framework is 320x240 and the size of input image in our proposed dilated fully convolutional network is 80x60. We refer to stochastic gradient descent (SGD) as the optimizer with the mini-batch size of 16 in the experiment. And we set the learning rate as 0.1 and the momentum as 0.9 for both global coarse-scale stage and local fine-scale stage.

5 Experiment Results and Analysis

5.1 Parameters Comparison

Fig. 3 demonstrates the effectiveness of dilated convolutions on decreasing the amount of parameters. "Total params" in (a) and (c) represents the amount of total parameters of the coarse-scale stage based on VGG-16 network and our proposed network respectively. (a) and (d) present the total parameters in the whole framework of VGG-16 network and our proposed architecture. Since the coarse-scale network is more complex with more layers than the fine-scale network, majority of parameters are generated from the coarse-scale stage, which can be observed from the comparison between total parameters in (a) and (c).

```
coarse_8 (Dense)              (None, 4800)
activation_16 (Activation)    (None, 4800)
reshape_1 (Reshape)           (None, 60, 80)
=================================================
Total params: 197,966,336
```

(a)

```
model_1 (Model)                 (None, 60, 80)
reshape_1 (Concatenate)         (None, 60, 80)
=================================================
Total params: 27,823,425
```

(b)

```
fine_4_linear (Activation)      (None, 60, 80, 1)
fine_4_reshape (Reshape)        (None, 60, 80)
=================================================
Total params: 198,085,773
```

(c)

```
fine_4_linear (Activation)      (None, 60, 80, 1)
fine_4_reshape (Reshape)        (None, 60, 80)
=================================================
Total params: 27,942,862
```

(d)

Figure 3: Comparison on the scale of parameters generated in CNN. (a) and (c) present the number of total parameters in the coarse-scale stage (Stack 1) and the whole framework based on VGG-16 network, (b) and (d) present the number of total parameters in the coarse-scale stage (Stack 1) and the whole framework based on our proposed network.

The experimental results demonstrate that our proposed network achieves over seven times reduction of the parameters amount both in the coarse-scale stage and the whole framework compared with the conventional VGG-16 network. Thus our proposed network can liberate considerable computation resources during the training process, which benefits from the significant parameters reduction. Besides, the limited scale of parameters makes the proposed network a viable candidate to be applied to other embedded architectures.

5.2 Depth Estimation Results

We propose a dilated fully convolutional neural network which is designed upon the architecture in [18], and evaluate its depth estimation performance compared with the conventional VGG-16 network. The experimental results are presented in Fig. 4.

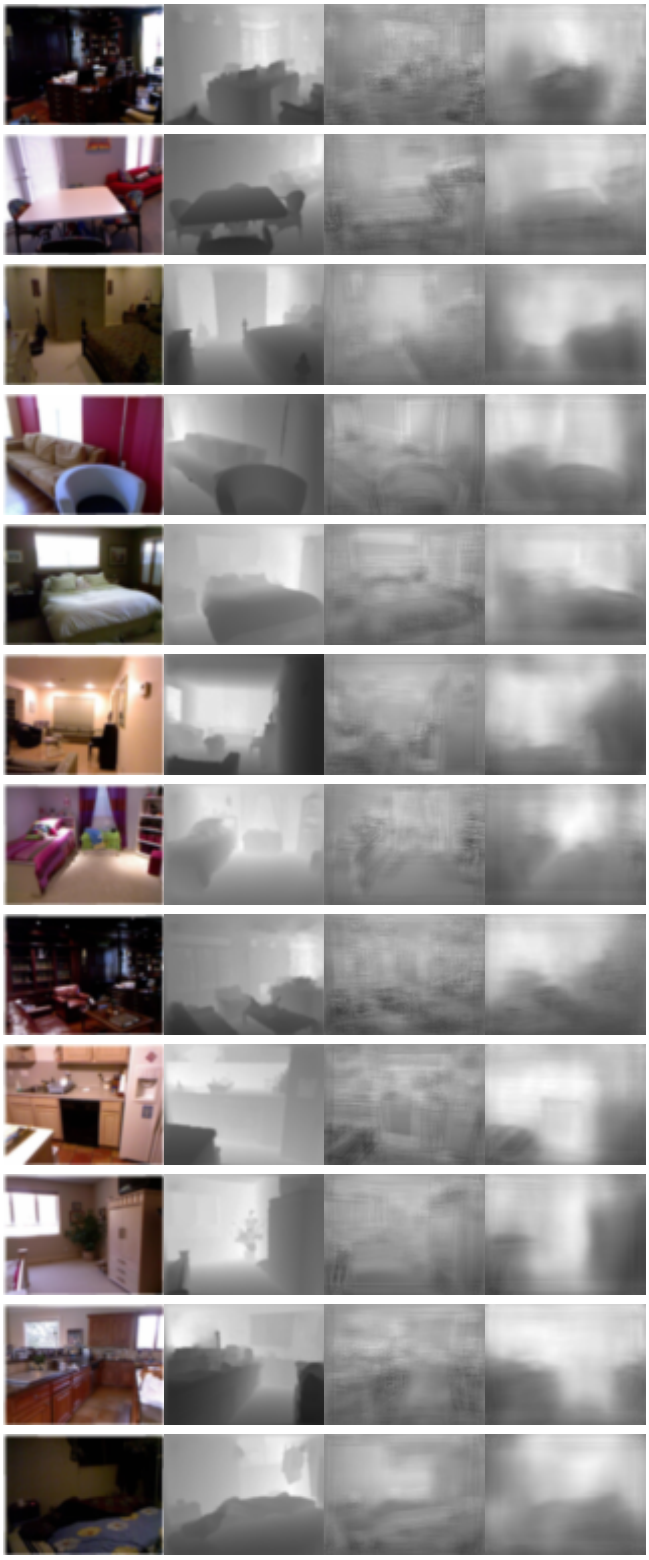


Figure 4: Example depth estimations. (a) and (b) represent the input images and the ground truth images respectively, (c) refers to the output depth estimation from VGG-16 network, (d) refers to the output depth estimation from our proposed network

The comparison between output images and ground truth images demonstrates that the dilated convolutions in our architecture can still obtain the related details between local pixels in a patch and the global view of a scene even without the fully connected layers.

Since fully connected layers learn all combinations of the features of the previous layers, the VGG-16 module achieves a good performance in some local spatial details. Comparing the predicted depth information based on VGG-16 module and our network, it is observed that our architecture can't perform as good as VGG-16 module when predicting depth maps around the object contours. The ambiguous pixels are generated during the training process, where majority of the output depth information will be blurred by the missing pixels around objects boundaries and reflective surfaces. Several techniques have been proposed to predict the boundary type of the superpixel edges, such as the discrete-continuous depth estimation approach in [8]. However, our model focuses on the enhancement over a global view of the scene and ignores the boundary performance. It's obviously observed that our network outperforms traditional VGG-16 network a lot globally, with much higher resolutions and more realistic representation of the relationships between objects and the environment.

6 Conclusion

In this paper, we proposed a dilated fully convolutional neural network to predict the depth information from a single image. The network is designed based on the multi-scale deep network in [18], which contains two stages: the coarse-scale network and the fine-scale network. The coarse stage predicts the global depth information, which is refined locally in the fine-tune stage. We replace the convolutional layers and fully connected layers in the coarse stage with dilated convolutions and convolutional layers respectively. By implementing the dilated convolutions, the network can reduce the amount of parameters significantly without the resolution reduction, which benefits from the exponential expansion of receptive fields in dilated convolutions. And the experiment results demonstrate that our proposed network achieves the state-of-the-art performance on depth estimation for NYU Depth V2 datasets. The output depth information outperforms VGG-16 network with higher resolutions and more realistic representation of the relationships between objects and the environment, while generating much fewer parameters and releasing more memory resources.

Predicting the depth information around the objects boundaries is a weak point of our proposed network. In future work, we will incorporate specific techniques into our network and effectively enhance the depth prediction performance around objects boundaries. Besides, our network only evaluates the visual performance on a single dataset, which is insufficient to demonstrate all the superiority. We plan to further evaluate the predicted depth information on several effective image criteria, which can represent the strong points that are not presented visually. And we will apply the dilated convolutions to more advanced CNNs to further improve the depth prediction performance.

References

[1] Y. Hua, Y. Liu, B. Li, M. Lu, "Dilated Fully Convolutional Neural Network for Depth Estimation from a Single Image," in 2019 International Conference on Computational Science and Computational Intelligence (CSCI), 612–616, IEEE, 2019, doi:10.1109/CSCI49370.2019.00115.

[2] K. He, J. Sun, X. Tang, "Single image haze removal using dark channel prior," IEEE transactions on pattern analysis and machine intelligence, **33**(12), 2341–2353, 2010, doi:10.1109/TPAMI.2010.168.

[3] B. Li, W. Zhang, M. Lu, "Multiple Linear Regression Haze-removal Model Based on Dark Channel Prior," in 2018 International Conference on Computational Science and Computational Intelligence (CSCI), 307–312, IEEE, 2018, doi:10.1109/CSCI46756.2018.00066.

[4] A. Saxena, M. Sun, N. A. Y, "Make3d: Learning 3d scene structure from a single still image," IEEE transactions on pattern analysis and machine intelligence, **31**(5), 824–840, 2009, doi:10.1109/TPAMI.2008.132.

[5] R. Hadsell, P. Sermanet, J. Ben, A. Erkan, M. Scoffier, K. Kavukcuoglu, U. Muller, Y. LeCun, "Learning long-range vision for autonomous off-road driving," Journal of Field Robotics, **26**(2), 120–144, 2009, doi:10.1109/IROS.2008.4651217.

[6] M. Ye, E. Johns, A. Handa, L. Zhang, P. Pratt, G.-Z. Yang, "Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery," arXiv preprint arXiv:1705.08260, 2017, doi:10.31256/HSMR2017.14.

[7] K. Tateno, F. Tombari, I. Laina, N. Navab, "Cnn-slam: Real-time dense monocular slam with learned depth prediction," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 6243–6252, 2017, doi:10.1109/CVPR.2017.695.

[8] M. Liu, M. Salzmann, X. He, "Discrete-continuous depth estimation from a single image," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 716–723, 2014, doi:10.1109/CVPR.2014.97.

[9] R. Szeliski, "Structure from motion," in Computer Vision, 303–334, Springer, 2011, doi:10.1007/978-1-84882-935-0_7.

[10] K. Karsch, C. Liu, S. B. Kang, "Depth extraction from video using non-parametric sampling," in European Conference on Computer Vision, 775–788, Springer, 2012, doi:10.1007/978-3-642-33715-4_56.

[11] I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari, N. Navab, "Deeper depth prediction with fully convolutional residual networks," in 2016 Fourth international conference on 3D vision (3DV), 239–248, IEEE, 2016, doi:10.1109/3DV.2016.32.

[12] Y. Wang, Y. Huang, W. Zheng, Z. Zhou, D. Liu, M. Lu, "Combining convolutional neural network and self-adaptive algorithm to defeat synthetic multi-digit text-based CAPTCHA," in 2017 IEEE International Conference on Industrial Technology (ICIT), 980–985, IEEE, 2017, doi:10.1109/ICIT.2017.7915494.

[13] Y. Wang, Z. Zhou, S. Jin, D. Liu, M. Lu, "Comparisons and selections of features and classifiers for short text classification," in IOP Conference Series: Materials Science and Engineering, volume 261, 012018, IOP Publishing, 2017, doi:10.1088/1757-899X/261/1/012018.

[14] Y. Wang, X. Zhang, M. Lu, H. Wang, Y. Choe, "Attention augmentation with multi-residual in bidirectional LSTM," Neurocomputing, **385**, 340–347, 2020, doi:10.1016/j.neucom.2019.10.068.

[15] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision, 2961–2969, 2017, doi:10.1109/ICCV.2017.322.

[16] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," arXiv preprint arXiv:1412.7062, 2014, doi:10.1109/TPAMI.2017.2699184.

[17] H. Jiang, E. Learned-Miller, "Face detection with the faster R-CNN," in 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 650–657, IEEE, 2017, doi:10.1109/FG.2017.82.

[18] D. Eigen, C. Puhrsch, R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," in Advances in neural information processing systems, 2366–2374, 2014.

[19] D. Eigen, R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in Proceedings of the IEEE international conference on computer vision, 2650–2658, 2015, doi:10.1109/ICCV.2015.304.

[20] C. Godard, O. Mac Aodha, G. J. Brostow, "Unsupervised monocular depth estimation with left-right consistency," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 270–279, 2017, doi:10.1109/CVPR.2017.699.

[21] F. Yu, V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv preprint arXiv:1511.07122, 2015.

[22] F. L. Kooi, A. Toet, "Visual comfort of binocular and 3D displays," Displays, **25**(2-3), 99–108, 2004, doi:10.1016/j.displa.2004.07.004.

[23] D. Scharstein, R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International journal of computer vision, **47**(1-3), 7–42, 2002, doi:10.1109/SMBV.2001.988771.

[24] Y. Boykov, O. Veksler, R. Zabih, "A variable window approach to early vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, **20**(12), 1283–1294, 1998, doi:10.1109/34.735802.

[25] M. Okutomi, T. Kanade, "A multiple-baseline stereo," IEEE Transactions on pattern analysis and machine intelligence, **15**(4), 353–363, 1993, doi:10.1109/34.206955.

[26] R. Memisevic, C. Conrad, "Stereopsis via deep learning," in NIPS Workshop on Deep Learning, volume 1, 2, 2011, doi:10.1.1.352.8987.

[27] O. Ozyesil, V. Voroninski, R. Basri, A. Singer, "A survey of structure from motion," arXiv preprint arXiv:1701.08493, 2017.

[28] P. F. Felzenszwalb, D. P. Huttenlocher, "Efficient graph-based image segmentation," International journal of computer vision, **59**(2), 167–181, 2004, doi:10.1023/B:VISI.0000022288.19776.77.

[29] B. Li, C. Shen, Y. Dai, A. van den Hengel, M. He, "Depth and Surface Normal Estimation From Monocular Images Using Regression on Deep Features and Hierarchical CRFs," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, doi:10.1109/CVPR.2015.7298715.

[30] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Communications of the ACM, **60**(6), 84–90, 2017, doi:10.1145/3065386.

[31] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[32] J. Long, E. Shelhamer, T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 3431–3440, 2015, doi:10.1109/TPAMI.2016.2572683.

[33] J. Zhu, R. Ma, "Real-time depth estimation from 2D images," 2016.

[34] Y. Li, X. Zhang, D. Chen, "Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes," in Proceedings of the IEEE conference on computer vision and pattern recognition, 1091–1100, 2018, doi:10.1109/ICCVW.2011.6130298.

[35] N. Silberman, D. Hoiem, P. Kohli, R. Fergus, "Indoor segmentation and support inference from rgbd images," in European conference on computer vision, 746–760, Springer, 2012, doi:10.1007/978-3-642-33715-4_54.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 808-814 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Discretisation of Second Order Generalized Integrator to Design the Control Algorithm of Unified Power Quality Conditioner

Mashhood Hasan[*,1], Bhim Singh[1], Waleed Hassan Alhazmi[2], Sachin Devassy[3]

[1]*Department of Electrical Engineering, Indian Institute of Technology (IIT Delhi), New Delhi, 110016, India*

[2]*Department of Mechanical Engineering, Jazan University, Jazan, 45142, Kingdom of Saudi Arabia*

[3]*CSIR-Central Electronics Engineering Research Institute, Power Electronics Group, Pilani, Rajasthan, 333031, India*

ARTICLE INFO

Article history:
Received: 23 December, 2020
Accepted: 08 March, 2021
Online: 31 March, 2021

Keywords:
AC mains
Discretization
Feedback unit
Power Quality
Voltage source converter
UPQC

ABSTRACT

In this paper, second order generalized integrator (SOGI) is discretized to design the control algorithm of unified power quality conditioner (UPQC). A UPQC is combination two voltage source converter (VSC) and VSC are connected with back to back DC link. The first one VSC is in series to maintain the desire voltage at point of common coupling (PCC) and second is connected with shunt VSC to share the reactive power demand of load. Moreover, it protects the AC mains from pollution of the load. A phase lock loop (PLL) based SOGI model is implemented to design an algorithm for UPQC under light polluted load. Whereas under highly polluted load the Laplace Transformation based PLL-SOGI model is fail to eliminate harmonics of current and voltage at PCC. Thus, a discretization of PLL-SOGI is needed to meet disadvantage. In this paper, reference current is generated under polluted load using discretization of PLL-SOGI model and compared it with actual current to pulse the gate of shunt VSC. Moreover, a feedback unit (m) is proposed to pulse the gate of series VSC under voltage sag/swell condition. A hardware setup is performed in the lab to verify the proposed algorithm under highly polluted load.

1. Introduction

This paper is extension of original work, existing in **2020 IEEE 9th Power India International Conference (PIICON)** [1]. In this paper, recent demand of electrical energy consumers and services are satisfied with international standard like Institute of Electrical and Electronics Engineers (IEEE) and International Electrotechnical Commission (IEC). It is done by using custom power devices (CPDs). The various kinds of CPDs are available in the present market while authors have chosen Unified Power Quality Conditioner (UPQC). This is unified controller to suppress the burden like current harmonics at source side and voltage harmonics at point of common coupling (PCC) of electrical distribution system [2]. Moreover, it suppresses the voltage related problems like voltage sag and swell at PCC [3]. A UPQC consist two voltage source converters (VSCs), associated back to back DC link. One VSC is connected in series w hereas second one is coupled in parallel to the load [4, 5]. The VSC is nothing without control algorithms which generates pulse . Thus, shunt and series part of VSCs require good control algorithms to improve the multiple power quality of the 3-phase AC primary distribution system [6-8]. The second order generalized integrator (SOGI) is the simple control approach which gives the information like time, magnitude and phase angle of the signals [9]. The SOGI have phase lock loop (PLL) combination and frequency locked loop (FLL) combination [10, 11]. The PLL based configuration of SOGI is discretization in the proposed work. The literature survey indicates many resemble works are reported in grid synchronization using SOGI [12, 13] to extract harmonics. In [14], SOGI based control approach to reduces sensors whereas power quality enhancement is reported in [15, 16]. A SOGI-PLL is presented in [17] to extract active and reactive power. A PLL based second and third order mixed generalized integrated is implemented in grid connected inverter [18]. As the order of equation is increased the further calculation require which again burden for programming to design a controller. Generally, PLL-SOGI and other modified SOGI are applicable to synchronize the voltages/currents signals [19, 20]. The phase of two periodic waves are minimized by PLL. Thus, two phases are synchronized exactly

[*]Corresponding Author: Mashhood Hasan, Email: mhasen@jazanu.edu.sa

in the proposed algorithms whereas back to back DC level voltage is increased to proper synchronization of two phases. The proposed configuration of SOGI is able to create two output signals. These output signals provide active and reactive components in the direct axis (α-axis) and quadrature axis (β-axis) respectively. Under polluted loads the point of common coupling (PCC) and AC mains get effected which produces harmonics. Thus, input currents/voltages signals of PLL-SOGI contains harmonics while output active components is passed through bandpass filter which reduces the harmonics components of current (i_α) and voltage (v_α) at certain level. Moreover, reactive component is passed through low pass filter at 90^0 degree of input signal. It gives reactive component (β) of current (i_β) and voltage (v_β). These voltages and currents signals are helpful to calculate real time active power ($p_L(t)$) and reactive power $q_L(t)$. While, researchers have found many ways to calculate real time $p_L(t)$ and $q_L(t)$ power [21-23]. Moreover, way of extraction of instantaneous power search accuracy of active and reactive component. Thus, in this work, PLL-SOGI is discretized using trapezoidal method. The trapezoidal methods measures accurate 90^0 degree lagging with respect to input signal under heavy polluted loads [24]. It helps to measure accurate instantaneous power of pL(t) and qL(t). The authors have following contributions in the present work which are given as follows,

- PLL-SOGI is discretized to design the control algorithm of UPQC which pulses the shunt part of VSC and series part of the VSC
- The reference signals of the AC mains currents are generated using simple mathematical calculation.
- A feedback unit 'm' is proposed to mitigate the voltage sag and swell at PCC.

Furthermore, in section 2, a brief information of model configuration is presented whereas section 3, control algorithm of UPQC is given. Results and discussion are shown in section 4 and in the last section 5 conclusion is presented.

2. Model Configuration

In Figure 1, a model arrangement of 3-phase, AC mains is connected with UPQC to mitigate the voltage sag, swell, reactive power and eliminate harmonics. The series part VSC of the UPQC is connected left from the PCC whereas shunt part VSC of UPQC connected at PCC. The back to back DC link voltage is maintained 180V for proper and fast working of the proposed algorithms. The DC link capacitor value is estimated nearly 3.3 mF. The AC mains supplies line voltage 110V at 50Hz to polluted load current with 20% THD and its apparent power is 819VA. The switching harmonics of series and shunt VSC are eliminated by R-C filters. The value of R and C are 10Ω and 10μF respectively. The series VSC is interfaced by series transformer whereas shunt VSC is interfaced by inductors value 8 mH. The control algorithm of series and shunt part are separately developed. In the control algorithm block, authors have contributed to develop an idea of algorithm for series VSC and shunt VSC.

3. Control Algorithms of UPQC

Two mathematical model descriptions are presented to design a control algorithm of UPQC. A UPQC has two VSC, thus it requires two separate algorithms.

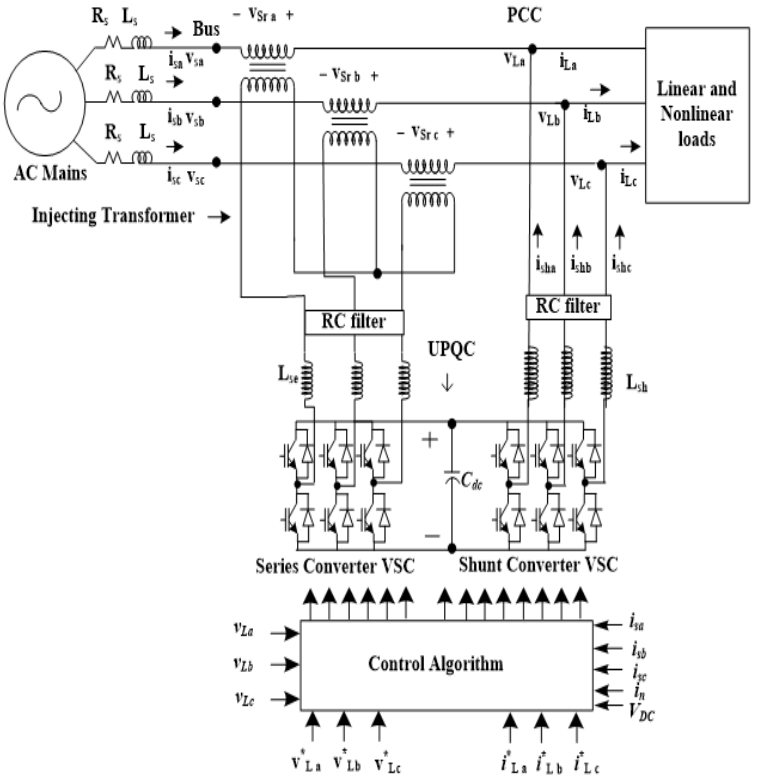


Figure 1: Schematic diagram of 3- Phase 3 Wire AC distribution system connected with UPQC

3.1. Design an Algorithm for Shunt VSC

The discretization of the input signals of SOGI is done to get the exact 90^0 degree phase shift of input signals. A trapezoidal method is suitable to design the algorithm. The H(z) is the discrete function of Laplace Function (LF) H(s), if it is satisfied by given rule.

$$H(s) = H(z)\Big|_{\frac{1}{s}=\frac{T_s}{2}\left(\frac{1+z^{-1}}{1-z^{-1}}\right)} \quad (1)$$

where z is z-transformation and T_s is the sampling frequency.

The simplest form of discrete form of SOGI structure [1] is given as,

$$v_\alpha(z) = \left(\frac{a_0(1+Z^{-2})}{1-(a_1)Z^{-1}+(a_2)Z^{-2}}\right)v(z) \quad (2)$$

The quadrature axis voltage (v_β) of general configuration of single phase PLL is estimated as follows,

$$v_\beta(z) = v_\alpha(z) * \frac{2}{T_s}\frac{1-z^{-1}}{1+z^{-1}} \quad (3)$$

where $a_0 = \frac{2k\omega T_s}{2k\omega T_s + (\omega T_S)^2 + 4}$, $a_1 = \frac{2(4-(\omega T_S)^2)}{2k\omega T_s + (\omega T_S)^2 + 4}$

and $a_2 = \frac{(2k\omega T_S - (\omega T_S)^2 - 4)}{2k\omega T_s + (\omega T_S)^2 + 4}$

The equation (2) and (3) are used to extract the discrete value of active part of voltage (v_α) and reactive part of voltage (v_β). These discrete values are helpful to get real time active power ($p_L(t)$) and reactive power ($q_L(t)$). The $p_L(t)$ and $q_L(t)$ are estimated as follows,

$$\begin{bmatrix} p_{La} \\ q_{La} \end{bmatrix} = \begin{bmatrix} v_{La.\alpha} & v_{La.\beta} \\ -v_{La.\beta} & v_{La.\alpha} \end{bmatrix} \begin{bmatrix} i_{La.\alpha} \\ i_{La.\beta} \end{bmatrix} \quad (4)$$

where p_{La} and q_{La} are real time active and reactive powers of phase 'a'. While, i_α active part and i_β are a reactive part of input AC current. The AC mains is designed for active power only thus, reactive part in equation (4) is taken as zero. Next stage, AC mains is delivered balanced power which is given as follows

$$p_s^* = \frac{p_{Ltotal}}{3} = \frac{(p_{La} + p_{Lb} + p_{Lc})}{3} \quad (5)$$

The equation (5) gives load demand of active power per phase. The reference of actual active and reactive part of reference current can find as

$$\begin{bmatrix} i_{sa.\alpha}^* \\ i_{sa.\beta}^* \end{bmatrix} = \begin{bmatrix} v_{La.\alpha} & v_{La.\beta} \\ -v_{La.\beta} & v_{La.\alpha} \end{bmatrix}^{-1} \begin{bmatrix} p_s^* \\ 0 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} i_{sa.\alpha}^* \\ i_{sa.\beta}^* \end{bmatrix} = \begin{bmatrix} v_{La.\alpha} & v_{La.\beta} \\ -v_{La.\beta} & v_{La.\alpha} \end{bmatrix}^{-1} \begin{bmatrix} p_s^* + p_{DC} \\ 0 \end{bmatrix} \quad (7)$$

The DC link of UPQC is controlled by a simple Proportional Integral (PI) controller. The input of PI is an error signal ($\varepsilon(n)$) of DC link voltage. And $\varepsilon(n)$ can be extracted by the difference of the reference voltage (v^*_{DC}) and actual sensed voltage ($v_{DC}(n)$). The power loss at nth sampling time is estimated using equation (8).

$$p_{DC}(n) = p_{DC}(n-1) + k_p\{\varepsilon(n) - \varepsilon(n-1)\} + k_i(n) \quad (8)$$

where kp and ki are gain constants of PI controller. While $\varepsilon(n)$ and $\varepsilon(n-1)$ sampling time.

Finally, using equation (7) reference currents (i_{sa}^*, i_{sb}^*, i_{sc}^*) can be evaluated as follows,

$$\begin{aligned} i_{sa}^* &= \frac{v_{La.\alpha}(n)}{v_{La.\alpha}^2 + v_{La.\beta}^2}\{p_s^* + p_{DC}\} \\ i_{sb}^* &= \frac{v_{Lb.\alpha}(n)}{v_{Lb.\alpha}^2 + v_{Lb.\beta}^2}\{p_s^* + p_{DC}\} \\ i_{sc}^* &= \frac{v_{Lc.\alpha}(n)}{v_{Lc.\alpha}^2 + v_{Lc.\beta}^2}\{p_s^* + p_{DC}\} \end{aligned} \quad (9)$$

However, the pulses of shunt VSC are generated by comparing the equation (9) signals and the signals of 3-phase AC mains currents (i_{sa}, i_{sb}, i_{sc},).

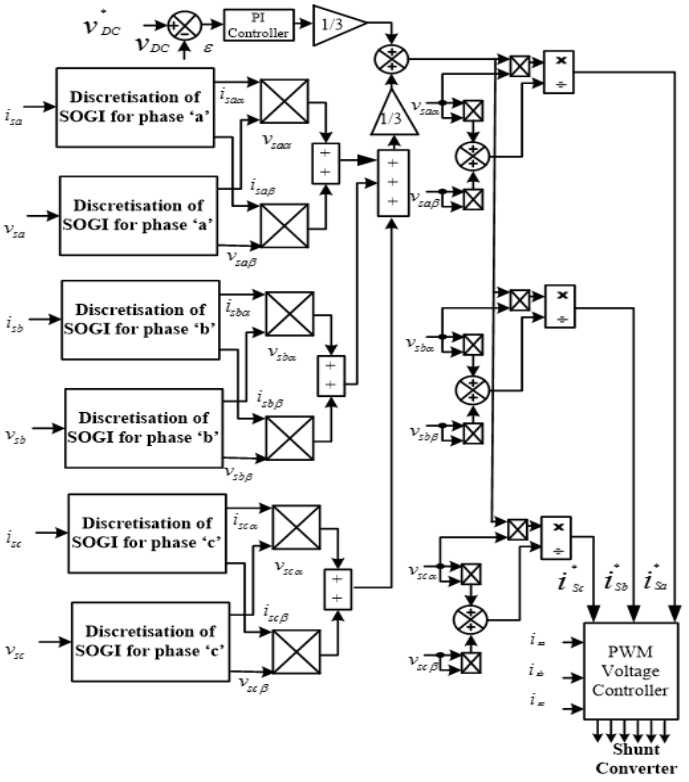


Figure 2: (a) Modified control Algorithm for shunt VSC

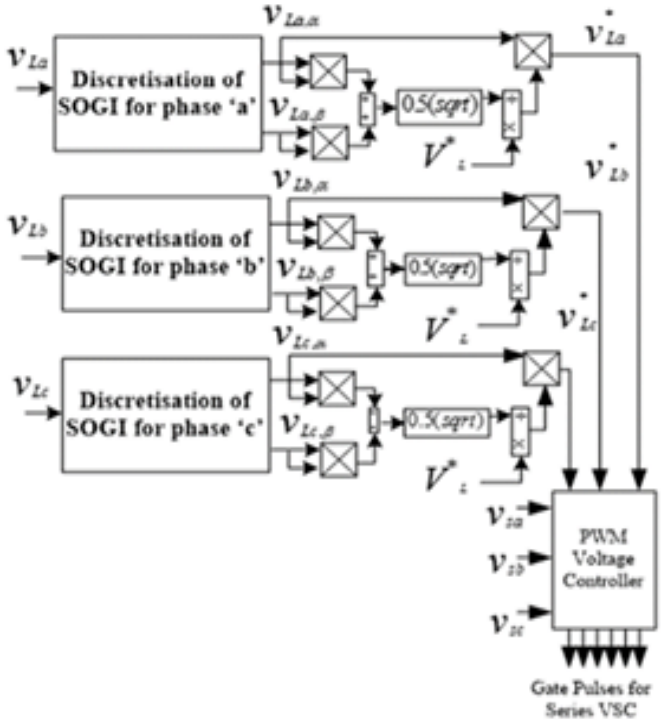


Figure 2: (b) Control Algorithm for series VSC

3.2. Design an Algorithm for series VSC

In Figure 2b, a feedback unit 'm' is played a role to design an algorithm of series VSC. Where m gives the proper feedback voltage under sag/swell condition of AC mains voltage. It maintains the desire PCC voltage. The actual PCC voltages are

extracted by a simple calculation which is square root sum of direct axis voltage ($v_{a,\alpha}$) and quadrature axis voltage ($v_{a,\beta}$). These values are further multiplying by 0.5 to get root mean square (rms) value. Moreover, a reference load voltage (VL*) is divided by the actual rms load voltage to obtain the m. Mathematically, m can be obtained as,

$$m = \frac{V_L^*}{v_{a,rms}} = \frac{V_L^*}{0.5 * \sqrt{v_{a,\alpha}{}^2 + v_{a,\beta}{}^2}} = \frac{V_L^*}{v_{La}} \quad (10)$$

Thus, under normal condition m predict a value equal to one to get desire voltage at PCC. This controller is an adaptive in nature. Similarly, m can predict a different correct value under abnormal condition of the AC mains. It is given as follows.

3.2.1 Voltage Sag Condition

In this case, assume, AC mains voltage of phase 'a' is running with voltage sag by factor x. Consequently, at PCC the voltage is reduced by factor x. The feedback unit is predicted m1 under voltage sag and it is calculated as follows,

$$m1 = \frac{V_L^*}{x * v_{a,rms}} = \frac{V_L^*}{0.5 * x * \sqrt{v_{a,\alpha}{}^2 + v_{a,\beta}{}^2}} = \frac{V_L^*}{x * v_{La}} = \frac{m}{x} \quad (11)$$

If factor x become one, the m1 is equal to m, hence it is proved by equation (11) that the proposed control algorithm is an adaptive in nature under voltage sag condition. Moreover, the generated reference voltage ($v^*_{La,}$) is equal to multiple of direct axis voltage ($v_{La,\alpha}$) and feedback unit (m1*x) under voltage sag condition. Now, v^*_{La} can find as,

$$v^*_{La} = m1 * x * v_{La\alpha} = m * v_{La\alpha} \quad (12)$$

3.2.2 Under Voltage Swell

In this case, assume, AC mains voltage of phase 'a' is running with voltage swell by factor y. Consequently, at PCC the voltage is reduced by factor y. The feedback unit is predicted m2 under voltage swell and it is calculated as follows,

$$m2 = \frac{v_L^*}{y * v_{a,rms}} = \frac{v_L^*}{0.5 * y * \sqrt{v_{a,\alpha}{}^2 + v_{a,\beta}{}^2}} = \frac{v_L^*}{y * v_{La}} = \frac{m}{y} \quad (13)$$

If factor y become one, the m2 is equal to m, hence it is proved by equation (13) that the proposed control algorithm is an adaptive in nature under voltage swell condition. Moreover, the generated reference voltage ($v^*_{La,}$) is equal to multiple of direct axis voltage ($v_{La,\alpha}$) and feedback unit (m2*y) under voltage swell condition. Now, v^*_{La} can find as,

$$v^*_{La} = m2 * y * v_{La,\alpha} = m * v_{La,\alpha} \quad (14)$$

In the same way, the above equation can extract for phase 'b' and phase 'c' of primary AC distribution system. The proposed control algorithm can work under abnormal condition of AC mains and it is good to opt for the UPQC.

4. Experimental Results

A 3-phase, 3-wire AC system is integrated with UPQC in the lab which is shown in Figure 3. It consists D-Space software to link with hardware and PC for evaluation the performance of the modified algorithm under polluted loads. In Table 1, model parameters are given. Whereas in Table 2 shows the THD and power factor under various load conditions. The hardware results are shown in Figures 5 and Figures 6&7 under steady state conditions and under dynamic condition respectively. Moreover, a comparative Matlab result is shown in Figure 4 to understand the value of modified algorithm.

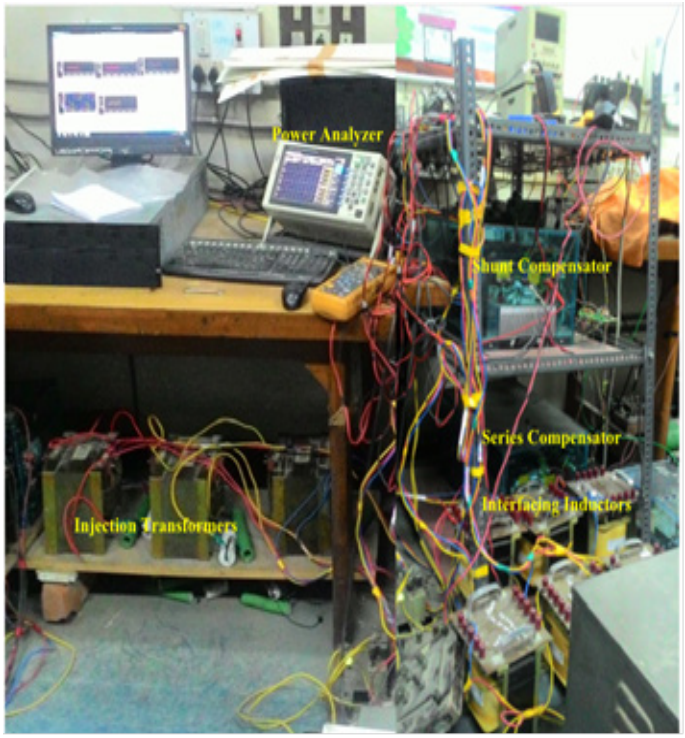


Figure 3: A lab model of 3-phase 3-wire AC system integrated with UPQC

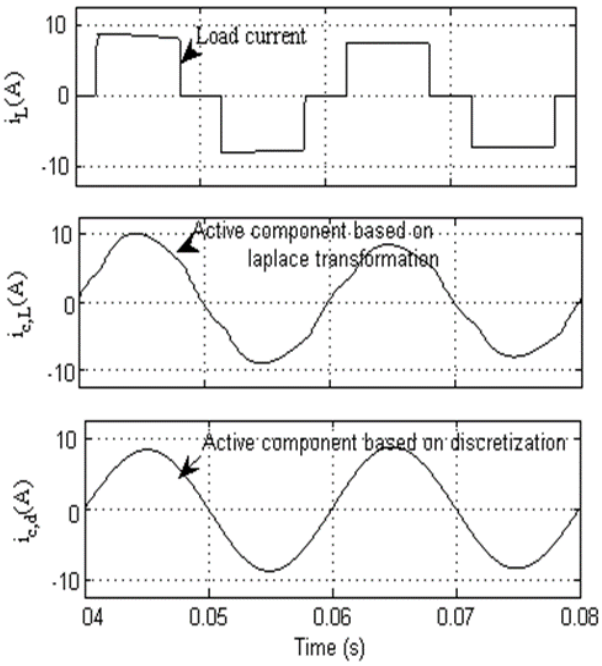
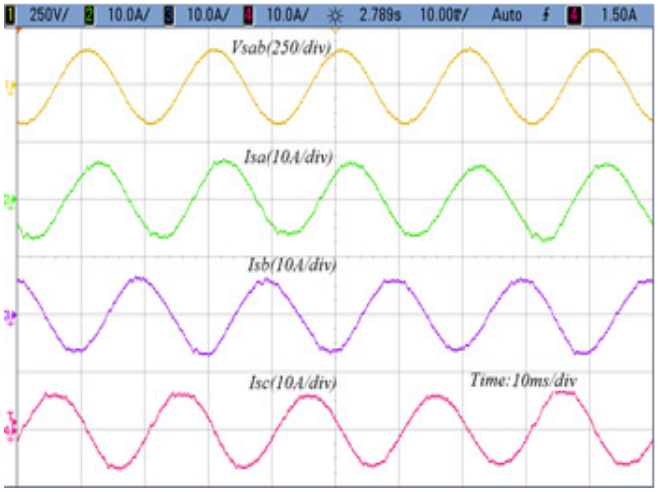


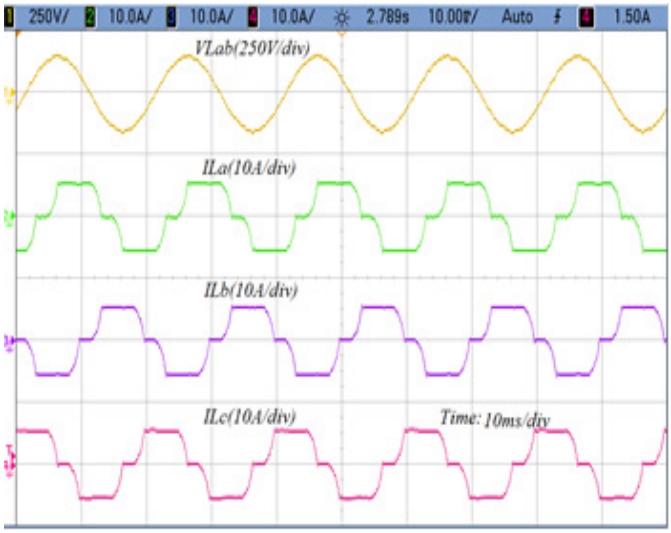
Figure 4 a comparatives results under quasi square wave demand of load current, (a) polluted load current waveform ($i_L(A)$), (b) AC mains current waveform under SOGI ($i_{c,L}(A)$) and AC mains current under modified SOGI ($i_{c,d}(A)$).

Table 1: Practical Parameters

Quantities	Value
PCC line voltage	110V
Frequency	50Hz
Load	819VA
Load current THD	20%;
DC-bus Voltage	180V
DC-bus Capacitor	3.3 mF
Shunt converter Interfacing Inductors	8 mH
Ripple Filter	10 µF, 10 Ω
Series Converter Interfacing Inductor	1.5 mH
DC-bus PI Gains	Kp = 4, Ki = 1
Gain constant	K=1
a_o	1.5698*10-4
a1	1.99984
a2	-0.999764



(a) The line to line voltage and Phase currents abc at AC mains under normal condition



(b) The line to line voltage and Phase currents abc at PCC under normal condition

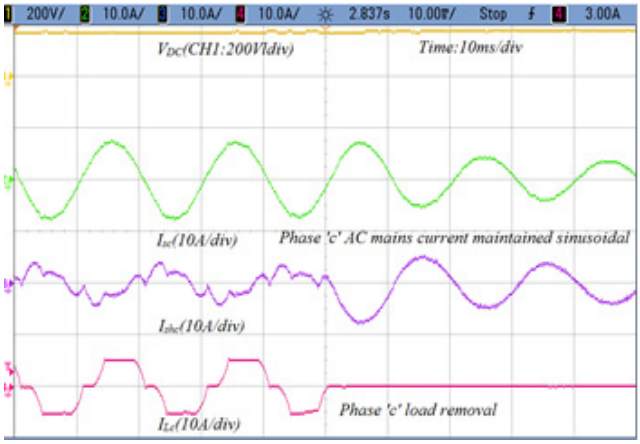
Figure 5: The AC mains voltage, currents, load voltage and load currents under normal condition.

4.1. Performance Under Normal Condition

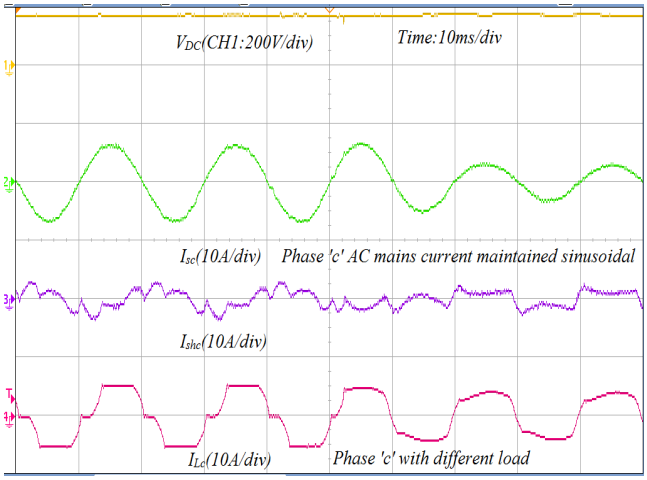
The power analyzer 3-phase 4 channel is used to show the results of modified algorithm. The line to line voltage v_{sab} and 3-phase AC mains current are indicated as i_{sa}, i_{sb}, and i_{sc} respectively. It is shown in Figure 5(a) whereas in Figure 5(b), the line to line load voltage v_{Lab} and single phase load current are indicated as i_{La}, i_{Lb}, and i_{Lc}. It is observed from the results that source AC mains are free from harmonics while load is under heavy polluted thus, shunt algorithm works properly. Whereas, series algorithm maintain desired voltage at PCC properly under steady state condition.

4.2. Performance Shunt VSC Under Dynamic Condition

In Figure 6(a), performance of modified control algorithm for shunt VSC is presented under dynamic load condition. In this case, phase 'c' of load is abruptly off. Thus, load current i_{Lc} is zero and at the same time shunt VSC become active and inject shunt current i_{shc} in phase 'c' to stable the AC mains source current i_{sc} . Whereas, in Figure 6(b), when load demands current in phase 'c' active, the shunt VSC injects harmonics to make supply pure sinusoidal AC mains current. It is seen that the modified control algorithm works properly to maintain AC mains source as an international standard and balanced. Moreover, the DC link voltage (V_{DC}) is unchanged under dynamics condition of load.



(a) Performance of shunt converter under zero load



(b) Performance of shunt converter under dynamic load

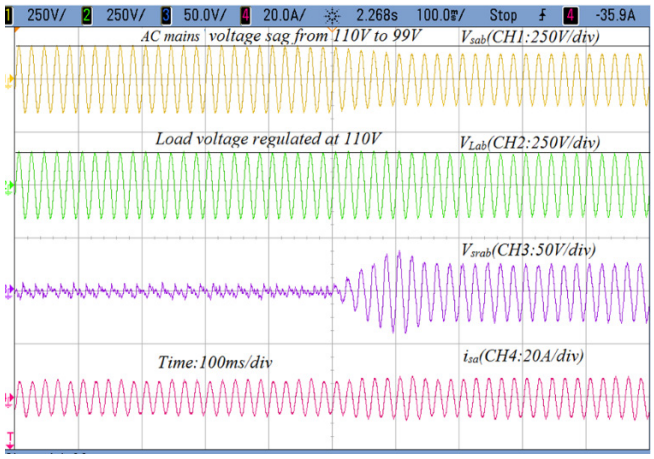
Figure 6: Current waveform of AC main current, shunt current and load current under dynamic condition

4.3. Performance Series VSC Under Dynamic Condition

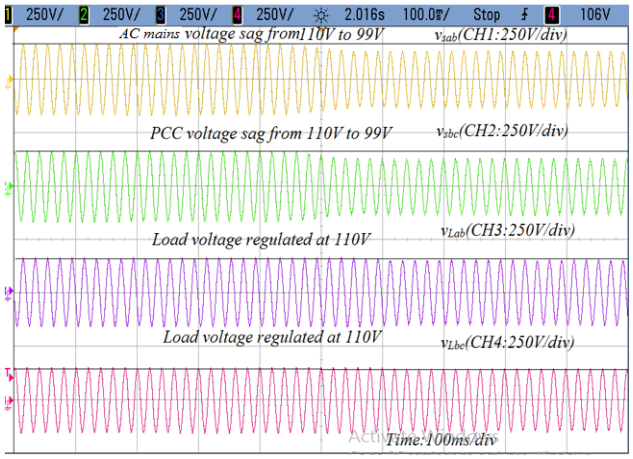
The performance of series VSC is shown in Figures 7(a)-(b) and Figures 7(c)-(d) under voltage sag and swell conditions. In Figure 7(a), AC mains line to line voltage v_{sab} is running under voltage sag while at PCC line to line voltage v_{Lab} is regulated to maintain the desire voltage of the load. The series voltage v_{srab} is added in series with the AC mains to regulate load voltage at 110V using series VSC.

In Figure 7(b), AC mains 3-phase voltages are regulated by series VSC. The line to line voltage v_{sab} and v_{sbc} are running under voltage sag condition whereas line to line load voltage v_{Lab} and v_{Lbc} are regulated at desire voltage 110V. Thus, the waveform of three phase voltage shows that the performance of feedback unit 'm' works satisfactory and series VSC works at desire level to mitigate voltage sag.

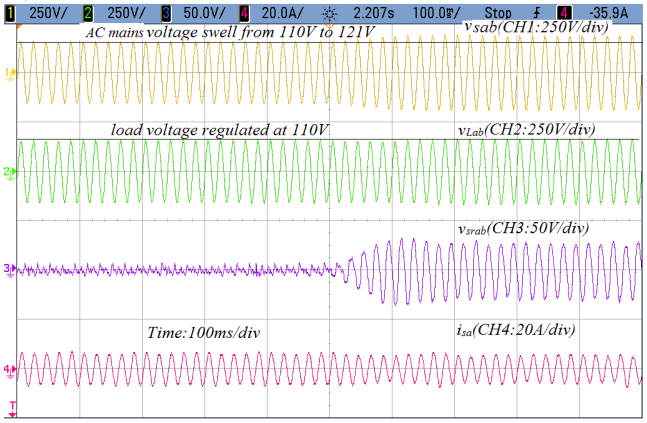
Moreover, the feedback factor 'm' is tested under voltage swell condition. In this case, series VSC injects voltage in feeder of AC mains to opposite phase of AC mains voltage. It suppresses the AC mains voltage and consequently load voltage is maintain at desire level 110V. In Figure 7(c), AC mains v_{sab} is running under voltage swell while at PCC v_{Lab} is regulated to maintain the desire voltage of the load. The series voltage v_{srab} is substract in series with the AC mains to regulate load voltage at 110V using series VSC. While, AC mains current i_{sa} reduces under voltage swell condition to balance the power of the AC mains.



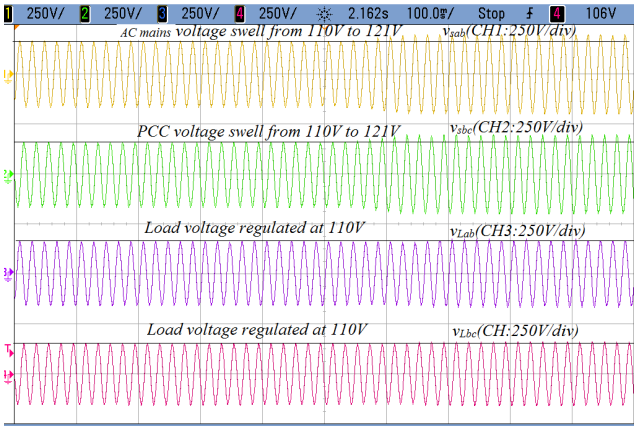
(a) Performance of series converter under voltage sag



(b) Regulated voltage of three phase load under voltage sag



(c) Performance of series converter under voltage swell



(d) Regulated voltage of three phase load under voltage swell condition

Figure 7: Voltage waveforms of AC mains voltage, load voltage and series converter voltage under dynamic condition

In Figure 7(d), AC mains 3-phase voltages are regulated by series VSC. The v_{sab} and v_{sbc} are running under voltage swell condition whereas v_{Lab} and v_{Lbc} are regulated at desire voltage 110V. Thus, the waveform of three phase voltage shows that the performance of feedback unit 'm' works satisfactory and series VSC works at desire level to mitigate voltage swell.

5. Conclusion

The discrete value of second order generalized integrator (SOGI) has realized to design control approach of UPQC which performance have been verified under heavy polluted load and under various conditions of load. Under normal condition, series VSC is inactive while shunt VSC eliminate the harmonics of the load and injects reactive power demand to the 3-phase load. And under dynamic condition, the shunt VSC shows the excellent performance whereas the series VSC works under voltage sag and swell conditions. The voltage sag from 110V to 99V while voltage swell from 110V to 121V are generated at AC mains while at load side (PCC) the voltage is regulated at 110V. The performance of proposed feedback unit 'm' for series part of UPQC has unique to suppress the sag and swell voltage of PCC and modified algorithm of shunt part of UPQC eliminates the harmonics of load and enhance the performance of AC mains. Thus, it has been seen that proposed algorithm of UPQC improve the multiple power quality of the 3-phase primary distribution feeder line of power system.

Table 2: THD and Power Factor Under Various Load Conditions

Parameters	Nominal	Sag	Swell
Load Current THD	20%	20.14%	19.80%
PCC Current THD	3.77%	2.88%	4.65%
Load Voltage THD	3.58	3.50%	3.55%
PCC Voltage THD	1.92	1.96%	1.9%
Power factor at grid	0.99	0.99	

References

[1] M. Hasan, B. Singh and S. Devassy, "Modified Control Technique for Series and Shunt Converters of UPQC to Enhance Power Quality," in 2020 IEEE 9th Power India International Conference (PIICON), SONEPAT, India, 2020, 1-6, doi: 10.1109/PIICON49524.2020.9113056.

[2] M. Hasan, A. Q. Ansari and B. Singh, "Parameters estimation of a series VSC and shunt VSC to design a unified power quality conditioner (UPQC)," 2015 39th National Systems Conference (NSC), Greater Noida, India, 2015, 1-6, doi: 10.1109/NATSYS.2015.7489111.

[3] S. Devassy and B. Singh, "Performance Analysis of Solar PV Array and Battery Integrated Unified Power Quality Conditioner for Microgrid Systems," in IEEE Transactions on Industrial Electronics, **68**(5), 4027-4035, 2021, doi: 10.1109/TIE.2020.2984439.

[4] M. Hasan, B. Singh and A. Q. Ansari, "An approach to minimize the VA size of UPQC-S and its performance comparison," 2018 IEEMA Engineer Infinite Conference (eTechNxT), New Delhi, 2018, 1-5, doi: 10.1109/ETECHNXT.2018.8385344.

[5] A. Q. Ansari, B. Singh and M. Hasan, "Algorithm for power angle control to improve power quality in distribution system using unified power quality conditioner," in IET Generation, Transmission & Distribution, **9**(12), 1439-1447, 2015, doi: 10.1049/iet-gtd.2014.0734.

[6] S. Chandrakala Devi, B. Singh and S. Devassy, "Modified generalised integrator based control strategy for solar PV fed UPQC enabling power quality improvement," in IET Generation, Transmission & Distribution, **14**(16), 3127-3138, 2020, doi: 10.1049/iet-gtd.2019.1939.

[7] M Hasan - Handbook of Research on Power and Energy System, Application of Power Electronics in Power Systems, 2018, page 31 DOI: 10.4018/978-1-5225-3935-3.ch010

[8] S. Devassy and B. Singh, "Implementation of Solar Photovoltaic System With Universal Active Filtering Capability," in IEEE Transactions on Industry Applications, **55**(4), 3926-3934, 2019, doi: 10.1109/TIA.2019.2906297.

[9] M. Karimi-Ghartemani, H. Karimi and M. R. Iravani, "A magnitude/phase-locked loop system based on estimation of frequency and in-phase/quadrature-phase amplitudes," in IEEE Transactions on Industrial Electronics, **51**(2), 511-517, 2004, doi: 10.1109/TIE.2004.825282.

[10] B. Trento, L. M. Tolbert and D. Costinett, "Grid synchronization using fixed filtering with magnitude and phase compensation," 2014 IEEE Energy Conversion Congress and Exposition (ECCE), Pittsburgh, PA, 2641-2647, 2014, doi: 10.1109/ECCE.2014.6953755.

[11] M. S. Reza, M. Ciobotaru and V. G. Agelidis, "Tracking of time-varying grid voltage using DFT based second order generalized integrator technique," 2012 IEEE International Conference on Power System Technology (POWERCON), Auckland, 1-6, 2012, doi: 10.1109/PowerCon.2012.6401432.

[12] T. Ngo, S. Biricik, and M. Basu, "A Self-tuning Grid Synchronization Method for Active Power Filters," Electric Power Components and Systems, **44**(17), 1-11, 2016, doi.org/ 10.1080/15325008.2016.1200696.

[13] W. Li, X. Ruan, C. Bao, D. Pan and X. Wang, "Grid synchronization systems of three-phase grid-connected power converters: A complex vector filter perspective," IEEE Trans. Ind. Elect., **61**(4), 1855–1870, 2014, doi.org/10.1109/TPEL.2011.2158238.

[14] S. Biricik, S. Redif and M. Basu, "Voltage Sensorless Control of Single-phase Active Power Filter Based on the Second-order Generalized Integrator Algorithm," *Electric Power Components and Systems*, **43**(7) 820–827, 2015, doi.org/ 10.1080/15325008.2014.100359.

[15] F. U. Nazir, N. Kumar, B. C. Pal, B. Singh and B. K. Panigrahi, "Enhanced SOGI Controller for Weak Grid Integrated Solar PV System," in IEEE Transactions on Energy Conversion, **35**(3), 1208-1217, 2020, doi: 10.1109/TEC.2020.2990866.

[16] P. Chittora, A. Singh and M. Singh, "Simple and efficient control of DSTATCOM in three-phase four-wire polluted grid system using MCCF-SOGI based controller," in IET Generation, Transmission & Distribution, **12**(5),1213-1222, 2018, doi: 10.1049/iet-gtd.2017.0901.

[17] S. Devassy, B. Singh. "Modified p-q Theory Based Control of Solar PV Integrated UPQC-S", IEEE Transactions on Industry Applications, **53**(5), 5031-5040, 2017, doi: 10.1109/TIA.2017.2714138.

[18] C. Zhang, X. Zhao, X. Wang, X. Chai, Z. Zhang and X. Guo, "A Grid Synchronization PLL Method Based on Mixed Second- and Third-Order Generalized Integrator for DC Offset Elimination and Frequency Adaptability," in IEEE Journal of Emerging and Selected Topics in Power Electronics, **6**(3), 1517-1526, 2018, doi: 10.1109/JESTPE.2018.2810499.

[19] C. M. Nirmal Mukundan, P. Jayaprakash, U. Subramaniam and D. J. Almakhles, "Binary Hybrid Multilevel Inverter-Based Grid Integrated Solar Energy Conversion System With Damped SOGI Control," in IEEE Access, **8**, 37214-37228, 2020, doi: 10.1109/ACCESS.2020.2974773.

[20] N. Kumar, I. Hussain, B. Singh and B. K. Panigrahi, "Implementation of Multilayer Fifth-Order Generalized Integrator-Based Adaptive Control for Grid-Tied Solar PV Energy Conversion System," in IEEE Transactions on Industrial Informatics, **14**(7), 2857-2868, 2018, doi: 10.1109/TII.2017.2777882.

[21] A. Nabae and T. Tanaka, "A new definition of instantaneous active-reactive current and power based on instantaneous space vectors on polar coordinates in three-phase circuits," in IEEE Transactions on Power Delivery, **11**(3),1238-1243,1996, doi: 10.1109/61.517477.

[22] H. Fujita and H. Akagi, "The Unified Power Quality Conditioner: The Integration of Series- and Shunt-Active Filters"; IEEE Trans. On Power Electronics, **12**(2), 315–322, 1998, doi:10.1109/63.662847

[23] M. Monfared, M. Sanatkar, and S. Golestan. "Direct active and reactive power control of single-phase grid-tie converters." IET Power Electronics, **5**, 1544-1550, 2012, doi.org/10.1049/iet-pel.2012.0131

[24] M. Ciobotaru, R. Teodorescu and F. Blaabjerg, "A new single-phase PLL structure based on second order generalized integrator," 2006 37th IEEE Power Electronics Specialists Conference, Jeju, Korea (South), 1-6, 2006, doi: 10.1109/pesc.2006.1711988.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 821-828 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Using Formal Methods to Model a Smart School System via TLA+ and its TLC Model Checker for Validation

Nawar Obeidat*, Carla Purdy

College of Engineering and Applied Science, University of Cincinnati, Cincinnati, OH 45221-0030, USA

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 24 March, 2021
Online: 04 April, 2021

Keywords:
TLA+
Verification and Validation
Safety

ABSTRACT

Formal methods are one of the efficient tools to verify and validate designs for different kinds of systems. Smart systems are attracting researchers' attention due to the rapid spread of new technologies all over the world. Modeling a smart system requires connecting heterogeneous subsystems together to build it. Our contribution to this work is in focusing on using formal methods to prove that a design model meets its specifications. We have chosen to design a smart school building system due to the lack of research in this particular area, and to prove that formal methods are appropriate for different systems applications. In this paper, we have used UML diagrams and the formal specification language TLA+ to design a smart school building system. We validate our design using the TLC model checker. The smart school system has many subsystems connected together including a secure access system, lighting control system, climate control system, and smoke detection system. Safety is a very important attribute in this system. Our goal is to have a smart system that satisfies its functional requirements as well as any non-functional requirements like safety. The system provides safety for employees and students in the smart school.

1. Introduction

Smart systems that have been designed and modeled by researchers include a smart school [1], smart library [2], a smart office [3], a smart home [4], and a smart campus [5]. The main focus of these systems was on safety and the ability to achieve specific goals. This paper is an extension of work originally presented at the 2020 IEEE 3rd International Conference on Information and Computer Technologies (ICICT), San Jose, CA, USA. [1]. In this paper, we modify the old smart school system in [1] using TLA+ tools to make it a more secure and safe system. To prove the design correctness, we validate our design using the TLC model checker.

We used both informal modeling methods like the Unified Modeling Language (UML) [6] and formal methods [7] to define system entities, system behavior, and sequence of actions. UML provides different kinds of modeling diagrams such as state diagrams, sequence diagrams, and object interaction diagrams. As well as diagrams, UML has notations and presentation conventions that have become common in the object-oriented domain and structured methods. UML can be defined as a set of graphical models which represents several properties of an object-oriented design [6]. The structural and behavioral models are UML's two most important model types. In this paper we choose to use behavioral models to represent the behavior of the smart school system.

To model a complete smart system, we need to integrate different subsystems together, which makes verification and validation of the whole system harder to accomplish. Formal methods are one approach that can provide verification and validation [7]. Formal methods are valuable due to their ability and effectiveness in designing systems which are close to bug-free. Many reputable companies have begun welcoming and using formal methods in the last decade [8, 9]. Furthermore, many researchers used formal methods to validate their design systems. Examples include the work done in [3, 10, 11, 12].

Formal methods support development, specification tools, and verification in both hardware and software systems [7]. Formal methods are techniques that are mathematically based and are used to prove that the system's specifications meet its implementation. Formal methods provide simplicity and remove complexity, which is an important factor in system development, as well as verifying different system attributes such as reliability, accuracy, correctness, security, and safety [7]. A formal verification scheme is used in formal methods to ensure the system will be correct

*Corresponding Author: Nawar H. Obeidat, University of Cincinnati, Cincinnati, OH 45221-0030, USA, obeidanh@mail.uc.edu

before accepting the design. All these characteristics of formal methods make them highly trusted compared to other verification methods [7]. Formal methods have been used for different systems applications to guarantee safety and correctness and have been extensively verified for many of these systems. The strength of formal methods lies in catching systems bugs and errors which cannot be caught by other verification methods. This makes them popular with big industrial companies seeking to verify and validate their complex systems. Modeling the system using an abstract mathematical model is the first step in applying formal methods. This first step describes the system's requirements using a formal specification language. In this work, we focus in ensuring both safety and security for our smart system since having a secure login system will enhance the overall system's safety as a result. We also add in additional safety features such as fire or smoke detection.

Formal methods use formal specification languages in the process of system analysis, requirements analysis, and system design [13]. Describing a system using a specification language is different from writing executable code using a programming language. Specification languages do not describe the "how?", they describe the "what?". Verifying the program correctness by creating proofs is one of the important features of specification languages. We can use formal system syntax, proof rules, and semantics rules when applying a specification language [13]. The language can be determined by the syntax and semantics, and the proof system is the result of the proof rules. Specification languages use expressions to represent specifications. These specification languages are used to stipulate design of hardware/software systems, describe a domain formally, or to give requirements recommendations to the system [13].

In this paper we will use the TLA+ specification language. TLA stands for "Temporal Logic of Actions". TLA+ is a high level mathematically based formal modeling language. It is used to model distributed and concurrent systems. It is also used to find design errors which are hard to find in the code level and hard to correct. The designer of TLA+ is Lamport, who wanted to describe distributed algorithms formally. He published his book *Specifying Systems* [13] in 2002. He describes TLA+ in this book, along with how to use TLA+, and how to use its efficient tools. TLA+ has modules that include specifications and can be reused independently. Most mathematicians consider TLA+ to be a standard basis to formalize specifications [13]. In TLA+, both properties and specifications of a system are written as logical formulas. Actions like hiding of the internal state, refinement, and composition of the system are performed using logical connectives of quantification, implication, and conjunction. In order to help a designer in the formal development process, TLA+ has supporting tools such as theorem provers and its powerful model checker TLC, which we used in this work to validate our model [13].

In this work, we will use the TLA+ formal specification language to model the smart school system. We describe the system's abstract model using UML. Extending work in [1], we improve the TLA+ model, and we use the powerful TLC model checker to verify and validate our system. Our modified smart school building system has various subsystems. Students, employees, and visitors must enter using a secure login sub-system to ensure security and safety. Each person must enter a correct username and password to enter the school building using the main door. Once the first allowed person enters the school, the lighting sub-system will work automatically, and so will the HVAC sub-system, which uses temperature sensors to sense the temperature and adjust it inside the building. The smoke detection sub-system will work all the time to sense any smoke and guarantee safety in the building and open all exit doors in case of fire.

In this paper, the related work is described in section 2, the UML models are described in section 3, the TLA+ formal specifications using TLA+ are described in section 4, the TLC verification model which represents the final result will be described in section 5, and finally the conclusion and future work will be described in section 6.

2. Related Work

Because of the massive technology evolution going on, smart systems implementation has become an attractive research area for researchers [14]. Many smart systems have been modeled by many researchers. For example, in [4], the author modeled a smart home system that uses a wifi network based on the AllJoyn framework. It uses asymmetric elliptic curve cryptography to apply authentications during system operation. The authentication process of this system allows the user to interact with the system and control it using an application program based on Android. Utilizing a mobile social network, a smart campus was proposed in [5]. The author set up a collaboration between a flexible system architecture and social interactions in the campus. His model addresses the server side represented by social connections and services and the client side represented by the mobile users. In [15], the author presented a prototype for a smart office system which was one of the pilot applications in the FP7 EU project ELLIOT (Experiential Living Lab for the Internet of Things). He used the LinkSmart semantic middleware for the solution he described. In [16], the author proposed a smart parking system. It uses parking destination and parking cost to reserves a parking spot. His smart system lowers the average parking time and cost to reserve a parking spot. For each decision, mixed-integer linear programming (MILP) was used. As a sub-system for a smart city, in [10], the author proposed a smart sewage system. He used UML, Nondeterministic Finite Automata (NFA), and TLA+ to model his smart sewage system. In [17], the author presented a survey of the enabling architecture, technologies, and protocols for an urban Internet of Things (IoT). He discussed best-practice technical solutions and guidelines for the Padova Smart City project in Italy [17].

Several researchers worked on modeling a smart campus system [18-20], focusing on a mobile-learning domain, cloud learning, E-learning, and an environmentally aware campus. As for modeling smart school buildings, a few researchers worked on modeling with focusing on power consumption and power management [21-24]. Our smart school building model is different because it has advanced features, including security and safety, compared to what was modeled so far in the same research area.

In smart systems, a main concept is having sub-systems which are connected together and work efficiently as one controllable system. Safety is one of the most important properties is any system, especially in a smart school system since it is dealing with human lives (students, teachers, and school employees). Security

also is an important property to have in a smart school system. It will enhance the overall system's safety. Most researchers modeled their systems focusing on either safety or security, but our model achieves a system which is both safe and secure.

Formal methods had been used to verify and validate many smart systems. Researchers used different formal specification languages to model their systems including TLA+ [13], VDM [25], Z [26], and Alloy [27]. In [11], the author modeled and verified a smart parking system using TLA+. In [3], the author designed and verified a smart office system using the VDM-SL toolbox. And in [28], the author used Alloy to validate his smart home system. In this work, we decided to use TLA+ to model and verify our smart school building system. We choose TLA+ over other specification languages for many reasons, which we list here.

In comparison to Alloy in modeling nested structures, TLA+ doesn't need too many layers of identification, which makes it simpler and more direct. TLA+ is also more expressive than Alloy. Although Alloy has its efficient Alloy Analyzer model checker that is faster than the TLC model checker and able to handle important large analyses that TLC currently is not able to handle, Alloy Analyzer crashes or hangs in some cases that are needed for larger systems. Also, using TLA+ allows users to trace every single state and get results for each one using its Trace Explore feature, which makes tracing bugs, finding bugs, and fixing them much simpler and easier. Alloy has the same advantage, but the results for systems with more than a few variables or a few time-steps are not clear-cut. Some other features make TLA+ a better choice than Alloy and other specification languages including VCC and Z. TLA+ is flexible when there is a need to support high-level functions and edit details to the specifications, unlike VCC that would require us to write "ghost code" which is a superset of the C programming language and Alloy that does not support high-level functions like recursive functions. One important feature in TLA+ is its powerful TLC model checker that can operate over massive state-spaces with reasonable throughput. The TLC model checker is also fast because of its ability to use multi-cores efficiently, unlike B, VCC, and Event-B model checkers, which are not able to use more than one core. TLA+ is also a very expressive language, and it supports the liveness property better than any other formal specification language [9].

In the following section, we will illustrate how to use UML as a first step to model our smart school building system.

3. UML Modeling for Smart School Building System

Unified Modeling Language (UML) is one of the most common languages used to represent the informal abstract model of a system. UML captures system properties and provides graphical notations. Smart school system outputs and inputs are shown in Figure 1 [1]. As shown in Figure 1, the smart school inputs are taken from a user or different kinds of sensors and reactions to these inputs appear as outputs from the system. The login input is handled by entering a valid username and password by the user, who can be a student, an employee, or a visitor and it may or may not open the main door based on the validation of the username and the password. The smoke sensor senses any smoke in the building and gives an alarm as an output, and in this work, we added a new output that in case of smoke all exit doors will be opened to increase the safety of our system. The temperature sensor senses the building's temperature and turns the heat or the AC on or keeps them off. The light sensor senses the natural light inside the building and adjusts the lights based on it. The system may have multiple sensors depending on the size, orientation, and architecture of the building. The outputs of the system depend on the inputs. For example, if the username and password were entered correctly as an input, the main door will automatically open to allow entrance to the school building as an output. If the HVAC sub-system receives an input from the temperature sensor that is the temperature is too low (e.g., 60 F), the output will be to automatically turn on the heat in the building. If the smoke detection sub-system receives an input from the smoke sensor that there is smoke in the building, the output will be to turn on the smoke alarm and open the exit doors to let everyone leave the school building immediately.

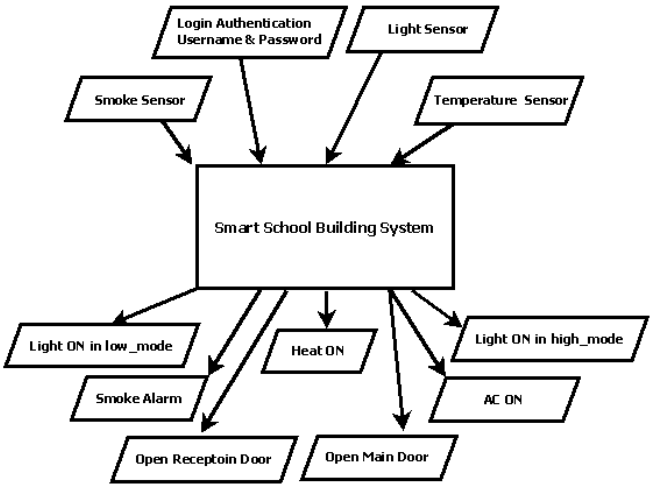


Figure 1: Inputs and Outputs of the Smart School Building System [1]

Figure 2 shows the UML use case diagram of the system. This diagram shows when the actors (student, employee, and visitor) login to the system by entering the username and password correctly, the main door will open, and they will have access to the sub-systems in the building.

Figure 3 [1] shows the UML activity diagram of the smart school building. The activity diagram will help in better understanding how the system works.

Figure 4 provides a UML sequence diagram of the smart school building system. It illustrates the sequence of all actions that happen in the system. As an extension of work in [1] and in order to increase system's security and eventually safety, the new smart school model requires that each person entering the school has to has a unique username and password. These usernames and passwords will be given by school or school district to each person. Visitors must request a username and password from the school before their visit, i.e., visitors to the school building must have a valid login username and password to be allowed to enter the building.

If the username and password are correct, the main door will automatically be opened. Once the first person enters the building, all systems, including the lighting sub-system and the HVAC sub-

system, will automatically start working. The smoke detection system will work all the time, even after school hours, to guarantee safety in the building in case of fire.

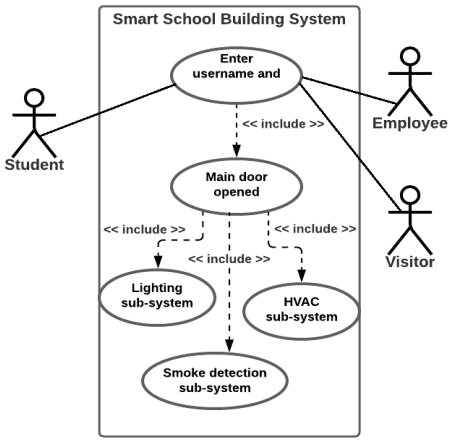


Figure 2: UML Use Case Diagram of the Smart School System

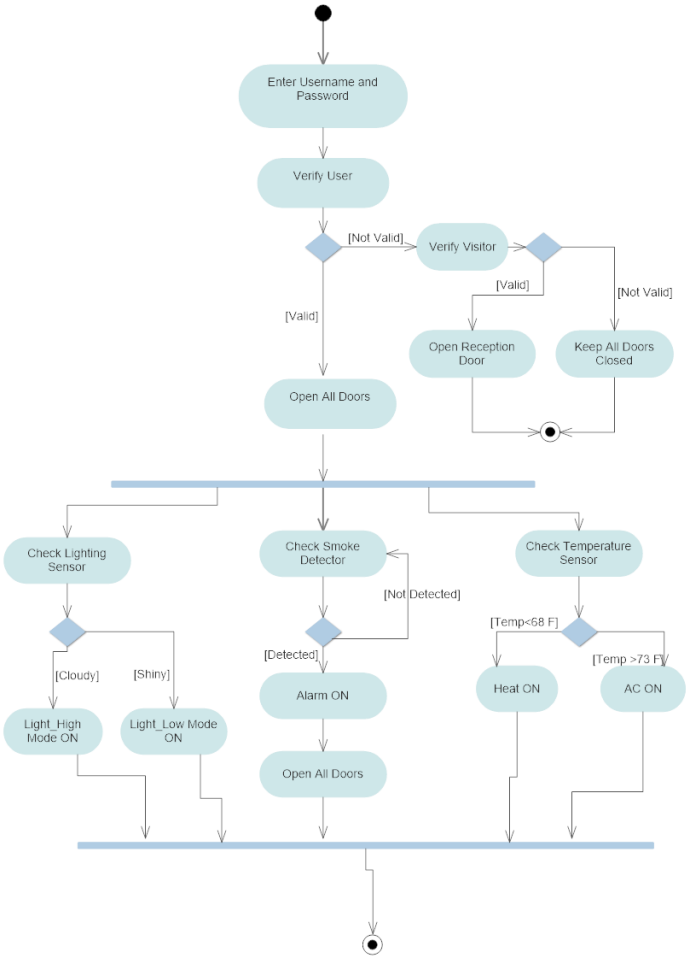


Figure 3: UML Activity Diagram of the Smart School Building System [1]

4. Formal Specification Using TLA+

This section illustrates the formal specification of our extended smart school system model. We use the TLA+ toolbox to write the system's specifications and the TLC model checker to validate our system. We represent all of the system's operations using TLA+.

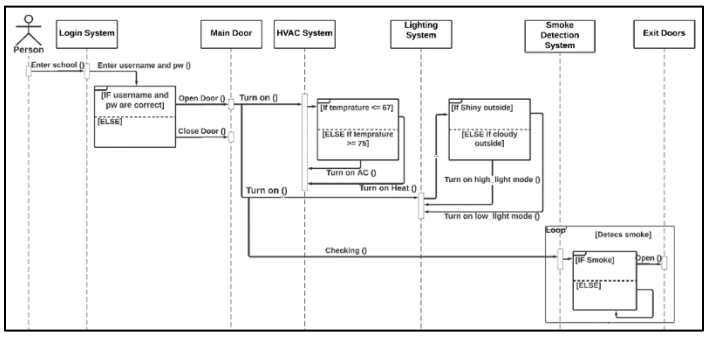


Figure 4: Sequence Diagram for the Smart School Building System

In TLA+, the system's specifications are called spec, and they are written in a module which we have called in our case *smartSchoolSystem*. In the module, we include the system's variables that we will use inside the module as shown in Figure 5.

MODULE *smartSchoolSystem*

VARIABLES *objects*, *person*, *lighting_sys*, *HVAC_sys*, *smoke_detect_sys*, *main_door*, *exit_doors*, *login_sys*, *username*, *password*, *outside_light*, *inside_temp*, *pc*

Figure 5: smartSchoolSystem Module Variables

The top module of the smart school system is represented by *smartSchoolSystem*. The module has the set of variables shown in Figure 5. For example, person variable represents anyone who is allowed to enter the building (employee, visitor, or student). The *main_door* variable represents the main door for people to enter the school building, and *exit_doors* represents the doors which will be opened in case of fire/smoke. The *username* and *password* variables represent the values of the person's username and password, and these values must be unique for each individual person.

In any TLA+ module, declaring the *Init* function, invariants, and *Next* function is a must. The *Init* function represents the initial values of the system's variables, the invariants represent the limitations and conditions in the system, and the *Next* function represents the next-state action.

In the *smartSchoolSystem* module, we declared the *Init* function with the range of all possible values of each variable in the module as shown in Figure 6. For example, *main_door* variable may take on either open or closed values only. Similarly, *person* variable may take either student, employee, or visitor values in this spec. The *pc* variable represents the current state, and *pc'* represents the next state as we will see later in the *spec*.

To ensure that the school building in our system is always safe in case of fire, we designed our system to have a smoke detection sub-system working all the time in and out of school hours. To apply this safety functionality in our system, we added a system invariant and called it *safe* as shown in Figure 7. This invariant

guarantees that the smoke detection system will be working all the time.

```
Init ≜ Global variables
  ∧ objects = {"person", "lighting_sys",
               "main_door", "HVAC_sys", "smoke_detect_sys", "exit_doors",
               "login_sys", "username", "password", "outside_light", "inside_temp"}
  ∧ person ∈ {"student", "employee", "visitor"}
  ∧ lighting_sys ∈ {"low_mode", "high_mode", "off", "on"}
  ∧ HVAC_sys ∈ {"heat_on", "AC_on", "off", "on"}
  ∧ smoke_detect_sys ∈ {"on"}
  ∧ main_door ∈ {"open", "closed"}
  ∧ exit_doors ∈ {"open", "closed"}
  ∧ login_sys ∈ {"ready", "reTry"}
  ∧ username ∈ {"correct", "wrong"}
  ∧ password ∈ {"correct", "wrong"}
  ∧ outside_light ∈ {"clear", "cloudy"}
  ∧ inside_temp = defaultInitValue
  ∧ pc = "enter_school"
```

Figure 6: Init Function

$$safe \triangleq smoke_detect_sys = \text{“on”}$$

Figure 7: System Invariant

The function *enter_school* is the function that represents any of the employees, students, and visitors entering the school. This function requires the person desiring entry to enter a valid username and password as input in order to open the main door for entrance as an output. If the username and/or the password is wrong, the main door will not open, and the system will ask the person to retry to enter valid values. Figure 8 shows the *enter_school* function.

```
enter_school == /\ pc = "enter_school"
                /\ IF person  = "student" /\ username = "correct" /\ password = "correct"
                   THEN /\ main_door' = "open"
                        /\ UNCHANGED login_sys
                   ELSE /\ IF person = "student" /\ username = "wrong"
                           THEN /\ main_door' = "closed"
                                /\ login_sys' = "reTry"
                           ELSE /\ IF person = "student" /\ username = "correct" /\ passw
                                   THEN /\ main_door' = "closed"
                                        /\ login_sys' = "reTry"
                                   ELSE /\ IF person  = "employee" /\ username = "corr
                                           THEN /\ main_door' = "open"
                                                /\ UNCHANGED login_sys
                                           ELSE /\ IF person = "employee" /\ usernar
                                                   THEN /\ main_door' = "closed"
                                                        /\ login_sys' = "reTry"
                                                   ELSE /\ IF person = "employee"
                                                           THEN /\ main_door'
                                                                /\ login_sys'
                                                           ELSE /\ IF person
                                                                   THEN /\
```

Figure 8: enter_school Function

```
Smoke ≜ ∧ pc = "Smoke"
        ∧ IF main_door = "open"
          THEN ∧ IF smoke_detect_sys = "smoke"
                 THEN ∧ exit_doors' = "open"
                 ELSE ∧ TRUE
                      ∧ UNCHANGED exit_doors
               ∧ pc' = "Smoke"
          ELSE ∧ pc' = "Light"
               ∧ UNCHANGED exit_doors
        ∧ UNCHANGED ⟨objects, person, lighting_sys, HVAC_sys,
                     smoke_detect_sys, main_door, login_sys, username,
                     password, outside_light, inside_temp⟩
```

Figure 9: Smoke Function

We set *smoke_detect_sys* to be *on* all the time as an invariant in our system as mentioned before. If there is smoke detected from the smoke detection sub-system, in this case *smoke* function will be *on* and all exit doors will open automatically to allow all people inside the school to leave immediately for their safety. OF course, the smoke detection sub-system will have a built-in alarm in case smoke is detected. Figure 9 shows the *Smoke* function.

Once the first person enters the school, the lighting and HVAC sub-systems will automatically start working. In the light function, the light sensor will start sensing the natural *outside_light* brightness to control the light inside the building. The *outside_light* could have two values, clear or cloudy. If it is clear outside, the lights inside the building will be turned on by the system in a *low_mode*. If it is cloudy outside, the lights inside the building will be turned on by the system in a *high_mode*. This lighting sub-system will help in managing and controlling power consumption in the system. Figure 10 shows the *Light* function. In practice there will be multiple light sensors for multiple sides of the building.

```
Light ≜ ∧ pc = "Light"
        ∧ IF main_door = "open" ∧ outside_light = "clear"
          THEN ∧ lighting_sys' = "low_mode"
          ELSE ∧ IF main_door = "open" ∧ outside_light = "cloudy"
                 THEN ∧ lighting_sys' = "high_mode"
                 ELSE ∧ TRUE
                      ∧ UNCHANGED lighting_sys
        ∧ pc' = "HVAC"
        ∧ UNCHANGED ⟨objects, person, HVAC_sys, smoke_detect_sys,
                     main_door, exit_doors, login_sys, username, password,
                     outside_light, inside_temp⟩
```

Figure 10: Light Function

The HVAC sub-system helps in controlling the temperature inside the school building. In the HVAC function, the temperature inside the building has been sensed by the temperature sensor and the HVAC sub-system acts based on that. If the *inside_temp* is more than 74 F, the HVAC sub-system will turn on the AC in the building. If the *inside_temp* is less than 69 F, the HVAC sub-system will turn on the Heat in the building. If the *inside_temp* is more than 69 F and less than 74 F, the HVAC sub-system will be turned off. Figure 11 shows the HVAC function.

```
HVAC ≜ ∧ pc = "HVAC"
       ∧ IF main_door = "open" ∧ inside_temp ≥ 74
         THEN ∧ HVAC_sys' = "AC_on"
         ELSE ∧ IF main_door = "open" ∧ inside_temp ≤ 69
                THEN ∧ HVAC_sys' = "Heat_on"
                ELSE ∧ IF main_door = "open" ∧ inside_temp > 69 ∧ inside_temp < 74
                       THEN ∧ HVAC_sys' = "off"
                       ELSE ∧ TRUE
                            ∧ UNCHANGED HVAC_sys
       ∧ pc' = "Done"
       ∧ UNCHANGED ⟨objects, person, lighting_sys, smoke_detect_sys,
                    main_door, exit_doors, login_sys, username, password,
                    outside_light, inside_temp⟩
```

Figure 11: HVAC Function

Figure 12 includes housekeeping functions which are essential for writing good specifications for a system. Some of these functions must be in any TLA+ spec, e.g., the *Next* function. The *Next* function enables collection and execution of all functions in the spec and moving to the next state in the system after initialization. The Termination function guarantees the

termination when *pc* reaches *Done* state. In TLA+, the Spec function is the main function that is responsible to run all system specifications in the main execution of the system.

$$Terminating \triangleq pc = \text{“Done”} \wedge \text{UNCHANGED } vars$$

$$Next \triangleq enter_school \vee Smoke \vee Light \vee HVAC \vee Terminating$$

$$Spec \triangleq Init \wedge \Box[Next]_{vars}$$

$$Termination \triangleq \Diamond(pc = \text{“Done”})$$

Figure 12: Terminating, Next, and Spec functions

In order to verify that the module is correct and there are no syntax errors in our system, we have used the TLA+ toolbox to write these specifications and save the module. Figure 13 shows the parsed model. As shown in the figure, the green box on the bottom right corner of the screen proves that this model is correctly parsed via the TLA+ toolbox. The next section will describe the system that will be verified through the TLC model checker.

```
155
156 Smoke == /\ pc = "Smoke"
157          /\ IF main_door = "open"
158                THEN /\ IF smoke_detect_sys = "smoke"
159                           THEN /\ exit_doors' = "open"
160                           ELSE /\ TRUE
161                                /\ UNCHANGED exit_doors
162                     /\ pc' = "Smoke"
163                ELSE /\ pc' = "Light"
164                     /\ UNCHANGED exit_doors
165          /\ UNCHANGED << objects, person, lighting_sys, HVAC_sys,
166                          smoke_detect_sys, main_door, login_sys, username,
167                          password, outside_light, inside_temp >>
168
169 Light == /\ pc = "Light"
170          /\ IF main_door = "open" /\ outside_light = "clear"
171                THEN /\ lighting_sys' = "low_mode"
172                ELSE /\ IF main_door = "open" /\ outside_light = "cloudy"
173                           THEN /\ lighting_sys' = "high_mode"
174                           ELSE /\ TRUE
175                                /\ UNCHANGED lighting_sys
176          /\ pc' = "HVAC"
```

Figure 13: TLA+ Parsed Model for Smart School System

5. Formal Verification Using TLC

In the previous section, we illustrated the smart school system specifications, and we parsed the module correctly using the TLA+ toolbox. We did this in order to verify the model is correct and to validate our work. We used the powerful TLC model checker to debug a TLA+ specification. It checks the specification's invariance properties of its finite state model [13]. TLC checks for deadlock and the system invariants. In our TLC model, we set the safety invariant to keep the smoke detection sub-system working all the time, as shown in Figure 14. This means that this invariant will guarantee that the smoke detector will work all the time, whether there is someone in the school or not, day and night, to enhance the safety in the school building. As shown in Figure 14 as well, the TLC model checker has a feature to check for a deadlock in the design. It's an optional feature. In this model, we choose to check for deadlock and it returns that there is no deadlock.

After setting-up the TLC model checker, we ran it to verify our *smartSchoolSystem* module. Figure 15 shows the TLC model checker while running.

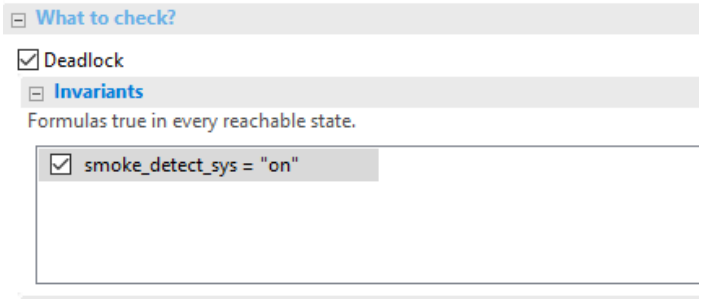


Figure 14: Safety Invariants Setup in TLC

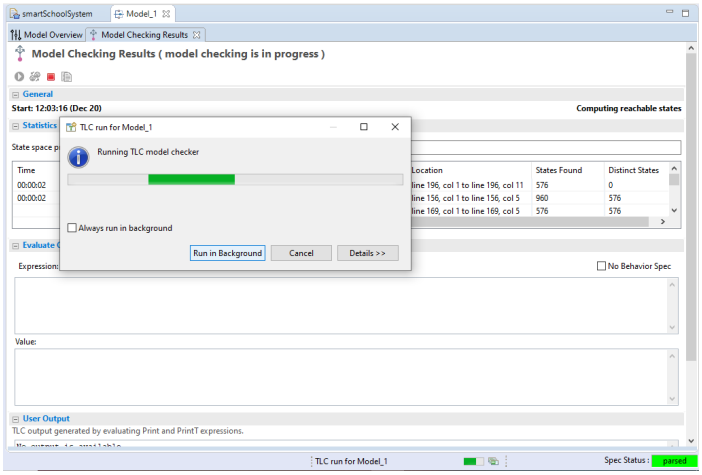


Figure 15: TLC Model Checker While Running

Figure 16 shows the final result for the *smartSchoolSystem* model. In order to validate a system using the TLC model checker, the model must be parsed, and the model checker should run to completion, and no errors should be detected. Our smart school system was verified and validated correctly using the TLC model checker since, as shown in Figure 16, the TLC model is parsed correctly with no errors, which proves our system's validation.

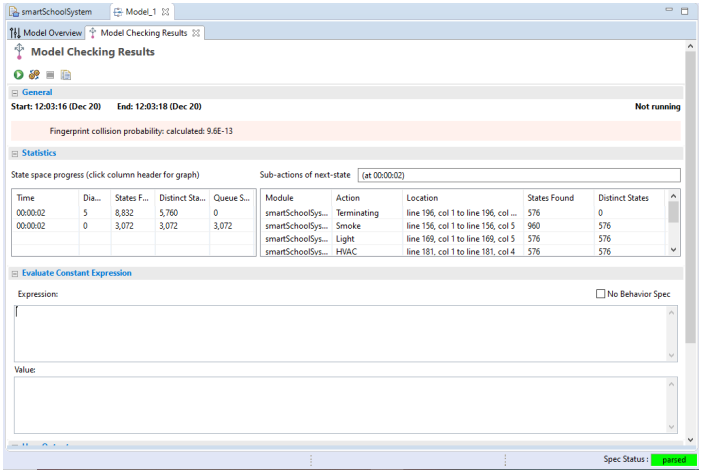


Figure 16: TLC Verification Model

6. Conclusion

This work is an extension of the work in [1]. We have modeled a smart school building system using UML and TLA+.

We used UML to define the system's components and to illustrate the sequence of actions in the system. TLA+ defines the system's specifications and the system's behavior. The TLA+ toolbox was used to capture the system's behavior and to parse the model. The final result was to verify the model using the TLC model checker. The model was successfully verified and validated with the TLC model checker. We used formal methods to validate our design and to make sure there are no errors in the design. In this design a failure may still happen**.** For example, if we simulated the design and built the school from this model, we could have a failure in the lighting system because in practice we will need a light sensor in each room. But our design assumes that the light will be the same in the whole building. This design is an initial and general design to show and explain our methodology, details such as the more complex lighting system would be needed in the final design.

To enhance security and safety in the system, the system requires each person who enters the building to login by entering a valid username and password. To enhance the safety as well, the smoke detection sub-system is working all the time. To control the power in the system, the lighting sub-system will use natural light when possible to reduce the power consumption in the building.

In future, we will work on improving the system by adding more sub-systems to it to enhance system security and safety.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Figure 2 and Figure 4 were made using chart-making tools at www.lucidchart.com and the authors would like to thank the website. Figures 5,6,7,9,10,11,12 were taken from the TLA+ tool that uses LaTeX translator to generate pretty-printed specs. Figures 8,13,14,15,16 were taken from the TLA+ toolbox. The authors would like to thank Mr. Leslie Lamport for the TLA+ tool which is essential in this work. The authors would also like to thank the University of Cincinnati for its support.

References

[1] N.H. Obeidat, C. Purdy, "Modeling a smart school building system using UML and TLA+," in Proceedings - 3rd International Conference on Information and Computer Technologies, ICICT 2020, Institute of Electrical and Electronics Engineers Inc.: 131–136, 2020, doi:10.1109/ICICT50521.2020.00028.

[2] P.S. Aithal, Smart Library Model for Future Generations (June 30, 2016). International Journal of Engineering Research and Modern Education (IJERME),1(1), 693-703. ISSN (Online): 2455 - 4200, (2016)., Available at SSRN: https://ssrn.com/abstract=2822978

[3] A. Rehman, S. Latif, N.A. Zafar, "Formal modeling of smart office using activity diagram and non deterministic finite automata," in 2019 International Conference on Information Science and Communication Technology, ICISCT 2019, Institute of Electrical and Electronics Engineers Inc., 2019, doi:10.1109/CISCT.2019.8777444.

[4] F.K. Santoso, N.C.H. Vun, "Securing IoT for smart home system," in Proceedings of the International Symposium on Consumer Electronics, ISCE, Institute of Electrical and Electronics Engineers Inc., 2015, doi:10.1109/ISCE.2015.7177843.

[5] Z. Yu, Y. Liang, B. Xu, Y. Yang, B. Guo, "Towards a smart campus with mobile social networking," in Proceedings - 2011 IEEE International Conferences on Internet of Things and Cyber, Physical and Social Computing, iThings/CPSCom 2011, 162–169, 2011, doi:10.1109/iThings/CPSCom.2011.55.

[6] A. Clark, A. Evans, "Foundations of the Unified Modeling Language.," In Proceedigs of the 2nd Northern Formal Methods Workshop. Springer, 1997.

[7] L. Pierre, T. Kropf, Correct Hardware Design and Verification Methods: 10th IFIP WG10.5 Advanced, 1999.

[8] S. Tasiran, Y. Yu, B. Batson, S. Kreider, Using formal specifications to monitor and guide simulation: Verifying the cache coherence engine of the Alpha 21364 microprocessor, 2002.

[9] C. Newcombe, "Why Amazon chose TLA+," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag: 25–39, 2014, doi:10.1007/978-3-662-43652-3_3.

[10] S. Latif, A. Rehman, N.A. Zafar, "Modeling of sewerage system linking UML, automata and TLA+," in 2018 International Conference on Computing, Electronic and Electrical Engineering, ICE Cube 2018, Institute of Electrical and Electronics Engineers Inc., 2019, doi:10.1109/ICECUBE.2018.8610971.

[11] S. Latif, A. Rehman, N.A. Zafar, "NFA based formal modeling of smart parking system using TLA +," in 2019 International Conference on Information Science and Communication Technology, ICISCT 2019, Institute of Electrical and Electronics Engineers Inc., 2019, doi:10.1109/CISCT.2019.8777445.

[12] J.-R. Abrial, "Formal methods in industry," in Proceeding of the 28th international conference on Software engineering - ICSE '06, Association for Computing Machinery (ACM), New York, New York, USA: 761, 2006, doi:10.1145/1134285.1134406.

[13] L. Lamport, Specifying Systems Preliminary Draft, 2001.

[14] R. Adner, D. Levinthal, "Demand heterogeneity and technology evolution: Implications for product and process innovation," Management Science, **47**(5), 611–628, 2001, doi:10.1287/mnsc.47.5.611.10482.

[15] K. Furdik, G. Lukac, T. Sabol, P. Kostelnik, "The Network Architecture Designed for an Adaptable IoT-based Smart Office Solution," International Journal of Computer Networks and Communications Security, **1**(6), 216–224, 2013.

[16] Y. Geng, C.G. Cassandras, "New 'smart parking' system based on resource allocation and reservations," IEEE Transactions on Intelligent Transportation Systems, **14**(3), 1129–1139, 2013, doi:10.1109/TITS.2013.2252428.

[17] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, "Internet of things for smart cities," IEEE Internet of Things Journal, **1**(1), 22–32, 2014, doi:10.1109/JIOT.2014.2306328.

[18] B. Hirsch, J. WP. Ng., "Education beyond the cloud: Anytime-anywhere learning in a smart campus environment." IEEE International Conference for Internet Technology and Secured Transactions. , 2011.

[19] Smart Campus: The Developing Trends of Digital Campus-- 《Open Education Research》2012 年 04 期, Mar. 2021

[20] M.R.M. Veeramanickam, M. Mohanapriya, "IOT enabled Futurus Smart Campus with effective E-Learning : i-Campus," GSTF Journal of Engineering Technology (JET), **3**(4), 2016, doi:10.5176/2251-3701_3.4.164.

[21] L. Pocero, D. Amaxilatis, G. Mylonas, I. Chatzigiannakis, "Open source IoT meter devices for smart and energy-efficient school buildings," HardwareX, **1**, 54–67, 2017, doi:10.1016/j.ohx.2017.02.002.

[22] D. Amaxilatis, I. Chatzigiannakis, G. Mylonas, Design and Implementation of a Platform for Smart Connected School Buildings, 2015.

[23] Y. Qu, H. Wang, S.M. Lun, H.D. Chiang, T. Wang, "Design and implementation of a Web-based Energy Management Application for smart buildings," in 2013 IEEE Electrical Power and Energy Conference, EPEC 2013, IEEE Computer Society, 2013, doi:10.1109/EPEC.2013.6802931.

[24] M. Brogan, A. Galata, The VERYSchool Project: Valuable EneRgY for a smart School-Intelligent ISO 50001 Energy Management Decision Making in School Buildings.

[25] P.G. Larsen, W. Pawlowski, "The formal semantics of ISO VDM-SL," Computer Standards and Interfaces, **17**(5–6), 585–601, 1995, doi:10.1016/0920-5489(95)00026-Q.

[26] The Object-Z Specification Language - Graeme Smith - Google Books, Mar. 2021.

[27] D. Jackson, "Alloy: A Lightweight Object Modelling Notation," ACM Transactions on Software Engineering and Methodology, **11**(2), 256–290, 2002, doi:10.1145/505145.505149.

[28] T. De Champs, B. Abdulrazak, H. Pigot, M. Ouenzar, M. Frappier, B. Fraikin, "Pervasive safety application with model checking in smart houses: The INOVUS intelligent oven," in 2011 IEEE International Conference on Pervasive Computing and Communications Workshops, PERCOM Workshops 2011, 630–635, 2011, doi:10.1109/PERCOMW.2011.5766965.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 829-836 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design and Development of an Advanced Affordable Wearable Safety Device for Women: Freedom Against Fearsome

Israt Humaira[1], Kazi Arman Ahmed[2], Sayantee Roy[2], Zareen Tasnim Safa[2], F. M. Tanvir Hasan Raian[1], Md. Ashrafuzzaman[1,*]

[1]*Department of Biomedical Engineering, Military Institute of Science and Technology, Mirpur Cantonment, Dhaka 1216, Bangladesh*

[2]*Department of Industrial and Production Engineering, Military Institute of Science and Technology, Mirpur Cantonment, Dhaka 1216, Bangladesh*

ARTICLE INFO

Article history:
Received: 16 December, 2020
Accepted: 28 January, 2021
Online: 04 April, 2021

Keywords:
Safety Device
Women
Surveillance
Wearable locket and bracelet
Bluetooth

ABSTRACT

Harassment and violence against women have become one of the social security problems in Bangladesh. In this paper, we aim to develop safety devices for women named BOHNNI and BADHON which resemble legitimate jewelry. We used a microcontroller for the hardware device to make it most decisive and less immoderate. BOHNNI, a locating device, is the imitator of a locket including a voice recognizer, Bluetooth, Arduino, GPS, and GSM module. BADHON which imitates a bracelet is a rescue device for the victim whenever she thinks of herself being in a very deliberate situation. Both devices are activated by the user's voice commands and also by a manual switch. The devices are aesthetically designed which will make the users enthusiastic to wear them. The device will generate messages to the predefined relative's numbers with the victim's location and relevant surrounding information. The device can also be used as a self-defending weapon as it can produce a shock up to 10 mA with an interval of two seconds which can temporarily paralyze or freeze a person. After calculating, we have obtained the lowest response time of BOHNNI and BADHON which is 1.95s, and the highest accuracy level of 91.67% in different situations that ensure the superlative performance level of our devices. We found our device as an all-in-one device that combines all the features in it regarding safety.

1. Introduction

Bangladesh has met the eligibility requirements for shifting from 'Least Developed Country' to 'Developing Country' status, the UN says [1]. Women's empowerment is also on the rise, but incidents of various forms of violence against women, including dowry torture, sexual assault, abduction, and rape continue to take place extensively in Bangladesh. The "State of Rights Implementation of Women Ready Made Garment Workers" study, conducted by Karmojibi Nari and Care Bangladesh, shows that about 12.7% of workers face sexual harassment in their workplaces [2]. Around 94% of women face harassment in public transport [3]. Bangladesh Mahila Parishad (October 2019) statistics indicated that in the last year at least 3,918 women have been victims of various forms of violence across the country [4]. [4]. Surviving in these situations becomes a great challenge for women. By adopting modern technology and devices to protect them from their oppressors, the defense strategy used by women needs to be revolutionized [5]. This current situation has encouraged us to come up with a solution to harassment and violence against women. This work has been done to ensure safety for women moving alone. It is not the only solution to a problem but also comes with wearable safety products. These are very handy safety devices that can be used by any woman. GPS and GSM modules are used to give alert calls and messages to the victim's previously established numbers and nearest police station numbers including the instant location and that location can be tracked immediately or later using our device "BOHNNI" and "BADHON". The devices that we have built can be activated by using vocalized passwords which ensures the surety of access to the devices in any critical situation. It is very easy to operate and anyone can operate it as it can be used manually or by using voice commands. This product is multi-functional. It makes it easier and faster to catch the criminals as the lowest response time of BOHNNI and BADHON is 1.95s and also works as a self-defender. The highest accuracy level of 91.67% was obtained in a

*Corresponding Author: Md. Ashrafuzzaman, Email: ashezaman@gmail.com

https://dx.doi.org/10.25046/aj060296

different situation that ensures the performance level of our devices. It is intelligent enough because it can understand the user's voice command and can be used by people of any age. Materials used there are recyclable. We have maintained the safety of the user while designing the device with polymer plastic material to protect the user from accidental shock. The notable accomplishment of our design is the shocking module which can give a shock to the harasser up to 10 mA with an interval of 2 seconds which help the victim to escape from the location. We believe that this device can bring a huge change in violence and harassment against women. It can be a blessing for women and also develop their social norms in our country.

2. Related Works

The project described in [6] developed a wearable device for the protection of women by analyzing physiological signals in association with body position. Galvanic skin resistance and body temperature are analyzed as the physiological signals. Real-time monitoring of data is obtained by wirelessly sending data to an open-source Cloud Platform. In critical situations, this device sends notifications to designated individuals. A proposal is available in [7] anticipated "FEMME" as a safety device specifically designed for women in situations of threat. Using Bluetooth, both the device and the smartphone are synced. With instant location information, it can record audio, give an alert call and message to the pre-set contacts. A hidden camera detector is a visible feature here as well. A smart Intelligent Security System was proposed in [8]. The system features a band on the wrist associated with a pressure switch. The screaming alarm and tear gas mechanisms are appointed for self-defending purposes. This system also sends the appropriate locations and messages to the emergency contacts and uses live streaming video to figure out the assaulter [8]. It was aimed to create a common system that incorporates a wearable computer and acts as a safeguard for women facing any danger in [9]. Body area sensors, Bluetooth communication, GPS, SMS and MMS, Internet connection, and a mobile database system are used for the composition of this system. Another proposal named "Suraksha" is developed in [10] to flash a warning to provide the police with an immediate location of the distressed victim so that the incident could be prevented easily. A smart intelligent device is proposed in [11] where a wearable smart band and a secret webcam linked via Bluetooth are integrated into the device. Women's information such as call log, messages, motion, pulse measurement, blood oxygen levels, heartbeat rating can be tracked by the application. It automatically generates signals to the predefined smartphones and the nearest police stations when the SOS present in the smart band is pressed continuously. A smart women's security system using Radio Frequency Identification (RFID) and Global Positioning System (GPS) was established in [12]. In scanning the information, the RFID reader is used and assigns that information to the microcontroller. In [13], the team developed a device that can use GPS to track the location of women and send emergency messages using GSM. The pulse sensor checks the victim's pulse and the device also sends the current GPS location to the ambulance every 10 seconds in the form of SMSs in abnormal health situations. The authors in [14] developed a device that will be clipped to the user's footwear and can be carefully triggered four times by tapping one foot behind the other. A device is developed in [15] that reads and creates patterns such as body temperature and pulse rate automatically during running, including temperature and pulse sensors. If the readings are higher than normal readings, more than one person will automatically get a call and message. In non-dangerous conditions, data is first collected by sensors to train the algorithm. A passive continuous monitoring system is proposed in [16] using both survivor-attached biosensors and machine learning techniques. According to the current status of wearable and biomedical device technology, the monitoring structure of the system supervises a lot of bio-signals. In [17] a system is designed to detect location with GPS and GSM mechanisms to pass the current location to any of the trusted contacts as a Google map link and services are provided from that moment on to track the locations to save the individual. A wearable system was designed in [18] that resembles a normal watch with a button that can be pressed in an unpleasant situation to activate the system. Using a Global Positioning System (GPS) sensor, the system monitors the victim's location. The system also includes a screaming alarm using a real-time clock. A smart band is developed in [19] that gets activated to send the GPS location to the ICE contacts and police control rooms by tapping on the screen twice. A micro USB charge is supported by the device and two metal points generate the shock on the top of the band screen. A project 'watch me' is developed in [20] for a purpose whenever a woman or child wearing a smartwatch called 'watch me' is exposed to sexual or vulnerable assault, the device works. It will automatically make a call to the registered contacts of the victim and will also detect the nearby police station via GPS / GSM to make a ring there so that police can arrive at the spot soon by tracking the GPS. The developers in [21] developed a female safety jacket that flashes a warning giving the police an instant location of the distressed victim so that the incident can be avoided and the culprit captured. An efficient woman's security recognition system has been introduced in [22] consisting of an ARDUINO NANO Controller, a Bluetooth module, a Taser, and a versatile Android. It monitors the victim's location using GPS and sends emergency messages via GSM. An electronic device including biometric touch sensors has been developed in [23] that generates extremely high voltage pulses at around 3 KV range. The high voltage causes body pain, numbness, and balance loss. From its use, no permanent injury is noticed. For women, a self-defense system using Raspberry-pi 0 controller and GSM, GPS, Temperature sensor, Heartbeat Sensor has been developed by [24]. The location of a place via GPS is indicated by a single click on this device and a message consisting of the location URL is sent to the registered contacts to assist them in dangerous situations. In terms of latitude and longitude, a rapid response mechanism was developed by [25] that helps women when they press the button attached to the device and the location information is sent as an SMS alert to a few predefined emergency contacts. A smart security wearable device for women was implemented by [26] in the form of a smart ring (SMARISA) comprising Raspberry Pi Zero, camera, buzzer, and button. The victim can activate the device by clicking on a button that fetches her current location and also captures the attacker's image.

For all sections of women in Bangladesh, a user-friendly mobile-based approach 'BONITAA' was developed by [27] with several unique features, primarily aimed at supporting victims of rape by facilitating their ways to obtain justice with legal solutions,

physical and mental support, etc. Another mobile app 'SafetiPin' was introduced by [28] where the device can gather crowd-source data and information on insecurity in cities using several features such as GPS navigation, digital camera, high-speed internet connectivity, and many more. The team in [29] created a women's security Android application. Using the Wikitude SDK, the women's safety app can be integrated with augmented reality that helps to place articles and reviews related to the current location. A smartphone application prototype was introduced in [30], which is primarily a safety decision aid designed to help women estimate the risk in their abusive relationships, set safety priorities, and develop a personalized safety plan. A "Virtual Friendly" device was created by [31], consisting mainly of GPS, GSM modem, microcontroller, RF transceiver, temperature sensor, voice recognizer to provide trouble-safety. The GPS receiver obtains location information in the form of latitude and longitude from satellites. The temperature sensor senses the temperature of the body and sends an SMS alert in the event of a low temperature. A portable device was developed by [32] incorporating a pressure switch that gets activated by compressing it. The pressure sensor senses the pressure instantly and a conventional SMS, with the location of the victim being sent to phone numbers stored in the device followed by a call to their parents/guardians [32]. A smart device is developed by [33] that can be clipped to a woman's footwear and can be discreetly triggered within 5 seconds by tapping her feet 10 or more times. It will also sense the unnatural rise or fall of her heartbeat and the values will be displayed on the smartphone of the victim.

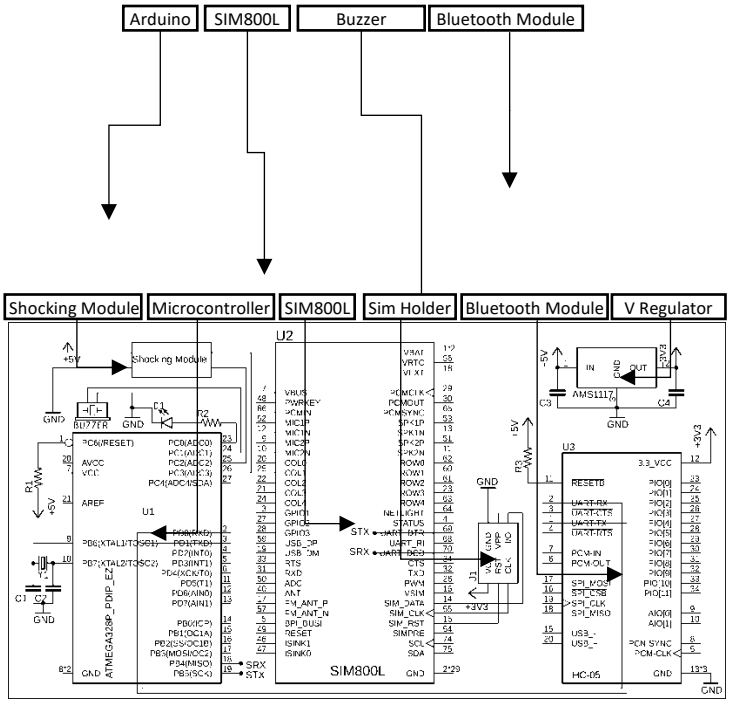


Figure 1: Circuit Diagram of BOHNNI and BADHON

3. Materials & Methods

3.1. Hardware Components

Materials that have been used to produce these devices are non-toxic, heat & current resistant. The core of the device is embedded with the PCB circuit board consisting of the circuit of GSM, Bluetooth, Arduino, voice recognition module, and camera module. The primary components of the circuits are the voice recognition module which is used to capture the user's voice & turning on the operation of the device, Arduino to send messages & location to the programmed numbers, GSM Module which is used to send messages & location to the programmed numbers, Bluetooth Module to send & receive a signal during the operation of the device and lastly the Camera Module to save pictures during the operation of the device.

3.2. Design

The circuit is designed in Proteus software. The microcontroller is the processing unit that processes all data from the sensor and voice recognition module. It also sends the necessary command to GPS and GSM units.

We have designed the initial outer designs of the device BOHNNI in SolidWorks that is represented below in Figure 2:



Figure 2: 3D Design of BOHNNI & BADHON

3.3. Fabrication of Prototype

'BOHNNI' and 'BADHON' are fabricated initially on bread-board connecting all parts including Arduino, LED, voice recognition module, Bluetooth module, GSM, and a GPS module. 'BADHON' consists of a shocking module, Arduino, a Bluetooth device, a video camera, and a buzzer. Initially, the video camera was not used in the prototype. During mass production or making a real-time product, the video camera will be included. A visual representation of our prototype 'BOHNNI' and 'BADHON' are given below in Figure 3 and Figure 4:

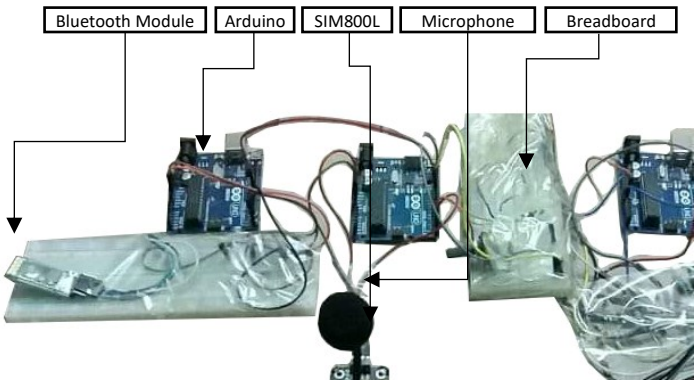


Figure 3: Prototype of BOHNNI

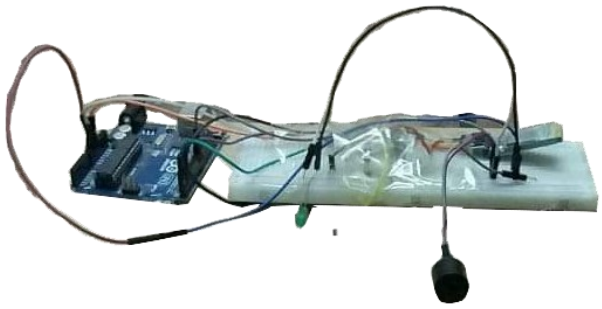


Figure 4: Prototype of BADHON

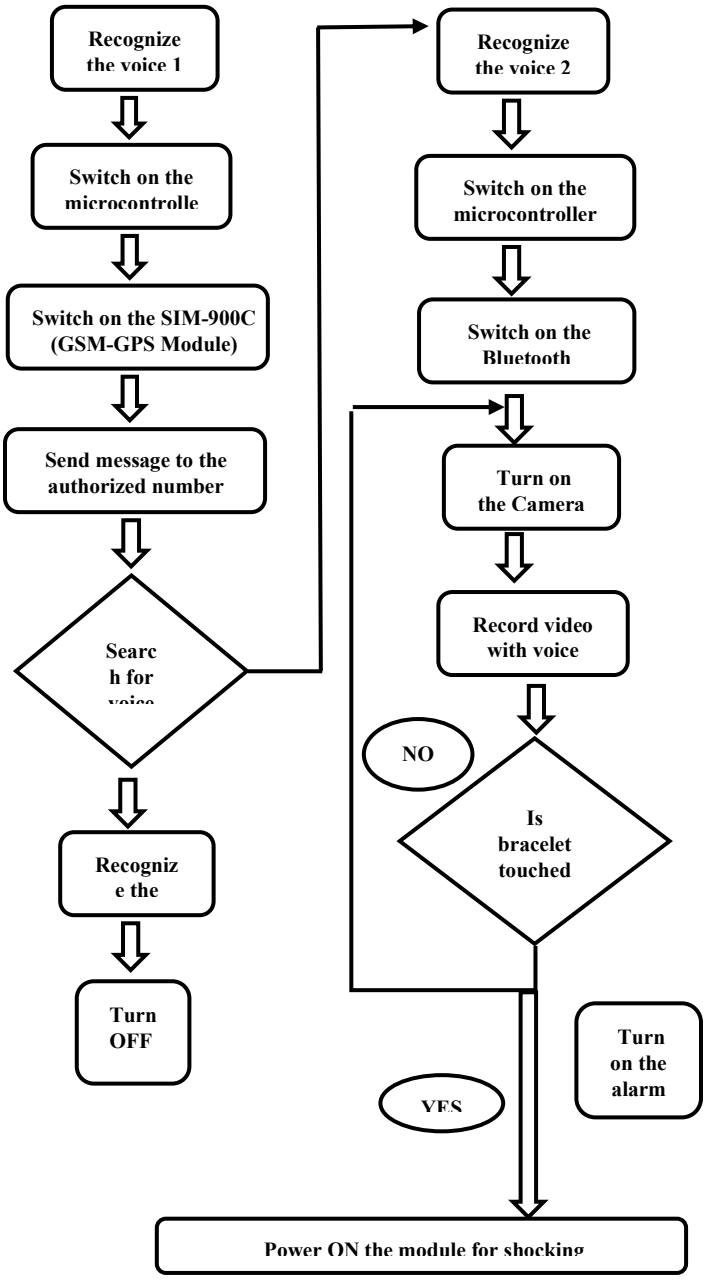


Figure 5: Flow chart of the operating process of BOHNNI and BADHON

3.4. Working Procedure

For WOMEN we are introducing a system that resembles some normal jewelry like locket (BOHNNI) and bracelet (BADHON). BOHNNI is a locating device including a voice recognizer, Bluetooth, Arduino, GPS, and GSM module. The device will be activated using the voice password set by the user. As it activates it will send messages to the user's family member's numbers which were previously set by her and the nearest police stations with the link of her location. This will provide vital protection to the victim. BADHON which reassembles a bracelet activates another voice password set by the user. It is a rescue device for the victim if she is in a very critical situation. This device includes a shocking module, Arduino, Bluetooth device, video camera, and a buzzer. Using another voice code in BOHNNI will activate the device BADHON. After it activates, the video camera will capture video of the victim's surroundings as it can be used as evidence in the future. The shocking system will be activated which can give a shock to the harasser up to 10mA (10 mA) with an interval of 2 sec which can make him paralyze or freeze his muscle for few minutes and help the victim to escape from the location. As the device will be made of polymer plastic the user will not be affected by the shock. It also has a buzzer that activates when it is touched. Then the buzzer will buzz 120 decibel sound to attract people from further. This device can also be used for children when aesthetically designed which will make the children enthusiastic to wear them. In this way, the guardian can always be aware of their child when they are not with them.

4. Results & Discussion

4.1. Design of Experiments

The design of the experiment runs by putting BOHNNI and BADHON together in different combinations. In table 1, for the first five rows, BADHON was placed on the right arm where BOHNNI was placed on the left arm, right wrist, left wrist, neck, and chest respectively. For the next four rows, BADHON was placed on the left arm where BOHNNI was placed on the right wrist, left wrist, neck, and chest respectively. Again, for the next three rows, BADHON was placed in the right wrist and BOHNNI was placed on the left wrist, neck, and chest respectively. For the last two rows, BADHON was placed in the left wrist and BOHNNI was placed in the neck and chest respectively. In this total of 14 rows, BOHNNI and BADHON were subjected to different combinations to check the response time of each for every possible combination. The experimental design has been formatted in Table 1 below.

Table 1: Different combinations and placements of BOHNNI and BADHON

Right Arm	Left Arm	Right Wrist	Left Wrist	Neck	Chest
BADHON	BOHNNI	--	--	--	--
BADHON	--	BOHNNI	--	--	--
BADHON	--	--	BOHNNI	--	--
BADHON	--	--	--	BOHNNI	--
BADHON	--	--	--	--	BOHNNI
--	BADHON	BOHNNI	--	--	--
--	BADHON	--	BOHNNI	--	--
--	BADHON	--	--	BOHNNI	--

--	BADHON	--	--	--	BOHNNI
--	--	BADHON	BOHNNI	--	--
--	--	BADHON	--	BOHNNI	--
--	--	BADHON	--	--	BOHNNI
--	--	--	BADHON	BOHNNI	--
--	--	--	BADHON	--	BOHNNI

Response times have been calculated by following the combinations of BOHNNI & BADHON's placement & placed in Table 2.

Table 2: Response time for different placement combinations of BOHNNI and BADHON

Right Arm	Left Arm	Right Wrist	Left Wrist	Neck	Chest	Total Response Time (s)
0.55	1.45	0	0	0	0	2
0.56	0	1.85	0	0	0	2.41
0.58	0	0	1.85	0	0	2.43
0.57	0	0	0	1.4	0	1.97
0.56	0	0	0	0	1.4	1.96
0	0.58	1.85	0	0	0	2.43
0	0.56	0	1.85	0	0	2.41
0	0.57	0	0	1.4	0	1.97
0	**0.55**	**0**	**0**	**0**	**1.4**	**1.95**
0	0	0.59	1.85	0	0	2.44
0	0	0.57	0	1.4	0	1.97
0	**0**	**0.55**	**0**	**0**	**1.4**	**1.95**
0	0	0	0.57	1.4	0	1.97
0	**0**	**0**	**0.55**	**0**	**1.4**	**1.95**

The total response time for BOHNNI and BADHON was obtained by adding the response time of BOHNNI and BADHON for each combination. From table 2, we have obtained the lowest response time that is 1.95s.

Findings from this data are the precise locations to place the device in the body. We found three locations with the minimum response time. Those combinations are-

- Right Wrist (BADHON) & Chest (BOHNNI)
- Left Wrist (BADHON) & Chest (BOHNNI)
- Left Arm (BADHON) & Chest (BOHNNI)

From these findings, we have designed our devices so that they can be worn on the wrist & chest to obtain the least response time.

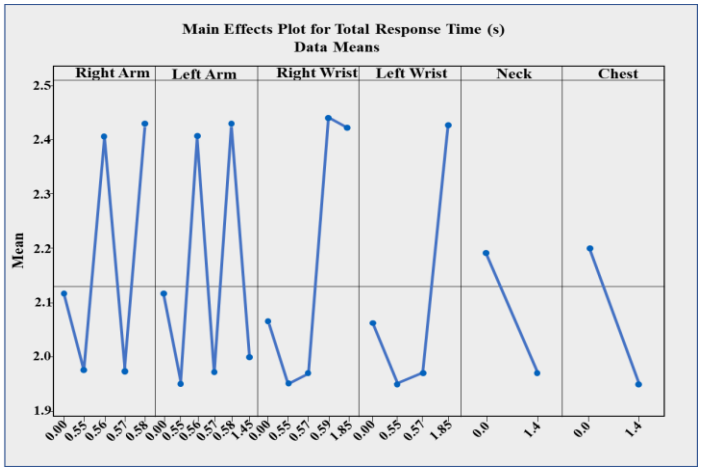


Figure 6: Main effect plots for Total Response Time

Table 3: The response of the device in different situations

Ser No	Different Situation	Tested Times	True (+)ve	False (-)ve	Sensitivity	% of Error
1.	Inside a room	12	11	1	0.92	**8.33%**
2.	Inside a bus	12	8	4	0.67	**33.33%**
3.	Inside a CNG	12	9	3	0.75	**25.00%**
4.	Inside a car	12	11	1	0.92	**8.33%**
5.	Walking on the street during morning	12	10	2	0.83	**16.67%**
6.	Walking on the street during noon	12	9	3	0.75	**25.00%**
7.	Walking on the street during night	12	10	2	0.83	**16.67%**
8.	Walking on the street during midnight	12	11	1	0.92	**8.33%**

% of Error= (1-Sensitivity) ×100 %

4.2. Accuracy calculation

Sensitivity analysis is the study of how to divide and allocate the uncertainty in the output of a mathematical model or system (numerical or otherwise) into its inputs to different sources of uncertainty. Uncertainty analysis, which focuses more on quantifying uncertainty and propagating uncertainty, is a related practice; ideally, uncertainty and sensitivity analysis should be performed in tandem. We have checked the sensitivity and percentage of error in various situations. The formula of Sensitivity (S_n) [34] is given below,

$$S_n = \frac{\textbf{True positive}}{\textbf{True positive} + \textbf{False negative}}$$

The obtained results using the Sensitivity formula are given in Table 3.

We have also obtained a bar diagram regarding the results. Here we have shown the percentage of error and accuracy in different situations.

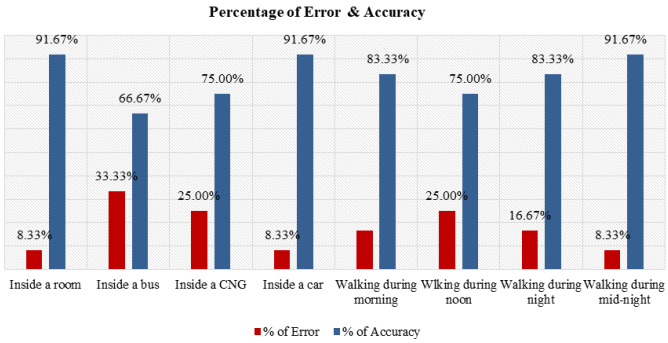


Figure 7: Bar chart of percentage & accuracy

From the chart, we found out 91.67% accuracy occurred during three situations. Those are: -

a) Inside a room
b) Inside a car
c) Walking on the street during mid-night

Other situation's accuracy is higher than the error percentage which indicates the feasibility of the device.

4.3. Comparative Analysis

We have made a comparative analysis with two of the existing works "Suraksha" [10] and "Watch Me" [20]. The major achievements of our design that distinguish itself from others are the shocking system which can give a shock to the harasser up to 10 mA with an interval of 2 sec which can make him paralyze or freeze his muscle for few minutes and help the victim to escape from the location and also the fact that we have taken the safety of the user under account while designing the device with polymer plastic material to protect the user from the shock.

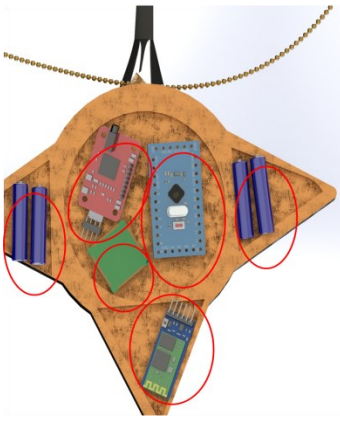
Table 4: Comparative analysis with existing works

Device	Suraksha	Watch Me	BOHNNI and BADHON
Country	India	India	Bangladesh
Features	Voice recognition, GSM/GPS module, Switch, force sensor.	Pulse rate sensor, motion sensor, GSM/GPS module, alarm.	Voice recognizer, GPS, GSM, buzzer, video camera, shocking system.
Product Components	Miniature device which can be embedded in jewelry or mobile phone etc.	Smartwatch	Wearable Locket, a bracelet made of polymer plastic.
Application	The device will activate by voice command or switch or if thrown by force then it will send an alert message to the police and registered number.	The sensor of the device will activate when the targeted heart rate for the targeted time is achieved, then will produce an alarm sound and send a notification to the cops and registered numbers.	The device will activate using the voice password set by the user. After activation, it will send messages to the user's emergency numbers and nearest police stations with the link of her location.
Limitations	There is no such feature to attack the attacker and also the device needs to be set with accessories.	The device is physically dependent.	Physical modifications can be done to ensure durability and the size of the locket can be reduced to some extent.

4.4. Mass Production Plan

Mass production plan for BOHNNI & BADHON starts with internal components placement. In BOHNNI, five components are needed to place into the inner circuit. From Fig 8 below, we can see the total weight of the components of BOHNNI is 26.1 gm & the outer portion of BOHNNI is made of a polymer of rubber and plastic for anti-shocking weighing 7.25 gm. The total weight of BOHNNI including all of its components is 33.35 gm which is much lower than the necklace that women usually wear.

From Fig 8, we can find the area of every component which defines the feasibility of mass production of this device. The device BADHON's weight & size is quite similar to the regular smartwatch and so we have not designed the positions of placement of its components but it will be made of polymer plastic to protect the user from getting sudden shocks.



Component Name	Quantity	Weight (gm)	Area (inch^2)
Arduino pro mini	1	1.8	0.7 × 1.3
Voice recognition module (Spiral)	1	1.8	3 × 2.5 × 0.4
Battery	4	16	0.132
GSM module	1	0.5	0.78 × 0.93
Bluetooth module	1	6	1.06 × 0.51

Figure 8: Inner components of BOHNNI including weight and size

5. Conclusions

In this twenty-first century, our country is experiencing a notable change in society because of women's empowerment. But the bitter truth is that these women are continuously being harassed every day and everywhere. So, women's safety issue has become a matter of great concern nowadays. For women, there are several safety devices available, but affordability for all has yet to be achieved. Keeping that in mind, we have designed our "BOHNNI" & "BADHON" to prevent harassment in our country. This device is simple structured, inexpensive, and feasible as well. This safety device is much more user-friendly & accurate during critical situations than other existing devices. It is very lightweight

and as it resembles jewelry, it can easily remain hidden from the attacker. We built the prototype & tested it during several situations. Our findings are the accuracy of the device which indicates the feasibility of using it. Our products- "BOHNNI" & "BADHON", will ensure a safe environment for women and children in all situations by providing the hidden camera, GPS and GSM module, shocking device, and voice recognizer. In our paper, we presented only the proto-type of a wearable safety locket and bracelet. Various health and activity monitoring sensor belts are already available on the market, but the design of an integrated safety device is completely different as very accurate calculations and designs are mandatory. Therefore, future work will aim to present both physical and appropriate modifications of the devices to improve their durability and prevent any damage to the hardware. With additional study and innovation, our project can be sufficiently competent to reduce the rate of crime against women and children.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to acknowledge the support of the Department of Biomedical Engineering, Military Institute of Science and Technology (MIST), Bangladesh.

References

[1] www.bdnews24.com, N.D.: Bangladesh eligible for UN 'developing country' status, https://bdnews24.com/bangladesh/2018/03/17/bangladesh-eligible-for-un-developing-country-status, last accessed 2020/10/29.

[2] Study: Rights of Bangladesh's female RMG workers still unmet, https://www.karmojibinari.org/dhakatribune/, last accessed 2020/10/29.

[3] Public Transport: 94pc women harassed, https://www.thedailystar.net/frontpage/public transport-94pc-women-harassed-1544506, last accessed 2020/10/29.

[4] How we violence against women in 2019,https://www.thedailystar.net/opiion/news/how-we-combated-violence-against-women-2019-1846438, last accessed 2020/10/29.

[5] D. Chitkara, N. Sachdeva and Y. Dev Vashisht, "Design of a women safety device," in IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Agra, **2016**, 1-3, doi: 10.1109/R10-HTC.2016.7906858.

[6] A. Jatti, M. Kannan, R. M. Alisha, P. Vijayalakshmi and S. Sinha, "Design and development of an IOT based wearable device for the safety and security of women and girl children," in IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, **2016**, 1108-1112, doi: 10.1109/RTEICT.2016.7808003.

[7] D. G. Monisha, M. Monisha, P. Gunasekaran, and S. Radhakrishnan, "Women safety device and application-FEMME," Indian Journal of Science and Technology, **9**(10), 1-6, 2016, doi:10.17485/ijst/2016/v9i10/88898.

[8] G. P. Miriyala, P.V.V.N.D.P Sunil, R. S. Yadlapalli, V. R. L. Pasam, T. Kondapalli, and A. Miriyala, "Smart Intelligent Security System for Women," International Journal of Electronics and Communication Engineering & Technology (IJECET), **7**(2), 41–46, 2016.

[9] S. Roy, A. Sharma, and U. Bhattacharya, "MoveFree: A ubiquitous system to providewomen safety," in 3rd International Symposium on Women in Computing and Informatics WCI '15, India, **2015**, 1-5. doi:10.1145/2791405.2791415.

[10] N. Bhardwaj and N. Aggarwal, "Design and Development of 'Suraksha'-A Women Safety Device," International Journal of Information & Computation Technology, **4**(8), 787-792, 2014.

[11] N. V. Kumar, and S. Vahini, "Efficient Tracking for Women Safety and Security using IOT," International Journal of Advanced Research in Computer Science, **8**(9), 328–330, 2017. doi:10.26483/ijarcs.v8i9.4915.

[12] S. M. Hussain, S. A. Nizamuddin, R. Asuncion, C. Ramaiah and A. V. Singh, "Prototype of an intelligent system based on RFID and GPS technologies for women safety," in 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, **2016**, 387-390, doi: 10.1109/ICRITO.2016.7784986..

[13] P. K. Verma, A. Sharma, D. Varshney, and M. Zadoo, "Women Safety Device with GPS, GSM and Health Monitoring System, **5**(3), 941-943, 2018.

[14] N. Viswanath, N. V. Pakyala and G. Muneeswari, "Smart foot device for women safety," in IEEE Region 10 Symposium (TENSYMP), Bali, **2016**, 130-134, doi: 10.1109/TENCONSpring.2016.7519391.

[15] Muskan, T. Khandelwal, M. Khandelwal and P. S. Pandey, "Women Safety Device Designed Using IoT and Machine Learning," in IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Guangzhou, **2018**, 1204-1210, doi: 10.1109/SmartWorld.2018.00210..

[16] I. Rodríguez-Rodríguez, J-. V. Rodríguez, A. Elizondo-Moreno, and P. Heras-González, "An Autonomous Alarm System for Personal Safety Assurance of Intimate Partner Violence Survivors Based on Passive Continuous Monitoring through Biosensors" Symmetry, **12**(3), 1-18, 2020. https://doi.org/10.3390/sym12030460.

[17] R. Velayutham, M. Sabari and M. S. Rajeswari, "An innovative approach for women and children's security based location tracking system," in International Conference on Circuit, Power and Computing Technologies (ICCPCT), Nagercoil, **2016**, 1-5, doi: 10.1109/ICCPCT.2016.7530325.

[18] M. K. Rajesh, K. Jyothsna, T. S. Aparna, T. Anjali, M. Meera, and S. D. Amrutha, "IoT-Based Women Security System," In: Ranganathan G., Chen J., Rocha Á. (eds) Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems, 89. Springer, Singapore. doi:10.1007/978-981-15-0146-3_134.

[19] S. Ahir, S. Kapadia, J. Chauhan and N. Sanghavi, "The Personal Stun-A Smart Device For Women's Safety," 2018 International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, **2018**, 1-3, doi: 10.1109/ICSCET.2018.8537376..

[20] A. Helen, M. F. Fathila, R. Rijwana and Kalaiselvi V.K.G., "A smart watch for women security based on iot concept 'watch me'," in 2nd International Conference on Computing and Communications Technologies (ICCCT), Chennai, **2017**, 190-194, doi: 10.1109/ICCCT2.2017.7972266..

[21] S. Gharge, M. Choudhary, S. Dubey, P. Gupta, and M. Neve, "Women Safety Jacket," International Journal for Research in Applied Science & Engineering Technology (IJRASET), **6**(3), 2606–2609 2018. doi:10.22214/ijraset.2018.3421.

[22] P. Saikumar, P. Bharadwaja and J. Jabez, "Android and Bluetooth Low Energy Device Based Safety System," in 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, 1180-1185. doi: 10.1109/ICCMC.2019.8819781..

[23] G. Toney, F. Jabeen and Puneeth S, "Design and implementation of safety armband for women and children using ARM7," in International Conference on Power and Advanced Control Engineering (ICPACE), Bangalore, 2015, 300-303. doi: 10.1109/ICPACE.2015.7274962.

[24] S. Patil, C. Pawar, C. Jadhav, and P. S. K. Gangurde, "A Novel Alerting System for Human Security," International Journal of Computer Applications, **182**(34), 16–19, 2018. doi: 10.5120/ijca2018918197.

[25] R. Sriranjini, "GPS and GSM Based Self Defense System for Women Safety," Journal of Electrical & Electronic Systems, **6**(2), 1-3, 2017. doi:10.4172/2332-0796.1000233.

[26] N. R. Sogi, P. Chatterjee, U. Nethra, and V. Suma, "SMARISA: A Raspberry Pi Based Smart Ring for Women Safety Using IoT," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, 2018, 451-454, doi: 10.1109/ICIRCA.2018.8597424..

[27] S. R. Mahmud, S. N. Tumpa, A. B. Islam, C. N. Ferdous, N. Paul and T. T. Anannya, "BONITAA: A smart approach to support the female rape victims," in 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, 730-733. doi: 10.1109/R10-HTC.2017.8289061..

[28] K. Viswanath and A. Basu, "SafetiPin: an innovative mobile app to collect data on women's safety in Indian cities," Gender & Development, **23** (1), 2015. doi: 10.1080/13552074.2015.1013669.

[29] P. Chaudhari, R. Kamte, K. Kunder, A. Jose, and S Machado, "'Street Smart': Safe Street App for Women Using Augmented Reality," In 2018 4th International Conference on Computing Communication Control and Automation (ICCUBEA), 1–6, Pune, India. doi: 10.1109/ICCUBEA.2018.8697863.

[30] M. Lindsay, J. T. Messing, J. Thaller, A. Baldwin, A. Clough, T. Bloom, K. B. Eden, and N. Glass, "Survivor Feedback on a Safety Decision Aid Smartphone Application for College-Age Women in Abusive Relationships,"

Journal of Technology in Human Services, **31**(4), 368–388, 2013. doi: 10.1080/15228835.2013.861784.

[31] B. Sumathy, P. D. Shiva, P. Mugundhan, R. Rakesh, and S. Prasath, "Virtual Friendly Device for Women Security," in International Conference on Physics and Photonics Processes in Nano Sciences, 1-12, 2019. doi: 10.1088/1742-6596/1362/1/012042.

[32] S. K. Punjabi, S. Chaure, U. Ravale, and D. Reddy, "Smart Intelligent System for Women and Child Security," in 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). 451–454, 2018. doi: 10.1109/IEMCON.2018.8614929.

[33] S. Bankar, K. Basatwar, P. Divekar, P. Sinha, and H. Gupta, "Foot Device for Women Security," in 2018 2nd International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 1-5, 2018. doi: https://doi.org/10.1109/ICCONS.2018.8662947

[34] S. M. Saklain Galib, S. M. Rabiul Islam, and M. A. Rahman, "A multiple linear regression model approach for two-class fNIR data classification," Iran Journal of Computer Science, 2020. doi: 10.1007/s42044-020-00064-0.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 837-845 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Study of Stirling Engine Efficiency Combined with Solar Energy

Oumaima Taki[1,*], Kaoutar Senhaji Rhazi[2], Youssef Mejdoub[2]

[1]*Laboratory of Networks, Computer Science, Telecommunication, Multimedia (RITM), CED Engineering Sciences, Higher School of Technology ESTC, Hassan II University, Casablanca, 20100, Morocco*

[2]*Laboratory of Networks, Computer Science, Telecommunication, Multimedia (RITM), Higher School of Technology ESTC, Hassan II University, Casablanca, 20100, Morocco*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 27 March, 2021
Online: 04 April, 2021

Keywords:
Stirling engine
External combustion
Heating
Solar energy
Mechanical energy
Renewable energy
Cycle efficiency

ABSTRACT

Fossil fuel can no longer supply the constantly spiking demands of energy around the world, hence the increasing research on renewable energies as an alternative. The Stirling Engine is an external combustion engine, giving us a wide range of heat sources: solar, nuclear. The Stirling engine makes best of use of solar sources in an environmentally friendly way. It has no emissions and live longer as compared to Photovoltaic cells. The Stirling engine can operate at Low Temperature difference, which makes it prominent. In order to study the efficiency of a conversion from thermal energy to work, we need to take into account the energy efficiency, which is a key parameter in Low Temperature Difference Stirling Engine, even if its efficiency is lower than those of high temperature Stirling engine. In this article, we are studying the efficiency of the Stirling engine as a first step using a parabolic mirror to focus the sun's radiation onto the engine. In this article, we are studying the efficiency of the Stirling engine as a first step, by making isothermal and adiabatic analysis of the engine to detail the operation throughout its process, and be able to act on the various input parameters that impact the value of the final yield, and in a second step, using a parabolic mirror to focus the sun's radiation onto the engine.

1. Introduction

Solar energy is an energy that falls into the category of renewable energies, because it is considered inexhaustible. Technologically, two ways are practiced in the use of direct solar energy; solar thermal energy and Photovoltaics. Regarding the solar thermal, it's a system that uses solar energy to produce heat by heating a fluid at more or less high temperature. We can therefore produce energy, like the case of classical thermal power stations. In this case, we are talking about thermodynamical central power plants. As to the Photovoltaics, it is a system which is composed of photovoltaic cells. It directly converts a part of solar rays to electricity with photovoltaic effect. A solar powered Stirling engine is a type of external combustion engine, which uses the energy from the solar radiation to convert solar energy to mechanical energy. The resulting mechanical power is then used to run a generator or alternator to produce electricity. Initially, Stirling engine was invented by Robert Stirling in the year 1816 [1].

Solar power generation could be accomplished using various methods, such as linear Fresnel systems, Parabolic through Solar tower systems, and most importantly Solar dish systems (Figure 2), which happen to be one of the most intuitive and efficient ways of concentrating solar heat on the receiver that drives the Stirling engine-generator unit. It is applied in several situations. Due to the available sizes of Stirling engines, this method is most useful in small capacity cases that do not exceed tens of kW.

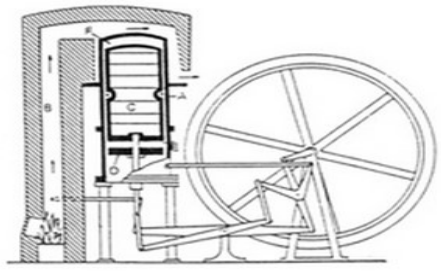


Figure 1: Stirling Engine. [1]

*Corresponding Author: Oumaima Taki, Email: oumaima.taki@ensem.ac.ma

https://dx.doi.org/10.25046/aj060297

Furthermore, we are studying the efficiency of the Stirling engine and comparing it with existent internal combustion engines to see if it is worth using it as an alternative. Furthermore, we are going to study the possibility to combine the Stirling engine with solar energy for a more environmental-friendly solution.

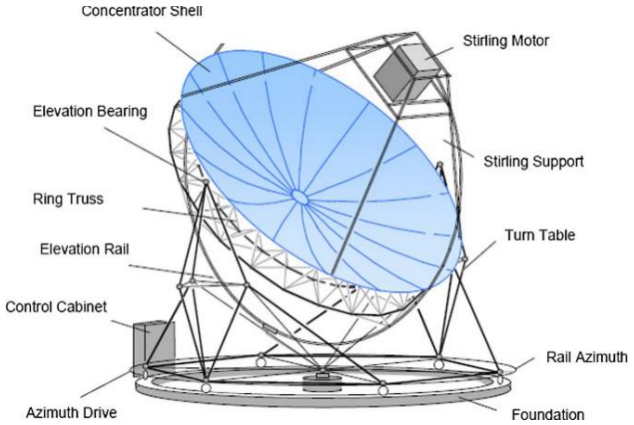


Figure 2: Design of the EURODISH System. [2]

2. Engine operation

A Stirling engine is a piston engine operating on the general principle of the Stirling cycle. The Stirling cycle and engine were defined in 1989 by the international scientific community as [3]:

"A Stirling cycle is defined as a process that occurs in any closed space containing a working fluid in which changes in volume induce cyclical changes in pressure of the fluid and its displacement in the closed space induce changes in cyclic temperatures in the fluid."

The Stirling engine offers the possibility of having one of the best efficiencies with less emissions unlike internal combustion engines. Its older models are less efficient and huge, but the current models are more developed, which improves efficiency, as well as the use of any external heat source for very high temperatures [4].

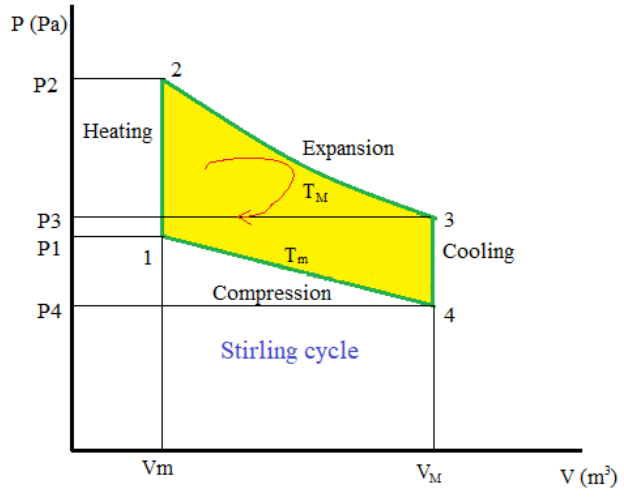


Figure 3: Stirling engine cycle. [1]

The theoretical Stirling cycle is similar to the Carnot cycle, except that the in th Stirling cycle the isochoric processes replace the adiabatic heating and cooling processes of the Carnot cycle. The Stirling cycle then involves four successive evolutions of an ideal gas between two heat sources that has constant temperatures Tc and Te, which in turn are separated by a perfect exchanger which has an isochoric process. When applying the first principle of thermodynamics, we get the same efficiency as the Carnot cycle [4].

The thermodynamic cycle can be plotted on a PV diagram that represents the variation of the pressure versus the volume. (Figure 3).

In a theoretical case, this thermodynamic cycle can be split into four reversible processes (Figure 3).

2→3: (heat transfer from an external source to the working fluid). The cold cylinder piston (the working piston) is at the top of its downstroke, while the hot cylinder piston (the expansion piston) is in the middle of its upstroke; the expansion piston moves down, while the working piston remains stationary. This is the engine time; the hot source supplies the gas with thermal energy, and the descent of the expansion piston drives the crankshaft. On the theoretical indicator diagram, this cycle time corresponds to curve 2-3. As the volume of the gas increases and its temperature is constant, the pressure of the gas in the hot cylinder decreases.

3→4: (heat transfer from the working fluid to the regenerator). The last stroke being completed, the cycle is returned to its initial state, the mechanical coupling between the two pistons is such that the working piston begins to rise, while the expansion piston goes down; during this double movement, the gas being hot, it gives up its heat to the regenerator and the gas cools as it passes from one cylinder to another. As its volume remains constant, its pressure decreases; which is represented by segment 4-3 of the theoretical diagram. The engine has returned to the starting point, the regenerator is ready to absorb heat again, and a new cycle can begin again.

4→1: (heat transfer from the working fluid to the cold source). the ingenious coupling between the pistons allows the expansion piston to be stationary while the working piston descends. The gas is compressed, but its temperature does not increase, because the compression takes place in the cylinder connected to the cold source. Energy is rejected to the cold source and the compression is isothermal; this time is represented by curve 4-1 on the theoretical indicator diagram.

1→2: (heat transfer from regenerator to working fluid) The expansion piston goes up and the working piston goes down, which allows the movement of the gas from the hot side, without changing the volume; segment 1-2 of the theoretical diagram is therefore vertical. Passing through the regenerator, the gas recovers the heat that was stored there and, at the same time, returns this element to its initial temperature. [1]

During this cycle, the system releases an amount of energy, which is needed later to heat the fluid, and restart the cycle as a loop.

Robert had the idea to use a regenerator to recover the transferred energy and then use it for heating. Ideally, the curve is elliptic, hence all amounts of energy are recovered (Figure 4).

We mostly find Stirling Engines in one of the three common configurations which are α, β and γ.

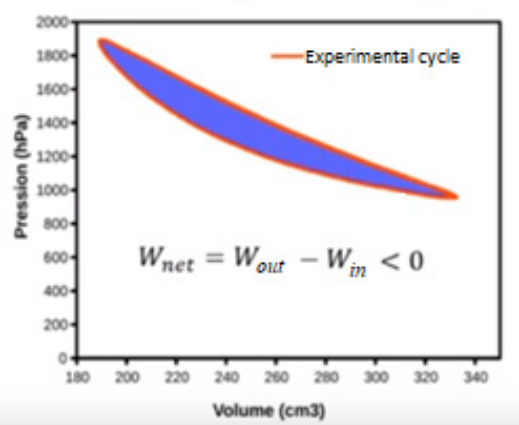


Figure 4: Stirling cycle curve. [1]

2.1. Alpha-type Stirling Engine

The α-type engine is composed of two separated cylinders (Figure 5). The two cylinders are exposed, respectively, to a hot and a cold temperature source. And to each cylinder, is sealed a "hot" piston and a "cold" piston. contains two separate power pistons in separate cylinders, a "hot" piston and a "cold" piston. It also has a pipe that connects the two cylinders. This pipe is usually filled by a regenerative material in order to enhance the thermal efficiency.

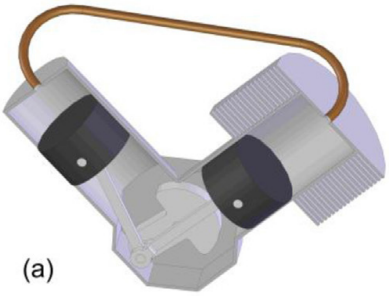


Figure 5: Alpha-type Stirling engine. [5]

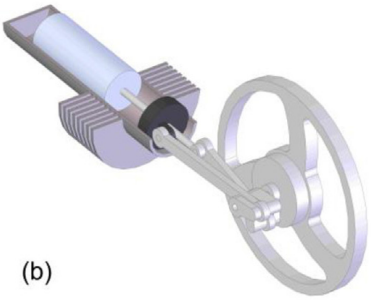


Figure 6: Beta-type Stirling engine. [5]

2.2. Beta-type Stirling Engine

In contrast with the α-type, β-type engine is composed of one unique cylinder with one piston sealed to it, and a displacer as shown in Figure 6. At the top of the cylinder, is placed a heat source, and a cold source at the bottom. The gas flow through the small clearance between the cylinder wall and the displacer; when it flows towards the hot end of the cylinder, then the expansion process occurs, and when it flows towards the cold end, then the compression process occurs . It is the displacer that allows the gas to move between the cold zone and the hot zone. The system is linked to a flywheel.

2.3. Gamma-type Stirling Engine

The γ-type engine is similar to a β-type engine, the main difference is that the cooling chamber is mounted in a separate cylinder as demonstrated in Figure 7, but it is still connected to the same flywheel.

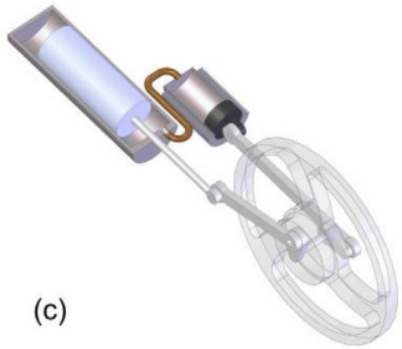


Figure 7: Gamma-type Stirling engine. [5]

2.4. Comparison between the three types

To make a comparison between the 3 types of architectures, the compression rate must be defined: The compression rate r_v is defined here as the ratio of the maximum volume VM by the minimum volume Vm that it will occupy during this same cycle:

$$r_v = \frac{V_3}{V_2} = \frac{V_4}{V_1} \quad (1)$$

For given hot and cold source temperatures, and for identical displacements, Alpha engines have higher compression ratios than those of type Beta, which in turn are slightly higher than those of Gamma engines. This has the consequence of being able to extract more power from an Alpha engine because it will run more quickly. The downside is that they require more design and manufacturing rigor.

3. Advantages and disadvantages

3.1. Advantages

- **Quiet operation**: In contrast with internal combustion engines, there is no relaxation in the atmosphere. With that absence of gas that will eventually escape, plus the absence of the open-close valves, this engine is quiet and has a reduced mechanical stress.
- **High efficiency**: Stirling engines have the best efficiency compared to an internal combustion engine, and it could even exceed 40 % as efficiency.
- **The multitude of possible "hot springs" and ecological aptitude:** Due to its heat supply method this engine can operate from any heat source.

- **Reliability and easy maintenance**: The technological simplicity of this engine allows engines to be very reliable and require little maintenance.
- **The long service life**: Due to its simplicity, the life of this engine is, in theory, longer than that of conventional engines. Indeed, it requires less maintenance and its replacement is much faster and less dangerous.
- **Reversible operation**: The Stirling cycle is reversible, when a Stirling engine is driven by another engine; it becomes a heat pump capable of working in cooling and heating mode.

3.2. Disadvantages

- **The price**: The main drawback of this engine is its manufacturing cost, which is about twice that of a diesel engine. Stirling engines require inlet and outlet heat exchangers, which contain the high temperature working fluid, and must withstand the corrosive effects of the heat source and the atmosphere.
- **Lack of flexibility**: Quick and efficient variations in power are difficult to achieve with a Stirling engine. This is more suitable for running at constant nominal power. This point is a big handicap for the automotive industry.
- **Height and weight:** External combustion, which requires heat exchangers at both hot and cold spots, makes the Stirling engine generally bigger in size and heavier than a generic internal combustion engine with the same power output.

3.3. Advantages of Stirling engine compared to an internal combustion engine

In comparison with a combustion engine, Stirling engine overtakes it on many levels, for example we can look at fuel flexibility; A Stirling engine does not require a highly refined liquid to operate, it can use a variety of liquid and gases, which makes it more flexible than a Diesel engine that required refined Diesel fuel. And because of the external-combustion process of the Stirling engine, it also burns any given fuel cleaner than an internal-combustion engine. In addition, Stirling engines can be balanced mechanically, making them less noisy by eliminating the mechanical vibration problems. On the other hand, internal-combustion engines have severe noise because of the periodic nature of their combustion and mechanical motion processes. We can enhance them to be more silent, by using mechanical isolation and acoustic design, but this would increase the cost of the engine, and would make it unpractical in some situations. [6].

3.4. Applications of the Stirling engine

Besides the academical use of the Stirling Engine, we can find this technology in various daily useful applications. The American Stirling Company offers one of these applications which is the wood stove Stirling fan (Figure 8): A silent fan that does not need electricity to move the heat from a wood stove, and have more heated area in the house instead of having only a restricted heated area near to the wood stove.

Another interesting application of the Stirling engine is the Combined Heat and Power systems (CHP) that can be very useful in businesses such as a commercial laundromat, since it generates electricity and utilizes the waste energy produce heat. There is also a smaller version of CHP systems, called micro CHP that have residential use [7].



Figure 8: The Stirling stove fan by American Stirling Company. [8]

The SAAB company [9], which is a Swedish company specialized in building submarines, uses the Stirling engine in their Gotland and Södermanland submarines classes, essentially because the Stirling engine is silent compared to Diesel engines. SAAB claims that the secret to their world's most silent submarine is Stirling engine based submarines do not need to surface and recharge the batteries, using the air-independent propulsion [9].

4. Enhancement of the Stirling Engine performance

To act on the Stirling engine performance, we are led to optimize the temperature of the cold and hot sources, to obtain an optimal temerature difference. We can also modify the geometry of the engine to keep the losses to an absolute minimum.

The role of the regenerator is to recover the heat from the cooling of the gas to heat it again. It therefore plays a key role in the operation of the engine.

Thus, it seems legitimate to seek to optimize the operation of the regenerator to improve that of the engine.

The MOD II automobile engine [10] that was produced in the 1980's [10] was among the most efficient Stirling engines, it reaches a maximum efficiency of 38.5% [10] , compared to a petrol engine which has a yield of (20-25%).

It was abandoned due to high development costs and fears of not being able to compete with internal combustion engines in terms of reactivity.

To reach a high efficiency in a Stirling engine we are led to use a regenerator; imperfect heat transfer that occurs between the engine and the source may lead to external losses of energy as well as internal losses.

The efficiency of an engine varies with the operating speed due to the different losses interactions

5. Modeling of the Stirling Engine

5.1. Efficiency Stirling engine:

r_v : Volume Ratio

λ : Temperature Ratio

S: Entropy

Cv : The molar heat of gas

T_i : Inlet temperature to Urieli's generalized cell

P_i : The pressure at a given time I

T_c : Compression temperature

T_{ck} : mean temperature of working fluid at cooler and compression space

T_e : Temperature at expansion side

T_d : Temperature at dead space

T_{he} : Temperature of working fluid at heater and expansion space.

T_k : Temperature of working gas at cooler.

V_c : Compression cylinder volume variation

V_d : Dead space volume

V_e : Expansion cylinder volume variation

V_r : Regenerator volume

w_c : Compression work

w_d : Work done from engine

w_e : Expansion work

w_s : West number

T_{max}: Maximum cycle temperature

T_{min}: Minimum cycle temperature

X : Distance

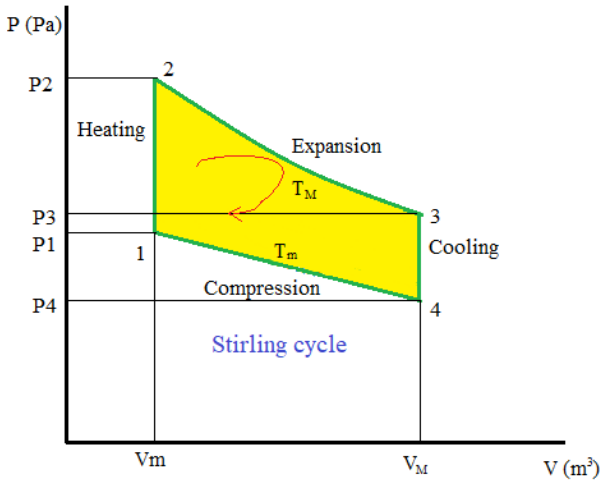


Figure 9: Stirling engine cycle. [1]

As the PV diagram for Stirling cycle shows (Figure 9):

4→1: The engine is operating with a minimum value temperature during the isothermal compression. And with a constant energy:

$$P_1 = \frac{P_4 V_4}{V_1} = P_4 r_v \quad (2)$$

$$T_1 = T_4 = T_{min}$$

With the provided heat (Q) = Recovered work (W_c)

$$Q = W_c = P_4 V_4 \ln\left(\frac{1}{r_v}\right) = nRT_4 \ln\left(\frac{1}{r_v}\right) \quad (3)$$

$$\text{Entropy change} = (S_4 - S_1) = = R\ln\left(\frac{1}{r_v}\right) \quad (4)$$

1→2 : In the isochoric heating phase, both the compression piston and the expansion piston moving respectively towards the regenerator and away from the regenerator, simultaneously, keeping volume between the two pistons constant. The working fluid is flows from the compression area to the expansion area, making its temperature gradually increasing from Tmin to Tmax. This gradual increase of the fluid temeprature while it pass through the regenerator creates a gradual increase of pressure. There is no work done and there is an increase in the entropy and the internal energy of the working fluid.

The volume remains constant throughout this process.

$$P_2 = \frac{P_1 T_2}{T_1} = \frac{P_1}{\lambda} \; ; \; V_1 = V_2 \quad (5)$$

If:

$$\lambda = \frac{T_1}{T_2} \quad (6)$$

With **λ** referring to the energy ratio defined by [11].

The provided heat Q will then be equal to:

$$Q = C_v(T_2 - T_1) = 0 \quad (7)$$

Since there is no work done.

$$\text{Change of entropy} = (S_2 - S_1) = R\ln\left(\frac{1}{\lambda}\right) \quad (8)$$

2→3: The temperature is constant during the isothermal relaxation, while the volume increases as well as the entropy. There is no change of energy.

$$P_3 = \frac{P_2 V_2}{V_3} = P_2\left(\frac{1}{r_v}\right) \quad (9)$$

$$T_2 = T_3 = T_{\max} \quad (10)$$

$$Q = W = P_2 V_2 Ln r_v = mRT_2 Ln r_v \quad (11)$$

$$(S_2 - S_3) = R\ln(r_v) \quad (12)$$

3→4: No work is provided during the isochoric cooling, the volume is constant.

$$P_4 = \frac{P_3 T_3}{T_4} = P_3\,\lambda \; (13) \quad (13)$$

$$V_3 = V_4 \quad (14)$$

$$Q = C_v(T_4 - T_3) \quad (15)$$

$$\text{And } (S_4 - S_3) = C_v \ln(\lambda) \quad (16)$$

Si

$$r_v = \frac{V_3}{V_2} = \frac{V_4}{V_1}$$

The provided heat will then be equal to $RT_2\,\mathrm{Ln}(r_v)$) and the released heat will be equal to $RT_4\mathrm{Ln}(r_v)$.

We can write the output as follows:

$$\eta = \frac{\mathrm{n}\left(RT_2\mathrm{Ln}(r_v) - RT_4\mathrm{Ln}(r_v)\right)}{\mathrm{n}RT_2\mathrm{Ln}(r_v)} \tag{17}$$

$$\eta = 1 - \frac{T_{min}}{T_{max}} = 1\text{-}\mathcal{Z} \tag{18}$$

We then find the ideal Carnot efficiency that corresponds to the possible maximum theoretical efficiency in a two heat sources engine.

5.2. Isothermal analysis

According [12], we can obtain an efficient heat transfer by assuming that the total mass of the working gas inside the engine remains constant:

$$M = m_c + m_k + m_r + m_h + m_e \tag{19}$$

And considering that the pression in all of the engine is constant, and:

$T_c = T_K$ et $T_h = T_e$

Based on the ideal gas equation, we have:

$$M = \frac{P}{R\left(\frac{V_c}{T_K} + \frac{V_K}{T_K} + \frac{V_r}{T_r} + \frac{V_h}{T_h} + \frac{V_e}{T_e}\right)} \tag{20}$$

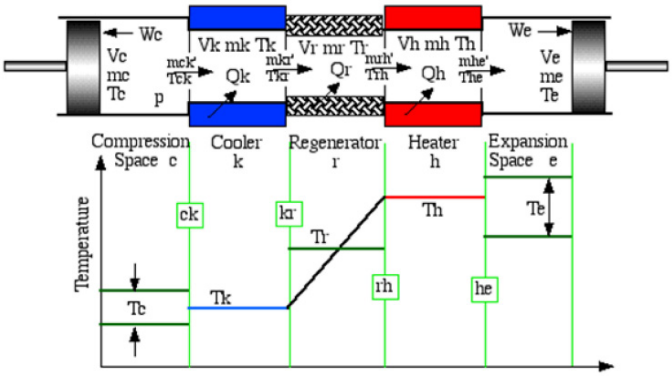


Figure 10: Ideal isothermal model. [13]

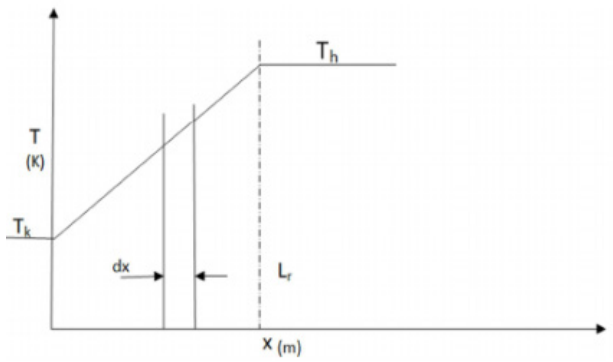


Figure 11: Linear profile of regenerator temperature. [13]

Then we can numerically define the assumption that the temperature profile is linear, by the equation of a straight line

$$T(x) = \frac{(T_h - T_k)x}{L_r} + T_K \tag{21}$$

The sum of mass of the gas is:

$$m_r = \int_0^{v_r} \rho dV_r \tag{22}$$

With ρ the density,

$$dV_r = A_r dx \tag{23}$$

Where dx is the derived volume for a constant free flow area, and

$$V_r = A_r L_r$$

And *Ar :*

$$A_r = \frac{V_r}{L_r} \tag{24}$$

By integrating (25), we obtain:

The definition of the regenerator effective average temperature (T), in terms of ideal gas equation:

$$m_r = \frac{V_r P}{RT_r} \tag{27}$$

Then, by comparing the equations (26) and (27) we obtain:

$$T_r = (T_h - T_k)/\ln(\frac{T_h}{T_K}) \tag{28}$$

Thus, the pression of the cycle can be written as:

$$P = MR\,(\frac{V_c}{T_K} + \frac{V_K}{T_K} + \frac{V_r\,\ln(\frac{T_h}{T_K})}{(T_h - T_k)} + \frac{V_h}{T_h} + \frac{V_e}{T_h})^{-1} \tag{29}$$

The total work of the cycle is the sum of the work of compression and expansion.

$$W = W_C + W_e \tag{30}$$

$$W = \int P dV_c + \int P dV_e \tag{31}$$

$$W = \int P(\frac{dV_c}{d\varphi} + \frac{dV_e}{d\varphi})d\varphi \tag{32}$$

5.3. Isothermal modeling

To study the heat transfer, it is primordial to consider the energetic equation of ideal gas.

In [14] the author's has modeled a generalized workspace cell, and as shown in Figure 12, it can be reduced to a workspace cell or a heat exchanger cell.

The enthalpy transfer out the cell (resp. into the cell) occurs with a mass flow rate m_0 at temperature T_0 (resp. a mass flow rate m_i at temperetaure T_i). The derivative operator is noted D and D_m refers to the mass derivative (d_m/ d_t).

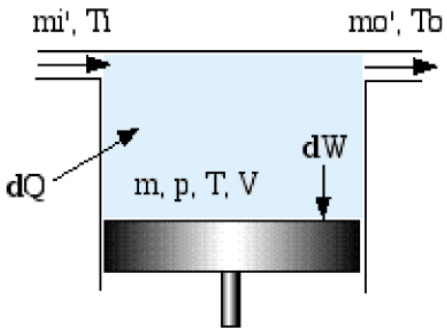


Figure 12: Generic cell diagram of the engine. [14]

The energy equation of the cell can be written as:

Heat transfer rate in the cell + Net enthalpy converted in thecell

=

Rate of work applied on the environment + Rate of increase of internal energy in the cell.

$$DQ + \left(C_p T_i m_i - C_p T_0 m_0\right) = DW_d + C_v D(mT) \quad (33)$$

where C_p and C_v are respectively the specific heat capacities of gas at constant pressure and constant volume. The equation below is the well-known classical form of the energy equation in which the terms of kinetic energy and potential energy have been neglected.

Then :

$$DQ = C_p T\,(m_0 - m_i) + DW_d + C_v TDm \quad (34)$$

For reasons of mass conservation, the difference $m_0 - m_i$ is simply the rate of mass accumulation inside of the cell

R= $C_p - C_v$

Therefore, we can write the equation as:

$$DQ = DW_d + RTDm \quad (35)$$

$$Q = \int DQ = \int DW_d + RTDm \quad (36)$$

5.4. Adiabatic modeling

The principle of this method uses a numerical resolution approach, it divides the volume of the engine into a certain number of control volumes, then it applies conversion equations to the momentum of the gas associated with its equation of state. The properties of the gas are considered uniform in each control volume. The interaction between volumes is taken into account when solving differential equations which is done simultaneously.

The complexity of the system to be solved and the calculation time depend on the assumptions used.

- Thermodynamic compression and expansion transformations are adiabatic,
- The gas pressure is uniform throughout the machine,
- He movement of the piston and of the displacer is sinusoidal,
- The working fluid follows the ideal gas law,
- The machine rotation speed is constant

The energy equation is applied to a generalized whole can be written as:

$$DQ + (C_p T_i m_i - C_p T_0 m_0) = DW_d + C_v D(mT) \quad (37)$$

The state equation is given by PV= nRT

$$\text{And} \quad C_p - C_v = R \quad (38)$$

Therefore:

$$C_p = \frac{R_\gamma}{\gamma - 1},\ C_v = \frac{R}{\gamma - 1}$$

With:

$\gamma = C_p/C_v$

By taking the logarithm on both sides of the equation and differentiating them, we get a differential form of the equation of state:

$$\frac{Dp}{p} + \frac{DV}{V} = \frac{Dm}{m} + \frac{DV}{V}$$

$$M = m_c + m_k + m_r + m_h + m_e \quad (39)$$

$$Dm_c + Dm_k + Dm_r + Dm_h + Dm_e = 0 \quad (40)$$

Since the volume and the temperature are constant, the differential equation is shortened to:

$$\frac{Dp}{p} = \frac{Dm}{m}$$

$$Dm_c + Dm_e + Dp\left(\frac{m_k}{p} + \frac{m_r}{p} + \frac{m_h}{p}\right) = 0 \quad (41)$$

$$Dm_c + Dm_e + Dp/R\left(\frac{V_K}{T_K} + \frac{V_r}{T_r} + \frac{V_h}{T_h}\right) = 0 \quad (42)$$

We apply the energy equation the equation that we obtain:

$$DQ_c - C_p T_{ck} m_{ck} = \mathrm{D}W_c + C_v \mathrm{D}(m_c T_c) \quad (43)$$

However, the compression space is adiabatic, DQ_c=0, plus the realised DW_c= pDVc, for continuity reasons, the accumulation rate of gaz Dm_c is equal to the mass difference of the gas given by m_{ck}.

$$C_p T_{ck} m_{ck} = \mathrm{pD}V_c + C_v \mathrm{D}(m_c T_c) \quad (44)$$

$$Dm_c = (\mathrm{pD}V_c + \frac{V_c Dp}{\gamma})/\mathrm{R}T_{ck} \quad (45)$$

$$Dm_e = (\mathrm{pD}V_e + \frac{V_e Dp}{\gamma})/\mathrm{R}T_{he} \quad (46)$$

Simplifying :

$$Dp = \frac{-\gamma p\left(\left(\frac{DV_c}{T_{ck}}\right) + (DV_e/T_{he})\right)}{\frac{V_c}{T_{ck}} + \gamma\left(\frac{V_k}{T_k} + \frac{V_r}{T_r} + \frac{V_h}{T_h}\right) + \frac{V_c}{T_{he}}} \quad (47)$$

We consider the continuity equation, given by:

$$\mathrm{Dm} = m_i - m_0 \quad (48)$$

We apply successively the equation above to each of the cells as shown in the figure:

$$m_{ck} = -Dm_c \quad (49)$$

$$m_{kr} = m_{ck} - Dm_k \quad (50)$$

$$m_{rh} = m_{kr} - Dm_r \quad (51)$$

$$m_{he} = m_{rh} - Dm_h \quad (52)$$

The total work done by the engine is the algebraic sum of the work done by the compression and expansion areas:

$$\mathrm{DW} = \mathrm{pD}V_c + \mathrm{pD}V_e \quad (53)$$

$$DQ + \left(C_p T_i m_i - C_p T_0 m_0\right) = (C_p pDV + C_v VDp)/R \quad (54)$$

Dans les espaces d'échangeur de chaleur, aucun travail n'est effectué, car les volumes respectifs sont constants.

$$DQ_k = \frac{V_k DpC_v}{R} - C_p(T_{ck}m_{ck} - T_{kr}m_{kr}) \quad (55)$$

$$DQ_R = \frac{V_r DpC_v}{R} - C_p(T_{kr}m_{kr} - T_{rh}m_{rh}) \quad (56)$$

$$DQ_h = \frac{V_h DpC_v}{R} - C_p(T_{rh}m_{rh} - T_{he}m_{he}) \quad (57)$$

The actual Stirling cycle engine is subject to heat transfer, internal heat losses and mechanical friction losses, to estimate hese losses [15] defined certain engine temperature ratios.

The ratio of the lower operating temperature to the upper operating temperature of the engine is defined by: £ $= \frac{T_k}{T_e}$;

The ratio of the cooler temperature to the heater temperature is defined as: € $= \frac{T_c}{T_h}$;

The ratio of the expansion area temperature to the heater temperature is noted as: ƨ $= \frac{T_e}{T_h}$. with α$=\frac{b}{a}$ and β= β$=\frac{c}{a}$, heat transfert coefficients.

Therefore, the average energy of the cycle is expressed by:

$$\mathrm{P} = Q_s - Q_R \quad (58)$$

$$\mathrm{P} = \mathrm{a}\,(T_h - T_e)\text{-}\mathrm{b}\,(T_c - T_k) \quad (59)$$

$$\mathrm{P} = \mathrm{a}\,T_h(1 + \alpha€ - ƨ - ƨ\,\alpha£) \quad (60)$$

The thermal efficiency must not exceed the efficiency of the Carnot cycle.

$$\frac{Q_s - Q_R}{Q_s - Q_T} = \frac{P}{Q_s - Q_T} \leq 1 - \frac{T_c}{T_e} = 1 - ƨ$$

This condition can be expressed as:

$$ƨ\,((\alpha+1)\,ƨ - \beta(1 - ƨ)^2) \geqq \alpha€ + ƨ \quad (61)$$

To obtain a maximum energy we can write :

$$ƨ = \frac{\alpha€ + ƨ}{(\alpha+1)\,ƨ - \beta(1-ƨ)^2} \quad (62)$$

By replacing (62) in (60)

The maximum power is:

$$\mathrm{Psi} = \frac{\alpha T_h[\alpha(ƨ - €)(1-ƨ) - \beta(1 + \alpha€)(1-ƨ)^2]}{(\alpha+1)ƨ - \beta(1-ƨ)^2} \quad (63)$$

$$ƨ_p = \frac{\alpha€ + \beta(\alpha€+1)}{\alpha + \beta(\alpha€+1)} \quad (64)$$

By differentiating (62) we obtain the point where the power is maximum:

$$ƨ_m = \frac{\alpha\beta(1-€) + \sqrt{(1+\alpha)(\alpha + \beta + \alpha\beta)}(\beta + \beta\alpha^2€^2 + \alpha€(1 + \alpha + 2\beta))}{\alpha + \alpha^2 + \beta + 2\beta\alpha + €\alpha^2\beta}$$

In [16] the author's has proven that the maximum of the values that we can obtain are independent of the regenerator's conductance value, and thus in the case of a uni-dimensional model of a Stirling engine.

6. Simulation and results

Now once we have characterized the engine geometrically, were all set to implement the isothermal model. The flow diagram in Figure 13 illustrates the steps of the model; we first use as an entry parameter the total mass of the initial engine m and an effective average pressure p.

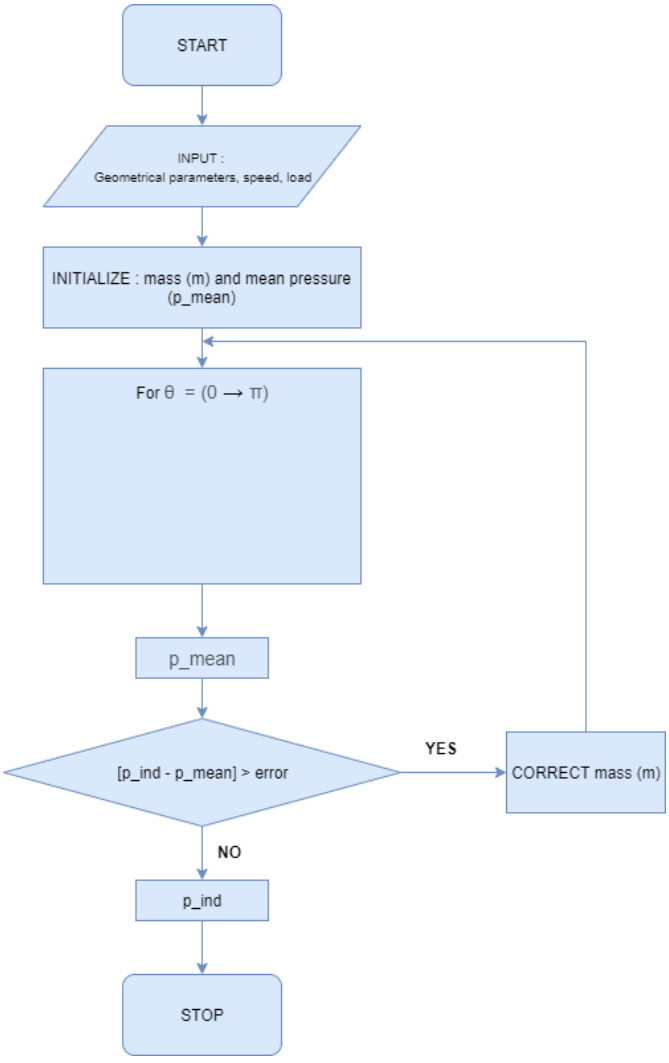


Figure 13: Isothermal model diagram.

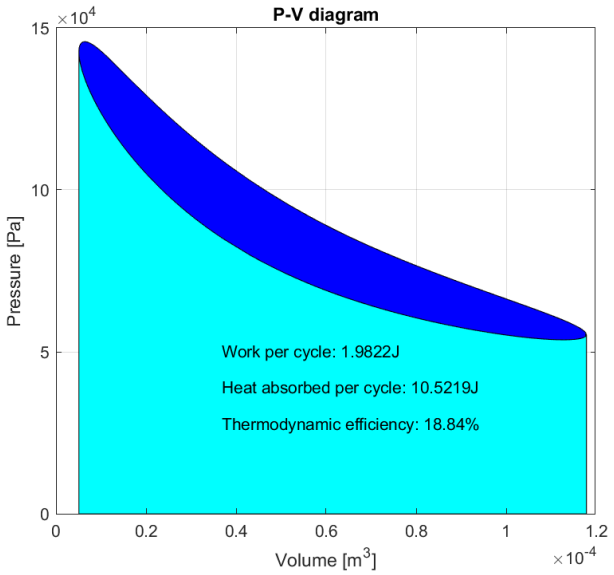


Figure 14: Pressure-volume diagram.

The algorithm then, for each value of θ, calculates the total average pressure as a function of the initialized mass m, and compares it with the reference pressure we defined at first. The algorithm iterate this comparison until we have convergence, i.e. the difference between the average pressure and the reference pressure is lower than the error.

The indicated work of the cycle is the obtained.

$$W = 1.9822 \text{ J}$$

In Figure 14 we observe that the curve is closed and cyclical, and in each cycle, the work is obtained by calculating the area inside the curve, while the area below the curve is the heat absorbed during that cycle. The thermodynamic efficiency is the ratio of the work on the amount of heat absorbed

7. Conclusion

The article studies the efficiency and the uses of the Stirling engine, and relies on the process of transforming renewable thermal energy into mechanical work. Renewable thermal energy is available at low cost for the long term and enviroment friendly. Such engine relying on this process is certainly interesting, even if it has a low thermal efficiency.

We have seen how the Stirling engine's perks can be used in different situations; it can be more advantageous than internal-combustion engine. Especially, at low maximum temperature and low temperature difference, the Stirling engine has virtually no substitute.

We can enhance the thermal efficiency by manipulating the shape of the displacer, the heat exchanger, the crank angle or use lf other working fluids. Stirling engine research and enhancements stimulate green education and can help reduce global warming and emissions.

This article is a first part of an modeling and simulation with MATLAB SIMULINK which will have as objective the improvement of the efficiency of the engine based on the difference of temperature, all while working on a concrete model.

References

[1] P. Gras, Le moteur Stirling et autres moteurs à air chaud, Decoopman Sciences & Techniques, 2010.

[2] R.M. Muthusivagami, R. Velraj, R. Sethumadhavan, “Solar cookers with and without thermal storage—a review”. Renewable and Sustainable Energy Reviews, **14**(2), 691-701, 2010, doi: 10.1016/j.rser.2008.08.018

[3] P. Stouffs, “Machine thermique non conventionnelles:état de l’art, application, problème à résoudre,” Journée d’étude de SFT (France), octobre 1999.

[4] B. Kongtragool, S. Wongwises, “Investigation on power output of the gamma-configuration low temperature differential Stirling engines”. Renewable Energy, **30**(3), 465–476, 2005, doi: 10.1016/j.renene.2004.06.003.

[5] Ahmed Abuelyamen, Rached Ben-Mansour, “Energy efficiency comparison of Stirling engine types (α, β, and γ) using detailed CFD modeling’’, International Journal of Thermal Sciences, **132**, 411-423, 2018, doi: 10.1016/j.ijthermalsci.2018.06.026.

[6] T. J. Marciniak, J. C. Bratis, A. Davis, and C. Lee. “An Assessment of Stirling Engine Potential in Total and Integrated Energy Systems’’ US . DEPARTMENT' OF ENERGY, 1979.

[7] S. Murugan n , Bohumil Horák , “A review of micro combined heat and power systems for residential applications’’ Centre ENET, VSB Technical University, 17. Listopadu, Ostrava 708 00, Czech Republic

[8] American Stirling Company, www.stirlingengine.com

[9] Daniel Nilsson, Senior Develoment Engineer SAAB, “Development of the Stirling AIP system’’, UDT 2020 Energy&Propulsion.

[10] N. Lanciaux, Contributionau développement d’un moteur Stirling, de la cogénération dans le bâtiment à l’autonomie énergétique, PhD Thesis Université d’Évry-Val d’Essone, 2015.

[11] G. Schmidt “Classic analysis of the functioning of the Stirling engine”. A report publishedin German Engineering Union, **XV**; 1871.

[13] M. T. García , E. C. Trujillo, J. A. Vélez Godiño and D. S. Martínez, “ThermodynamicModel for Performance Analysis of a Stirling Engine Prototype”, Energies, MDPI, Open Access Journal, **11**(10), 1-25, 2018

[14] C.J. Rallis, “A new ported constant volume external heat supply regenerative cycle”. in 12th Intersociety Energy Conversion Engineering Conference, 1977.

[15] J.R Senft , “Theoretical limits on the performance of Stirling engines”. Int J Energy Research, **22**(11), 991–1000,1998.

[16] D. Boer PCT, “Maximum obtainable performance of Stirling engine and refrigerators”, Journal of Heat Transfer, **125**(5), 911-915, 203ASME J Heat Transfer 2003, doi: 10.1115/1.1597618.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 842-849 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Diagnosis of Tobacco Addiction using Medical Signal: An EEG-based Time-Frequency Domain Analysis Using Machine Learning

Md Mahmudul Hasan[1, *], Nafiul Hasan[2], Mohammed Saud A Alsubaie[3], Md Mostafizur Rahman Komol[1]

[1]*Centre for Accident Research and Road Safety- Queensland, Queensland University of Technology, Brisbane, QLD 4059, Australia*

[2]*Department of Electrical and Electronic Engineering, Bangladesh Army University of Engineering and Technology, Qadirabad, 6431, Bangladesh*

[3]*Department of Mathematics, Taif University, Taif, 26513, Saudi Arabia*

ARTICLE INFO

Article history:
Received: 17 December, 2020
Accepted: 19 January, 2021
Online: 05 February, 2021

Keywords:
Electroencephalogram
Tobacco
Smoking
Machine Learning
Time-frequency domain

ABSTRACT

Addiction such as tobacco smoking affects the human brain and thus causes significant changes in the brainwaves. The changes in brain wave due to smoking can be identified by focusing on changes in electroencephalogram pattern, extracting different time-frequency domain features. In this aspect, a laboratory-based study has been presented in this paper, for assessing the brain signal changes due to the tobacco addiction. Four classifier models, namely, Logistic Regression (LR), K- Nearest Neighbor (KNN), Support Vector Machine (SVM) and Random Forest Classifier (RFC) were trained and tested for assessing the performance of the time domain, frequency domain and fusion of time-frequency domain features, with a five-fold cross-validation. Four different performance measures (sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve) were used to measure the overall performance, and the results suggested that the classifiers based on time-frequency domain features perform the best while using combinedly. Using the utilized fusion of the time-frequency domain features, the classification models can identify the smoker group with an accuracy ranged from (86.5-91.3%), where the RFC shows the best accuracy of 91.3%, which is higher than the three other classifiers models.

1. Introduction

This paper is an extension of work originally presented in International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2) [1]. The presented paper [1] utilized electroencephalogram (EEG) for the diagnosis of tobacco smoking based on only one machine learning model (artificial neural network), where the current article is expanded further to validate the EEG based diagnosis using multiple machine learning models. Also, this paper examines the utility of the time domain and frequency domain, individually and their combination on the EEG based tobacco addiction diagnosis.

In this modern era of life science, research in the field of neuroscience and cognitive engineering is flourishing with technological evolution. Electroencephalogram (EEG) is being used in this sector to understand sophisticated conditions of the brain as this is sensitive and susceptible to any action, especially for drug addiction, for example, alcohol, morphine, heroin or Cannabis addictions. Different stages of drug addiction can be determined only through EEG analysis, and this is very much necessary in the treatment of drug addiction. Some drug elements, such as nicotine is found in nature (from Nicotiana tabacum and Nicotiana rustica). They also have good effects on health; for example, nicotine is being used in the treatment of cognitive disorder and depression. Its impact on the body depends on the way of its metabolism and absorption, such as P-450 pathway degrades several body parts. At a level, it strengthens/weaken other taken drugs. A review work has done to analyze the pros and cons of nicotine, where limitation they noted are the analysis of the complex way of changing cognitive function and host inflammatory response [2]. By analyzing the EEG response, the changing pattern can be determined easily. Drug elements like

*Corresponding Author: Md Mahmudul Hasan, 2 Rochester Terrace, Brisbane, QLD 4059, Australia; Email: mahmudul.hasan.eee.kuet@gmail.com

tobacco, tar and nicotine affect on mood and behavior, which is controlled by human brain consisting of neurons [3].

2. Literature Review

A significant amount of public awareness against tobacco addiction have been conducted to date, but not all the public announcements were effective. To know the degree of effectiveness of public service announcement against smoking, a pilot study was conducted in Rome by collecting EEG, heart rate and galvanic skin response. Based on approach withdrawal, effort and emotional indexes, it was found that these parameters that show notable differences between effective and ineffective perception [4].

In addition to the public awareness, several studies have been conducted on identifying the physical and mental changes due to the smoking habit and potential treatments. An analysis was performed by [5], who had experimented on 21 male smokers. The authors had analyzed the EEG wave of the participants before and after the horizontal rotation treatment. They found that the treatment is beneficial in improving the EEG Alpha band, which reduces the smoking effect. An increment in alpha rhythm denotes higher relaxation and concentration ability of subjects [5]. Another study was conducted on 19 smoker participants. The functional magnetic resonance imaging (fMRI) and EEG analysis represented the effect of nicotine on the brain while doing oddball tasks, which is a response time task [6]. The study suggested that the integrated EEG-fMRI system is better identifying the brainwave changes due to smoking.

Though most of the studies has investigated the effect of tobacco smoking on the brain using the EEG, very few studies had done the differentiation of EEG characteristics using machine learning algorithms. In a study [7], the authors have done an experiment on 20 participants having 10 smokers and 10 non-smokers, to observe the changes in the EEG signal due to the smoking effect. The authors had used a Support Vector Machine (SVM) classifier based on Radial Basis Function (RBF) kernel and found that the the power spectral density (PSD) features performed better than the Fast Fourier Transform (FFT) features. Another study developed an EEG-based architecture to determine the effect of tobacco in the brainwaves, using 3 participants in a laboratory-based work [1].

The authors extracted time, frequency domain features from the EEG and showed that the frequency domain features, especially the power spectral density (PSD) and the Fast Fourier Transform (FFT) are most sensitive to the smoking condition than the time domain features for the smoker detection. The study showed promising results but utilized only one classifier (ANN) for the classification approach and only one performance measure (mean square error) for the assessment of the system [1]. However, considering the sensitivity and specificity metrics are most important for an EEG based detection system. As higher sensitivity with lower specificity leads to the higher false detection and the opposite trend causes the missing of a lot of positive states (in this case, smokers), a compromise between the two metrics is crucial.

This paper embodies a methodology for the diagnosis of tobacco smoking, based on the time-frequency analysis of the EEG signal. Also, the analysis was conducted to find efficient analyzing model and feature. BIOPAC® system was used for data collection, and Acknowledge-4.1® package [8] was used for extracting features. Later, python 3.6.9 version was used in Google Colab platform for data analysis. Most importantly, four supervised classification models, namely, Logistic Regression, K- Nearest Neighbor (KNN), Support Vector Machine (SVM) and Random Forest Classifier (RFC) were trained and tested for evaluating the performance of each of the EEG rhythm, with a five-fold cross-validation. Moreover, four different performance measures (sensitivity, specificity, accuracy and area under the receiver operating characteristic curve) were utilized to examine the performance of the diagnosis system. The following part of this paper is organized as follows- a brief methodology, including experimental design and tools, then the result section with the findings. Last, the paper was concluded, followed by a short discussion on the outcomes.

3. Methodology

3.1. Experimental Design

The methodology for building an electroencephalogram based tobacco addiction diagnosis system is shown in Figure 1. After selecting three participants, EEG were obtained by the BIOPAC® system. After that the removal of noise and artifacts were considered.

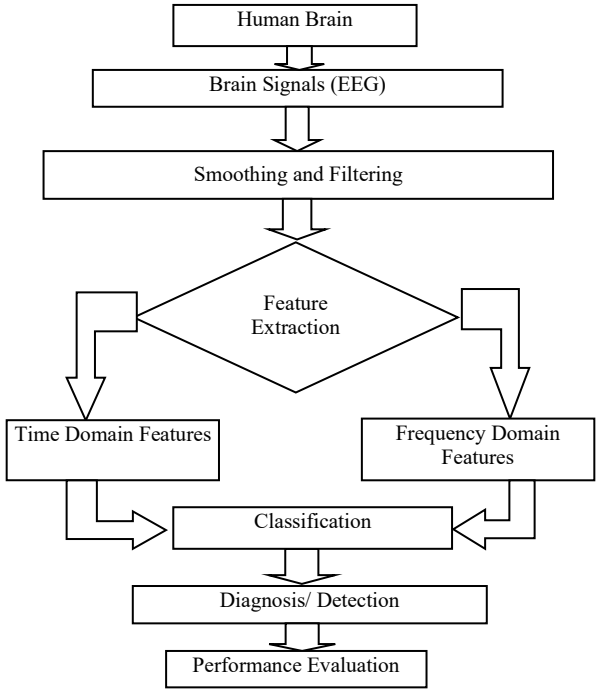


Figure 1: Block diagram for the proposed EEG based smokier / nonsmoker detection system

Afterwards, eight features were extracted for each domain, and the selected features were supplied towards the machine learning tools. Four different classification models, namely K-nearest neighbours (KNN), support vector machines (SVM), logistic regression (LR) and random forest classifier (RFC) models were developed in python 3.6.9 platform. The evaluation of the best classifier and domain was done by comparing their performance.

3.2. Experimental Equipment

3.2.1. Hardware tool

BIOPAC® MP 36 device was used at Biomedical Engineering lab, KUET for the experimentation purpose. The experimental setup is shown in Figure 2 [8].

3.2.2. Pre-processing and feature extraction software tool

Feature extraction was an important step in this study, which was done using the BIOPAC® student Lab Pro and Acqknowledge 4.1® software Google Colab research platform with the python 3.6.9 for developing the Machine learning-based classification models [9].

3.3. Participants

In total, three subjects participated in this experimental study, who were male, healthy and not suffering from psychological illness. The subjects were instructed to close their eyes while solving some simple arithmetic questions for 20 minutes. An interpreter used to ask the arithmetic question to the participants. In that respect, the brain response (EEG signal) is a result of cognitive event-related potential (ERP) type. A total of ten trials were taken for each subject. The cognitive event was selected because in that case, the participants can easily concentrate and in eye closed condition, there are no eye blinking/ EOG artefacts [9].

Electrodes placement was configured on the right central (C4), and the right occipital (O2) position as the regions are responsible for problem solving and cognitive function, respectively. It is to note that, the authors have selected the cognitive task from their experience from previous pilot studies where the cognitive task gives clear signals with less noise [10, 11].

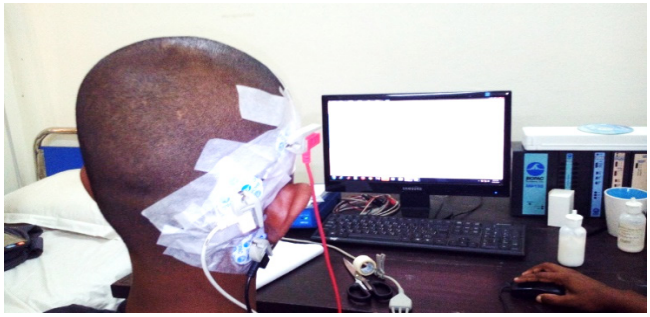


Figure 2: Representational view while conducting the experiment in BME lab, KUET

3.4. Experimental Procedure

3.4.1. Signal Preprocessing

The primarily obtained EEG signals could have contained noise due to muscle movement, eye blinking and hand movement. Along with line frequency was 50 Hz. To pre-process raw EEG was gone through bandpass finite impulse response (FIR) filter with a range of 0.5 to 44 Hz, as it removes the non-linear trends of the signals. Later the signal was further smoothed, taking a moving average over a short period of the signal.

3.4.2. Feature Extraction

Various features were extracted in time and frequency domain, including standard deviation (STDDEV), maximum value (E_{max}), skew (sk), kurtosis (k), power spectrum density (PSD) mean, PSD max, Fast Fourier Transform (FFT) mean, FFT max (total eight feature) were extracted for each subject using the Acknowledge 4.1® software. For feature extraction was done using a 5-second epoch length.

3.4.3. Feature Scaling

The time and frequency domain features, which were extracted from the EEG signal, have different range in their magnitude. As the machine learning models work with various features putting them in the same matrix, it is essential to put all the features in a same range, which is referred to as feature scaling. Two common types of feature scaling is done in preliminary data: standardization and normalization [12, 13]. MinMaxScaling was done in this study in python 3.6.9 platform, as a part of the normalization process. *MinMaxscaler()* function from *sklearn* library was used fo this purpose. Here the data is shrunk within a range between [-1,1].

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

The formula of min-max scaling can be given by equation (1). Here, x_{new} is the normalized value of a feature point x, within a range x_{min} and x_{max} [14].

3.5. Classification

For the diagnosis of tobacco smoking, classification is the main and last step, which is done through machine learning. Machine learning is the application of artificial intelligence, which provides a system capable of learning nature from a given data. There are three categories of Machine Learning models and applications, supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is extensively used for the classification and regression problem [15]. Previous studies worked with EEG have used supervised learnings, especially K-Nearest Neighbour [16], Support Vector Machine [17], Random Forest Classifier [18] and Logistic Regression [19]. Based on the previous studies, these four classifiers were chosen for the data classification in this research.

3.5.1. Logistic Regression (LR)

Logistic regression is a supervised learning model, which works based on the linear method, and the predictions are made using a logistic or sigmoid function σ(t). The sigmoid function is 'S' patterned curve that takes a real number and maps within a range between 0 and 1. Equation (2) represents the sigmoid function.

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}} \quad (2)$$

Considering two types of variables, dependent and independent, Logistic regression predicts the dependant variable based on the independent variable. The 'C parameter' was tuned here in the Logistic regression model to reduce overfitting [20,21].

3.5.2. K-nearest neighbours (KNN)

KNN is a supervised learning algorithm, and a non-parametric method where k nearest training examples in the feature space is taken as input and neighbours vote do the classification generally used for classification and regression. At the very starting point,

KNN read the value of K, type of distance D and test data; then it finds the K nearest neighbours D to the test data and thus sets the maximum label class of K to test data. The same process is gone through an iterative process named looping. In details, its algorithm initializes the value of K from 1 (setting as initial iteration value). After loading data, iteration from initial K =1 (generally) to the total number of training data point while distances specifically Euclidean distance between test data and each row of training data is measured and sorted in ascending order to get topmost k rows from the sorted array and the most frequent class is returned as the predicted class [22]. The value of K was tuned, and the k for best efficiency was chosen in the classifier model in this research to reduce overfitting.

3.5.3. Random Forest Classifier (RFC)

Ensemble learning models, such as Random forests are made of individual decision trees with a logic of group of weak learners to finally make a strong learner while the decision trees operate as divided or conquer. A class is predicted from every decision tree and a final class is predicted by model depending on their vote [22]. Two parameters were tuned in the RFC models in this study, namely, '*n_estimate*', which implies the number of trees in the forest and 'max-depth' which signifies the depth of each tree.

3.5.4. Support Vector Machines (SVM)

An SVM is a supervised learning algorithm, which aims to obtain a hyperplane classifying the data point (data points can be at any side of hyperplane) in feature dimensional space while depending on both linear and non-linear regression. Data points distance across to hyperplane are called support vector whose detection can exchange hyper plane's location [22]. The model used a Gaussian kernel for SVM classifier in this research due to the non-linear trend of the dataset. Two parameters- 'C' and 'gamma' was adjusted within a set of values using the grid search algorithm to reduce overfitting.

3.6. Performance Measures

3.6.1. Sensitivity or True Positive Rate (TPR)

True positive rate or Sensitivity is the proportion of the true positives (desired factor), which is correctly identified from the given test set [23]. The definition of sensitivity can be provided by equation (3), where TP = True Positive and FN = False Negative. In this study, sensitivity is the measure of the proportion of successfully identifying a smoker.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+ FN}} \tag{3}$$

3.6.2. Specificity or True Negative Rate (TNR)

True negative rate Specificity is the proportion of true negative (undesired factor) in which was correctly excluded from the given test sets [23]. The definition of specificity can be provided by equation (4), where TN = True Negative and FP = False Positive. In the case of this study, specificity is the measure correctly identifying a non-smoker. In this study, accuracy is the proportion of successful identification, either smoker or non-smoker.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+ FP}} \tag{4}$$

3.6.3. Accuracy

The overall accuracy is the proportion of true results (either true positive or true negative) in an experiment [23,24]. The definition of accuracy can be provided by equation (5), given that TP = True positive, TN= True Negative, FP= False Positive and FN = False Negative. In this study, accuracy is the proportion of the successful identification, either a specific person or not being that person.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{5}$$

3.6.4. Area under the receiver operating characteristic (ROC) curve (AUC)

ROC is the plot of the sensitivity (true positive rate) against the (1- specificity) or false positive rate, where all the possible combination of TPR and FPR are plotted, showing the trade-off between them [23–25]. As sensitivity and specificity are two major parameters of performance measures, AUC under ROC always provides a compromise between them. Though there are few methods for validation, five-fold cross-validation was done in this study while evaluating the performance measures. The mean value and the standard deviation (SD) were noted, considering the five experimental validations. Thus, the mean sensitivity, specificity and AUC was calculated from the obtained confusion matrix.

4. Results

4.1. Data visualization (Box plot and violine plots)

Data visualization is an important part to observe the data arrangement. Given data points found from the selected features were plotted in box and violin plots to observe the range of each of the features. The following Figure 3 and Figure 4 shows that the time and frequency domain features are having a versatile variation in the range. Range of the difference features varies among themselves either in the time domain or in the frequency domain, and therefore, the feature scaling was performed.

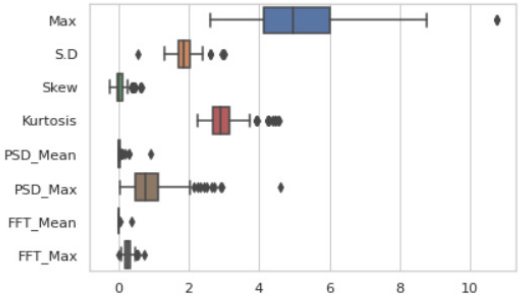


Figure 3: Box Plot of part EEG data showing varying magnitude of differnt features

4.2. Classification Performance

All the features were scaled and were supplied towards the machine learning models after necessary parameter tuning. Four different performance measures were evaluated, namely, sensitivity or true positive rate (TPR), specificity or True negative rate (TNR), accuracy and area under the receiver operating characteristic (ROC) curve (AUC). The obtained results are listed in the Table 1.

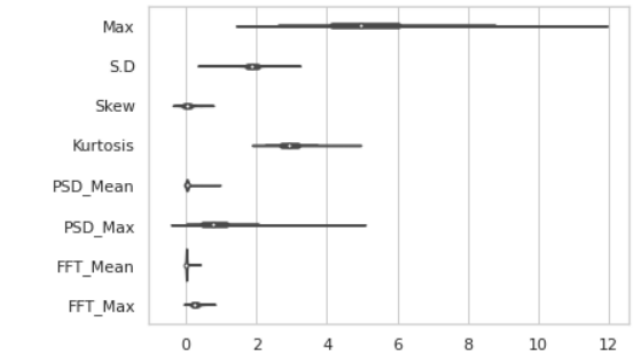


Figure 4: Violine Plot of part EEG data showing varying magnitude of differnt features

Table 1: Performance measures (mean value) for EEG based smoker detection using four different classifiers, five-fold cross-validation

Domain	Performance Measures	LR	KNN	RFC	SVM
Time Domain	Sensitivity	58.9	56.07	63.6	53.1
	Specificity	55.2	57.54	61.3	55.3
	Accuracy	57.4	55.53	62.4	59
	AUC	53.7	54.04	65	60.1
Frequency Domain	Sensitivity	70.7	77.14	81.2	70.9
	Specificity	70.7	71.35	83.7	76.3
	Accuracy	76.1	73.75	82.9	72.9
	AUC	71.4	77.59	83.2	77.1
Time-Frequency Domain	Sensitivity	89.2	83.75	94.3	85
	Specificity	86.3	91.25	92.1	87.2
	Accuracy	**87.2**	**87.5**	**91.3**	**86.5**
	AUC	78.3	80.7	92	88.2

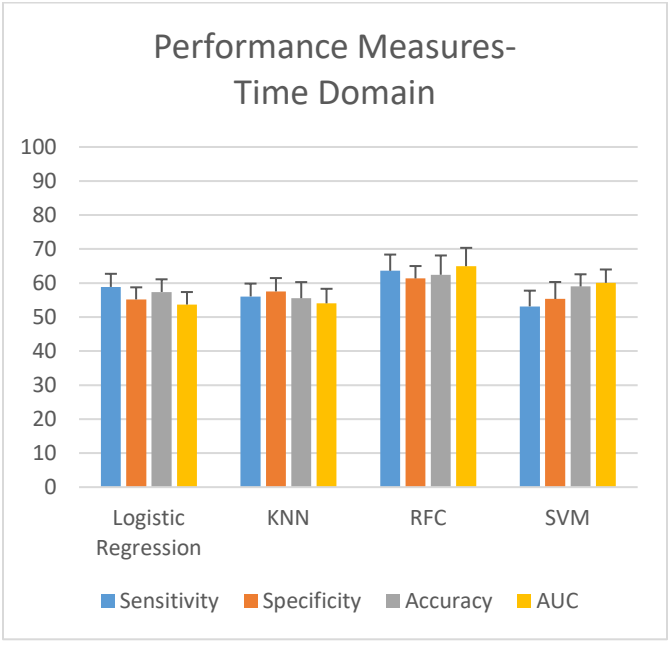


Figure 5: Performance measurement of time domain

4.2.1. Scenario-1: Classification using Time Domain Features

The following Figure 5 shows the plots of the performance measures (mean ± SD) obtained from the classification of smokers and non-smokers using the time domain features from four different classifiers, LR, KNN, SVM and RFC, respectively. The plots show that the gap between sensitivity and specificity is highest in LR (3.68%) and lowest in the case of KNN (1.44%). Overall, RFC gives an accuracy of 62.4%, which performs the best.

4.2.2. Scenario-2: Classification using Frequency Domain Features Domain

The following Figure 6 shows the plots of the performance measures (mean ± SD) obtained from the classification of subjects addicted to smoking using the frequency domain features from four different classifiers, LR, KNN, SVM and RFC, respectively. The plots show that the gap between sensitivity and specificity is higher in KNN (5.8%) and SVM (5.4%) and lowest in the case of LR (0.05%). Overall, RFC gives an accuracy of 82.9%, which performs the best.



Figure 6: Performance measurement of frequency domain

4.2.3. Scenario-3: Classification using the Time-Frequency Domain Features Domain Features

The following Figure 7 shows the plots of the performance measures (mean ± SD) obtained from the classification of subjects addicted to smoking using the time domain features from four different classifiers, LR, KNN, SVM and RFC, respectively. The plots show that the gap between sensitivity and specificity is the highest in the case of SVM (7.5%) and lowest in the case of RFC (2.2%) and ANN (2.2%). Overall, RFC gives an accuracy of 91.3%, which performs the best.

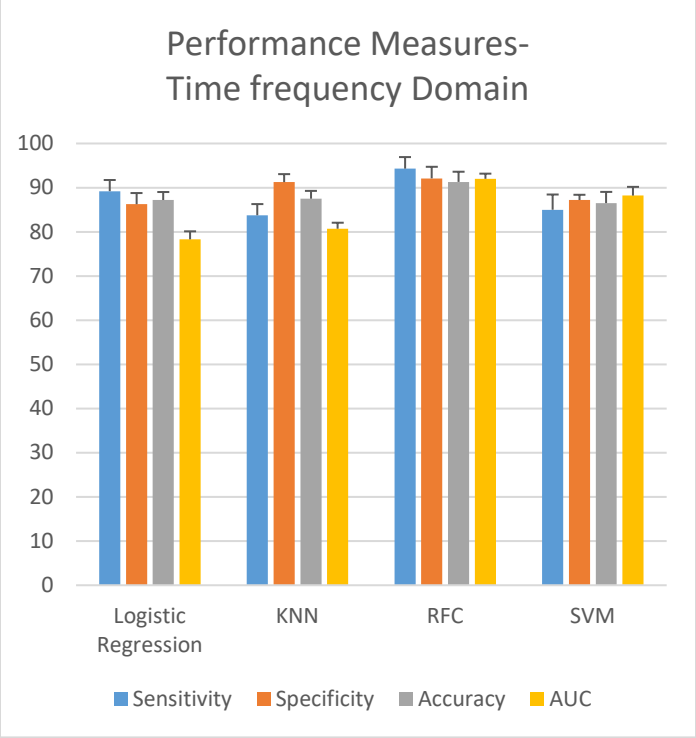


Figure 7: Performance measurement of the time-frequency domain

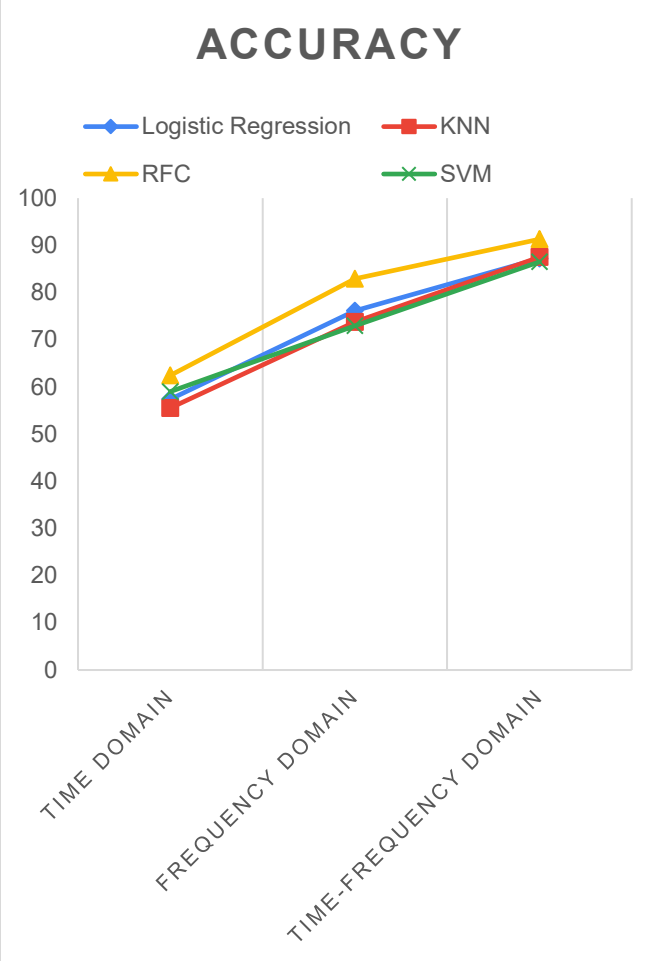


Figure 8: Comparison of accuracy metrics for four different classifiers

4.3. Choosing the best Scenario

The overall accuracy was considered as the reference metrics while finding out the best scenario, as it is difficult to compare different classifiers using several measures. The plots of the accuracy for four different classifiers corresponding to different domain are shown in the Figure 8 below. From the given figure, it is evident that the accuracy for random forest classifier is better than any other domains for all the four classifiers.

The accuracy plots also reveal the relative comparison among the time domain features, frequency domain features and the effect of the fusion of both time and frequency domain. It is evident that the frequency domain features perform better than the time domain features for all of the classifiers. Again, the time-frequency fusion outperforms the previous scenarios when the time or frequency domain feature were used individually. So, in the rest of the paper, the combined-time frequency domain features will be considered for further analysis.

4.4. Choosing the best classifier

The plots for the area under the ROC curve for the classifiers built using the time-frequency domain of EEG features are shown in the Figure 9 below. The figure illustrates that the RFC classifiers show the best compromise between sensitivity and specificity, with covering the highest area under the ROC curve (AUC= 0.92%).

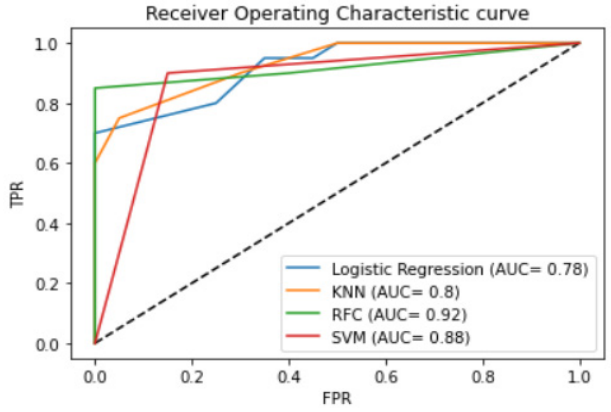


Figure 9: Comparison of AUC for four different classifiers for time-frequency domain

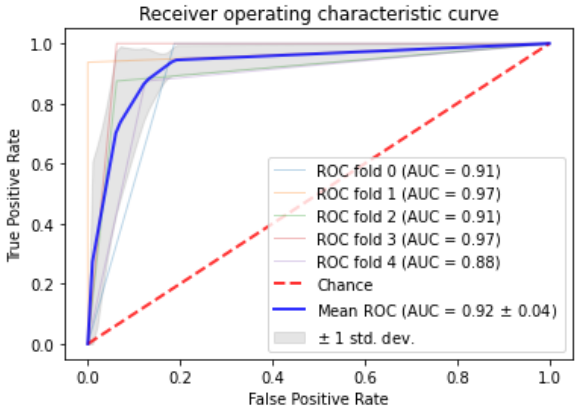


Figure 10: ROC Plots for time-frequency domain-based RFC model with 5-fold CV

Area under the AUC curve plots for 5 different experiments in 5-fold cross-validation with time frequency domain features using RFC classifier is shown in Figure 10. The AUC ranged from (0.78-0.92) for all the models with the time-frequency domain features, with a mean of 0.92 and 0.04 standard deviation. This signifies that the time-frequency domain shows an excellent performance than the other domains of EEG signal in tobacco smoking diagnosis.

5. Discussion

Four different classifiers were used in this study for assessing the performance of the EEG domains for the diagnosis of tobacco smoking. The results revealed that the time-frequency domain performs the best among the other domains. It also revealed that the maximum performance was obtained using the RFC Classifier, with a sensitivity, specificity and accuracy of 94.3%, 92.1% and 91.3%, respectively. Moreover, RFC based model time-frequency domain-based dataset shows promising AUC (0.92), which is a good compromise between sensitivity and specificity.

The finding of this study is consistent with some of the previous studies [1,7]. The study [7] achieved an accuracy ranged from 97.33-97.50%, while using the frequency domain features, such as, PSD and FFT features of EEG signal, which also supports the finding of the current study. Though the study of [7] used only RBF kernel based SVM, the current study validated the other classifiers which support the findings. In another study [1], where authors found the time-frequency domain as the best performing domain, though they have used only one classifier (ANN), and one performance metrics (Accuracy). Contrarily, the time domain performed the worst (RFC accuracy= 62.4%) in their study, while using a holdout approach for validation. The current study used a five-fold cross-validation and validated the outcome of the previous research with multiple classifiers.

Based on the results and analysis in the study, the following research implications and recommendations can be provided. First, using the frequency domain features is always recommended to diagnose tobacco smoking using the EEG signals. However, using time-frequency fusion is highly recommended as this combination provides a higher classification performance. Second, the given outcomes could be correlated with the drug-related impairment in the human brain, which could provide further insight into the correlation between the two addiction.

There are some feasibility issues with the study. First, using EEG sensors to detect the smoking habit could be a cumbersome procedure. However, the research could add value to observe the effect of different cigarettes depending upon their nicotine levels and their impact on the human brain. Second, the use of EEG as ground truth in tobacco-related experiments. If the setup is available in the lab, the EEG measurement could be used as ground truth when observing the effect of Tobaccos smoking on other factors. Third, the impact of drug addiction is more severe in the human brain than smoking. This experiment can add value to the relative comparison of the drug and tobacco addiction. However, the scope of the paper is not out of limitations. Inter-individual difference among participants is another factor, which is needed to be considered. As the paper represented a novel methodology of EEG based diagnosis of tobacco smoking, more research is required to find out the feasibility in real-world conditions as well.

6. Conclusion

To develop an EEG based diagnosis of tobacco addiction, an analysis was done in this study to find out the feasibility of the time and frequency domain features using this proposed model. Here, ultimate results were obtained after applying several steps- feature scaling, tuning of classifiers, and finally with five-fold cross-validation of the developed models. The research investigation found that the combination of the time-frequency domain features with RFC classifier showed the best accuracy while the time domain features showed the lowest accuracy. This analysis shows that time-frequency domain shows the best Accuracy with SVM (86.5 %), LR (87.2%), KNN (87.5%) and RFC (91.3%), Time-domain shows the lowest accuracy with all classifiers while the ffrequency domain shows higher accuracy than time domain, but still, this is less than the combined time-frequency domain performance. Among all the classifiers, RFC showed the best Accuracy and SVM showed the lowest accuracy. However, the experiment could be done on more number of participants to validate the model based on lave one participant out approach. Also, efficiency will increase with the addition of more EEG channels which could be considered for future implications.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors want to thank the faculties, staff, and HDRs for providing the signal recording facilities at Biomedical Engineering Laboratory, KUET.

References

[1] M.M. Hasan, N. Hasan, A. Rahman, M.M. Rahman, "Effect of Smoking in EEG Pattern and Time-Frequency Domain Analysis for Smoker and Non-Smoker," in 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), IEEE: 1–4, 2019. doi: 10.1109/IC4ME247184.2019.9036492

[2] C.N. Metz, P.K. Gregersen, A.K. Malhotra, "Metabolism and biochemical effects of nicotine for primary care providers," Medical Clinics, **88**(6), 1399–1413, 2004.

[3] R.C. O'reilly, Y. Munakata, Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain, MIT press, 2000.

[4] G. Cartocci, E. Modica, D. Rossi, A.G. Maglione, I. Venuti, G. Rossi, E. Corsi, F. Babiloni, "A pilot study on the neurometric evaluation of 'effective' and 'ineffective' antismoking public service announcements," in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE: 4597–4600, 2016.

[5] Z.M. Hanafiah, K.F.M. Yunos, Z.H. Murat, M.N. Taib, S. Lias, "EEG brainwave pattern for smoking behaviour after horizontal rotation treatment," in 2009 IEEE Student Conference on Research and Development (SCOReD), IEEE: 559–561, 2009. doi: 10.1109/SCOReD247184.2019.9036492

[6] T. Warbrick, A. Mobascher, J. Brinkmeyer, F. Musso, T. Stoecker, N.J. Shah, G.R. Fink, G. Winterer, "Nicotine effects on brain function during a visual oddball task: a comparison between conventional and EEG-informed fMRI analysis," Journal of Cognitive Neuroscience, **24**(8), 1682–1694, 2012.

[7] L.C. Chin, A.M. Zazid, C.Y. Fook, V. Vijean, S.A. Awang, M. Affandi, L.S. Chee, "Differentiate Characteristic EEG Tobacco Smoking and Non-smoking," in Journal of Physics: Conference Series, IOP Publishing: 12055, 2019, doi:10.1088/1742-6596/1372/1/012055.

[8] D. Acquisition, Analysis with BIOPAC MP Systems, AcqKnowledge 4 Software Guide. pdf, Biopac Systems, Inc. ISO 9001: 2008.

[9] N. Hasan, M.M. Hasan, M.A. Alim, "Design of EEG based wheel chair by using color stimuli and rhythm analysis," in 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE: 1–4, 2019, doi:10.1109/ICASERT.2019.8934493.

[10] M.M. Hasan, M.H.A. Sohag, M.E. Ali, M. Ahmad, "Estimation of the most effective rhythm for human identification using EEG signal," in 2016 9th International Conference on Electrical and Computer Engineering (ICECE), IEEE: 90–93, 2016, doi:10.1109/ICECE.2016.7853863.
[11] M.M. Hasan, M. AshfaqulIslam, S.A. Imtiyaz, M. MahbubHasan, "Presumption method for detecting and analyzing human mental behavior by employing EEG signal," in 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), IEEE: 519–523, 2018.
[12] M. Wester, W. McMullen, A. Macy, "Biopac Student Lab Pro, Hardware and Software Reference Manual, ver. 2.1, BioPac Systems," Inc., Santa Barbara, CA.[Accessed on 26 April 2018], 1997.
[13] B.B. flow Monitor, Biopac Systems, Inc. ISO 9001: 2000.
[14] C. Colantuoni, G. Henry, S. Zeger, J. Pevsner, "SNOMAD (Standardization and NOrmalization of MicroArray Data): web-accessible gene expression data analysis," Bioinformatics, **18**(11), 1540–1541, 2002, doi:https://doi.org/10.1093/bioinformatics/18.11.1540.
[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, "Scikit-learn: Machine learning in Python," The Journal of Machine Learning Research, **12**, 2825–2830, 2011.
[16] F. Riaz, A. Hassan, S. Rehman, I.K. Niazi, K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," IEEE Transactions on Neural Systems and Rehabilitation Engineering, **24**(1), 28–35, 2015, doi:10.1109/TNSRE.2015.2441835.
[17] S. Bhattacharyya, A. Khasnobish, S. Chatterjee, A. Konar, D.N. Tibarewala, "Performance analysis of LDA, QDA and KNN algorithms in left-right limb movement classification from EEG data," in 2010 International conference on systems in medicine and biology, IEEE: 126–131, 2010.
[18] H. Lee, S. Choi, "Pca+ hmm+ svm for eeg pattern classification," in Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings., IEEE: 541–544, 2003.
[19] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," Computer Methods and Programs in Biomedicine, **108**(1), 10–19, 2012, doi:https://doi.org/10.1016/j.cmpb.2011.11.005.
[20] J. Kim, J. Lee, C. Lee, E. Park, J. Kim, H. Kim, J. Lee, H. Jeong, "Optimal feature selection for pedestrian detection based on logistic regression analysis," in 2013 IEEE International Conference on Systems, Man, and Cybernetics, IEEE: 239–242, 2013.
[21] H. Rajaguru, S.K. Prabhakar, "Non linear ICA and logistic regression for classification of epilepsy from EEG signals," in 2017 international conference of electronics, communication and aerospace technology (ICECA), IEEE: 577–580, 2017.
[22] K. AlSharabi, S. Ibrahim, R. Djemal, A. Alsuwailem, "A DWT-entropy-ANN based architecture for epilepsy diagnosis using EEG signals," in 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), IEEE: 288–291, 2016, doi:10.1109/ATSIP.2016.7523093.
[23] W. Zhu, N. Zeng, N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," NESUG Proceedings: Health Care and Life Sciences, Baltimore, Maryland, **19**, 67, 2010.
[24] A.-M. Šimundić, "Measures of diagnostic accuracy: basic definitions," Medical and Biological Sciences, **22**(4), 61, 2008.
[25] D. Justin, R.S. Concepcion, A.A. Bandala, E.P. Dadios, "Performance Comparison of Classification Algorithms for Diagnosing Chronic Kidney Disease," in 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE: 1–7, doi:10.1109/HNICEM48295.2019.9073568.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 845-852 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Updated Analysis of Business Continuity Issues Underlying the Certification of Invoicing Software, Considering a Pandemic Scenario

Nelson Russo[*, 1], Leonilde Reis[2]

[1]School of Science and Technology, University of Trás-os-Montes and Alto Douro, Vila Real, 5000-801, Portugal

[2]School of Business Administration, Polytechnic Institute of Setúbal, Setúbal, 2914-503, Portugal

ARTICLE INFO

Article history:
Received: 06 October, 2020
Accepted: 11 November, 2020
Online: 27 November, 2020

Keywords:
Business continuity
Invoicing software
Software certification
Electronic invoicing
Pandemia

ABSTRACT

Portuguese organizations that have invoicing software, certified by the Tax and Customs Authority, need to comply with technical requirements that involve business continuity and disaster recovery. The recent tax legislative changes created conditions for the dematerialization of documents, allowing waiving invoice printing, encouraging the adoption of an electronic invoicing and document archiving system. The pandemic situation boosted the need for organizations to integrate this paradigm in their business processes. However, there are some constraints in the implementation of these requirements, due to technical issues, interpretation of tax legislation or the selection of frameworks or good practices for Information and Communication Technologies. The objective of the work is to present a set of concerns underlying the design of a business continuity plan, supported by current tax legislation, by standards and codes of good practice. In view of the constraints of Portuguese business capacity, it is also presented a minimum solution that meets the legal, regulatory, good practices and conceptual requirements of Information and Communication Technologies for initiating the design of a Business Continuity Plan. The method used in this investigation was based on the analysis of international standards ITIL, ISO, CMMI, COBIT through the assertive interconnection with the subject under study with the dispositions stated in the Portuguese legal framework in the field of invoicing. The main result was the conception of a decision support process for designing a guide, concerning the optimization of the business continuity plan design process. In face of the problematic in study, it is considered that the main expected results were achieved, by fostering the design of Business Continuity Plan in Portuguese organizations, in order to reduce the gap between the practices currently in place and the requirements underlying certification, as a way to prepare organizations to deal with disruptive events in invoicing business processes.

1. Introduction

This work is an extended version of the paper [1] originally presented in 2019 at the 14th Iberian Conference on Information Systems and Technologies (CISTI), where some guidelines for the conception of o Business Continuity Plan (BCP) were presented, having in consideration the organization capabilities and the tax and legal obligations.

Since this is a new study, there are limited implementation results to present. Nevertheless, the companies questioned referred to the need to implement electronic archive to reduce costs, optimize business process and archiving procedures and find information faster. This extended version attends those concerns and goes a step further by integrating those requirements in the analysis, providing a new tool to cope with constraints during the pandemic situation.

The above direction presents further the Business Continuity (BC) components and updates the guidelines for the BCP. Now it is induced a key perspective in the development of the guidelines, including the forced paradigm of remote work and the constraints for business operations, consequence of a pandemic situation. Organizations must endeavor and adapt to resist the changes or

*Corresponding Author: Nelson Russo, al70461@utad.eu

thrive with the opportunities and continue the digital transformation.

Adding to this situation, in recent years, the software development industry has been faced with legislative changes, which include more functional, and tax or fiscal requirements. By legal imposition, there was an expansion on the range of organizations covered by the obligation to have Invoicing Software (ISw) certified by the Autoridade Tributária e Aduaneira (AT), the Portuguese Tax and Customs Authority.

The resulting legislative simplifications created conditions for the dematerialization of documents, providing for the possibility of dismissing invoice printing, encouraging the adoption of an electronic invoicing and electronic document archiving system.

Thus, measures of sustainability and cost reduction were incorporated, namely, the reduction of file space, reduction of paper, consumables and printing hardware, allowing also to expand the optimization and automation of invoicing and archive procedures. These measures may be a stimulus for organizations to invest in the development and use of new technological instruments, incorporating a philosophy of innovation and sustainability.

In this sense, having in mind that ISw are one of the components of Information Systems (IS), they are subject to a specific framework in the tax area and to a set of functional requirements that aim to ensure data integrity, security of information and business continuity. In this work, the diplomas that frame and describe the technical requirements that the ISw must observe to obtain a certificate of conformity will be presented.

Whether due to the investment capacity, the integrated knowledge of the requirements involved, the interpretation of the diversity of existing standards and good practices, the perception of imminent risk or simply the definition of priorities, some organizations have not yet started designing the BCP or have the notion that technology will simply solve problems during, and after, a disruptive event in their activity. A BCP does not necessarily need to have a high degree of complexity, in the case of small organizations. It is commonly discussed that the pertinence of analyzing risks, understanding how to continue the business in the event of a disaster and to recover from that disaster are crucial activities in business continuity and that should be understood as adding value to organizations and not just as a requirement to meet to be compliant. However, the failure to define a BCP makes it difficult to comply with tax legislation in the context of the certification of ISw, increasing the duration in time of the certification process, usually in scenarios where organizations, for commercial reasons, need quickly to obtain the certificate issued by AT. In addition, the present pandemic situation triggered the need for organizations to transform digitally and to integrate their business software with electronic invoicing capabilities, especially for online sales.

Thus, the need arose to create a guide for the design of a BCP, to reduce the gap in organizations, between the perception of business continuity requirements and the requirements underlying the certification of an ISw, translating them into concrete proposals with due legal and fiscal justification, standards and good practices in the area of Information and Communication Technologies (ICT).

2. Literature review

The accomplished literature review presents business continuity as a program to be followed by organizations that intend to be prepared. It included the review of regulations in the field of the thematic, as well as the analysis of Data Centers and legislative needs underlying the certification of ISw.

2.1. Business Continuity

Organizations can be thought of as a set of value chains or business processes transversal to the various organic units that compose them [2]. Economic, technological and human uncertainties have long presented organizations with the possibility of crises arising, impeding their ability to operate and maintain their business processes and, finally, survive [3].

Organizations increasingly realize the importance of adopting proactive approaches, it is essential to safeguard business processes and the relationship between them, in order to achieve business objectives [4]. Thus, organizations must prepare themselves in the event of disruptions to their productivity and competitive capacity, especially in business processes, supported by ICT services, made available to the organization.

With this in mind, organizations must develop a set of policies and procedures to minimize the impact of these disturbances. The importance of these measures is emphasized, whose impact may affect productivity and competitive capacity, especially in the case of organizations whose dependence on ICT services is marked. This is the concept underlying the BC, in which an organization must have the strategic and tactical capacity, to plan and respond to business incidents and interruptions [5], in order to continue business operations, at an acceptable level predetermined.

One of the accepted responses in BC is Disaster Recovery (DR), defined as the process performed by an organization, to resume business services, after a disruptive event [6]. In this sense, the authors advocate a DR planning approach [7], favoring the recovery of business processes, especially those supported by ICT service.

However, organizational resilience can be understood, as the ability to increase the probability of continuity of any system against disruptive incidents, whether are internal or external variations, changes, disturbances, disruptions and surprises [4]. This concept of resilience also covers effective planning for the relaunch of business processes in the short term, with the BCP, in addition to the long-term restoration, with the Disaster Recovery Plan (DRP) of interrupted operations, following disruptive events.

Being prepared for these events requires proactive planning of the organization's resources so that it can deal with disasters effectively and efficiently [8]. Therefore, the organization must plan its actions before a disaster occurs, taking into account the crisis management phases, resulting from a disaster, as proposed in Figure 1.

Thus, a BCP tries to eliminate or reduce the impact of a disaster and is designed to avoid or mitigate risks and reduce the time

necessary to restore conditions to a state of normal operation [10], aiming to preserve the assets of an organization [11].

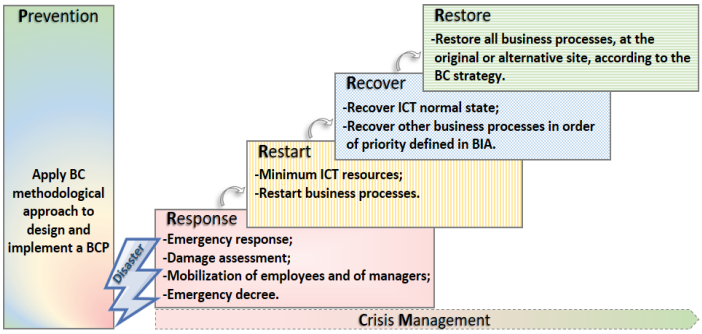


Figure 1: Stages of a business interruption. Adapted from [9].

Some vital activities are identified by Hiles [11], who considers the development of the BCP as a project, with a defined scope, with the analysis of risks and identification of the organization's specific key threats. Through the systematic process of Business Impact Analysis (BIA), vulnerabilities are revealed and it is determined and evaluate the potential effects, or impact on business, of an interruption. The definition of the strategy and requirements for technological resources is a success factor and it is essential to test, exercise and provide training on the BCP [11].

Figure 2 shows the essential components of the Business Continuity Management (BCM), according to the literature review and the standards presented in this work: the contribution of the International Organization for Standardization (ISO) through ISO 22301:2019, ISO 22313:2020 and ISO/IEC 27031:2011, the National Fire Protection Association (NFPA) emergency preparedness standard called NFPA 1600. Includes the vision of the Information Technology Infrastructure Library (ITIL), the perspective of the Control Objectives for Information and related Technology (COBIT) and Capability Maturity Model Integration (CMMI).

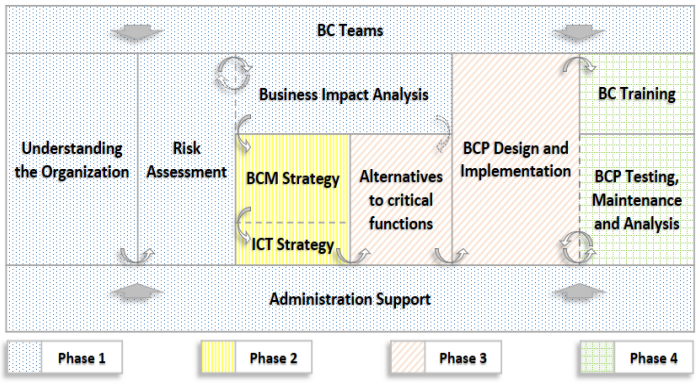


Figure 2: BCM components. Adapted from [12]

Figure 2 presents a sequence of action and interrelationship of the components of the BCM. It also presents the time phasing, in the methodological approach for the design and maintenance of a BCP [12], suggesting the set of components that must be completed, to start another phase.

2.2. Data Center

The Data Center infrastructures must be seen as a guarantee of fail-proof and insurance of business continuity, but if there is a disruptive event, it should be at least as short as possible. The Uptime Institute created the Tier Classification System standard, which is the global language of Data Center performance [13] as a means of effectively evaluating it in terms of business requirements for system availability. It offers a consistent method of comparing unique and customized installations based on the expected performance or uptime of the site's infrastructure.

2.3. Standards

A set of standards was reviewed in order to incorporate good practices on the business continuity topic.

NFPA 1600 2019

Widely used by entities, NFPA 1600 2019 was adopted by the United States Department of Homeland Security as a voluntary consensual standard for emergency preparedness [14].

The standard consists of 10 chapters and 13 annexes in which a common set of criteria for all hazards is established in disaster, crisis and emergency management and business or operations continuity programs. It follows the Plan-Do-Check-Act approach, providing the fundamental criteria of preparation and resilience, including the planning, implementation, execution, evaluation and maintenance of programs for prevention, mitigation, response, continuity and recovery [15]. The standard may be applicable to public, private, non-profit and non-governmental entities.

ISO 22301:2019

The ISO 22301:2019, on the theme “Security and resilience - Business continuity management systems - Requirements”, is a standard for implementing a BC Management System and continuously improving the BC resources based on priorities and feedback management. It specifies the requirements for implementing, maintaining and improving a management system to protect against, reduce the likelihood of, prepare for, respond to and recover from disruptions when they arise [16].

ISO 22313:2020

The ISO 22313:2020 standard, with the theme “Security and resilience - Business continuity management systems - Guidelines on the use of ISO 22301”, provides guidelines and recommendations for the application of the requirements of the Business Continuity Management System (BCMS) indicated in ISO 22301. The guidelines and recommendations are based on international good practices. This standard is applicable to organizations that implement, maintain and improve a BCMS [17], who seek to ensure compliance with the stated business continuity policy, who need to be able to continue delivery of products and services, at acceptable predefined capacities following a disruption and who seek to improve their resilience through the effective application of the BCMS.

ISO/IEC 27031:2011

The International Standard ISO/IEC 27031:2011 entitled “Information technology - Security techniques - Guidelines for information and communication technology readiness for business continuity”. This standard describes the concepts and principles of ICT readiness for business continuity, and provides a framework of methods and processes to identify and specify all aspects (such

as performance criteria, design, and implementation) for improving an organization's ICT readiness to ensure business continuity [18].

2.4. Reference Standards and Good Practices

Within the scope of standardization and improvement of business processes, there is a concern with the application of good practices. Thus, due to its relevance, this chapter presents the set of frameworks and good practices ITIL, COBIT and CMMI.

ITIL

ITIL 4 brings ITIL up to date by re-shaping much of the established ITSM practices in the wider context of customer experience, value streams, and digital transformation, as well as embracing new ways of working, such as Lean, Agile, and DevOps. ITIL 4 provides the guidance organizations need to address new service management challenges and utilize the potential of modern technology. The guidance provided can be adopted and adapted for all types of organization and service. It is designed to ensure a flexible, coordinated and integrated system for the effective governance and management of Information Technology-enabled services [19].

One to the Service management practices, under ITIL management practices, is Service continuity management. Its purpose is to ensure that the availability and performance of a service are maintained at sufficient levels in case of a disaster [19].

COBIT

COBIT 2019 was designed and created by Information Systems Audit and Control Association (ISACA). Is a framework for the governance and management of information and technology, aimed at the whole enterprise. COBIT 2019 is composed of 40 ICT processes in two fundamental areas, Governance and Management, divided into four domains. COBIT describes how an enterprise can design a customized and tailored governance solution for enterprise Information and Technology (IT) and should build its governance system for enterprise ICT [20].

COBIT defines the components to build and sustain a governance system: processes, organizational structures, policies and procedures, information flows, culture and behaviors, skills, and infrastructure. It defines the design factors that should be considered by the enterprise to build a best-fit governance system and addresses governance issues by grouping relevant governance components into governance and management objectives that can be managed to the required capability levels [20].

It was important in this work to characterize the focus on strategic alignment between ICT and business solutions; in the generation of value, but optimizing costs and ICT services; risk management, disaster recovery and continuity of operations.

CMMI

CMMI V2.0 is an integrated set of best practices that enable business to improve performance of their key business processes. It helps a business to understand its current level of capability and performance and its practices can guide improvement [21].

CMMI deals with various Capability Areas like “Managing Business Resilience”, that addresses the ability to anticipate, prepare for, and respond to interruptions in order to continue operations. Incorporates Practice Areas namely “Incident Resolution and Prevention” to resolve and prevent disruptions to sustains service delivery levels. “Risk and Opportunity Management” in order to identify, record, analyze, and manage risks and opportunities and “Continuity” to plan mitigation activities for disruptions to operations so that work can continue or resume [21].

3. Tax legislation

Portuguese tax legislation defines technical requirements for ISw [22] to guarantee data integrity, information security or business continuity and are described, briefly, in the legislation presented below. Some requirements establish the conditions to be observed when scanning and archiving in electronic format, intended to replace the respective paper archive. It also regulates the technical conditions for the issue, upon acceptance by the recipient, of invoices issued electronically and for their conservation and archiving, as well as for the existence and conservation of backup copies of the data supporting the ISw [23].

3.1. Decree Law No. 28/2019

The main objectives of Decree Law no. 28/2019, of February 15, of the Presidency of the Council of Ministers, are to promote legislative simplification and harmonize divergent rules on the conservation of documents for tax purposes. In this context, Decree Law no. 28/2019 created the conditions for the “Paperless invoice” and apposition of Quick Response (QR) code.

It provided for the possibility of waiving the printing of invoices and also strengthened the conditions for the dematerialization of documents, encouraging the adoption of an electronic invoicing system and electronic document archive, being an asset for the development of a more ecological vision, based on reducing the use of paper, inks and printing toners. The conditions created allow organizations to reduce costs with the fulfillment of tax obligations, stimulating the development and use by organizations of new technological instruments, incorporating a philosophy of innovation and debureaucratisation. For this purpose, a substantial reform of the rules applicable to the archiving of books, records, databases and supporting documents of accounting was introduced [23].

3.2. SAF-T (PT)

The Standard Audit File for Tax Purposes - Portuguese Version (SAF-T (PT)) is a standardized tax audit file for data export in Extensible Markup Language (XML) format, mandatory for ISw, created in 2007, with the publication of Ordinance 321-A/2007, of March 26, of the Ministry of Finance and Public Administration.

Its objective is to allow an easy export, at any time, of a predefined set of accounting, invoicing records, and relevant tax documents, in a legible and common format, regardless of the ISw used, without affecting the internal structure of the ISw database or its functionality [24]. The SAF-T (PT) file is intended to facilitate the collection and processing, in electronic format, of the relevant tax data. It allows simplifying procedures and boosting the use of new technologies [24], which may be the file that contains

the list of the relevant tax documents registered on archive. The most current SAF-T (PT) data structure is set out in Annex I to Ordinance No. 302/2016, of December 2, of Finance.

3.3. Decree Law No. 198/2012

The Decree Law no. 198/2012, of 24 August, of the Ministry of Finance, the starting point for the known E-Fatura service, creates a regime that regulates, namely, the electronic transmission of data elements of invoices and other documents with tax relevance, including printing of transport documents, to AT. Establishes measures to control the issuance of invoices and other documents of tax relevance, defines the means for its communication to AT and creates a tax incentive when the acquirer requires such documents [23].

It should be noted that this communication can be made through electronic transmission in real-time integrated with a certified ISw using a web service provided by AT or by sending the summary file, based on the file SAF-T (PT), exported periodically (usually monthly) by a certified ISw.

3.4. Invoicing Software Certification

Compliance with the provisions of Ordinance No. 363/2010, of 23 June, allows obtaining the title of an ISw certified by AT [25]. AT certification is important to foster compliance with the law and legal technical requirements and good ICT practices.

Taxable persons, whose obligation to issue an invoice is subject to the rules established in Value Added Tax Code (VATC), are obliged to use, exclusively, ISw that have been subject to prior certification by AT, whenever have had, in the previous calendar year, a total turnover exceeding 50,000 euros or are using an ISw or have organized accounting.

Thus, if organizations do not choose to use an ISw that have already been certified and intend to use those that they have produced, they should request their certification. To start the certification process, organizations must submit, on the Finance Portal, an official "Modelo 24" form.

The issuance of the certificate may be preceded by compliance tests to observe the requirements resulting from Portuguese tax legislation and a set of defined technical requirements [25]. It should be noted that the certified version of an ISw must comply with the corresponding requirements, which is the responsibility of the ISw producer, even after the compliance tests and throughout the life of the ISw [23].

The compliance tests are conducted by AT, being responsible for analyzing, clarifying and conducting the face-to-face compliance testing meeting an AT Inspector, with functions assigned to certification of ISw.

Since the beginning of the COVID-19 pandemic declaration, faced with the impossibility of free movement and other restrictions, AT has prepared a videoconference solution that would allow to give support to software producers and to safely perform compliance tests. After all the relevant requirements have been verified by videoconference, a meeting would be scheduled to conduct compliance tests, in a face-to-face meeting. Thus, it was possible to keep the certification process active.

AT maintains, on the Finance Portal, an updated list of programs and respective certified versions, as well as the identification of the software producers, proving that during the COVID-19 confinement months, several ISw were certified.

3.5. Ordinance No. 363/2010

The preamble to Ordinance No. 363/2010, of 23 June, states that the increasing use of ISw has undeniable advantages in terms of the speed of information processing. However, it introduces new risks in terms of fiscal control, due to the possibility of subsequent adulteration of recorded data, enhancing situations of tax evasion [25].

In this perspective, the aforementioned ordinance defines rules so that the ISw observe the requirements that guarantee the inviolability of the information initially registered, allowing, consequently, that only the ISw that respect such requirements can be used, after certification by AT. The certification of ISw depends on the cumulative verification of several requirements, among them must be able to export the SAF-T (PT) file, have access control to ISw and respect the technical requirements defined in Order No. 8632/2014.

3.6. Order No. º 8632/2014

Order No. 8632/2014, of July 3, from the Ministry of Finance, was published to comply with the provisions of Ordinance No. 363/2010, which states that one of the requirements for obtaining ISw certification is to observe the technical requirements approved by the AT. Thus, ISw even if already certified, must comply with its approximately 100 technical requirements in different areas, including backup policies, data recovery, access controls or document signing and security [22].

4. Methodology

The research-action methodology in the realistic perspective was the main axis of the investigation protocol. It helped to reduce the complexity of the problem, due to the variety of categories of organizations and their ICT strategies [26], as well as business specificities. It would be a hindrance to integrate the requirements and the hypotheses for solving the problem in a single approach [27]. Thus, a questionnaire was created for ICT managers as a way of capturing relevant information about the invoicing solution implemented in the organization. From the data analysis, it's inferred a BCP proposal that integrates the researcher's knowledge about the problem, the applicable legal and normative requirements, the constraints and the various conceptual technological alternatives of resolution in order to outline a sufficiently comprehensive guide for most of the organizations in Portugal.

5. Results

The guide for the design of a BCP is outlined, starting with a questionnaire with the main questions, structured in specific knowledge areas that the ICT managers must ask in the analysis of the problem, as shown in Figure 3. It culminates in a better understanding of the business processes and procedures involved that will be the target of the BCP, regarding procedures to be designed for disaster recovery and business continuity of the ISw. It enhances the understanding of ISw's data backup policies and

approaches to the implementation of a technological conceptual solution capable of eliminating vulnerabilities and mitigating the effects of a disruptive event in the invoicing business processes.

Structurally, the questionnaire contains a set of multiple-choice questions, each contributing to the perception of the organization invoicing solution and to the compliance with specific requirements defined in the tax legislation, standards and good practices researched.

Figure 3 shows the decision support process with the questions that make up the questionnaire. The word “invoice” refers to all relevant tax documents, including, for example “orders” or, “credit note”. The order in which the questions are answered is not important, since the software that supports the guide adapts to each new question answered. Three questions allow diverting the flow of the program, as its answer invalidates the following interconnected questions:

9. "Is there an Invoicing Software?" - if not, it is not needed to ask about the producer responsible for the ISw and its architecture;

13. "Is the Invoicing Software connected to a central software?" - if not, there is no need to conclude about the location of the central database in the client/server configuration;

15. “Are invoices issued electronically?” - if not, there is no need to question if, or how Electronic Data Interchange (EDI) is used.

All other questions follow a common reasoning flow and are positioned according to the theme of the question, which may be important to request answers to the questionnaire from another employee of the organization. It is important to note that the guide was specially designed for organizations that have an ISw, and it is not an objective to find proposals outside this scope. The questions areas are:

- Organization: Allows the organization to be classified according to the criteria of Small and Medium Enterprises (SME) [28], to know the sector of activity in order to understand the investment capacity [29] and the total turnover to infer, and also about the obligation to have an ISw certified by AT;
- Invoicing volume: Allows to verify the invoicing issuance days; whether the organization has continuous invoicing (for example for scheduling jobs in periods of less activity) or only during normal working hours, as well as the average number of documents issued per day. It aims to reveal the periodicity of backup copies and more appropriate techniques, taking into account the conclusions about investment inferred in the “Organization” area;
- Printing: The objective is to understand whether organizations print the documents issued and keep the accounting archive in paper, in order to suggest scanning and archiving in electronic format, as well as the possibility of recovering documents in case of corruption of the database;
- Invoicing Software: Find out if the organization uses an ISw, even if not required, inferred by the conclusions in "Organization". Know if the ISw was developed in-house or acquired in order to understand the decision-making capability on requirements and technical capability on the ISw. It is important to understand if the ISw is desktop-based, web-based or a terminal application in order to infer the technological infrastructure associated with the physical location of the database;
- Electronic issuance: It aims to infer the processing of the documentation produced and propose the electronic issuance according to the technique used and with the conclusions in “Invoicing volume”;
- Manual issuance: Intends to assess the need to comply with the requirements set out in point 2.4 of Order 8632/2014 [22]. Propose business continuity in case the ISw is inoperable, taking into account the conclusions in “Invoicing volume”.
- Electronic Archive: Intends to assess the need to comply with the requirements set out in Decree Law No. 28/2019. Aims to propose a reduction of archiving costs, the efficiency of archiving procedures, especially those supported on remote work business processes and improve business continuance readiness.

The flow of decisional process and the questions that are part of the questionnaire are presented in Figure 3.

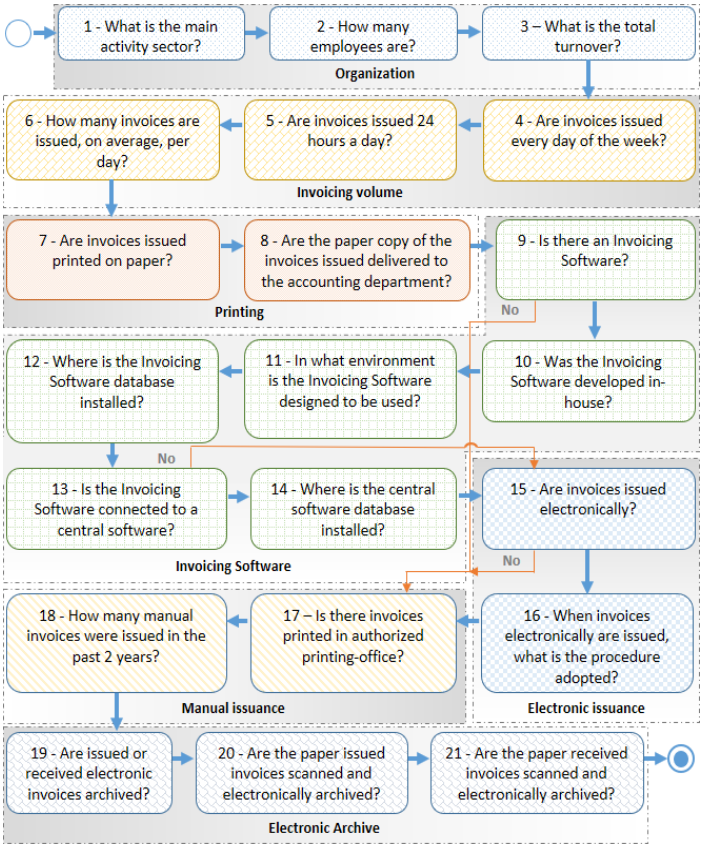


Figure 3: Decision support process for the design of the guide.

The questionnaire presented in Figure 3 has the basic or relevant questions to ask, and the multiple-choice answer will drive the decisional process conclusions. For example, when questioned about the connection to a central software, the answer will have a variety of options to try to determine the architectural solution implemented.

5.1. Discussion

Stablished or entrepreneurial entities, when starting or reformulating their commercial activity and business processes, should take into account the potential of legislative considerations mentioned in this work. The focus on online businesses, amplified by the increased use of telework, resulting from the measures to contain this recent pandemic, allows some management, optimization and control functions to be performed remotely. The effects are reflected in the need to optimize business processes, the time spent on activities, but above all in technological innovation that mitigates the gaps underlying this new vision of the business and digital world. This new vision will depend on the dynamization of ICT in business processes and the quick response to this desire, where there is a need to make quick and effective decisions, but with the necessary efficiency so that the change is adequate.

5.2. Future work

The aim is to design an approach that allows the questionnaire to be published online and to obtain more data for treatment, in order to refine the design of a BCP supported by technological solutions.

Nevertheless, for questionnaire evolution, some other requirements need to be addressed, especially the ones who improve the security of the software system and some other ones that enhance business continuity and disaster recovery.

Another research direction is to design a case study, to evaluate the business processes dynamization or optimization and obtain the return on investment achieved. The study should occur before and after the implementation of electronic invoicing and/or electronic archiving in the organization, in a set of preselected business processes. The target organization is one that intents to transform its business processes for telework, triggered for example, by restrictions due to pandemic situations.

6. Conclusions

The previous identification of the constraints felt by the organizations in the design of a BC solution allowed to conduct the literature review and the identification of the issues, considered relevant, in the context of supporting the design of a guide for the design of BCP. The review of the literature on the theme of BC, including standards and good practices, and the relevant applicable tax legislation was important in this work. The systematization of information in this field adds value to the scientific community, as well as to users who are part of this type of problem.

On the other hand, the change in the paradigm of issuing and communicating invoices and tax relevant documents triggered the need for organizations to prepare for the challenge of innovation or technological adequacy for electronic invoice issuing, as well as for the electronic archive. The current pandemic situation has accelerated this need for digital transformation.

The tax legislation points to the requirements to be met and provides for the possibility of issuing documents by electronic means, the possibility of their electronic transmission to AT, and the possibility of their electronic archive. Understanding what are the correct procedures and what are the requirements involved in the certification of an ISw in Portugal it is considered relevant for organizations that develop the software, as well as for organizations that decide to purchase a solution. This work aimed to help the identification of the main regulatory documents in the thematic, as well as the legal characterization of invoicing in Portugal, in order to support organizations in the design or adaptation of the ISw. The implementation of electronic invoicing and electronic archiving procedures can enhance business continuity and boost the innovation that leads to new software products, digital transformation, marketing approaches, or new ways of making business.

The questionnaire made it possible to systematize the variables involved underlying the tax requirements, the existing technological infrastructures, and the invoicing and archiving procedures implemented. Through a process of analysis of the answers, vulnerabilities can be perceived, as well as the needs for business continuity and disaster recovery. It is considered that the constraints perceived by organizations were reduced by improving the understanding of the variables that they must observe, the areas in the legislation and regulations that they must integrate, as well as the needs of human and technological resources, according to their investment capacity.

This work encourages the adoption of concrete strategies and procedures for business continuity, where invoicing and electronic archiving may be relevant, for organizations with an ISw, certified or not by AT.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] N. Russo, L. Reis, "Análise da Problemática Subjacente à Certificação de Programas Informáticos de Faturação," in CISTI'2019 - 14th Iberian Conference on Information Systems and Technologies. Coimbra, Portugal, 2019. doi:10.23919/CISTI.2019.8760638.

[2] R. Medina-Mora, T. Winograd, R. Flores, "Action Workflow as the Enterprise Integration Technology," Bulletin of the Technical Committee on Data Engineering, IEEEE Computer Society, **16**(2), 1993.

[3] B. Herbane, "The evolution of business continuity management: A historical review of practices and drivers," Business History **52**(6), 978-1002, 2010. doi:10.1080/00076791.2010.511185.

[4] N. Sahebjamnia, S. A. Torabi, S. A. Mansouri, "Building organizational resilience in the face of multiple disruptions," International Journal of Production Economics, **197**, 63-83, 2018. doi:10.1016/j.ijpe.2017.12.009.

[5] R. K.Ramakrishnan, S.Viswanathan, "The importance of Business Strategy in Business Continuity Planingem" in A. Hiles, The definitive handbook of Business Continuity Management Third Edition. West Sussex, England: John Wiley & Sons, Ltd, 2011. doi:10.1002/9781119205883.ch3.

[6] M. M.Al-shammari, A. A. Alwan, "Disaster Recovery and Business Continuity for Database Services in Multi-Cloud" in 1st International Conference on Computer Applications & Information Security (ICCAIS), 1-8, 2018. doi:10.1109/CAIS.2018.8442005.

[7] B. Herbane, D. Elliott, E. M. Swartz, "Business Continuity Management: time for a strategic role?," Long Range Planning, **37**(5), 435-457, 2004. doi:10.1016/j.lrp.2004.07.011.

[8] N. Sahebjamnia, S. A. Torabi, S. Mansouri, "Integrated business continuity and disaster recovery planning: Towards organizational resilience," European Journal of Operational Research, **242**(1), 2015. doi:10.1016/j.ejor.2014.09.055.

[9] A. Guindani, "Gestão da Continuidade dos Negócios," Integração, 2008.

[10] V.Cerullo, M. J. Cerullo, "Business Continuity Planning: A Comprehensive Approach" Journal Information Systems Management, **21**(3), 70-78, 2004. doi:10.1201/1078/44432.21.3.20040601/82480.11.

[11] A. Hiles, "An introduction to business continuity planning" in A. Hiles, The Definitive Handbook of Business Continuity Management, xxiii-xxvii, John Wiley & Sons, Ltd, 2007.
[12] N. Russo, L. Reis, "Methodological Approach to Systematization of Business Continuity in Organizations," Handbook of Research on Multidisciplinary Approaches to Entrepreneurship, Innovation, and ICTs, **24**, 2020. doi:10.4018/978-1-7998-4099-2.ch010.
[13] Uptime Institute, "Tier Classification System", Uptime Institute LLC, September 26, 2020, retrieved 2020-10-12, [Online] Available: https://pt.uptimeinstitute.com/tiers.
[14] National Fire Protection Association, "NFPA 1600® Standard on Continuity, Emergency, and Crisis Management", April 28, 2020, retrieved 2020-04-28, [Online] Available: https://www.nfpa.org/codes-and-standards/all-codes-and-standards/list-of-codes-and-standards/detail?code=1600.
[15] NFPA 1600, "NFPA 1600® Standard on Continuity, Emergency, and Crisis Management 2019 Edition". National Fire Protection Association, 2019.
[16] ISO 22301:2019, "Societal security - Business continuity management systems – Requirements", Switzerland: ISO, 2019.
[17] ISO 22313:2012, "Societal security - Business continuity management systems – Guidance", Switzerland: ISO, 2012.
[18] ISO/IEC 27031:2011, "Information technology - Security techniques - Guidelines for information and communication technology readiness for business continuity", ISO/IEC, 2011.
[19] ITIL, "ITIL Foundation ITIL 4 Edition". Londres: AXELOS, 2019.
[20] ISACA, "COBIT 2019 Framework - Governance and Management Objectives". Schaumburg, IL 60173, USA: ISACA, 2018.
[21] CMMI Institute, "CMMI Model V2.0". CMMI Institute, 2018.
[22] Order No. 8632, "Despacho n.º 8632/2014, de 03 de julho". Lisbon, Portugal: Ministry of Finance, 2014.
[23] Decree Law No. 28, "Decreto-Lei n.º 28/2019, de 15 de fevereiro". Lisbon, Portugal: Presidency of the Council of Ministers, 2019.
[24] AT, SAF-T PT (Standard Audit File for Tax purposes) - Portuguese Version, September 19, 2020. http://info.portaldasfinancas.gov.pt/pt/apoio_contribuinte/SAFT_PT/Paginas/news-saf-t-pt.aspx.
[25] Ordinance No. 363, "Portaria n.º 363/2010, de 23 de junho". Lisbon, Portugal: Autoridade Tributária e Aduaneira, 2010.
[26] KPMG, "Risco e Resiliência - Estudo de Continuidade do Negócio em Portugal", 2018.
[27] R. Yin, Case Study Research - Design and Methods (volume 5, fourth ed). California: Sage Inc, 2009.
[28] Commission, COMMISSION RECOMMENDATION of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. Official Journal of the European Union. The Commission of the European Communities, 2003.
[29] N. Russo, L. Reis, "Caracterização da faturação em Portugal: sob a perspetiva da certificação de programas informáticos de faturação," Revista de Ciências da Computação, **14**, 67-84, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 846-859 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Application-Programming Interface (API) for Song Recognition Systems

Murtadha Arif Bin Sahbudin*,Chakib Chaouch, Salvatore Serrano, Marco Scarpa

Department of Engineering, University of Messina, Messina, 98166, Italy

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 09 February, 2021
Online: 04 April, 2021

Keywords:
Song Recognition
Mobile API
Signal Processing
Database Clustering
IoT

ABSTRACT

The main contribution of this paper is the framework of Application Programming Interface (API) to be integrated on a smartphone app. The integration with algorithm that generates fingerprints from the method ST-PSD with several parameter configurations (Windows size, threshold, and sub-score linear combination coefficient). An approach capable of recognizing an audio piece of music with an accuracy equal to 90% *was further tested based on this result. In addition the implementation is done by algorithm using Java's programming language, executed through an application developed in the Android operating system. Also, capturing the audio from the smartphone, which is subsequently compared with fingerprints, those present in a database.*

1 Introduction

An audio representation includes a recording of a musical piece's output. Digital sound recordings are based on the analog audio signal being sampled. Sampling is achieved by capturing the signal amplitude at a specified sampling rate and storing them in binary format. In terms of recording efficiency, the sampling rate and the bit rate (number of bits used to store each sample) are the two most important variables. Audio CDs use a 44.1 kHz sampling rate or 44,100 samples per second, and each sample uses 16 bits, which mainly an industry standard.

Common audio streaming sites host millions of audio files, and thousands of broadcast stations transmit audio content at any given time. The ever-increasing amount of audio material, whether online or on personal devices, generates tremendous interest in the ability to recognize audio material. It achieves this by using identification technology that seeks to work at the highest degree of accuracy and specificity.

Songs recognition identifies a song segment either from a digital or an analog audio source. Song rankings are based on radio / TV broadcasting or streaming; copyright protection for songs or automatic recognition of songs that a person wishes to identify while listening to them are different applications of such a system. Important information such as song title, artist name, and album title can be provided instantly. To create detailed lists of the particular content played at any given time, the industry uses audio fingerprinting systems to monitor radio and TV broadcast networks. Through automatic fingerprinting devices, royalties' processing relies on the broadcasters who are required to produce accurate lists of content being played.

Given the high demand for an application, several approaches have already been studied based on song fingerprinting recognition, such as [1]–[4] and [5]. Nowadays, the state of the art of recognition techniques are those developed by Shazam [6], [7] and SoundHound [8], and the detection system by [9]. These services are widely known for their mobile device applications.

Audio streams of many broadcast channels or recordings of different events are typically analyzed using fingerprint systems for media monitoring. As these systems work on massive quantities of data, the data models involved should be as small as possible, while the systems need to efficiently run on massive and growing reference databases. Besides, high robustness criteria are determined by the application for media monitoring. Although the sensitivity to noise may not be the primary concern for this use case, the systems need to identify audio content that different effects may have changed.

Android is the most popular mobile operating system globally, despite the presence on the world market of the likes of Apple iOS. It is mainly used for smartphones and tablets, but thanks to its characteristics, it is also extended to other devices, such as laptops, cameras, and IoT devices. Developing applications requires Android Software Development Kit (SDK), which contains all the tools to create and run new software, such as debuggers, libraries,

*Corresponding Author: Murtadha Arif Bin Sahbudin, Department of Engineering, University of Messina, Email: msahbudin@unime.it

and emulators. The integrated development environment (IDE) officially supported for this purpose is Android Studio, released and available for free from the official developer site for Android and the SDK. Applications are Java-based and are distributed through self-installing packages, i.e., Android application package (APK) files (a variant of the format JAR), which contain all the components and resources of the created software, including source code, XML, images, and binary files.

In particular, this research deals with the process and the theoretical notions that lead to the generation of linekeys previously mentioned in [10], and [11]. Music recognition is a method of identifying a segment of an audio signal from a digital or analog source; this process uses the power spectral density (PSD) to process the acquired data to obtain complete information about the audio signal source. The fingerprint generation takes place on the client-side, in this case, the Android mobile application. The communication protocol, shown in Fig. 1, depicts the process of communication and interfacing with the recognition server with database collection.

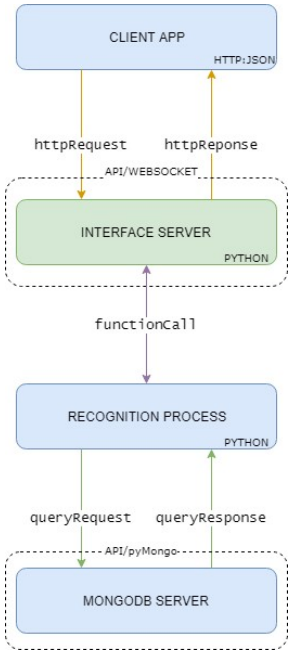


Figure 1: Client and server protocol interface

2 Challenges

The number of songs in the music industry has recently increased significantly, according to a report in [12]. With massive databases, the management and identification of songs using a conventional relational database management system have become more difficult. For large datasets, a common linear search technique that checks the existence of any fingerprint in an array one at a time has a noticeable decrease in efficiency [13]. The stored information, therefore, needs a scalable database system to meet the execution time, memory use, and computing resources for recovery purposes, which is suggested in [14].

Song recognition systems usually operate on vast amounts of data and are expected to meet several robustness requirements depending on the actual use case. Robustness to different kinds of lossy audio compression and a certain degree of noise would seem to be the minimum requirement. Systems designed to detect short audio segments' microphone recordings involve high background noise robustness, such as noise and distortion, or even multiple songs played in the surrounding.

It is crucial to have robust and quick recognition for effective song information retrieval. Major consumers need details about trending tracks, airtime schedules, and song versions, such as music labels, manufacturers, promoters, and radio stations. They, therefore, demand an application that is capable of generating information that is fast and precise.

In the field of real-time song recognition, the entertainment industry, particularly in music, the extensive collection of digital collections, and the commercial interest are opening new doors to research. In 2017, the global digital music industry expanded by 8.1 percent, with total revenues of US$ 17.3 billion, according to IFPI 's Global Music Report 2018 [15]. For the first time in the same survey, 54 percent of the revenue alone comes from digital music revenue.

However, the most challenging application for bringing new songs to listenership is still the FM frequency radio station. The FM frequency channel for music broadcasting in European countries is still actively reliant on radio stations [16]. Radio stations and music companies have been working to advance music industry data analytics by creating ways of analyzing broadcast songs through new services and platforms.

It is an interest to broadcasters and advertisers to measure radio audience size and listening patterns over a broadcast radio station to achieve a source of revenue [17]. However, based on demographics and psychographics (psychological criteria) of the target audience of the station, variations in the region of station promotional material can be predicted [18].

In addition to robustness criteria, the seriousness or effect of incorrect results must be considered, and the necessary performance recognition characteristics of fingerprinting systems must be taken into account. For instance, if a song recognition system is used, an unidentified match is missing, the user can waste storage space. However, on the flip side, a specific song recognized as false match systems that report false positives should be avoided.

Most critically, false positives are expensive for large-scale media monitoring; revenue might be attributed to the wrong artist. False negatives, another form of error, may lead to hours of unidentified material that will have to be checked with manual effort. Any form of error would increase the maintenance cost of a system.

To overcome it, we have proposed a more scalable big data framework using fingerprint clustering. Besides, a new recognition algorithm also was required for the new clustered collection. We also compared the performance from both the legacy system (non-clustered) and the new clustered database.

We define extensions of scale modifications that are likely to be encountered when developing a framework for monitoring FM radio broadcasting stations—investigating our dataset of the reported output of radio segments through the percentage of accuracy. This estimation will serve as the appropriate gold standard throughout this study, i.e., a device should be robust to at least this range of scale noise but may be needed to cope with even more severe distortions.

3 Problem Statement

Fingerprint databases, recognition process instances, and FM transceivers are the legacy system in commercial use. The problems of an audio recognition system in normal use can be described from the following aspects:

a) Near Similarity: This occurs when virtually the same audio fingerprints are produced by two or more perceptually different audio recordings, leading to serious problems in the recognition process. Therefore, the key goals when developing an audio fingerprinting algorithm are to keep the probability of collision as minimal as possible.

b) In-variance to noises and spectral or temporal distortions: Audio signal is usually degraded to some degree due to some kinds of sounds and vibrations when captured or playing in actual environments. The audio fingerprints of the damaged audio signal can be the same as those of the original signal. Important features are still unchanged. In cases of this fingerprinting technique, high robustness must be obtained.

c) Minimum length of song track needed for identification: Due mainly to time and storage limitations, making the entire of an unknown audio track in real-time music recognition is still impractical. Nevertheless, it is ideal that only a few seconds of the track is required to find the unknown audio.

d) Retrieval speed and computing load: Recognition results can be provided in a few seconds in most real-time applications. However, with the increase in song recordings in the audio reference database, locating the matching object correctly in real-time becomes very difficult.

A fingerprint is a type of distinct digital representation of the waveform of a song. A fingerprint can be obtained by collecting significant characteristics from the various audio properties. Without sacrificing its signature, the created fingerprint may also be segmented into several parts. Moreover, fingerprints can be processed at a much smaller scale relative to the audio waveform's initial form.

The items below are several criteria that should take into consideration for a robust audio fingerprint:

- Consistent Feature Extraction: The key feature of fingerprint generation is that it can replicate an audio fingerprint identical to that of a music section.
- Fingerprint size: Fingerprint file size has to be small enough so that more music collections can be stored in the database. In reality, a lightweight fingerprint offers effective memory allocation during processing.
- Robustness: Even if external signal noise has affected the source audio, fingerprints may be used for identification.

4 Related Works

We provided our outstanding contributions to the academic literature that satisfy all the success criteria listed above. We show that, despite our large comparison sets, our method is efficient and that there is an extensive search problem caused by the invariances of the hashes and their robustness in signal modifications. Mainly, we designed the proposed device for low-cost hardware, demonstrated its capabilities, and avoided costly CPU processing.

Despite studies on the identification of songs and fingerprints by other researchers such as Shazam [7] and SoundHound [8]. We understand the clustering design using K-means for an experiment in the real fingerprint database with a set of 2.4 billion fingerprints provided by the company as datasets. We want to emphasize that a database of this scale seldom appears in the research literature, but there is a chance that it exists. The next critical aspect was the audio recognition system. Although initial K-means computing for compilation is resource-intensive, we achieved significant speed efficiency at the end of the day [19]. Also, the proposed architecture and algorithm will lead to a new insight into song recognition.

Moreover, we had introduced an IoT-based solution to song recognition in a cloud environment in this study. We have developed a recognition system to integrate audio streams from remote FM Radio stations [20]. We conducted a song recognition technique based on the K-modes clustered cloud database of MongoDB [11]. We supported various collections of fingerprint length tests to ensure the best accuracy and reliability of the test.

The new fingerprint extraction technique's significant findings focused on Short Time Power Spectral Density (ST-PSD) was also implemented [10]. Later of which binary encoding group's attributes lead to the reliability of K-modes. Besides, this study clarified the identification methodology primarily through hamming distance measure in the predetermined cluster table. The findings were given by sampling 400 random 5-second queries from the initial song set in the experiment. Using this method, the optimal chosen combination of parameters is the identification ratio of 90%.

We have already introduced extracting fingerprints from audio in our previous works based on the ST-PSD calculation. We introduce several significant and remarkable improvements to the previously proposed algorithm to enhance the robustness of the linekeys to temporal shift and the consistency of the fingerprints for perceptually distinct audio signals in this paper. The proposed fingerprints are based on calculating the audio signal's short time spectral power density ST-PSD obtained on the Mel frequency scale.

According to our tests and performance measurements, the framework can be used specifically for different areas of fingerprinting applications. In addition to typical fingerprint applications, these are, for instance, the identification of audio copies and media tracking, copyright identification of songs.

5 K-means Clustering in MongoDB Database Methodology

K-means clustering is used for non-classified results, which performs an unsupervised algorithm for many data. The fingerprints are grouped into subgroups by the K-means classification. As such, objects in the same category (clusters) are more similar to each other. Whereas K-means is often used for high-dimensional data classification, we take advantage of the centroid value as a distance point when executing the nearest computation in this experiment. The basis for using K-means is that we need to create relatively uniform dataset-size clusters.

The K-Means clustering is method of partitioning n data, into k

clusters in which each data is associated with the cluster with the nearest mean (intra-cluster distance). Thus, several different distinct clusters are formed. Therefore, the primary goal of K-means is to minimize intra-cluster distances. This is done by determining the J index as Equation (1) follows:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \|x_i^{(j)} - c_j\|^2 , \quad (1)$$

where c_j is the mean value of the j-th cluster with $i \leq j \leq n$, and $x_i^{(j)}$ represents a fingerprint that falls into the j cluster. The c_j is usually called *centroid*s. Sets of k clusters, $k+1$ boundaries and k centroids are discovered by the K-means algorithm, reducing the J optimization index. More precisely, a fingerprint partition set of $S = \{S_1, S_2, \ldots, S_k\}$ resulting from fingerprint partitions is computed from:

$$\underset{S}{\operatorname{argmin}}\, J \quad (2)$$

As reference points, the c_j centroids of the clusters are used. Nevertheless, we are expected to specify the number of k clusters subsequently computed. Figure 2 provides an overview of the clustering phase in this implementation, where the original fingerprint collection stored in MongoDB was initiated.

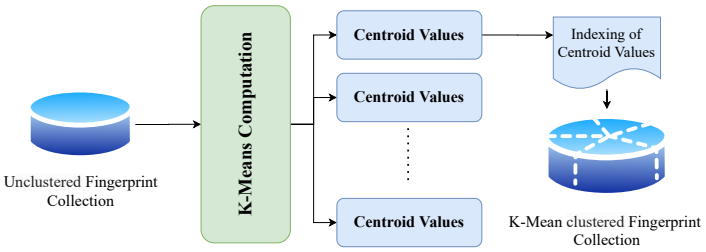


Figure 2: K-means clustering implementation overview

A reduced data sample of $2 \cdot 10^6$ (random segment) was selected from the overall selection based on the fingerprint distribution analysis. There is no particular data restriction that the K-means can carry out. However, this depends on computational power, time constraints, and other hardware specification.

Furthermore, once the calculation is completed, the centroids' values are generated and represent a key value for the MongoDB cluster array. Next, the nearest distance fingerprints were determined to the centroid value and transferred the data to the subgroup collections.

5.1 Stepwise Implementation of K-means

The experimentation aimed to produce several $k = 10,000$ clusters with key centroids values as reference. In the next phase, song fingerprints are distributed into the new clustered collections.

Step 1: Batch Data Processing

The original approach was to separate the first instances into blocks by introducing a segmentation of data while handling many data for K-mean computation. Next, using the *pickle* data structure in python, the trained model was stored from the computation. Subsequently, once all fingerprint datasets were collected, we used the previous model values and incrementally updated them. Code Listing 1 shows the steps of the python *scikit-learn* library [21] implementation algorithm for K-means. Below describe the steps of the data training sequence:

1. As the first elements of *n_cluster* centres, the algorithm selects 10,000 points.
2. Then, each 100,000 of linekeys that were loaded into memory and loaded to the K-means calculation using these data. Each cluster center was recomputed as the average of the points in that cluster.
3. Finally, the function at item 2 is repeated until the clusters converge. If there is no further change in the assignment of the fingerprints to clusters, the algorithm converges.

```
def dumpFitKmean(n_clusters,i,chunk,chunksize):
if i==1:
X = chunk
mbk = KMeans(init='K-means++',n_clusters=int(
    n_clusters), n_init=1, random_state=42)
else:
X = chunk
mbk = getFitKmean(i-1)
mbk.fit(X)
pickle.dump(mbk,open("pickelDump"+str(i)+".p", "wb"))

def getFitKmean(i):
dumpFile = "pickelDump"+str(i)+".p"
mbk = pickle.load(open(dumpFile, "rb"))
return mbk

if __name__ == "__main__":
i=1
chunksize = 100000
for chunk in pd.read_csv(filename, chunksize=chunksize
    , header=None):
dumpFitKmean(n_clusters,i,chunk,chunksize)
i=i+1
```

Listing 1: A Python implementation of K-means computation

Step 2: Building Cluster Segmentation in MongoDB

Once the centroid values were formed, the distance to each centroid was calculated for each fingerprint $f^{(i)}$ in order to find its partition affiliation by Equation (3)

$$p^{(i)} = \underset{j=1}{\overset{k}{\operatorname{argmin}}} \|f^{(i)} - C_j\| \quad (3)$$

In order to obtain the distance for the fingerprint $f^{(i)}$, the Euclidean distance was performed between $f^{(i)}$ and any centroid value C_j. From the distance function list, argmin was used to get the minimum distance. As a result, acquired the fingerprint association $p^{(i)}$ with its corresponding cluster $S_{p^{(i)}}$.

The code in Listing 2 demonstrates the steps of the algorithm implemented in python to obtain an aggregation of fingerprint-clusters.

```
def centroid_linekeys_distance(self,linekeys,
    indexCentroids,centroids):
distance = abs((int(linekeys)-int(centroids)))
```

```
result = distance,centroids,indexCentroids
return result

def get_CentroidsDB(self):
result = self.collection_centroids.find()
result = [[document['_id'],document['centroids']] for
    document in result]
result = sorted(result,key=lambda x: x[1])
return result

def update_CentroidsDB(self,sortedDistance,songNo,
    linekeys,linekeysPosition):
self.collection_updateCentroids = self.db['
    zentroid2song'+str(centroidIndex)]
self.collection_updateCentroids.insert({'SongNo':
    songNo, 'linekeys':linekeys, 'linekeysPos':
    linekeysPosition } )
```

Listing 2: Python code fingerprint-cluster association

5.2 Song Recognition and Information Retrieval

During the recognition phase, the fingerprint sequence in the clustered fingerprints database must be identified. The algorithm performs a sequential search window in two stages. First, allocate each fingerprint in the query sequence to the nearest centroid value. Then, perform an in-depth search within the corresponding cluster values and obtain a set of candidates. Here, by applying a similar algorithm to the clustering method to ensure the accuracy of the result.

The benefit of the implemented cluster framework is that a binary search technique can be used according to [22]. The first step is to find the position of a specific fingerprint query value within the sorted array. This method examines and searches the nearest key value for the median centroid key value of the selected fingerprint in each step.

5.2.1 Real-Time Slide Window

Figure 3 illustrates the method used to classify a song based on an input fingerprint source. As seen, the fingerprint query stream was chunked into significant portions of the windows. This sliding window approach is a common information retrieval technique used to detect matching sequences. Each window represents a part of the time in the current album. As mentioned above, each fingerprint represents an instant δ of the audio source. Configurable fingerprints window sizes, e.g. 500, 1000, 2000, 3000, 5000 or 6000 which translate into the actual time section, are essential. Here, it is suggested that the window size does not exceed 6000 fingerprints, as this represents around 60 seconds of actual audio.

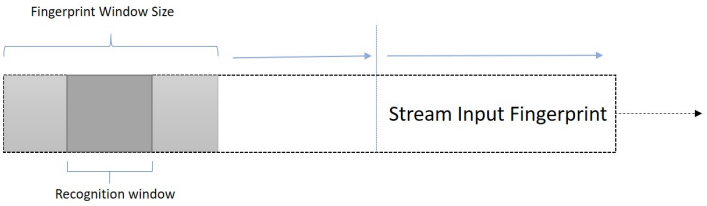


Figure 3: Fingerprint Windows Slides Recognition

Each fingerprint is not unique, so that it could appear in some songs, and then a set of song candidates will be provided. However, in the quest for a result, we can exclude those candidates and obtain a winner. It was next, segmenting a sub-window of 3/4 from the fingerprints used in the window. This sub-segment window aims to reduce the time for recognition. Also, the selected fraction size weight is appropriate to reflect the full fingerprint in the window.

5.2.2 Fingerprint Cluster Identification and Recognition

There are two stages of the recognition process for each fingerprint query for the search's effective result. Firstly, to determine the cluster $p^{(i)}$ representing the nearest centroid as in Equation (4) . The value of the fingerprints $q^{(i)}$ is compared to the value of each centroid C_j. Therefore, returns the cluster $p^{(i)}$ from the set of partitions to get the nearest reference.

$$p^{(i)} = \underset{j=1}{\overset{k}{\operatorname{argmin}}} \|q^{(i)} - C_j\| \tag{4}$$

Finally, Equation (5) returns the minimum value *argmin* between the fingerprint query $q^{(i)}$ with the fingerprint set $f_j^{p^{(i)}}$ within the defined cluster $p^{(i)}$. This results in a single result of $wf^{(i)}$ of the song details (title).

$$wf^{(i)} = \underset{j}{\operatorname{argmin}} \|q^{(i)} - f_j^{p^{(i)}}\| \tag{5}$$

Each winner of the fingerprint $wf^{(i)}$ will be associated with a set of songs stored in the collection containing the $wf^{(i)}$ in their sequence. For these songs, the column list entry is set to "1" with an entry for each song to be recorded. For this reason, the following subsection 5.2.3 will be clarified by maintaining a record of N columns forming a matrix called **songs**.

Code Listing 3 shows the steps implemented in python for fingerprint recognition.

```
def centroid_linekeys_distance(linekeys,indexCentroids
    ,centroids):
distance = abs((linekeys-centroids))

def get_linekeys_match(centroidIndex,linekeys):
collection_centroidsValues = db['zentroid2song'+str(
    centroidIndex)]
result =  collection_centroidsValues.find({'linekeys':
    linekeys})

if __name__ == "__main__":
for lk in inputLinekeys:
sortedDistance = min(distance)
```

Listing 3: Python fingerprint recognition

5.2.3 Candidates Scoring

From the previous step, the information for each song was obtained from a single fingerprint question. However, it is crucial to assess the overall result inside the fingerprint sequence in the defined window. Therefore, using the Equation (6) to determine the frequency of $F_k^{(n)}$ of the song results in a N row fingerprint window. The highest score value of the song would be as winning candidates for the window section.

$$F_k^{(n)} = \sum_{i=1}^{N} \mathbf{songs}_{k,i} \tag{6}$$

where n is the index of the current window. Thus evaluating the winner song using Equation (7).

$$ws^{(n)} = \underset{k}{\operatorname{argmax}} F_k^{(n)} \tag{7}$$

As an initiation a counter is set to the value 1.Each winning song that is $ws^{(n)} = ws^{(n-1)}$ will increase the counter value. Therefore, as soon as the winning song $ws^{(n)} \neq ws^{(n-1)}$ evaluates the W number of subsequent windows for the same winner it will reset the counter back to 1. Therefore the length of the airtime song is evaluated as $d = N \cdot W \cdot \delta$.

6 IoT Based Song Recognition

As shown in Figure 4, the system framework is planned to access the FM Frequency source at a different location by deploying a low-cost IoT system receiver. This FM receiver streams the audio to a cloud instance that converts it to an audio file that is then processed for recognition. Each module of software and hardware configuration components must be integrated to achieve this implementation. Each of these components is discussed in the next section, which includes the IoT system 6.1, the communication protocol 6.2, the recognition server 6.3, and the clustered database.

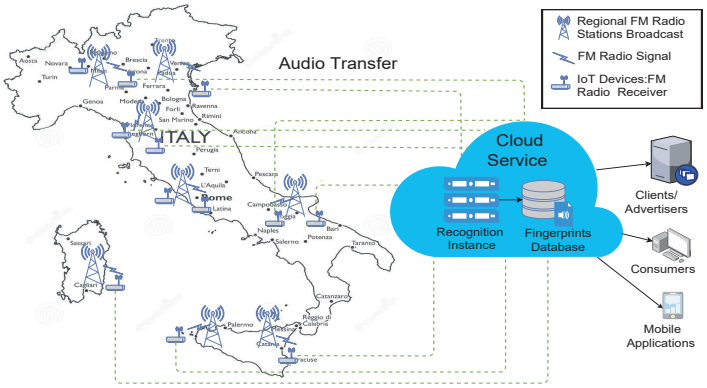


Figure 4: Overview of IoT based Song Recognition Framework

6.1 IoT Device Design

Software-defined radio (SDR) offers a level interface that allows access to filters, mixers, amplifiers, modulators/demodulators on software, meaning on computer embedded systems. This implementation uses RTL-SDR USB dongles based on the RTL2832U chipsets [23] that can read frequencies between 24 and 1,766MHz. The Raspberry Pi is a single-board computer made on a single board and, therefore, small and cheap. It can be used for light programming or, as in this project, to create devices dedicated to the Internet of Things or home automation with various sensors. The operating system used in the project is Raspbian, an official distribution of Raspberry Pi, based on Debian Linux and suitably adapted to the Raspberry Pi. The operating system was downloaded via NOOBS to a 16GB MicroSD.

The FM radio receiver is built using a Raspberry Pi and a dongle, which converts the analog audio signal into a digital audio stream. Python libraries were used for the application, more specifically, the library to launch RTL-FM. The python syntax with the options required to play an FM radio station is shown in Listing 4.

```
import subprocess, signal, os
def newstation(station):
global process, stnum

part1 = "rtl_fm -f "
part2 = "e6 -M wbfm -s 200000 -r 44100 | aplay -r
    44100 -f S16_LE"
cmd = part1 + station + part2
print 'Playing station :', station

# kill the old fm connection
if process != 0:
process = int(subprocess.check_output(["pidof","rtl_fm
    "]))
print "Process pid = ", process
os.kill(process,signal.SIGINT)

# start the new fm connection
print cmd
process = subprocess.Popen(cmd, shell=True)

def setvolume(thevolume):
os.system('amixer sset "PCM" ' + thevolume)
print 'volume = ' , thevolume

process = 0

while True:
answer = raw_input("Enter a radio station (i.e. 107.9)
    or volume (i.e. 50%): ")
if answer.find('%') > 0:
setvolume(answer)
else:
newstation(answer)
```

Listing 4: Python RTL-FM Frequency Configuration

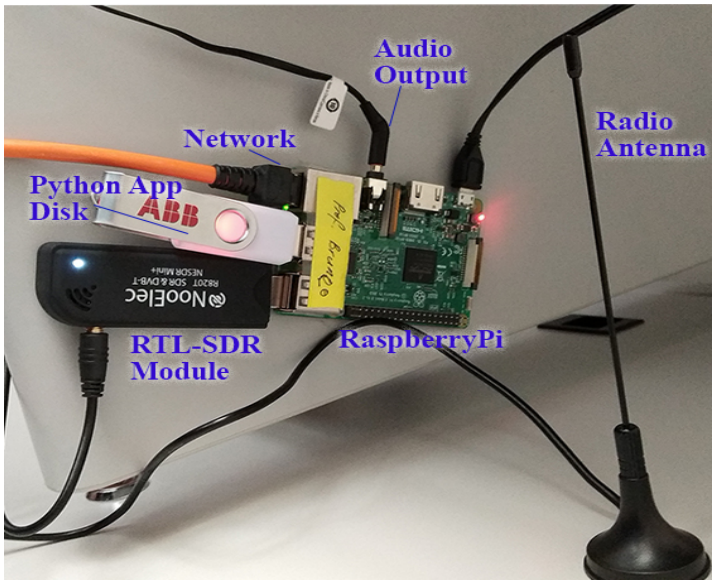


Figure 5: Raspberry Pi with RTL-DSR module Hardware

Figure 5 shows the actual physical Raspberry Pi device configured with the RTL-SDR hardware module.

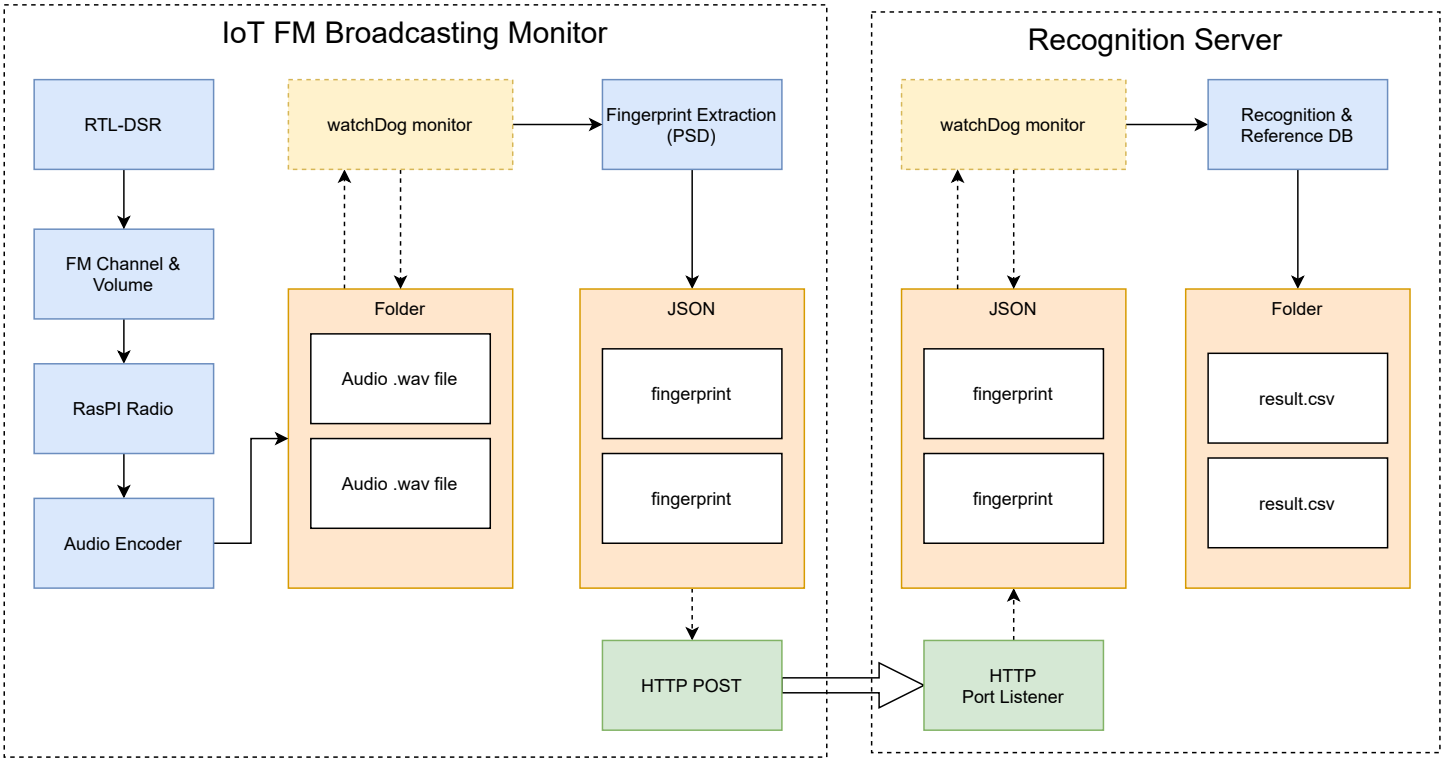


Figure 6: Overview of Framework of IoT device protocol to Recognition Server

6.2 Communication Protocol

Once established the FM frequency channel, we stream the Raspberry Pi output using Real-Time Streaming Protocol (RTSP) with a specific TCP/IP port to maintain end-to-end connection using a python RTSP libraries [24]. The RTSP server will process incoming requests for streams.

Figure 6 shows the component of the Raspberry Pi IoT device and the recognition server communication protocol. The IoT server's final packets consisted of 5 seconds of fingerprint conversion, which is encapsulated as a JSON string as a packet. Additionally, the audio conversion fingerprint algorithm is embedded inside the IoT device; thus, a smaller size packet of string could be transferred much faster. Therefore, the recognition of the fingerprint will be done on the server instance once receiving the packets. The listening port is open as a service for both sides with a directory observer to monitor any incoming packets as demonstrated in code Listing 5.

```
def postJson(linekeys,newFilename):
print "posting"
data = {}
data['fingerprintFile'] = newFilename
data['linekeys'] = map(str,linekeys.tolist())
print data

req = urllib2.Request('http://localhost:8008')
req.add_header('Content-Type', 'application/json')
response = urllib2.urlopen(req, json.dumps(data))

class Watcher:
DIRECTORY_TO_WATCH = "/home/pi/Desktop/PiRadio/
    inputfile/"
def __init__(self):
self.observer = Observer()

def run(self):
event_handler = Handler()
self.observer.schedule(event_handler, self.
    DIRECTORY_TO_WATCH, recursive=True)
self.observer.start()
try:
while True:
time.sleep(5)
except:
self.observer.stop()
print "Error"
self.observer.join()
```

Listing 5: HTTP Protocol of Fingerprint Exchange

6.3 Recognition Server

A fingerprint is a digital vector that identifies a piece of song signal and can outline the content of that piece of the song. The fingerprint is achieved by removing the main characteristics from audio by choosing distinctive features. Also, fingerprints for storage will minimize the size of the original song with the standard structure.

In this phase, as shown in Listing 6, the streaming audio is converted to *.wav* file in such a way it could be further converted into a sequence of fingerprints.

```
part1 = "rtl_fm -f "
part2 = "e6 -M wbfm -s 200000 -r 44100 | sox -t raw -e
    signed -c 1 -b 16 -r 44100 - inputfile/record"+
    str(numberCount)+".wav"
cmd = part1 + station + part2
```

```
def getArrayKey(filename,noOfSplit,windowlength,step):
fftsize=128
yFinal=[]
D, fs = sf.read(filename)
r = D.size
c = 0
if c > 1:
D = multiToMono(D)
D = np.array_split(D,noOfSplit)
```

Listing 6: Conversion Recording Stream to Audio .wav

The features below are the qualities of an audio fingerprint that we take into consideration:

- Consistent Feature Extraction: The essential part of the fingerprint generation is that it can reproduce a similar audio fingerprint given only a segment of the audio.
- Fingerprint size: The fingerprint's size has to be small enough, so a more exhaustive song collection can be archived in the database. Furthermore, a well-compressed fingerprint uses efficient memory allocation during processing.
- Robustness: Fingerprint can be used for recognition even if the source audio has been affected by external signal noise.

7 Mobile Music Recognition App

The Android application detects and captures audio through the device's microphone, the audio signals and then converts them into fingerprints. The application was developed and compatible with API version 19 (KitKat) and earlier version. It comprises five object classes that includes *FFT, PSD, Fingerprint, MainActivity, RecordAudio, and ToolsAudio*. The following will discussed on class *MainActivity* which provides the function *RecordAudio*, as in the UML Figure 7

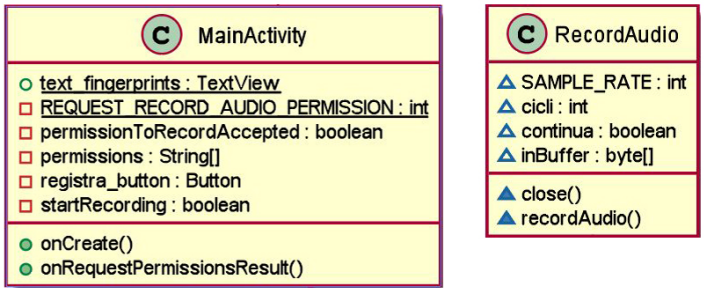


Figure 7: Class MainActivity for RecordAudio

The class *MainActivity* is identified as *Activity*, which is essentially a window containing the application user interface, consisting of a file XML relative to the layout displayed from a class.

Java programming is used to define their behavior; its purpose is to interact with users and goes through various standard functionality cycles, as in Figure 8. When an activity runs, three methods are invoked to interact directly with the user: *onCreate, onStart, onResume*; when instead Android puts the activity to rest, the methods invoked are *onPause, onStop, onDestroy*.

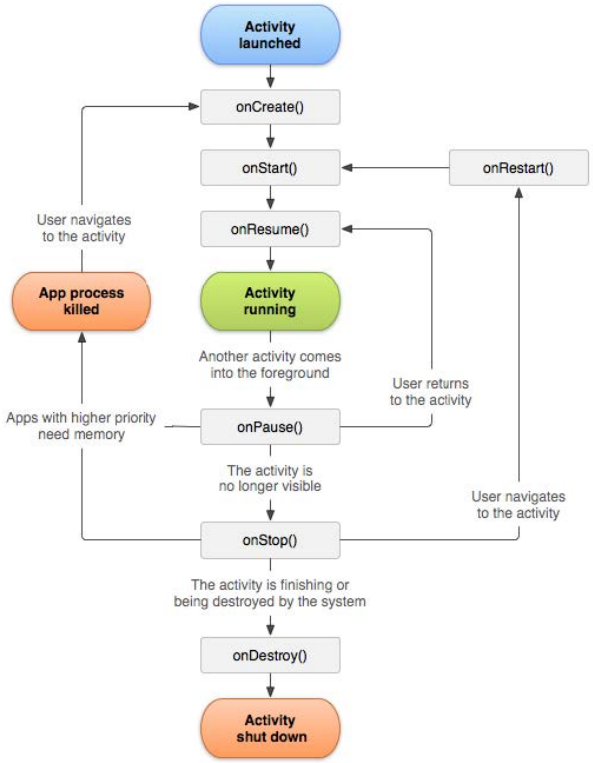


Figure 8: Android Java classes function cycle

Therefore, the graphic interface interacts with the class provided by *MainActivity*; which overrides its default method internally, *onCreate* is invoked when the activity is created.

To manage the user's interaction with the button, it requires to associate a listener method defined as *onClickListener* , through the method *setOnClickListener*.The class *onClickListener* is an abstract however once an instance of it has been created, the method has been redefined through *onClick*.

The *onClick* as in Listing 7, will provide instructions to be executed when the button is clicked. The process flow implemented is as follows: when the "Register" button is press, as shown in Figure 9, the variable is checked *startRecording* to check whether the registration service is started or not. Therefore, the method will be start *recordAudio* of the class *RecordAudio* to record audio signals. Next, the label "Stop" button initiates the stop service. This basic function is almost similar to Shazam's mobile app.

```
register_button.setOnClickListener (new View.
    OnClickListener(){
 public void onClick(View v){
  if (startRecording) {
   au.recordAudio();
   register_button.setText ("Stop");}
  else {
   au.close();
   register_button.setText("Register");
  }
  startRecording =! startRecording;}
});
```

Listing 7: Java button.setOnClickListener

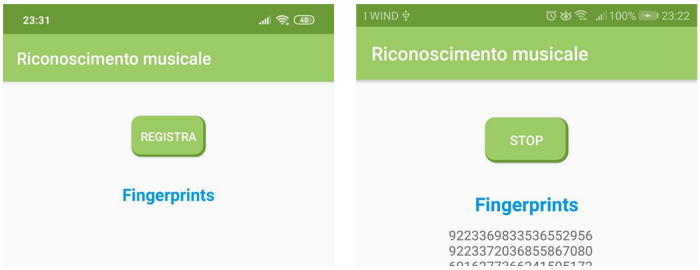


Figure 9: User interface for Start and Stop recoding Audio

The final method implemented was the *requestPermissions*, which allows the application to request user permission for using the internal microphone. This method is related to *onRequestPermissionResult*, which manages the results granted the use of the physical module from the mobile.

7.1 Audio Recording

The implementation of audio detection is done by the class RecordAudio. For this purpose, the use of libraries is particularly important:

- *android.media.MediaRecorder*, for recording an audio stream;
- *android.media.AudioRecord*, to manage audio resources;
- *android.media.AudioFormat*, to access audio formats and channel configurations.

The method *recordAudio* as shown in Listing 8, it is invoked by the *MainActivity* as previously described. The invocation will create a thread that initially does the following:

- the priority of the thread for the acquisition of audio signals is set;
- the bufferSize, the maximum size of bytes that the thread can detect at each cycle;
- an object *AudioRecord* is created, which sets different specifications, such as the audio source data (DEFAULT that is the microphone), the sampling frequency (44100), the number of channels (MONO), the coding PCM (16 bit), and the bufferSize;
- the method invoked *startRecording* on the newly created object to initialize and start the audio capture session.

```
void recordAudio() {new Thread(new Runnable(){
  @Override
  public void run(){
   android.os.Process.setThreadPriority(
   android.os.Process.THREAD_PR IORITY_AUDIO);
   int bufferSize =
   AudioRecord.getMinBufferSize(SAMPLE_RATE,
   AudioFormat.CHANNEL_IN_MONO,
   AudioFormat.ENCODING_PCM_16BIT);
   byte [] audioBuffer = new byte [bufferSize];
   AudioRecord record = new
   AudioRecord(MediaRecorder.AudioSource.DEFAULT,
   SAMPLE_RATE,
   AudioFormat.CHANNEL_IN_MONO,
   AudioFormat.ENCODING_PCM_16BIT,
   bufferSize);
   record.startRecording();
```

Listing 8: Mobile Audio Recording

To read input data from smartphone microphone, the class *AudioRecord* is used in a while loop read in which is recorded in a byte array. Therefore, data is not immediately sent to the class Fingerprint for their conversion into linekeys, but an adequate number of iterations is expected for a given number of byte to be read. Thus the transformation of a static vector is always performed by creating a thread, as already discussed. The size of the data to be converted has been chosen so that there are no errors in the process they will be requested, especially in the method buffer of the class Fingerprint.

7.2 API Design

As part of this project, for communication purposes between client and server, an API has been developed that uses REST's architectural style. The web server was developed in Python. It implements an HTTP server and contains the *pyMongo* API, which is necessary for interfacing with the database. The project foresees the MongoDB non-relational database for the fingerprint storage similarly used in work in [19]. The first step was to install the MongoDB database and all dependencies (Python libraries) required by the source code. MongoDB is a NoSQL type DBMS that stores JSON format. A cURL is a command-line tool for transferring data over the network using an HTTP protocol. The advantage lies in its independence from the programming language used. It is proved to be particularly useful for testing the interaction with the server, and therefore the appropriately agreed HTTP request.

7.2.1 Fingerprint Packets Structure

The JSON format for the exchange of data between client and server is constructed. The file is structured in such a way that it has two fields. Firstly, the name-value pair which indicates the *fingerprintFile* ID, therefore of the form *"name":"value"*. The value is a progressive integer that identifies the ID of the JSON request containing the fingerprints. Note that the server requires this to be a string type.

Next, the values contain linekeys, concatenated one after the other. The number of linekeys present in the array is not fixed, but the code must be implemented to parameterize the linekey number, i.e., generalized to number possible values. This is a design parameter, and according to the specifications, it has been set to 50 linekeys.

The following in Listing 9 shows an example of a JSON file, in its complete form, which has 4 set linekeys (fingerprints) and the *fingerprintFile* ID with a value of 2.

```
{
 "linekeys":["13844060856490393600","
      8645919525052907526" ,"6054232079929966592","
      7274368929366016"],
 "fingerprintFile":"2"
}
```

Listing 9: JSON Fingeprint packets

7.2.2 Web Server and API Protocol

A Python code in the file has been implemented *server.py* API that allows communication with the server. The code is isolated from communication with the MongoDB database for the time being and shows the response of HTTP requests arriving from clients. Some of the classes and their relationships are shown in Figure 10.

The following Listing 10 shows the code fragment related to the HTTP web server's implementation where the parameter server address is defined. By default, the port number set is 8009.

```
    def run(server_class=HTTPServer, handler_class=
        Server, port=8009)
    server_address = ('', port)
    httpd = server_class (server_address, handler_class)
    print('Starting httpd on port% d ...' % port)
    httpd.serve_forever()
    print('Started httpd on port% d ...' % port)
```

Listing 10: Python HTTP Class

The code Listing 11 is related to the structure of HTTP header request accepted by the server. The method *headerSet()* handles the parameters of the HTTP request, and the way to response in case errors occur. For example, if the body contained a format other than JSON, a status code would be sent 400.

```
    def headerSet(self):
    self.send_response(200,"ok")
    self.send_header('Content-type','application/json')
    self.send_header('Accept','application/json')
    self.send_header('Access-Control-Allow-Origin','*')
    self.send_header('Access-Control-Allow-Credentials'
        ,'true')
    self.send_header('Access-Control-Allow-Methods','
        GET,POST,OPTIONS,HEAD,PUT')
```

Listing 11: Python HTTP Header Request

Next, the deserialization of the JSON file begins: a dictionary data structure is used for data storage. During the test phase, the cURL application was used to initiate the command from client to the server, an example of a request is shown in Listing 12

```
    curl --data "{\"linekey\":\"3441945721\", \"
        requestId\":\"007\"}"
    --header "Content -Type: application/json"
    http://localhost:8009
```

Listing 12: cURL Application Request

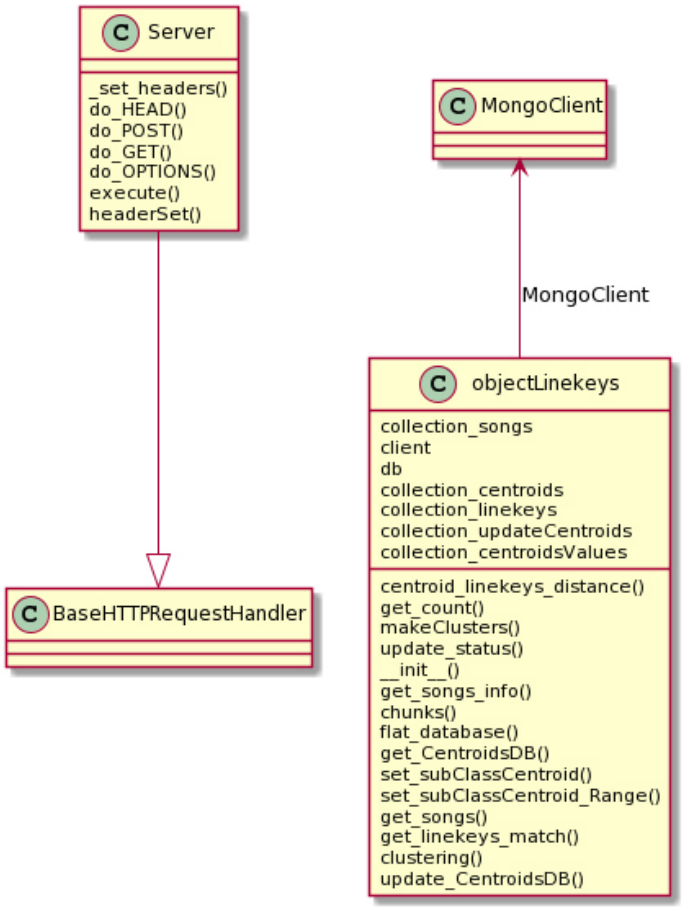


Figure 10: Web server UML classes

8 Evaluation and Performance

The first stages of testing and running the app were carried out using the Android Virtual Device (AVD) Google Pixel 2. Furthermore, the server has been configured to provide the client's data (to verify the correctness). Figure 11 shows the server running while listening on port 8009.

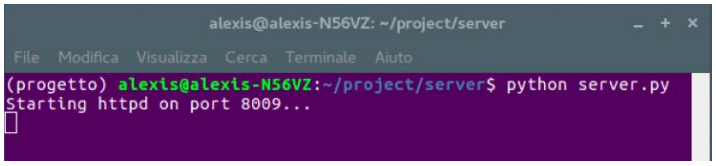


Figure 11: Initiation of server listener

The server open to listening for incoming HTTP requests from android clients. The layout of the android app is shown in Figure 12. As soon as the user taps the button "Register," linekey generation begins. Once the application has reach 50 linekeys, the packets will be sent to the webserver for recognition.

Execution continues and, if the HTTP request was successful and the packet is valid, the server status code is displayed "200 OK". Figure 13 shows that the linekeys have been received while the confirmation is shown on the right.

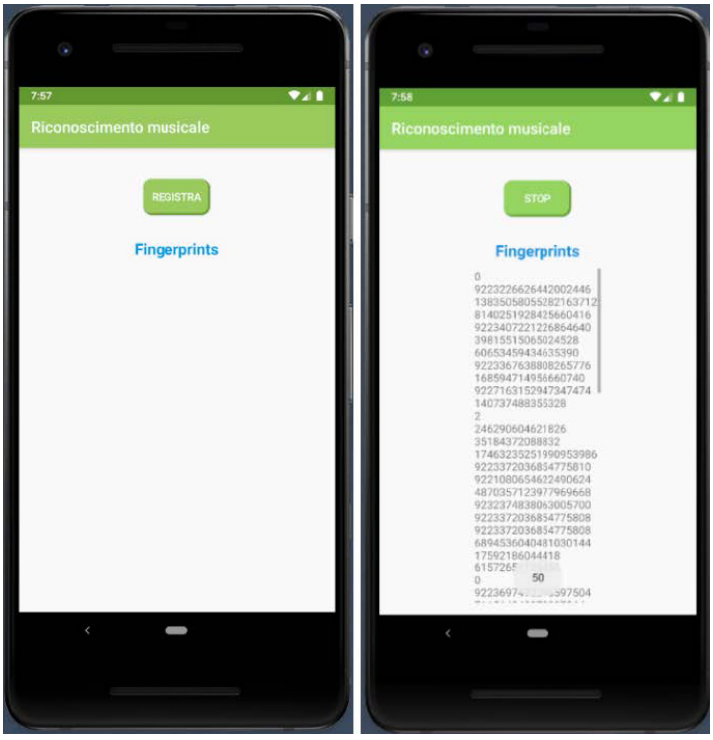


Figure 12: Listening for incoming HTTP requests

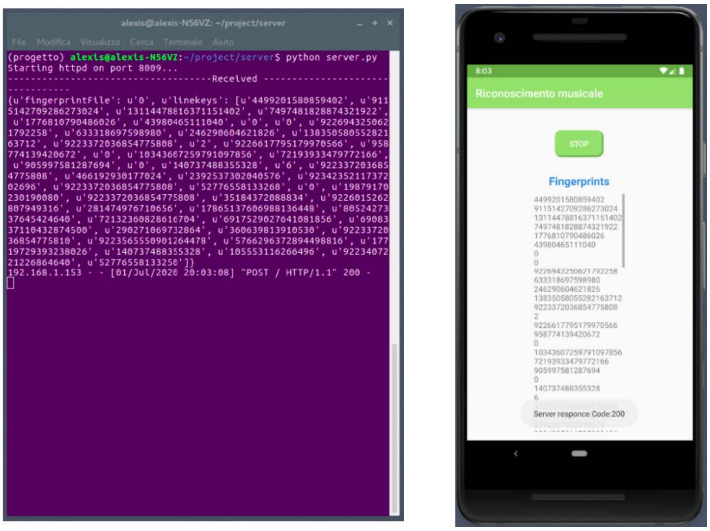


Figure 13: Successful HTTP requests

In any case, the generation of the linekeys proceeds in real-time and received queued HTTP requests follow one after the other until the user immediately ends the generation execution by tapping on the "Stop" button. A web server's public IP address is used in the final phase, which resides in the department lab. The app was installed on a physical Android smartphone using the APK generated by Android Studio for the real purpose.

8.1 Performance Evaluation: ST-PSD Fingerprint

This section provides the evaluation of the proposed method ST-PSD fingerprints. As noted earlier, to verify the proposed hamming distance computing algorithm for music recognition (a method of exploring the Hamming distance for audio fingerprinting systems), an evaluation was carried out using real music data.

The proposed approach's efficiency was evaluated by experiments using a set of 100 real songs, with 4 real queries for each song, selecting the starting point of the piece of audio to be recognized at random. The random position was selected in the sample domain of the audio so that the extraction of the line keys is not associated with the reference keys found in the collection. Also, each piece of audio corresponds to a length of 5s.

8.1.1 System parameters

Candidates Scoring The linear combination of two sub-score metrics was considered in order to determine a score for each song. The first score is calculated by combining a T_d threshold, i in the fingerprint for each linekey. Provided the matrix $\mathbf{H_d}$ and the matrix T_d, a second matrix of $\mathbf{A}$ has been constructed in Equation (8) where:

$$\mathbf{A_{i,j}} = \begin{cases} 1, & \text{if } H_d(i,j) < T_d. \\ 0, & \text{if } H_d(i,j) \geq T_d. \end{cases} \tag{8}$$

then evaluate the first sub-score in Equation (9) for each song as:

$$B_j = \frac{1}{W} \sum_{k=0}^{W-1} A(i,j) \quad j = 1, \ldots, \sigma \tag{9}$$

A third matrix $\mathbf{C}$ as in Equation (10) was considered for the second sub-score where each variable contains a normalized value in the range $[0-1]$ approaching 0 when the Hamming distance is large and approaching 1 when the Hamming distance is low:

$$C_{i,j} = \begin{cases} \left(1 - \frac{H_d(i,j)}{T_d}\right), & \text{if } H_d(i,j) < T_d. \\ 0, & \text{if } H_d(i,j) \geq T_d. \end{cases} \tag{10}$$

then evaluate the sub-score for each song as in Equation (11):

$$D_j = \frac{1}{W} \sum_{k=0}^{W-1} C_{i,j} \quad j = 1, \ldots, \sigma \tag{11}$$

The final score for each song will be evaluated as Equation (12):

$$\mathrm{FS}_j = \alpha B_j + (1-\alpha) D_j \tag{12}$$

where α is a parameter in the range $[0, 1]$ to be chosen to maximize recognition performance.

The winner song for fingerprint F will be the song with the higher FS_j, by obtaining its index using the relation in Equation (13)

$$\mathrm{WS} = \arg\max_{j=0}^{M-1} \mathrm{FS}_j \tag{13}$$

The main focus is to use the optimal parameters to perform the proper classification of audio samples. For this reason, a range of potential parameter configurations and corresponding findings have been investigated. Mainly, taking account the size of fingerprint in terms of the number of line keys (W), the size of the threshold used to create the $\mathbf{A}$ and $\mathbf{C}$ matrix (Td) and the value of the coefficient α used to combine the two sub-scores. The following table 1 indicates all the parameters used to test the performance measure.

Table 1: Parameters used for performance evaluation

α	[0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
W	[10, 20, 30, 40, 50]
T_d	[10, 15, 20, 25, 30]

8.1.2 *Evaluation steps*

Figure 14 indicates a bar graph in which the height of each bar for the various fingerprint sizes used is proportional to the recognition ratio. Mainly, the bar shows the results taking into account the combination of parameters (α and T_d), which allows better results in terms of the recognition ratio. The values of the parameter combination are displayed within each bar.

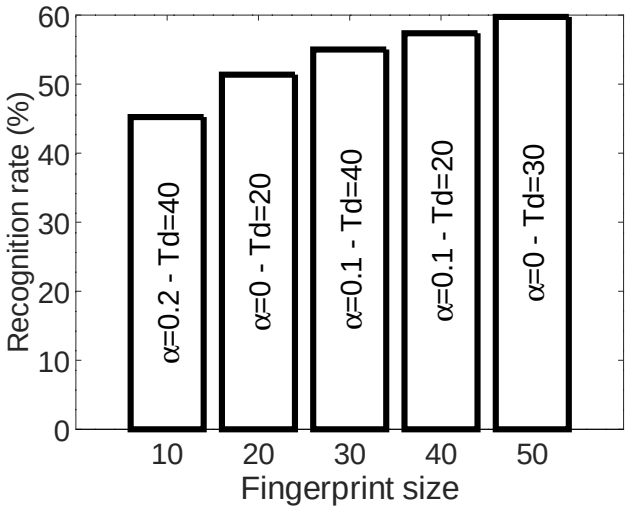


Figure 14: Recognition rate for different fingerprint size using the better combination of parameters α and T_d

While the identification rate does not seem so high, it must be considered that this rate is compared to a single fingerprint. Using a song piece larger than the fingerprint size for song recognition, the song recognition rate can be improved using fingerprints obtained by a sliding window added to the linekeys sequence. Even so, the fingerprint recognition rate found implies that, on average, 2 would be automatically labeled with the right song, taking a sequence of more than 4 fingerprints. Exploiting this result will dramatically increase the song recognition rate, which will subsequently be clarified.

For that we proposed to extract the fingerprints by applying a sliding window that overlaps the $W-1$ linekeys with the adjoining linekeys. In this way, the fingerprint produced in each step will be equal to the production step of each linekey (i.e., 100ms per linekeys). Taking into account an audio piece of 5 seconds length which equals to 50 windows (called fingerprint) and denoted as $W = 50$. Thus it's possible to produce a sub-query fingerprint size for evaluation which is equal to $W = 10$, $W = 20$, $W = 30$, $W = 40$ and $W = 50$ respectively.

To improve the proposed recognition system's efficiency, a further discrimination variable was added, based on the distance between the two highest F-Score values obtained for each fingerprint in the search process. Especially by evaluating the distribution of this distance (called Δ) when the fingerprint was correctly recognized, and the fingerprint was misclassified.

Figure 15 shows the accuracy in terms of

$$\frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total population}}$$

varying the value of threshold T_Δ in a specific selected 4 windows range with better combination of parameters. Therefore, the best performance using the parameter combination equal to ($W = 10, \alpha = 0.2, T_d = 40$) and a threshold $T_\Delta \approx 0.01$. The accuracy in this case will approach 90%.

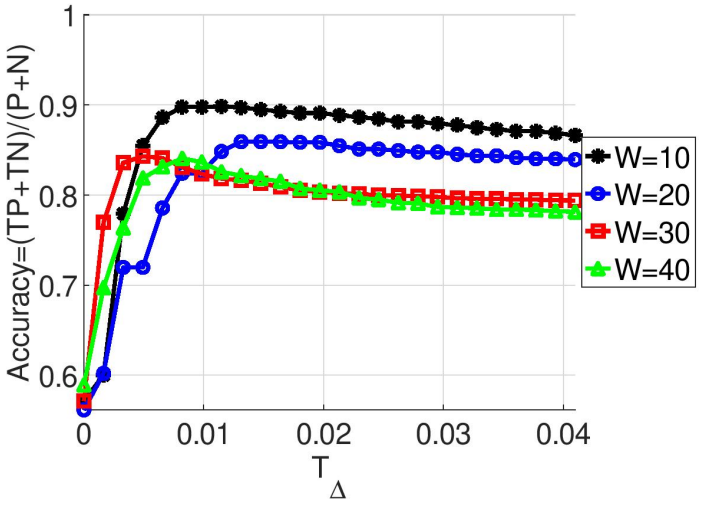


Figure 15: Accuracy varying the T_Δ threshold for different better combination of parameters [$W = 10, \alpha = 0.2, T_d = 40$, $W = 20, \alpha = 0, T_d = 20$, $W = 30, \alpha = 0.1, T_d = 40$, $W = 40, \alpha = 0.1, T_d = 20$]

During the recognition step, if the condition $\Delta < T_\Delta$ is true for a given fingerprint, the fingerprint may ignore the relative classification. To recognize a song played in an audio piece, count all the classification for which the condition $\Delta \geq T_\Delta$ is checked.Using this approach for each song in the collection will produce a counting value, called $CS_i, \quad \forall i \in [1, \ldots, \sigma]$. Song j is considered recognized for the specified piece of audio if the counter value associated with song j is greater then the counter value associated with all other songs as specified in Equation (14):

$$CS_j > CS_i, \quad \forall i \neq j \tag{14}$$

Using this approach, the selected combination of parameters ($W = 10, \alpha = 0.2, T_d = 40$) and a threshold $T_\Delta = 0.01$ resulting a recognition ratio equal to 100%.

Furthermore, as in Figure 16, this proposed approach was set for comparison with the landmark-based approach by [3]. Using the same data set audio query and creating a landmark song database with $10/sec$ fingerprint hashes. The landmark-based approach obtains 85.25% which is marginally lower compared to the best output accuracy of 90% with a parameter combination equal to ($W = 10, \alpha = 0.2, T_d = 40$) and a threshold of $T_\Delta \approx 0.01$.

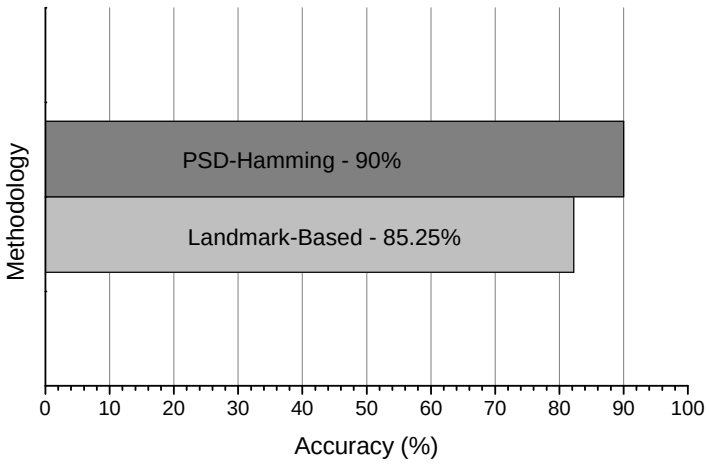


Figure 16: Accuracy comparison between PSD-Hamming and Landmark-Based

9 Conclusions

In conclusion, this research demonstrates the achievement of the project's objectives in the context of musical recognition. Client-server communication has been implemented, managing to work on the server-side with the programming language Python and client-side with language Java on the Android platform.

Also, structuring the information in JSON format is a basic design for this implementation. The client is now able to forward the generated fingerprints to a dedicated server. The graphical interface design for Android enables the audio signal acquisition and displays the screen's generated fingerprints. The implementation has been extended to show the screen feedback on communication with the server by displaying appropriate message notifications. The fingerprints acquired by the server then perform the recognition with that stored database collection.

We believe it would be worthwhile to obtain a more extensive song collection for future analysis and experiments in this direction and be used as a commercial system. This would allow us to gain insight into the types of effects and combinations that prevent an automated recognition system from correctly identifying certain audio query portions. We hope that our contributions can support the active field of audio fingerprinting research. The attention to detail in our evaluation methodology, together with provable reference data collections, will allow our system to serve as a basis for further research in this field.

Acknowledgment We would like to thanks MIUR as the funder of the Industrial PhD program. We are also very grateful for the continuous support from research team members at MDSLabs, University of Messina.

References

[1] C. Bellettini, G. Mazzini, "A framework for robust audio fingerprinting." JCM, **5**(5), 409–424, 2010.

[2] J. Deng, W. Wan, X. Yu, W. Yang, "Audio fingerprinting based on spectral energy structure and NMF," in Communication Technology (ICCT), 2011 IEEE 13th International Conference on, 1103–1106, IEEE, 2011.

[3] D. Ellis, "The 2014 labrosa audio fingerprint system," in ISMIR, 2014.

[4] M. Suzuki, S. Tomita, T. Morita, "Lyrics Recognition from Singing Voice Focused on Correspondence Between Voice and Notes." in INTERSPEECH, 3238–3241, 2019.

[5] Y. Shustef, "Music recognition method and system based on socialized music server," 2015, uS Patent 9,069,771.

[6] A. L.-C. Wang, C. J. P. Barton, D. S. Mukherjee, P. Inghelbrecht, "Method and system for identifying sound signals," 2014, uS Patent 8,725,829.

[7] A. L.-C. Wang, J. O. Smith III, "Systems and methods for recognizing sound and music signals in high noise and distortion," 2018, uS Patent 9,899,030.

[8] B. Mont-Reynaud, A. Master, T. P. Stonehocker, K. Mohajer, "System and methods for continuous audio matching," 2016, uS Patent 9,390,167.

[9] A. M. Kruspe, M. Goto, "Retrieval of song lyrics from sung queries," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 111–115, IEEE, 2018.

[10] M. A. B. Sahbudin, C. Chaouch, M. Scarpa, S. Serrano, "Audio Fingerprint based on Power Spectral Density and Hamming Distance Measure," Journal of Advanced Research in Dynamical and Control Systems, **12**(04-Special Issue), 1533–1544, 2020, doi:10.5373/JARDCS/V12SP4/20201633.

[11] C. Chaouch, M. A. B. Sahbudin, M. Scarpa, S. Serrano, "Audio Fingerprint Database Structure using K-modes Clustering," Journal of Advanced Research in Dynamical and Control Systems, **12**(04-Special Issue), 1545–1554, 2020, doi:10.5373/JARDCS/V12SP4/20201634.

[12] Y. Murthy, S. G. Koolagudi, "Content-Based Music Information Retrieval (CB-MIR) and Its Applications toward the Music Industry: A Review," ACM Computing Surveys (CSUR), **51**(3), 45, 2018.

[13] E. Olteanu, D. O. Miu, A. Drosu, S. Segarceanu, G. Suciu, I. Gavat, "Fusion of speech techniques for automatic environmental sound recognition," in 2019 International conference on speech technology and human-computer dialogue (SpeD), 1–8, IEEE, 2019.

[14] C. Sreedhar, N. Kasiviswanath, P. C. Reddy, "Clustering large datasets using K-means modified inter and intra clustering (KM-I2C) in Hadoop," Journal of Big Data, **4**(1), 27, 2017.

[15] P. Domingo, F. Moore, Global Music Report 2018 - ANNUAL STATE OF THE INDUSTRY, IFPI, 2018.

[16] L. Hallett, A. Hintz, "Digital broadcasting–Challenges and opportunities for European community radio broadcasters," Telematics and Informatics, **27**(2), 151–161, 2010.

[17] M. C. Keith, The radio station: Broadcast, satellite and Internet, Focal Press, 2012.

[18] R. F. Potter, "Give the people what they want: A content analysis of FM radio station home pages," Journal of Broadcasting & Electronic Media, **46**(3), 369–384, 2002.

[19] M. A. B. Sahbudin, M. Scarpa, S. Serrano, "MongoDB Clustering using K-means for Real-Time Song Recognition," in 2019 International Conference on Computing, Networking and Communications (ICNC), 350–354, IEEE, 2019, doi:10.1109/ICCNC.2019.8685489.

[20] M. A. B. Sahbudin, C. Chaouch, M. Scarpa, S. Serrano, "IoT based Song Recognition for FM Radio Station Broadcasting," in 2019 7th International Conference on Information and Communication Technology (ICoICT), 1–6, IEEE, 2019, doi:10.1109/ICoICT.2019.8835190.

[21] http://scikit-learn.org/stable/modules/clustering.html/k-means.

[22] B. Saini, V. Singh, S. Kumar, "Information retrieval models and searching methodologies: Survey," Information Retrieval, **1**(2), 20, 2014.

[23] C. Laufer, The Hobbyist's Guide to the RTL-SDR: Really Cheap Software Defined Radio: A Guide to the RTL-SDR and Cheap Software Defined Radio by the Authors of the RTL-SDR. com Blog, CreateSpace Independent Publishing Platform, 2018.

[24] M. Stewart, "RTSP Package," 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 850-861 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Multi-Layered Machine Learning Model For Mining Learners Academic Performance

Ossama Embarak*

Computer Information Science, Higher Colleges of Technology, Abu Dhabi, 51133, UAE

ARTICLE INFO

Article history:
Received: 19 December, 2020
Accepted: 24 January, 2021
Online: 05 February, 2021

Keywords:
Mining students academic performance
Machine learning
Academic Layered Factors Model
At-risk Students
Recommendation systems for education
Personalized Learning

ABSTRACT

Different colleges and universities have different approaches to dealing with low-performance learners. However, in most cases, analgesics do not deal with root problems. This research suggests a model of three layers of variables sequentially adaptable to a deep-root issue. The suggested model can identify early pupils who could be at risk because of inaccurate or lack of match sequences and suggest rehabilitation. The approach proposed was implemented at three levels. First, we examined the personality type for 180 learners from different majors: Security and Forensics, Networking, and Application Development, using the MBTI test. Second, we build a knowledge matrix for courses by dividing each learning outcome into its knowledge segments. Then, we build the skills matrix for courses by decomposing each learning outcome into its skills segments. We then use machine learning (SVM, DT and association rules) algorithms to mine student performance on a smaller scale of knowledge and skills, taking into account their personality types instead of measuring an entire course's holistic performance. Finally, we developed a system of recommendations to detect performance deviations in knowledge and skills and provide adaptive learning materials that fit the examined students' personality. The proposed approach demonstrates its validity and effectiveness. However, it needs regular updates on learners' performance, which could be automated and linked to evaluation tools. The framework also has a minor impact on learners' privacy since it exposes individual personalities to their advisors.

1. Introduction

This paper is an extension of work initially presented in 2018 Fifth HCT Information Technology Trends (ITT), IEEE [1]. On average, students have different talents, distinct attitudes, and different levels of enthusiasm in their educational programs. Although various students may share a major, it is, however, important to analyze students in many aspects, including the knowledge and skills they have learned. The students have a wide range of learning styles that reflect how they can analyze, plan, and respond to the learning environment [2]. Some students may be comfortable thinking of hypotheses and abstract ideas, while others feel better being a little more concrete with data. Some of the student body prefers active learning, interaction by taking notes and seeing the material presented [3]. It shows how different the students are to each other even within the same major. There are three learners of the course. One learner who has begun to learn by auditory means, another who has begun with tactile activities and finally another learner who is visually dominated [4]. The challenge of the day is to equip our students with the information and skills they will need to accomplish their goals into their future professional careers, regardless of their personal preference. There are many majors in computing that require different knowledge and essential skills than those required for their future careers. To memorize everything here is not exactly effective; i.e., memorization is inefficient. An inability to memorize is a type of learning disorder. Students with such a condition do not expend much effort in studying or preparing for exams. There are two different behaviours that those students might do. Some students might adopt a more in-depth approach when understanding the meaning, whereas the others might make a more strategic approach to understand such type of knowledge, get the best knowledge of their peers and gain related skills [5,6]. As it should be, learners often have their own personal values that affect their own personal ability to absorb new knowledge and learn new functional skills [7].

Knowing students' personality patterns is extremely valuable in finding out their skills; the Myers-Briggs Type Indicator (MBTI) is very much in use. as a means to do just that. Neural dynamics testing techniques, as used in MBTI, helps with understanding the form of human personality. The four main personality traits are eye contact, emotional perception of facial expressions, sensitivity to language, and sensitivity to others' physical appearance [8].

The first pattern is the orientation of energy: Extroversion (E) is preferable to those who prefer to be energized for situations, people and things (that is, for the external world)

*Corresponding Author: Ossama Embarak, oembarak@hct.ac.ae

compared with their complement, preferring to direct their energy towards information, beliefs, ideas and explications (that is, the internal world) (I).

The second pattern is the information and things an individual wants to deal with; Those who prefer to deal with facts and describe clearly situations are sensitive (S). They prefer to deal with ideas, expectations and unknown factors that are not obvious, rather than their additions dependent on intuition. (N).

The third pattern is the sort of decision-making, People would rather make decisions based on objective logic and analysis- Thinking (T) than those who prefer to make decisions based on the values they believe important — feelings Feelings (F).

Fourth, how people make their planning, Those who prefer well-designed and planned lifestyles judge (J) against those who like flexibility, flow and react to things, i.e. perception (P) [9].

- ***Extraversion(E) vs Introversion (I)***
- ***Sensing (S) vs Intuition(N)***
- ***Thinking(T) vs Feeling(F)***
- ***Judging vs Perceiving(P)***

A person may be more inclined to be extroverted than an introvert. Likewise, a person could be more sensing than intuitive, more thoughtful than feeling and judging than perceived. The table below decipher examples of detected personalities.

Table 1:Sample types of personality

Personality Type	ISTJ	ISFJ	ESFP	INTP	ENTJ
Extraversion(E)	×	×	√	×	√
Introversion (I)	√	√	×	√	×
Sensing (S)	√	√	√	×	×
Intuition(N)	×	×	×	√	√
Thinking(T)	√	×	×	√	√
Feeling(F)	×	√	√	×	×
Judging(J)	√	√	×	×	√
Perceiving(P)	×	×	√	√	×

Each personality fits more than the others to a particular career; for example, the ISTJ personality type could fit more into the Systems administrator.

I: Propensity to think about things in the mind
S: Prefer ideas with practical applications
T: Take decisions on an impersonal basis, using a logical reasoning
J: Prefer detailed step-by-step instructions

Combining these four couples generates 16 alternative personality types that enable scientists to understand individuals and advance in a specific area through measurement (s). Students with a preference for introversion, probably doing a lot of thorough work or considering a problem, can find it difficult to live in a noisy or interactive environment [10]. Consequently, they can tend to pick less active and interrupted concentrations. The measurement of computer skills in vital terms and concepts, which are understandable and learning for students during their studies. We have to measure skills such as troubleshooting, network setup, audit protection framework, etc. In addition, it is possible to develop applications, systems, networks, etc.

Maintaining a system to ensure students' personality types is extremely important for understanding their emotions, thinking, and behaviour. For example, understand if a student prefers to work alone or prefers to work with others. Prefers an extremely organized and fixed career, or someone who would like a flexible, open schedule, which allows him or her to be spontaneous. This information helps to determine which career(s) are appropriate for the preferences of students. There are no special advantages or disadvantages over these patterns as they complement each other. If a student has been detected as an Extroversion with a specific percentage, it is an Introversion of the supplement percentage; therefore, each pair is valid for each person but with different strength. The networking employee is expected to be self-driven to deal with network issues and not wait for outside directions. Extraverts tend to be more active when learning, while introverts tend to be more reflective when learning.

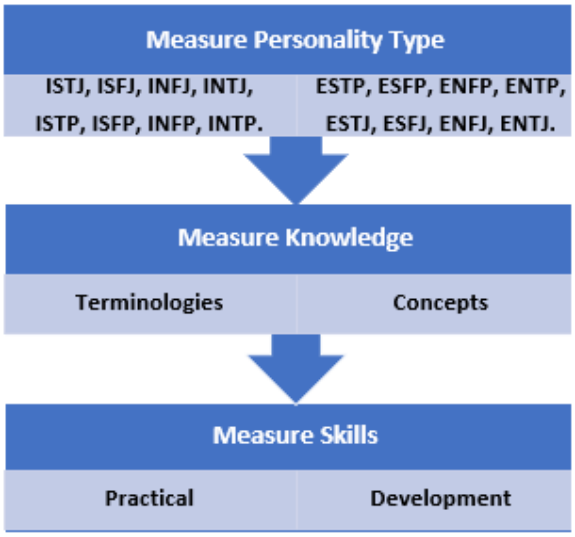


Figure 1: Learner comprehension layers

3. Literature Review

We cannot think of a single method of teaching suitable for all pupils of the same major. Most academic institutions follow the same method, which relies on the delivery of lectures, where students have to absorb the material and then repeat it in the review. This approach does not solve the gaps in student skills, expertise and competencies. Different learning styles define students they prefer to interpret knowledge in different ways [11]. Several models were developed to classify the way people learn; Jung's psychological theory, also known as the Myers-Briggs Type Indicator, is the most famous model (MBTI) [12]. This test provides ample information about a person's personality and the consequences of his or her education. The engineering department of the University of Western Ontario has demonstrated that in the first year of engineering courses men introverts, intuition, thinkers and judges are more likely to excel than extraverts, felt, sensed, and interpreted people at the lower end of their range [13]. They used an MBTI test and performance test on 119 students to determine whether that particular group of seniors would perform very well later in a class like that. The type of personality that is more likely to drop out of the curriculum is recommended to be one that is not overly concerned with grades.

Sensors personality ratings were significantly lower than Intuitors in courses with few high abstractions, an especially important practical question posed by the tests. Introverts may be well-suited for team activities, but extraverts will enjoy working together. Persons who score highly in the Intuitive

cluster score higher on creativity and problem-solving skills than those with a Sensing personality [14]. For example, in understanding academic success, another model created in recent years is the Felder and Silverman [15]. Its purpose is to answer four questions, which are as follows.

On the student's preferred topic, what information do they preferentially perceive? For the student who has a sensory personality, they prefer to be focusing on interacting with what they see, hearing the sounds, and using their physical sensations as a way of experiencing. Those who think more specifically are more comfortable with their abstracts (theories, math, memories, thoughts, and are more likely to solve the problems faster and innovative. What kind of information are our senses perceived most effectively? That is visual personality (exhibited in the eyes) or verbal personality (expressed through speech). What process is the student used to incorporate this information into their outside learning? Actively through participation in physical activity, or through meditation. How does the student characteristically get from "not understanding" to "understanding?" The classroom needs to be set up with logical steps in place, with students thinking linearly and being able to only work with a partial understanding (logic step), or where students think globally and applying their knowledge until they fully understand the picture and have a holistic perspective (global system step) [16]. The Kolb Model assumes that every class possibly includes students of every category. This leads to the belief that these learners are often more likely to become better learners, leading them to perform better [17].

The key issue is how to find an appropriate method of learning for new students when the knowledge is lacking. There are numerous approaches that solutions tend to be used in solving a new state: a cold start problem, [18] Develop the algorithm that finds the closed likeminded learner, and so it generates its recommendations accordingly. [19] Researchers found that students with characteristics that esteem the parental world among students in technology majors were more likely to be of an ESFJ, ISFJ, ESFP, INTJ, and ISTP. On the other hand, students majoring in management information systems tend to be more ESTJ and ISTJ. In an article by [20]ISTJ, INTJ, and ISFJ are the most common personality types found in the computing industry. A study that was done by [21] shows that the relationship between the students' personality and their academic performance depends on the correspondence between the students' personality and the chosen major. Researchers have found, first of all, that system analysts are more extroverted than other thought-provoked people, that computer designers are more introverted and less extroverted than other thought-provoked people, and that programmers are more introverted and less extroverted than other thought-provoked people.

4. The Proposed Approach

The details of the proposed approach are to be applied in four phases. In the first phase, we build the courses' knowledge and skills matrix by decomposing the course's learning outcomes into its knowledge and skills segments. This will help apply machine learning algorithms to mine learners' performance on a smaller scale of concept segments rather than on the entire course's performance or an entire block of concepts. In the second phase, we collected the personalities of students using the MBTI test. In the third level, we assess learners' success (Knowledge & Skills) considering their personality and concentration. Finally, apply a recommendation framework to enhance the performance of learners using machine learning.

4.1. Construct a matrix of knowledge and skills

Analyze courses in terms of skill and knowledge. The starting point was to decompose the courses into their relevant knowledge and skills, representing all the concepts used in each CLO course and laid out in the PLO curriculum, as seen in the following analysis for CIS1403, fundamental to the programming course.

Table 2: Fundamental of programming knowledge & Skills

CLO Number	Description	knowledge	Skills
CLO1	Apply basic programming concepts to write simple programs that use data types, variables, constants, expressions, and statements, focusing on these constructs in Java.	Just remember data types. Recognize the appropriate parameters and values. Please, understand expressions.	Apply the data types. Read through the used data types Create a fully working application in which all variables are properly typed.
CLO2	Write programs that take control flow decisions by means of conditional statements and iterative decisions.	Understand how expressions work. Learn to understand Iteration statements.	Make use of iterations to accomplish control flow. Use conditionals to control the flow of a program.
CLO3	Write programs using unidimensional arrays for initialization, access, transversal, and searching for different data types.	Understand one-dimensional array	Write programs using single-dimensional arrays Write programs that access, update, and search one-dimensional arrays.
CLO4	Develop programs using built-in and user-defined functions with parameter lists.	Understand different user-defined functions {void and return}	Analyze user-defined functions Write functions with parameters
CLO5	Organize and document the program's source code in line with industry standards and best practices.	Recognize coding standers Recognize coding best practices	Apply coding best practices Analyze a programming problem Write a fully working program with function calls

The method retains four key components: intelligence, abilities, attitude and soft skills of the learner. The developed machine learning model used by these components to assess learners of the greatest importance, learners' key causes of At-risk, and we strive to calculate their employability ratio in the light of these core components.

4.2. Evaluation of student personalities

Understanding the personality type will help clarify the preferences, and how or why students may be different. Personality types are useful for understanding how a person can learn, manage, monitor, communicate, collaborate, compromise and cope with stress. Understanding the behaviours of learners helps prepare students to improve their knowledge and skills. This helps to apply a self-awareness tool to recognize strengths, consider and manage growth needs, and establish a career and personal development strategies.

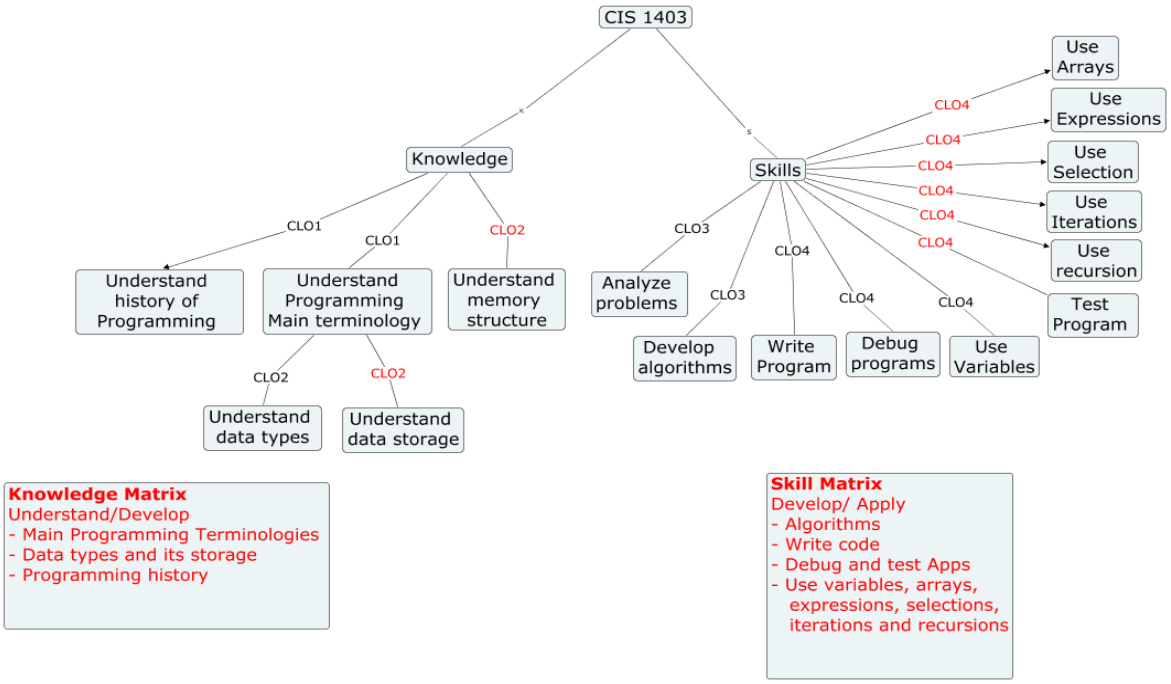
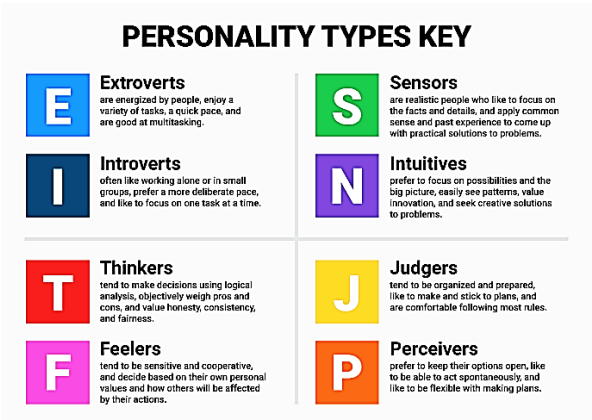


Figure 2: Course knowledge and skills decomposition

We collected data about students' personality using https://www.16personalities.com/, where students' future profession can be categorized into MBTI Sixteen Personality Types shown below [22].



ISTJ	**ISFJ**	**INFJ**	**INTJ**
factual	detailed	committed	independent
practical	traditional	creative	visionary
organized	service-minded	determined	original
steadfast	devoted	idealistic	global
ISTP	**ISFP**	**INFP**	**INTP**
logical	caring	compassionate	independent
realistic	adaptable	original	theoretical
adventurous	gentle	creative	analytical
self-determined	harmonious	empathetic	reserved
ESTP	**ESFP**	**ENFP**	**ENTP**
activity-oriented	enthusiastic	creative	enterprising
versatile	friendly	versatile	outspoken
pragmatic	cooperative	perceptive	challenging
outgoing	tolerant	imaginative	resourceful
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
logical	thorough	loyal	logical
systematic	responsible	verbal	strategic
organized	detailed	energetic	fair
conscientious	traditional	congenial	straightforward

Figure 3: Personality types main features

4.3. Build a recommendation system

Build a recommendation framework that considers learners' personalities from on the one hand and the knowledge and skills they have from the other, as shown in the figure below.

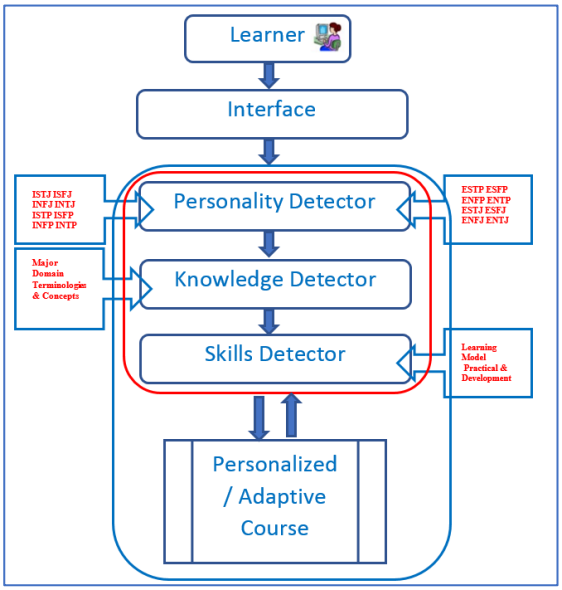


Figure 4: Three-layered course personalization model.

5. Research Procedures

For this research, 180 students were asked twice to use different platforms to take the Myers-Briggs Style Test. Eliminating all irregularities, only correct and clean data for treatment has been processed. Data collection incorporates the Type of Personality and the Personality Code; Extraversion (E), Introversion (I), Sensing (S), and Intuition (N) (P).

Variables

Three factors influence the development of students and thus their academic status in this research. First, we will consider the

personality type created through a four-letter combination, each of which represents a set of independent characteristics. : extroverted versus introduced, sensing versus intuitive, thought versus feeling and judging versus prospecting. The second factor is the students' knowledge based on each class: the third factor is the students' skills that meet their concentration needs. We can determine if a student is in good academic standing by considering their first-year grades, time in school, etc.

5.1. Create Learners profiles

We have developed student profiles that incorporate their personality style and academic knowledge and skills. The figure below shows the types of students, which is used to analyze whether the student's personality affects their achievements in the three majors of computer networking.

Table 3: Learners measured personality

Catalog	Academic Standing	Cgpa	Major	Emsat English Overall Score	Emsat Math	High School Average	Personality Type	Extraverted	Introverted	Intuitive	Observant	Thinking	Feeling	Judging	Prospecting	Assertive	Turbulent
201810	Very Good	3.12	CSF	1175	1225	86.24	ESTP-T	56	44	45	55	60	40	44	56	39	61
201620	Near At-Risk	2.29	CSF	1175	0	84.8	INFJ-T	12	88	55	45	43	57	61	39	42	58
201910	Near At-Risk	2.08	COM	1300	300	87.7	ISFJ-A	39	61	47	53	43	57	57	43	51	49
201910	Distinctive	3.81	COM	1050	850	88.57	INFJ-T	15	85	64	36	19	81	60	40	26	74
201920	Good Standing	2.6	COM	1075	725	89.59	ENFP-T	54	46	52	48	44	56	49	51	47	53
201810	At-Risk	1.8	CIN	1025	425	87.2	ENTJ-T	61	39	52	48	58	42	56	44	44	56
201520	Near At-Risk	2.34	FJM	0	0	81.5	ISFP-T	47	53	41	59	44	56	46	54	37	63
201820	Very Good	3.13	COM	1125	550	82.31	ESTJ-A	83	17	42	58	51	49	67	33	69	31
201810	Distinctive	3.67	CSF	1025	850	93.96	INFJ-T	43	57	53	47	43	57	56	44	46	54
201810	At-Risk	1.82	CIA	1150	375	87.6	INTJ-T	37	63	51	49	64	36	67	33	49	51
201810	Good Standing	2.66	CIN	1125	450	88.3	INFJ-T	39	61	74	26	56	44	56	44	42	58
201910	Good Standing	2.71	COM	1100	0	91.3	ESTP-T	56	44	45	55	28	72	71	29	56	44
201910	Near At-Risk	2.44	COM	0	0	78.6	ISFJ-T	32	68	30	70	49	51	69	31	42	58
201810	Good Standing	2.65	CSF	1150	300	87.3	ISFP-A	39	61	37	63	47	53	49	51	56	44
201810	Good Standing	2.79	CSF	1250	1050	88.85	INTP-T	39	61	55	45	56	44	42	58	49	51
201720	Near At-Risk	2.45	CIN	700	300	90.5	INTP-T	49	51	53	47	53	47	49	51	46	54
201820	At-Risk	1.84	COM	1100	325	76.92	ESFP-A	29	71	46	54	53	47	64	36	43	57
201810	At-Risk	1.78	FJW	1100	300	89.5	ESFJ-T	56	44	45	55	43	57	53	47	44	56
201910	Very Good	3.11	COM	1075	850	89.78	INFP-T	54	46	32	68	18	82	60	40	36	64
201810	Good Standing	2.76	CIA	1100	300	93.6	ISFJ-T	57	43	49	51	44	56	63	37	43	57

The following two tables show a sample of students' knowledge and skills performance in the programming course.

Table 4: Learners knowledge performance

Catalog	Academic Standing	Cgpa	Major	Remember Data Types	Understand Variables, Constants	Understand Expressions	Understand Selection Statements	Understand Iteration Statements	Understand One Dimensional Array	Understand Different User Defined Functions {Void And Return}	Knowledge Perforamnce
201810	Very Good	3.12	CSF	0.9	1.8	1.8	2	2	1.91	1.91	1.76
201620	Near At-Risk	2.29	CSF	0.925	1.85	1.85	1.775	1.775	1.89	1.89	1.71
201910	Near At-Risk	2.08	COM	0.9125	1.825	1.825	1.85	1.85	1.74	1.74	1.68
201910	Distinctive	3.81	COM	0.92	1.84	1.84	1.755	1.755	1.8	1.8	1.67
201920	Good Standing	2.6	COM	0.925	1.85	1.85	1.72	1.72	1.53	1.53	1.59
201810	At-Risk	1.8	CIN	0.725	1.45	1.45	1.6	1.6	1.9	1.9	1.52
201520	Near At-Risk	2.34	FJM	0.975	1.95	1.95	1.075	1.075	1.734	1.734	1.50
201820	Very Good	3.13	COM	0.87	1.74	1.74	1.24	1.24	1.83	1.83	1.50
201810	Distinctive	3.67	CSF	0.7275	1.455	1.455	1.46	1.46	1.89	1.89	1.48
201810	At-Risk	1.82	CIA	0.7275	1.455	1.455	1.46	1.46	1.89	1.89	1.48
201810	Good Standing	2.66	CIN	0.9375	1.875	1.875	1.12	1.12	1.7	1.7	1.48
201910	Good Standing	2.71	COM	0.8225	1.645	1.645	1.185	1.185	1.89	1.89	1.47
201910	Near At-Risk	2.44	COM	0.8225	1.645	1.645	1.185	1.185	1.89	1.89	1.47
201810	Good Standing	2.65	CSF	0.7875	1.575	1.575	1.72	1.72	1.442	1.442	1.47
201810	Good Standing	2.79	CSF	0.8625	1.725	1.725	1.32	1.32	1.63	1.63	1.46
201720	Near At-Risk	2.45	CIN	0.8	1.6	1.6	1.28	1.28	1.8	1.8	1.45

201820	At-Risk	1.84	COM	0.7625	1.525	1.525	1.52	1.52	1.46	1.46	1.40
201810	At-Risk	1.78	FJW	0.6	1.2	1.2	1.68	1.68	1.7	1.7	1.39
201910	Very Good	3.11	COM	0.7	1.4	1.4	1.44	1.44	1.6838	1.6838	1.39
201810	Good Standing	2.76	CIA	0.75	1.5	1.5	1.56	1.56	1.43	1.43	1.39
201820	Near At-Risk	2.16	COM	0.8	1.6	1.6	1.005	1.005	1.81	1.81	1.38

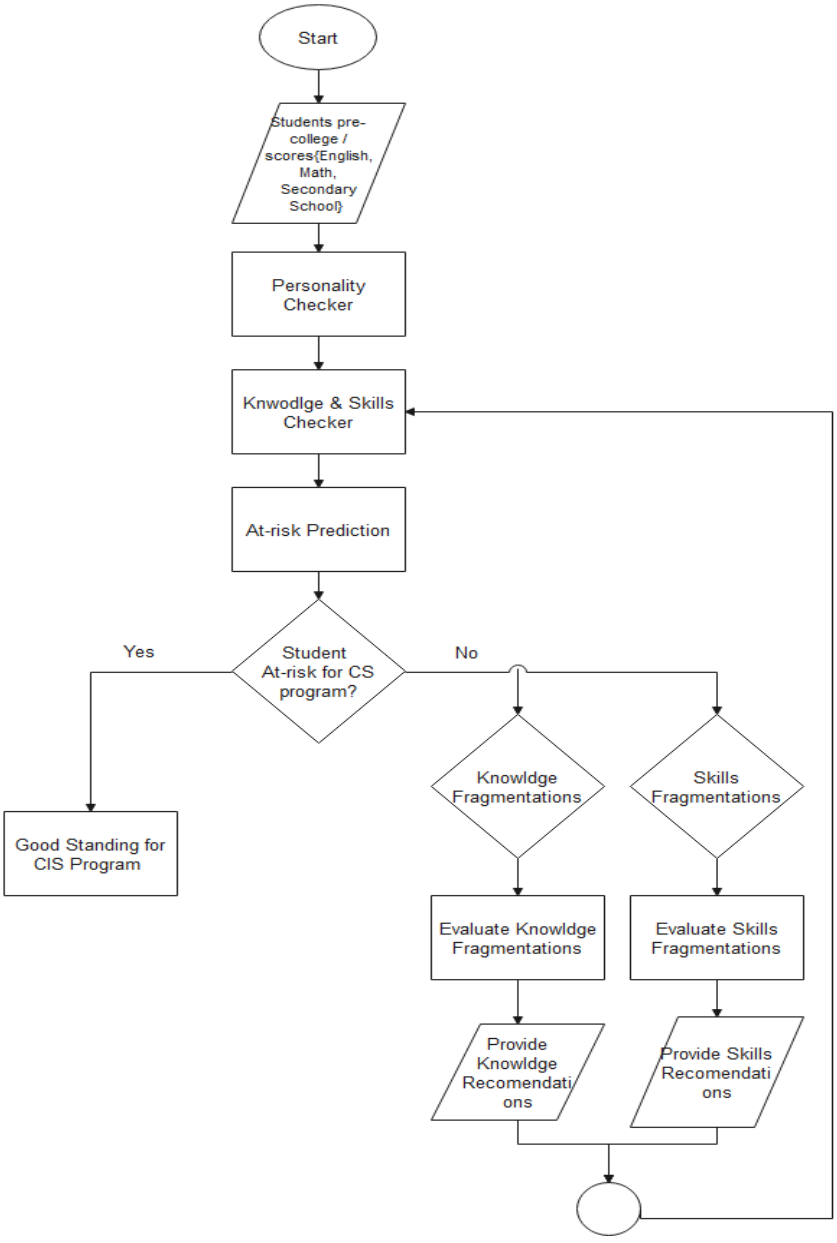


Figure 5: Flow chart for detecting the best program

Table 5: Learners skills performance

Catalog	Academic Standing	Cgpa	Major	Apply Coding Best Practices	Analyse User Defined Functions	Analyse A Programming Problem	Apply Data Types	Evaluate The Used Data Types	Create A Fully Working Application Using Proper Variables Data Types	Apply Control Flow Via Conditional Statements	Apply Control Flow Via Iterations	Write A Program Using Selection And Iteration	Write Programs That Employ One-Dimensional Arrays Initialization	Write Programs That Employ One-Dimensional Arrays Accessing,	Write Functions With Parameters	Write A Fully Working Program With Function Calling	Skills Perforamnce
201810	Very Good	3.12	CSF	2.01	3.82	3.82	2.865	4.775	2.865	3	3	5.73	6	5.73	4.02	4.02	3.97
201620	Near At-Risk	2.29	CSF	2.67	3.78	3.78	2.835	4.725	2.835	2.663	2.663	5.67	5.325	5.67	5.34	5.34	4.10
201910	Near At-Risk	2.08	COM	2.664	3.48	3.48	2.61	4.35	2.61	2.775	2.775	5.22	5.55	5.22	5.328	5.328	3.95

201910	Distinctive	3.81	COM	2.673	3.6	3.6	2.7	4.5	2.7	2.633	2.633	5.4	5.265	5.4	5.346	5.346	3.98
201920	Good Standing	2.6	COM	2.175	3.06	3.06	2.295	3.825	2.295	2.58	2.58	4.59	5.16	4.59	4.35	4.35	3.45
201810	At-Risk	1.8	CIN	2.588	3.8	3.8	2.85	4.75	2.85	2.4	2.4	5.7	4.8	5.7	5.175	5.175	4.00
201520	Near At-Risk	2.34	FJM	0.981	3.468	3.468	2.601	4.335	2.601	1.613	1.613	5.202	3.225	5.202	1.962	1.962	2.94
201820	Very Good	3.13	COM	2.481	3.66	3.66	2.745	4.575	2.745	1.86	1.86	5.49	3.72	5.49	4.962	4.962	3.71
201810	Distinctive	3.67	CSF	1.806	3.78	3.78	2.835	4.725	2.835	2.19	2.19	5.67	4.38	5.67	3.612	3.612	3.62
201810	At-Risk	1.82	CIA	1.806	3.78	3.78	2.835	4.725	2.835	2.19	2.19	5.67	4.38	5.67	3.612	3.612	3.62
201810	Good Standing	2.66	CIN	2.265	3.4	3.4	2.55	4.25	2.55	1.68	1.68	5.1	3.36	5.1	4.53	4.53	3.42
201910	Good Standing	2.71	COM	2.358	3.78	3.78	2.835	4.725	2.835	1.778	1.778	5.67	3.555	5.67	4.716	4.716	3.71
201910	Near At-Risk	2.44	COM	2.358	3.78	3.78	2.835	4.725	2.835	1.778	1.778	5.67	3.555	5.67	4.716	4.716	3.71
201810	Good Standing	2.65	CSF	1.8	2.884	2.884	2.163	3.605	2.163	2.58	2.58	4.326	5.16	4.326	3.6	3.6	3.21
201810	Good Standing	2.79	CSF	1.695	3.26	3.26	2.445	4.075	2.445	1.98	1.98	4.89	3.96	4.89	3.39	3.39	3.20
201720	Near At-Risk	2.45	CIN	2.454	3.6	3.6	2.7	4.5	2.7	1.92	1.92	5.4	3.84	5.4	4.908	4.908	3.68
201820	At-Risk	1.84	COM	1.845	2.92	2.92	2.19	3.65	2.19	2.28	2.28	4.38	4.56	4.38	3.69	3.69	3.15
201810	At-Risk	1.78	FJW	1.665	3.4	3.4	2.55	4.25	2.55	2.52	2.52	5.1	5.04	5.1	3.33	3.33	3.44
201910	Very Good	3.11	COM	2.19	3.368	3.368	2.526	4.21	2.526	2.16	2.16	5.051	4.32	5.0514	4.38	4.38	3.51
201810	Good Standing	2.76	CIA	1.86	2.86	2.86	2.145	3.575	2.145	2.34	2.34	4.29	4.68	4.29	3.72	3.72	3.14
201820	Near At-Risk	2.16	COM	2.538	3.62	3.62	2.715	4.525	2.715	1.508	1.508	5.43	3.015	5.43	5.076	5.076	3.60
201810	Near At-Risk	2.42	CSF	2.055	3.02	3.02	2.265	3.775	2.265	2.34	2.34	4.53	4.68	4.53	4.11	4.11	3.31

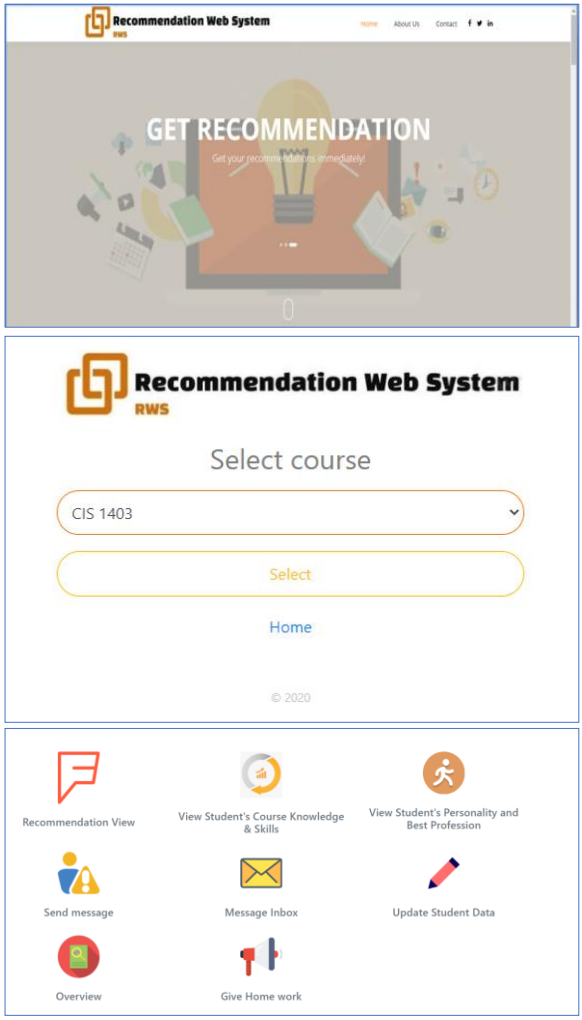


Figure 6: Recommendation Education system for At-risk learners

In order to find the best predictor for student performance, we have used different machine learning algorithms (Decision Tree, Random Forest, SVM and Association rules); then we evaluated and refined the used algorithm. The main aim is to develop a primitive platform that can be used to develop an education recommendation system to support At-risk learners 24/7 and teachers to coach students with poor knowledge, skills, and to provide them with the right advice.

Applying online coaching to at-risk students needs to build a recommendation framework that predicts at-risk students considering their knowledge and skills performance in segments of concepts rather than an entire course, considering their personality style. This helps to give coaching in the form of advice crumbs, as shown in the implementation figure 9 below, which may be tutorials, segment coding, exercises, peer discussions, etc.

The association rules are used to created recommendations considering student knowledge and skills performance.

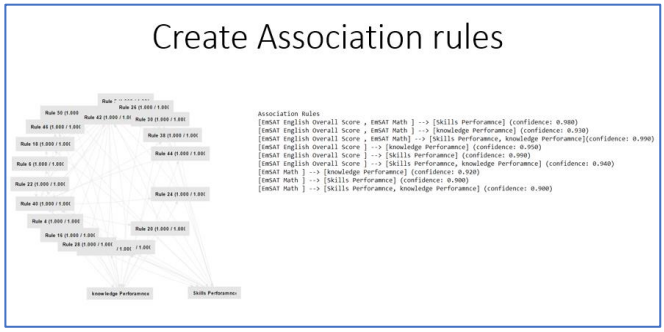


Figure 7: Applied Association rules for creating recommendations

5.2. Explore profiles of students

First, we purified information from all NA instances, measured the mean and variance and the range for each pattern.

As shown in the table below, we measured the mean and variance and the three majors' GPA quartile.

Table 6: Descriptive statistical of the GPA - CIS three-concentrations

Major	count	mean	std	min	25%	50%	75%	max
Applications Development	32.00	3.10	0.40	2.15	3.00	3.20	3.25	3.90
Networking	56.00	2.69	0.61	1.40	2.27	2.56	3.20	3.74
Security and Forensics	126.00	2.69	0.54	1.30	2.40	2.68	3.00	3.68

In the application development department, the students typically have the highest average GPA, the Security and Forensics students have the second-highest average GPA, and the Networking students have the lowest average GPA. In the three conditions, we have been able to identify different personalities. Security and forensic students share the introverted sensing function type, which means they are more likely to feel, and judge - ISFJ.

It should be clear that, as demonstrated below, we have security students with various types of personalities, however, the GPA scores declined slightly accordingly.

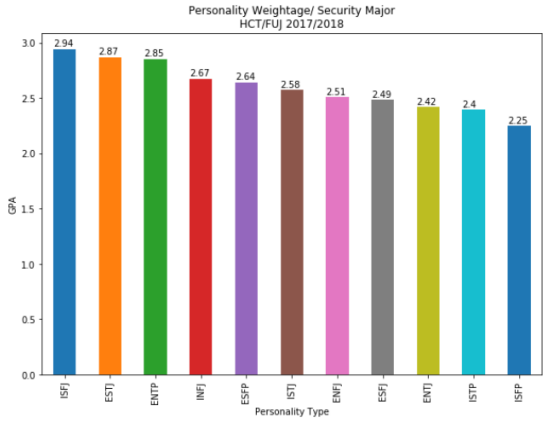


Figure 8: Personality type's weight-Security and Forensics Major.

Students form Networking Major are more ISTJs, which means more Introduction (I), Sensing (S), Thinking (T), and Judging (J). As shown in the below figures, major students in the App development are more ISTJ.

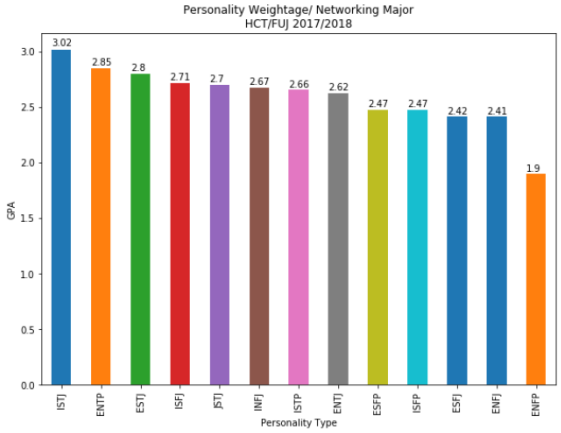


Figure 9: Personality type-Networking major.

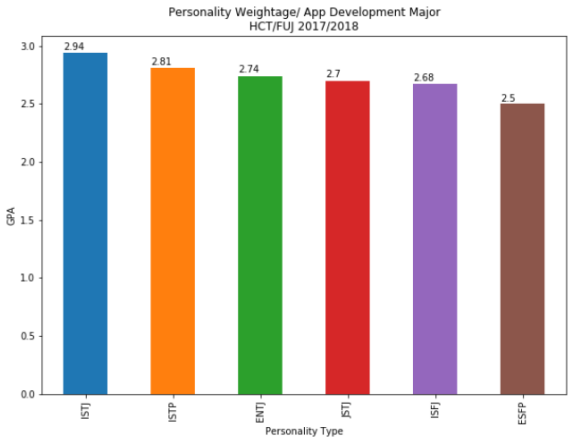


Figure 10: Personality type – App Development major.

6. Experimental Results

In the following tables, we summarized the statistics on student success.

Table 7: Four courses performance statistics for personality types - Security and Forensics major

	GPA	CIS 1403	CIS 2403	CIN 2003	CSF 3103	Extraverted	Introverted	Sensing	Intuitive	Thinking	Feeling	Judging	Prospecting
count	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00
mean	2.44	1.84	1.92	2.18	2.81	42.78	57.22	56.96	43.04	52.70	47.30	59.08	40.92
std	0.30	0.47	0.50	0.51	0.48	20.61	20.61	8.62	8.62	10.82	10.82	16.78	16.78
min	1.87	1.33	1.52	1.00	2.00	7.00	30.25	40.50	34.00	36.00	35.50	32.00	24.50
25%	2.30	1.53	1.68	2.05	2.80	30.58	39.00	56.00	38.67	44.00	39.41	42.00	29.75
50%	2.50	1.73	1.75	2.33	2.84	34.71	65.29	60.00	40.00	57.67	42.33	67.19	32.81
75%	2.53	2.03	2.00	2.38	3.06	61.00	69.42	61.33	44.00	60.59	56.00	70.25	58.00
max	2.92	2.92	3.20	2.85	3.57	69.75	93.00	66.00	59.50	64.50	64.00	75.50	68.00

Table 8: Four courses performance statistics for personality types-Networking major

	GPA	CIS 1403	CIS 2403	CIN 2003	CSF 3103	Extraverted	Introverted	Sensing	Intuitive	Thinking	Feeling	Judging	Prospecting
count	13.00	13.00	13.00	13.00	13.00	13.00	13.00	13.00	13.00	13.00	13.00	13.00	13.00
mean	2.50	1.80	1.79	2.56	2.71	51.81	48.19	53.47	46.53	49.63	50.37	50.01	49.99
std	0.40	0.53	0.49	0.56	0.53	21.06	21.06	7.47	7.47	11.28	11.28	18.07	18.07
min	1.90	1.00	1.30	1.71	1.90	25.33	19.00	41.10	38.00	34.10	33.95	24.25	24.50
25%	2.24	1.57	1.60	2.29	2.26	31.78	30.25	46.00	40.25	39.00	40.54	36.00	31.38
50%	2.50	1.71	1.68	2.40	2.88	55.57	44.43	56.38	43.62	51.00	49.00	46.00	54.00
75%	2.79	2.00	1.94	2.70	3.06	69.75	68.22	59.75	54.00	59.46	61.00	68.62	64.00
max	3.13	3.18	3.22	4.00	3.43	81.00	74.67	62.00	58.90	66.05	65.90	75.50	75.75

Table 9: Four courses performance statistics of personality types-App Development major.

	GPA	CIS 1403	CIS 2403	CIN 2003	CSF 3103	Extraverted	Introverted	Sensing	Intuitive	Thinking	Feeling	Judging	Prospecting
count	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
mean	2.96	3.07	2.98	2.14	1.88	29.68	70.32	43.94	56.06	44.54	55.46	37.97	62.03
std	0.34	0.15	0.21	0.45	0.14	8.85	8.85	12.80	12.80	10.81	10.81	13.60	13.60
min	2.52	2.90	2.74	1.60	1.74	20.33	60.14	30.57	40.00	30.88	43.71	26.60	43.00

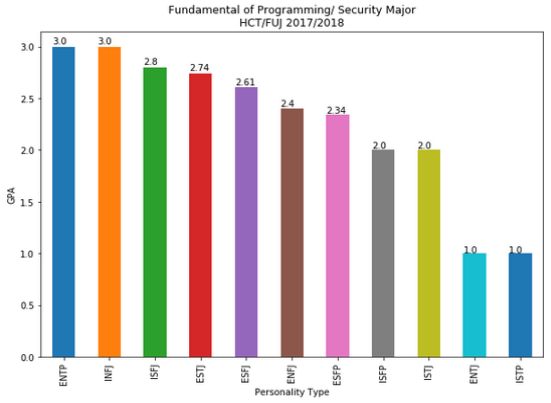
25%	2.79	2.98	2.86	1.86	1.78	23.53	64.58	35.79	49.25	39.22	49.18	29.15	57.04
50%	3.02	3.07	2.98	2.17	1.90	29.27	70.73	42.60	57.40	45.50	54.50	34.14	65.86
75%	3.19	3.16	3.10	2.45	2.00	35.42	76.47	50.75	64.21	50.82	60.78	42.96	70.85
max	3.27	3.23	3.22	2.60	2.00	39.86	79.67	60.00	69.43	56.29	69.12	57.00	73.40

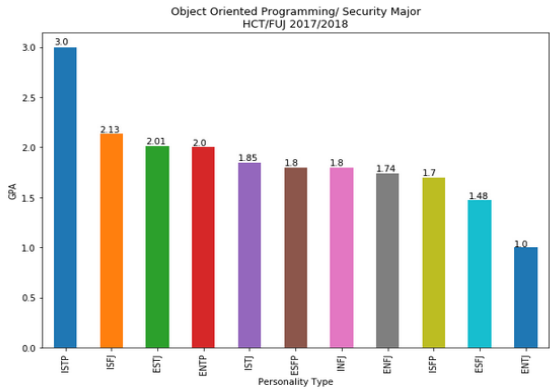
The count feature shown in the above table should be clear, referring to each major's personality styles. In addition, the mean, minimum and standard deviation of GPA values from each course is calculated, and the four courses and personalities are also calculated. Apart from CSF 3103, the course for the security major, the App development students did not receive the results of NaN but took the networking course CIN-2003. We evaluated students' performance with their personalities in the four classes. In every course, according to the college grading system, we determined the GPA value of students.

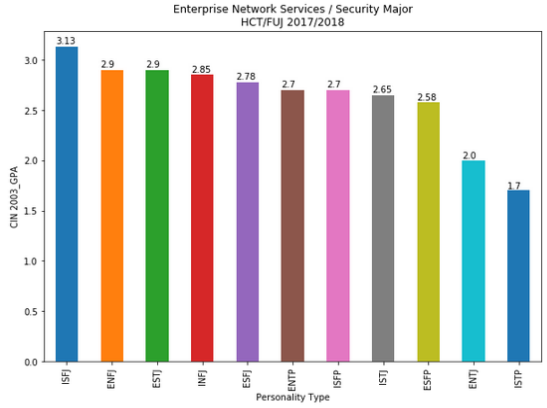
Table 10: College GPA grades conversion values

Score	F	D	D+	C-	C	C+	B-	B	B+	A-	A
GPA	0	1	1.3	1.7	2	2.3	2.7	3	3.3	3.7	4

a) The personalities patterns in the four courses for Security and Forensics major.







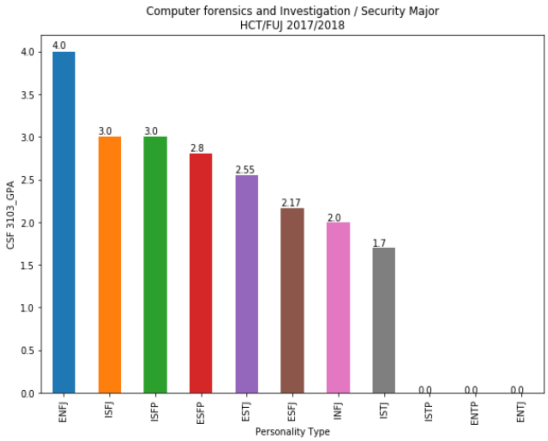
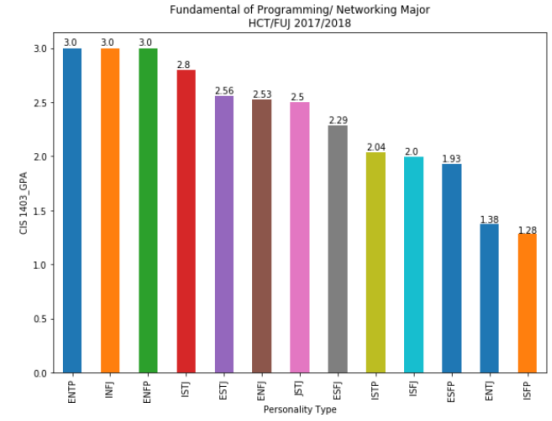
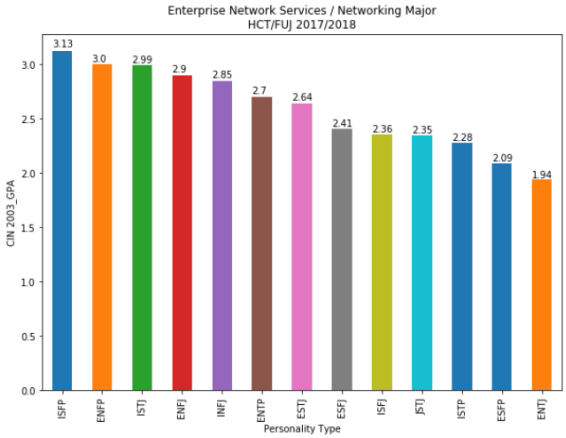
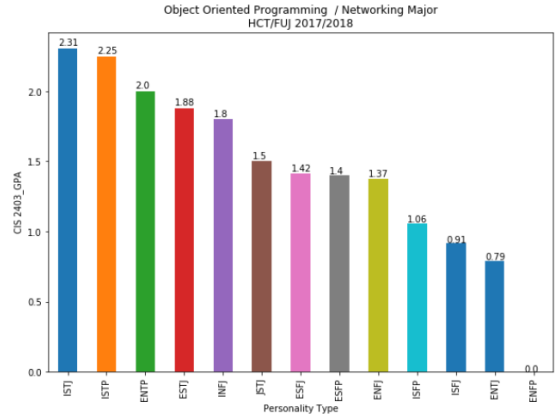


Figure 11: The impact of personality as a measure for four courses on security and forensics major.

b) The personalities patterns in the four courses for Networking major.







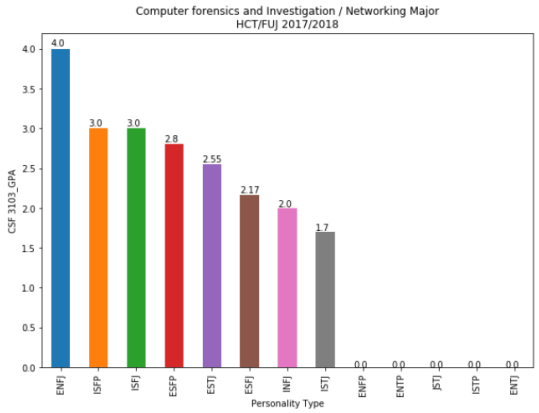


Figure 12: Impact of personality for Networking major as a measure on the four courses

c) The personalities patterns in the four courses for App Development major

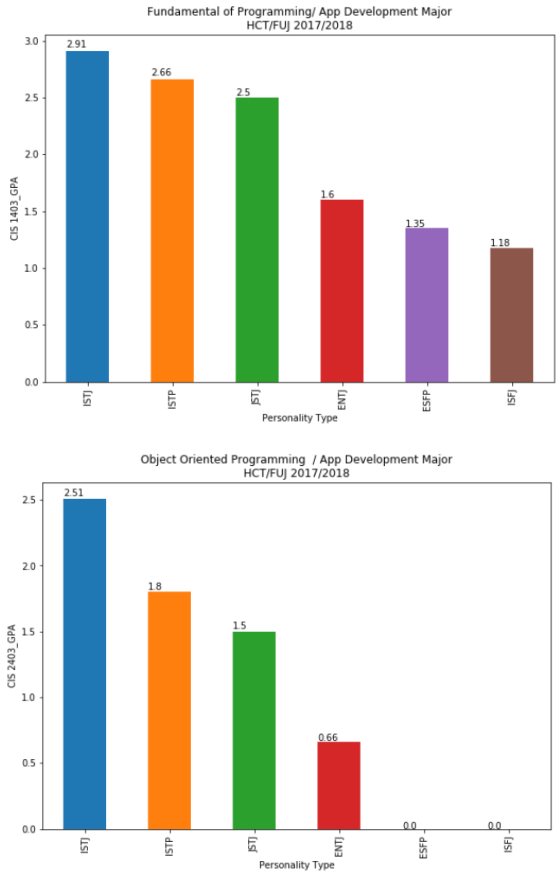


Figure 13: The personality impact on App Development as measured on the four courses.

In each major and different course, the following table summarizes the results of the best personalities.

Table 11: Personality performances in the four courses.

	GPA	Fundamentals of Programming	Object-Oriented Programming	Enterprise Network Services	Computer Forensics and Investigation
Security and Forensics	**>=3.0**	ENTP, INFJ	ISTP	ISFJ	ENFJ, ISFJ, ISFP
	>=2.5	ISFJ, ESTJ, ESFJ		ENFJ, ESTJ, INFJ, ESFJ, ENTP, ISFP, ISTJ, ESFP	ESFP, ESTJ
	>=2.0	ENFJ, ENFP, ISTJ, ISFP	ISFJ, ESTJ, ENTP	ENTJ	ESFJ, INFJ
Networking	**>=3.0**	ENTP, INFJ, ENFP		ISFP, ENFP	ENFJ, ISFP, ISFJ
	>=2.5	ISTJ, ESTJ, ENFJ, ISTJ		ISTJ, ENFJ, INFJ, ENTP, ESTJ	ESFP, ESTJ
	>=2.0	ESFJ, ISTP, ISFJ	ISTJ, ISTP, ENTP	ESFJ, ISFJ, ISTJ, ISTP, ESFP	ESFJ, INFJ
App Development	**>=3.0**				
	>=2.5	ISTJ, ISTP, ISTJ	ISTJ	ISTJ, ISTP	
	>=2.0			ISTJ	

Based on their personality, we may describe the conditions for a major choice to be made, or we have the ability to predict a student's GPA (grade point average). Take an example such as this study by the hospital and try to analyze it with the below algorithm.

If *"ISFJ" or "ESTJ"*

➔ *The best major is Security & Forensics*

ELSE IF *"ISTJ" or "ENTP"*

➔ *The best major is Networking*

ELSE IF *"ISTJ" or "ISTP"*

➔ *The best major is App Development.*

In the case of teaching computers to "understand" the best major, we can also rely on specific patterns, the systematic way, in the algorithm below.

If "S" and "J"

➔ *choose Security & Forensics*

ELSE IF ("S" or 'N') AND "T"

➔ *The best major is Networking*

ELSE IF "S" AND "T"

➔ *The best major is App Development.*

It is important to have measured students' knowledge and skills to recognize getting into trouble students' academic status and their needs. Students are given an assessment in four courses that cover OOP, Enterprise Networks and Services. A breakdown of students' knowledge from various academic majors was compiled, and the numbers are listed below.

The factors that we looked at were the students' different academic performance, as well as the skills they demonstrated in the OOP course.

Table 12: The knowledge of OOP course over the three majors.

Knowledge	Object-Oriented Programming					
	GPA	Concepts of object-oriented programming	Terminology of methods	Terminology of functions	Terminology of inheritance	Terminology of encapsulation
Security and Forensics	>=3.0	82%	77%	76%	81%	83%
	>=2.5	0%	0%	0%	0%	0%
	>=2.0	74%	71%	71%	76%	75%
Networking	>=3.0	0%	0%	0%	0%	0%
	>=2.5	0%	0%	0%	0%	0%
	>=2.0	77%	68%	74%	69%	67%
App Development	>=3.0	0%	0%	0%	0%	0%
	>=2.5	77%	72%	76%	68%	81%
	>=2.0	0%	0%	0%	0%	0%

Table 13: OOP skills performance in the three major categories.

Skills	Object-Oriented Programming					
	GPA	Implement OOP tools	Use advanced features	Develop programs	Use Collections Framework	Test and evaluate a program
Security and Forensics	>=3.0	87%	79%	77%	76%	79%
	>=2.5	0%	0%	0%	0%	0%
	>=2.0	77%	65%	68%	72%	81%
Networking	>=3.0	0%	0%	0%	0%	0%
	>=2.5	0%	0%	0%	0%	0%
	>=2.0	80%	66%	71%	73%	72%
App Development	>=3.0	0%	0%	0%	0%	0%
	>=2.5	79%	69%	77%	80%	83%
	>=2.0	0%	0%	0%	0%	0%

7. Discussion and Conclusions

This study proposes a model for evaluating university students' academic status and GPA performance ambitions. Faculty and staff call upon these three factors in decisions: personality preferences, perceived knowledge and skills as well as what they expect from each student. In their attitudes, knowledge and skills, computing students are not identical. We see the disparities between students' grades because the universities lack an understanding of these disparities between learners' abilities and a special opportunity for creating a truly unique learning method that matches their individuality. This makes it difficult to apply the same philosophy and expects all students to be highly qualified.

This system allows for specific algorithms to be included, in the students' guidance, academic major choice or even the selection process. In a class on security, when we use sections for personality types, we can use an algorithm to predict the student's ranking automatically.

IF "Security & Forensics major."
IF "ENFJ" or 'ISFP' or ISFJ'
➔ *Expected high grade*
ELSE IF 'ENTJ' or 'ENPT'
➔ *Expected fail case*
ELSE
➔ *Expected low grade*

We did not look at all scenarios, only example algorithms.

IF "OBJECT-ORIENTED COURSE"
IF "ISTP" or "ISTJ"
➔ *Expected high grade*
ELSE
➔ *Expected low grade*

In general, consider how OOP awareness is summarized in Table 9. Security students have the ability to implement higher object-oriented concepts than those with a non-security background. Again, we want to emphasize the personality that makes these differences, not the selected major. This clarifies that major security students could perform well than other major students because of the difference in a student's personality that corresponds to the offered programming course. See the algorithm below that demonstrates it.

For a Major
If Personality 'ISTJ' & 'FUNDAMENTAL OF PROGRAMMING'
➔ *Expect GPA >=2.0*

In view of the basics, the students in App developers major with the 'ISTJ' personality achieve round 2.91 GPA, while networking students are expected to achieve 2.8 GPA and security and forensics students 2.0 GPA. The differences are seen in the second and third levels as a result of other factors: knowledge and skills.

In the following sections, the benefits of implementing the three-layer model are listed.

1. Maintain high-quality graduates with an appropriate level in their technical knowledge and skills.
2. Minimize cases at risk.
3. Keep an eye on signs of inattentiveness and poor academic performance.
4. A decrease in the number of delayed graduate students.
5. Avoid the disproportionate dropout from prospective students in response to dismissal or academic warnings.
6. Make sure to have a high level of student retention.
7. Provide tools to assist individuals to become more effective in selecting courses and getting the right major.
8. Keep clear knowledge of each student's academic progress in case of a student falling behind.

8. The Conclusion & Future Work

Educational institutions have a tremendous burden of handling students with low academic performance (At-risk students). Many approaches support this group of pupils, such as psychological therapy, a proper timetable for vulnerable pupils, recall, personal training, mock-tests, direct private education, or success centres. However, these methods are not enough to solve the issue, since other factors could also influence the learner's success, which could be their family difficulties, cognitive style, prior academic performance, and the foundation of the college level. Learners personalities are the most important factor that needs a great deal of attention to understand students' individuality and natural abilities. This greatly helps to adapt the

materials supplied to their perceptions and offers solid support during the study programme. This paper proposes a recommendation system based on three key angles: personalities of the learner, skills of the learner and academic knowledge of the learner. The proposed layered model helps predict at-risk students, provide better advising, enhance learners' performance and skills, discover their employability competencies, and support learners with a recommendation in terms of tutorials, exercises, etc. during their study period. Based on the association rule, the proposed recommendation framework is structured to predict student success in different college programs and predict their academic status based on their personality analysis and academic performance. It is highly recommended that the admission and consultancy framework be maintained at the university level to fill the fragmentation of students' knowledge and skills. This improves student success and clearly identifies the study program's weaknesses and encourages students to address any observed fragmentations.

This pilot study offers further insight as to the strengths and weaknesses of the three specific majors who receive the most job postings, "Application Development, Networking and Security and Forensics." In addition to measuring a student's actual academic performance, the system also distributes personalized advice based on the student's digital profile. Collecting data from multiple learners from different programs allows us to understand what the program is trying to accomplish and what students need to know. The data completed throughout the learning process helps guide and get the learners to further levels of understanding. Future research will continue to create more trust in this advice through the study's continuation to include students from various other programs, like Business Solutions, Applied Media and Engineering. The proposed and implemented solution is very successful and leads to quick results, but it does not offer as much information because it does not account for the students' learning progress. Advisors constantly have access to learners' information about their personalities which affect their privacy.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by H.C.T. Research Grants (H.R.G.) [Fund No: 113108]. We want to express our gratitude to the Higher Colleges of Technology for the financial grant.

References

[1] O.H. Embarak, "Three Layered Factors Model for Mining Students Academic Performance," in 2018 Fifth HCT Information Technology Trends (ITT), IEEE: 219–226, 2018. https://doi.org/10.1109/CTIT.2018.8649491

[2] F. Martin, D.U. J O.L. Bolliger, "Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment," ERIC, **22**(1), 205–222, 2018.

[3] M. Salam, N. Ibrahim, M. J I.J. of I. Sukardjo, "Effects of Instructional Models and Spatial Intelligence on the Mathematics Learning Outcomes after Controlling for Students' Initial Competency," ERIC, **12**(3), 699–716, 2019.

[4] R.A. J O.J. of S.S. Faisal, "Influence of Personality and Learning Styles in English Language Achievement," **7**(8), 304–324, 2019.

[5] C.Y.R. Loh, T.C. J J. of E. Teo, S. Policy, "Understanding Asian students learning styles, cultural influence and learning strategies," **7**(1), 194–210, 2017.

[6] E. Care, H. Kim, A. Vista, K. J C. for U.E. at T.B.I. Anderson, "Education System Alignment for 21st Century Skills: Focus on Assessment," 2018.

[7] J. Halberstadt, J.-M. Timm, S. Kraus, K. J J. of K.M. Gundolf, "Skills and knowledge management in higher education: how service learning can contribute to social entrepreneurial competence development," 2019.

[8] M. Gerlach, B. Farb, W. Revelle, L.A.N. J N. human behaviour Amaral, "A robust data-driven approach identifies four personality types across four large data sets," **2**(10), 735–742, 2018.

[9] A. Baumert, M. Schmitt, M. Perugini, W. Johnson, G. Blum, P. Borkenau, G. Costantini, J.J.A. Denissen, W. Fleeson, B. J E.J. of P. Grafton, "Integrating personality structure, personality process, and personality development," **31**(5), 503–528, 2017.

[10] A. Kamal, S. J E. Radhakrishnan, I. Technologies, "Individual learning preferences based on personality traits in an E-learning scenario," **24**(1), 407–435, 2019.

[11] H.T.-C. in human behavior, undefined 2016, "Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities," Elsevier,.

[12] R. Stein, A.B. %J S. Swan, P.P. Compass, "Evaluating the validity of Myers-Briggs Type Indicator theory: A teaching tool and window into intuitive psychology," **13**(2), e12434, 2019.

[13] Z.T. Ardakani, L. Samadi, "The Relationship between Introversion as Personality Trait and Job Satisfaction among EFL Teachers."

[14] L.S. Byrd, A study of an arts integration curriculum and its impact on academic achievement, University of South Alabama, 2019.

[15] S. J E.L.T. Alnujaidi, "The Difference between EFL Students' Preferred Learning Styles and EFL Teachers' Preferred Teaching Styles in Saudi Arabia," **12**(1), 90–97, 2019.

[16] H. J D.S. Demirkan, "An inquiry into the learning-style and knowledge-building preferences of interior architecture students," **44**, 28–51, 2016.

[17] G. J C. Falloon, Education, "Using simulations to teach young students science concepts: An Experiential Learning theoretical analysis," **135**, 138–159, 2019.

[18] O. Embarak, M. Khaleifah, A. Ali, "An Approach to Discover Malicious Online Users in Collaborative Systems," in International Conference on Emerging Internetworking, Data & Web Technologies, Springer: 374–382, 2019.

[19] J. Reynolds, D.R. Adams, R. Ferguson, P. Leidig, "The personality of a computing major: It makes a difference," in 2016 Proceedings of the EDSIG Conference. ISSN, 2473–3857, 2016.

[20] M.Z. Tunio, H. Luo, W. Cong, Z. Fang, A.R. Gilal, A. Abro, S. %J I.A. Wenhua, "Impact of personality on task selection in crowdsourcing software development: A sorting approach," **5**, 18287–18294, 2017.

[21] S.J.L. Foster, M.R. %J I.J. of E.S. Huml, "The relationship between athletic identity and academic major chosen by student-athletes," **10**(6), 915, 2017.

[22] Myers Briggs Assessment - Learn the Power of Personality Preferences.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 860-870 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Load Balancing Techniques in Cloud Computing: Extensive Review

Ahmad AA Alkhatib*, Abeer Alsabbagh, Randa Maraqa, Shadi Alzubi

Alzaytoonah University of Jordan, Amman, 11183, Jordan

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 24 March, 2021
Online: 04 April, 2021

Keywords:
Cloud computing
Load Balancing
Software Defined Network
Static Algorithms
Dynamic Algorithm

ABSTRACT

It has become difficult to handle traditional networks because of extensive network developments and an increase in the number of network users, and also because of new technologies like cloud computing and big data. Traditional networks are experiencing an increase in VM load and in the time taken for processing tasks. Hence, it has become essential to modify the traditional network architecture. A notion called Load balancing techniques that increases the conformance of network management was presented recently to deal with this problem. The critical need for load balancing emerges due to network resources limitations and requirements fulfillment that facilitates traffic distribution through various resources to enhance the efficiency and reliability of network resources. This task has been carried out by several researchers before, who have presented various algorithms with their benefits and shortcomings. The focus of this research is on the notion of cloud computing load balancing and on the advantages and disadvantages of a chosen load balancing algorithm. Furthermore, it examines the metrics and issues of these algorithms.

1 Introduction

Cloud Computing has universally greater interest in web technologies currently. With the increasing demands of the cloud, popular website's servers are getting overloaded, in order to fulfill users requirement load balancing is one of the promising solutions. Load balancing is the procedure of sharing the load between multiple processors in a distributed environment to minimize the turnaround time taken by the servers to cater service requests and make better utilization of the available resources. cloud computing incorporates and enhances the advantages of a few existing distributed computing technologies. Network computing refers to a "model of distributed computing that employs geographically distant resources and hence, allows users to gain access to computers and data and to handle their accounts from varied locations". Virtualization is another technology that hides the physical properties of computing resources to obscure complexity when dealing with applications, systems or end-users [1]–[3].

Cloud computing initially represented online business applications, referred to as application service provision (ASP). Later on the term was used more extensively following the distribution of services by the company, where the name given to each service provider was based on the service they offered to the customer [4].

Since its introduction in late 2006, several definitions of cloud computing have been formulated. It can essentially be defined as a framework for allowing on-demand network access to a shared set of online configurable computing resources (like network, storage, server, services and applications) [5],[6].

The definition of cloud computing provided by HP is as follows: "Everything as a Service" [7]. On the other hand, Microsoft considers cloud computing to signify "Cloud and the Client" [8]. According to T-Systems, cloud computing refers to "renting infrastructure, software and bandwidths under specified service conditions. It should be possible to modify these components routinely on the basis of the customer requirements and should be widely available and secure. Furthermore, there are 2-end service level agreements (SLAs) and use-dependent service invoices that are part of cloud computing" [9].

It is possible to distinguish cloud computing into three models, i.e. SaaS (SOFTWARE–AS-A SERVICE), which provides a user interface to the user which they can access through a browser or desktop application; PaaS (PLATFORM-AS-A SERVICE, which offers applications development tools and a programming language execution environment to the user; and finally IaaS (INFRASTRUCTURE- AS-A-SERVICE), which offers virtual computerized resources to the user that they can manage in accordance with their requirements. These models can also be classified in accordance with the services they offer [10], [11].

- Database as a Service (DaaS)
- Storage as a Service (SaaS)
- Network as a Service (NaaS)

*Corresponding Author: Ahmad AA Alkhatib, Ahmad.Alkhatib@zuj.edu.jo

- Expert as a Service (EaaS)
- Communication as a Service (CaaS)
- Security as a Service (SECaaS)
- Monitoring as a Service (Maas)
- Testing as a Service (TaaS)

The Cloud Computing categories are shown in Figure 1.

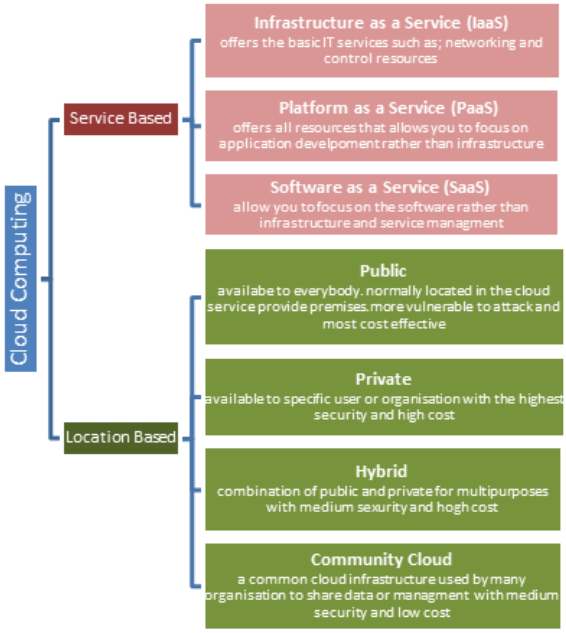


Figure 1: Cloud Computing categories

Cloud computing is becoming quite popular; hence, there has been a significant increase in the amount of processing being carried out in the clouds. However, cloud computing is experiencing several issues in providing the required services [11], including:

- Load Balancing
- Performance analysis and modeling
- Throughput and response time
- Security and privacy
- Resource management
- QoS

The structure of the rest of the paper is as follows: the Load Balancing Concept and Challenges described in Section 2. Section 3 presents the load balancing techniques classification . A few static and dynamic load balancing algorithms are discussed in Section 4. Section 5 Discussion and comparison with the latest studies in this field are compared and Lastly,section 6 gives a conclusion of the study .

2 Load Balancing

There are a series of nodes in the cloud that are connected to one another. Here, one of the nodes is either selected randomly or on the basis of an algorithm to fulfill requests made by users. There may be different data and hence, there is different load on every node, and each node in a cloud can have unequal load of tasks based on the quantity of work requested by clients. Load balancing is a very significant issue that made it imperative for cloud computing to distribute load on the resources available so as to achieve lower response time. This makes sure that no VMs in the system are over-loaded at any time [12]. In addition, balance workload continues to be an issue in the cloud as they are unable to determine the quantity of demand that is obtained in the cloud [13].

Load balancing concepts face several challenges because there are various physical and logical issues that may have an impact on the technique being used. A few of these issues are listed in Table I [11]:

Table 1: LOAD BALANCING ISSUES

Graphical Distribution node	There is geographical distribution of data centers in clouds. On the other hand, distributed nodes are considered as a single system without taking into account various factors like communication delay, networking delay and distances between nodes, resources and users.
Virtual Machine Migration	Multiple VM are allowed by the VM concept on the same Physical Machine. Due to the distinct VM structure, the physical machine may become overloaded.
Algorithm Complexity	To ensure that it does not affect the efficiency of the cloud, load balancing algorithm should be simple and concise.
Heterogeneous nodes	In these times, users have different requirements. It is essential to have heterogeneous nodes to fulfill users' need for services. The load balancing decision is influenced by the heterogeneity of the nodes.
Single point of Failure	In general, load balancing algorithm is typically carried out on a central node to assign tasks. The entire computing fails in case the central node fails.
Load Balancer Scalability	Computing power, topology, storage, etc. determine the response time of load balancing

3 Load Balancing Techniques Classification

The classification of the algorithms depends on the existing status of the system and involves two categories [14], [11]: static algorithms that obtain information pertaining to the system and identify the existing resources for using before commencing. The load is distributed on the existing VM till the work concludes, which is preferably used when the VM capabilities are similar to one another. This type of algorithm has various techniques, for example round

robin, Min-Min, weight, Max-Min, honey bee foraging, throttled, etc. The issue faced with static algorithms is that they depend on static information which is unable to effectively demonstrate dynamic load variations taking place within the VMs [15]. Dynamic algorithms are the second kind of algorithms and are different from static algorithms in the sense that they exhibit flexibility. These algorithms work on the basis of rescheduling the tasks allocated to them over the available VM while carrying out the work. Better performance is exhibited by static algorithms in contrast to the dynamic algorithm, whereas with respect to the competitive ratio, dynamic algorithms have a larger ratio compared to the static algorithm [16]. The typical structure of load balancing in cloud environments is demonstrated in Figure 2, whereas Figure 3 shows the classification of the load balancing methods. Several parameters are used to assess load balancing algorithms, which are regulated by various policies. The metrics are presented in Table 3 [15], [17], while an outline of the load balancing policies is given in Table 4.

Table 2: Load Balancing Metrics

Metrics	Description
Performance	It refers to ensuring the consistency and efficiency of the algorithms carried out
Response time	It refers to the time taken by the system to give a reaction to the user request. he more rapid the response time, the higher the satisfaction obtained by the user.
Throughput	The number of tasks carried out for each unit of time.
Scalability	The ability of the algorithm to handle the tasks of the user in an efficient and effective way.
Fault Tolerance	The robustness of the algorithm to solve issues and errors so that its performance can be regained
Migration Time	The time needed to shift the load from a certain VM to another in accordance with the VM loading during situations of overload or under-load.
Resource Usage	Determines the optimum use of computing resources in the data center that should be employed by the algorithm
Makespan	Time needed to carry out and treat a given set of tasks

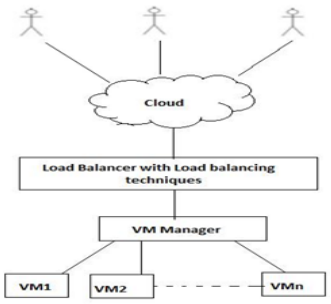


Figure 2: General Structure of Load balancing in Cloud Environment [18]

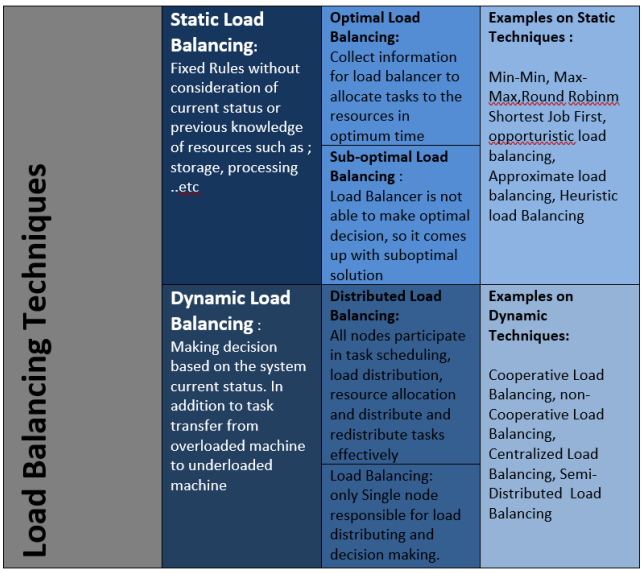


Figure 3: Load balancing techniques classification

Table 3: Load Balancing Policies

Selection Policy	The required tasks that need to be transferred from a certain VM to another are specified in this policy. This is done depending on the extent of overhead that is to be transferred
Location policy	In this policy, the tasks are sent to the free, under-loaded and available VMs so that they can be fulfilled. The desired VM is determined in this policy on the basis of the availability of the required services so that the task can be transferred in accordance with the available technique, like Negotiation, Random and Probing.
Transfer policy	In this policy, the conditions for shifting the task from a local VM to another local or remote VM are determined. Two types of tasks are used in this regard: the existing tasks and the last task received.
Information policy	This policy deals with keeping information pertaining to the resources safe in the system to ensure that other policies can benefit from them while making decisions.

4 Load Balancing Common Algorithms

A few static and dynamic load balancing algorithms will be presented in this section, with benefits and drawbacks for each algorithm:

4.1 Round Robin Algorithm

This static algorithm is one of the simplest methods used following the selection of the existing VMs. One of these VMs is randomly chosen by the data center unit to commence its operations. In addition, it is organized in a circular manner. After this, every VM that gets a request is shifted towards the final part of the list [19], [20][21],.

There is, however, no interaction between this algorithm and the distinct abilities of the VMs because the tasks are equally divided on the VMs, even though they have distinct abilities (Figure 4) [15]. Hence, Weight Round Robin algorithm is used to enhance this algorithm, which assigns weight for each VM on the basis of its abilities, and then offers the ability to each node to have a specific number of tasks on the basis of the weights allocated to each VM. This algorithm is considered similar to Round Robin algorithm from the perspective of time division because it distributes the time as a circular. However, the distinction is that the tasks are offered to VM on the basis of particular restrictions, for example checking weights [22], [23].

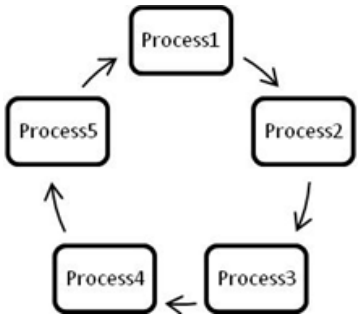


Figure 4: Round Robin Load Balancer [19]

4.2 Throttled Load Balancer (TLB)

A table is generated in this algorithm that includes the virtual machines as well as the existing state (available/busy). If a specific task is allocated to a virtual machine, a request is made to the control unit within the data center, which will look for the ideal VM suit with respect to their abilities to achieve the required task [24]. The load balancer will send -1 back to the data center if an appropriate VM is not available [15]. Figure 5 presents a demonstration of Throttled Load Balancer.

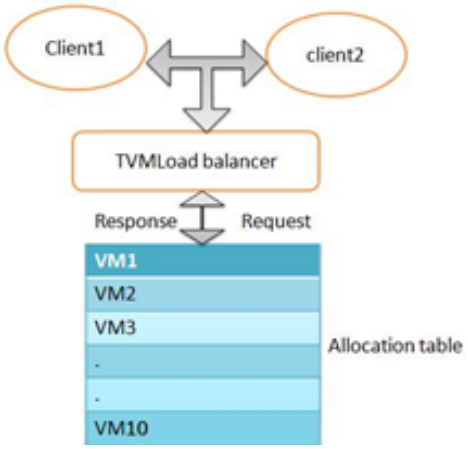


Figure 5: Throttled Load Balancer

In contrast, the process of looking for the ideal virtual machines always takes place from the start of the table each time; therefore, certain VMs are not employed. A modified throttled algorithm was put forward by [25], which works by changing the process of VM selection. After receiving the subsequent request, the VM at index adjacent to the VM already allocated is selected based on the state of the VM. In addition, [22] presented an efficient throttled algorithm, which consisted of 3 algorithms, i.e. throttled algorithm, ESCE (Equally Spread Current Execution algorithm) and Round Robin. This algorithm is considered as an advancement of the main throttled algorithm, where the enhanced algorithm employed the data structure to maintain information regarding the VMs. In contrast, Hash Map index is used to search for VM and allocated tasks where it works at a faster pace compared to the throttled algorithm.

Furthermore, Divide-and-Conquer and Throttled algorithm (DCBT) approach is taken to be the hybrid approach because it gives preference to the order when allocating the VM. This is done to achieve the highest usage of resources by reducing the overall time taken to execute the task. In addition, this algorithm seeks to update the table by dividing the independent jobs equally among all the VMs. Hence, it works faster than the throttled algorithm [26].

4.3 Min-Min Load Balancing Algorithm

This algorithm is easy to use and works at a faster pace [27]. In addition, it improves performance and consists of a series of tasks. The time taken to execute the task is computed and allocated to VMs on the basis of the smallest completion time for the existing tasks. The process will continue till it is ensured that each task has been allocated to VM [16]. Because of the existence of a greater number of smaller tasks, this algorithm performs better compared to if there were bigger tasks. However, this will lead to starvation because of giving priority to smaller tasks and deferring the bigger tasks [21].

4.4 Max Min Load Balancing Algorithm

As stated by [28], [29], this algorithm is quite similar to the Min-Min Load Balancing, based on the calculation time. In this algorithm, all existing tasks are sent to the system, after which the calculation is carried out for determining the least time to complete each of the given tasks. The selected task then has the maximum time to be completed will be allocated to the relevant machine [11]. A comparison of the performance of this algorithm with the Min-Min algorithm shows that the Max-Min algorithm is better because there is just one large task in the set, which means that the Max-Min algorithm will carry out the shorter tasks alongside the larger task [15].

4.5 Opportunistic Load Balancing Algorithm

This algorithm is a type of static algorithm that is not capable of describing the existing workload of the VM; therefore, it allocates the tasks randomly to all nodes in the system to ensure they are all working [30], [31], and [32]. Tasks are accomplished at a slow pace through this algorithm as it does not compute the existing implementation time [33]. Hence, it provides incorrect outcomes for the load balance [15].

4.6 Ant Colony Optimization

This algorithm that is based on ACO allows for distributing the workload efficiently between the nodes of a cloud. The working of this algorithm is influenced by the Ant concept, where it looks for a new pathway if it comes across an obstacle and allocates a new route between the nodes [34]. The algorithm functions by first initializing the table, carrying out the data flow and achieving the threshold level for the required nodes. When the flow passes through the nodes, the algorithm examines the node; if it is under load, then it uses the highest trailing pheromone (TP) that provides the route for the under-loaded node. The table is then updated till the node achieves the threshold limit. Nonetheless, if the threshold limit is attained, it uses the Foraging Pheromone (FP) to examine new sources of food and update the table till it attains an under-load, after which it reassigns resources. This cycle keeps on occurring till the completion of the process.

4.7 Honey Bee Algorithm

The working of this algorithm is influenced by the way bees behave when looking for honey, because its main objective is to divide the workload on the VM, considering the lack of excessive resource utilization and lack of under-usage of resources. This algorithm works by choosing a VM that fulfills two main requirements. Fewer tasks are allocated to this VM compared to those assigned to other VM machines. The time taken by VM to process falls within the average processing time taken by all other VMs [35], [23].

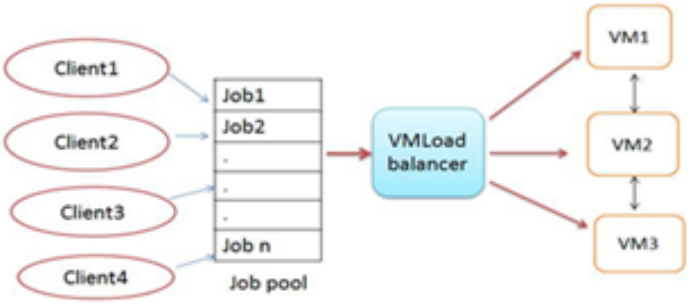


Figure 6: Active Monitoring Load Balancerr

4.8 Active Monitoring Load Balancer (AMLB)

It is a type of dynamic load technology [21]. This technology obtains information relevant to each VM and to the number of requests that are presently allocated to each of them [24]. The Data Center Controller (DCC) scans the VM index table after receiving a new request to determine the VM that is least loaded or idle. First-come-first serve concept is employed by this algorithm to allocate load to the VM that has the smallest index number for more than two servers [21].

The VM ID is sent back by the AMLB algorithm to the DCC which then sends the request to the VM represented by that ID. The AMLB is informed about the new allocation by the DCC and it is sent the cloudlet [24].

Once the task is completed, the information is sent to the DCC and the VM index table is reduced. When a new request is received, it goes over the table again using load balancer and then the process allocation occurs. Figure 6 presents an illustration of AMLB.

4.9 Genetic Algorithm

In the process of genetic algorithms, scheduling takes place on the basis of the biological notion of population generation. The purpose of using this algorithm was to enhance the distribution of load within the cloud, where the algorithm starts functioning using the preliminary population procedure. The initial population consists of the set of all individuals who are involved in determining the optimal solution. Each solution in the population is referred to as an individual, and each individual is described as a chromosome to make it appropriate for carrying out genetic operations [36]. It was presumed by the authors in [37] that a comparison of the newly developed population will be carried out with the previous one. The solutions are then chosen to obtain solutions (offspring) on the basis of its fitness function [38]. The fitness function is used to determine the quality of individuals in the population with respect to the specified optimization goal [37].

For every chromosome, the fitness function is computed, after which the most appropriate results are chosen to be used as parents. The operation then commences with a crossover process; a part of each parent is used to create a new child; this child is then improved by employing a mutation process. This process keeps on occurring till the best results are obtained [39]. The basis of genetic algorithms is randomness; however, it is not the same as the random search in that it approves the most suitable individuals within a population. The crossover rate and mutation probability values have a significant impact on the performance of the genetic algorithm [38].

The general scheme of the genetic algorithm is shown in Figure 7.

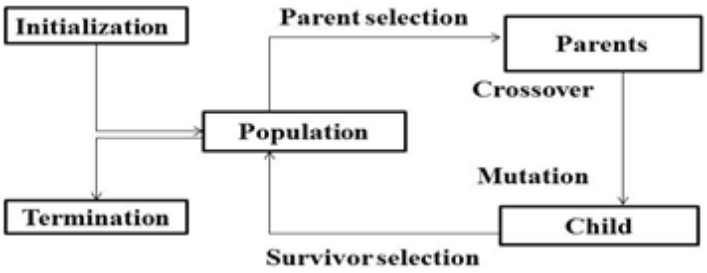


Figure 7: General scheme of genetic algorithm [40]

4.10 First Come First Serve

This algorithm works by distributing new tasks to resources that have the least waiting time (i.e. resources that have the least number of tasks). Here, there is sequential execution of the tasks, with the work commencing from the first task till its completion, after which the subsequent task from the queue is carried out. This algorithm is used as it allocates tasks to the virtual machines without considering the specifications of the virtual machines available and also the time that will be taken by the tasks in queue when allocating new tasks [41].

4.11 Generalized Priority Algorithm

The mechanism of action of this algorithm is to give priority to the tasks according to their size (Higher missions size take higher priority), and the default machines are given priority based on the power of the processor, then the appropriate virtual machine is chosen for the priority of the required task [41]. An example on GPA suppose you have 6 VMs represented by their Id and processor speed VMs = 0, 850 , 1, 1000, [5], 3, 150, 4,250, 5,750. Here the priority goes to VM 4 will because it has the highest speed, then priority is given to VM 2 and then VM 6 and so on [41]. This algorithm works as in Figure (pseudo):

```
Pseudo code:
For each Datacenter create a VM list
   Find cost processor, processing speed, memory and storage of server (VM)
End loop
For each
Find length of array
          Sort cloudlet array where computational power is key

For each VM find index in LB table where VM allocation =0
Fill array cloudlet Where array length = MIPS count
Find VM for suitable cloudlet
Find matched VM id
End loop
End loop
Update Datacenter list
Update VM list
```

Figure 8: K. Generalized Priority Algorithm Pseudo code

4.12 NBST Algorithm

Initially, this algorithm assumes that the number of instructions required for the tasks is on the waiting list and their length is known, then it arranges all available virtual devices according to their implementation speed MIPS (Million instructions per second) and arranging the cloudlets according to their length. Where this algorithm depends on dividing the number of VM and cloudlets into two parts after arranging them in descending order until access to the VM or cloudlets one at most, then assigning the VM groups to the cloudlets groups [42].

4.13 Load Balancing Technique based on Cuckoo Search and Firefly

The cuckoo search is based on the brooding behavior of cuckoo bird, where this bird laid their eggs on other bird's nest, the best quality eggs are carried forward to the next generation and the worst nests are abandoned [43]. Cuckoo Search with Firefly is a hybrid algorithm is used for loud balancing in a cloud computing environment; this algorithm schedules the tasks and allocates the overloaded VMs tasks to the under-loaded VMs. This algorithm finds the best VM in less time and improves the loud balance efficiency and avoids the task imbalanced situations in the entire system [44]. The LB technique starts with scheduling the tasks using Round Robin method, then calculate the capacity and the load of each VM, then it identifies the overloaded and under-loaded tasks using Cuckoo Search with Firefly to start the migration of tasks from overloaded VMs to under-loaded VMs [44]. It avoids the task imbalanced situations in the entire system. Also this method has the advantages of high interchange rate and less number of tuning parameter. This method has the advantages of high interchange rate and less number of tuning parameter. The proposed method migrates 2 task, while existing method requires 7 task for migration.

4.14 predicted load balancing algorithm based on the dynamic exponential smoothing

This algorithm predicts the load of the VMs by using the dynamic exponential smoothing model. Where the dynamic balancing algorithms focus on the fastest response speed algorithm and the least connection algorithm [45].

Exponential smoothing is one of prediction algorithms based on time series, which takes advantage of all historical data, and differentiates them through the smoothing factor to allow recent data make a greater impact on the analytical value than long-term data, [46].

The allocated node has the shortest corresponding time, when a new request arrives to the server node that has a little connection numbers, the dynamic algorithm takes the load characteristics of the server node as the reference, makes the adjustment with every weighted factor, and reflects the real-time load under the support of weighted value, by that the algorithm helps to find the corresponding smoothing coefficient with the VM load time series of current phrase, and helps to make prediction with the load value at the next moment of this VM[45]. The load predicted balancing algorithm developed by the improvement that permits to construct the dynamic smoothing exponent and to obtain the load prediction with a higher accuracy through the comparison of short load time series. Also it helps to provide more accurate service demand for users and enhance the resource utilization ratio of the server node to a higher level [45]. The accuracy of the model prediction can be affected, because of the selection standard of α exponential smoothing is not very clear [45].

5 Discussion

A comparison of the load balancing algorithms is presented in this section. Each load balancing algorithm is described in specific identifiers with its main properties and limitations.

5.1 Round Robin

- *Algorithm_Name: [Round Robin].*
- *Algorithm_Type: [Static].*
- *Algorithm_Overhead: [No Overhead].*
- *Algorithm_Degree_of_complexity: [Simple].*
- *Algorithm_Strength_points :*
  - *Simplicity*
  - *Easy to install*
- *Algorithm_Limitations:*
  - *Cannot be improved any more*
  - *No multi-tasking capabilities*

  - *Could lead to over head*
  - *Does not take capacity importance and task duration into account*
- *Algorithm_Performance: [Low Performance].*
- *Algorithm_Throughput: [Low Throughput].*
- *Algorithm_Fault_tolerance: [No FT].*

5.2 Throttled

- *Algorithm_Name: [Throttled].*
- *Algorithm_Type: [Static].*
- *Algorithm_Overhead: [No Overhead].*
- *Algorithm_Degree_of_complexity: [Medium].*
- *Algorithm_Strength_points :*
  - *The data center searches for the best VM that fits its capabilities with the task required*
  - *Maximization of resource usage*
  - *Increase average of execution time*
- *Algorithm_Limitations:*
  - *Does not define time limits*
  - *Could result in idle VMs*
- *Algorithm_Performance: [medium Performance].*
- *Algorithm_Throughput: [high Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.3 Min-Min

- *Algorithm_Name: [Min-Min].*
- *Algorithm_Type: [Static].*
- *Algorithm_Overhead: [Medium Overhead].*
- *Algorithm_Degree_of_complexity: [Simple].*
- *Algorithm_Strength_points :*
  - *Quick*
  - *Easy to install*
- *Algorithm_Limitations:*
  - *Leads to starvation*
- *Algorithm_Performance: [Increased performance for smaller tasks].*
- *Algorithm_Throughput: [Improved Throughput].*
- *Algorithm_Fault_tolerance: [No FT].*

5.4 Max-Min

- *Algorithm_Name: [Max-Min].*
- *Algorithm_Type: [Static].*
- *Algorithm_Overhead: [Medium Overhead].*
- *Algorithm_Degree_of_complexity: [Simple].*
- *Algorithm_Strength_points :*
  - *Execute the large tasks in parallel with short tasks*
- *Algorithm_Limitations:*
  - *Leads to starvation.*
- *Algorithm_Performance: [Better performance than min-min].*
- *Algorithm_Throughput: [Improved Throughput].*
- *Algorithm_Fault_tolerance: [No FT].*

5.5 Opportunistic

- *Algorithm_Name: [Opportunistic].*
- *Algorithm_Type: [Static].*
- *Algorithm_Overhead: [Low Overhead].*
- *Algorithm_Degree_of_complexity: [Simple].*
- *Algorithm_Strength_points :*
  - *The advantage is quite simple and reach load balance*
- *Algorithm_Limitations:*
  - *No fairness in task assignment*
  - *Tasks are planed slowly because it does not determine the nodes current execution time*
  - *the whole completion time (Make span) is very poo*
- *Algorithm_Performance: [poor Performance].*
- *Algorithm_Throughput: [Limited Throughput].*
- *Algorithm_Fault_tolerance: [No FT].*

5.6 Ant Colony Optimization

- *Algorithm_Name: [Ant Colony Optimization].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [High Overhead].*
- *Algorithm_Degree_of_complexity: [High complexity].*
- *Algorithm_Strength_points :*
  - *The style of living ants depends on the*
  - *transition from VM to another VM.*
  - *Better performance*
  - *Reduce operation time*
  - *Can be combined with heuristic search algorithm*
- *Algorithm_Limitations:*
  - *Complex and load on the system unit.*
  - *Untrusted in reality*
- *Algorithm_Performance: [High Performance].*
- *Algorithm_Throughput: [High Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.7 Honey Bee

- *Algorithm_Name: [Honey Bee].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [High Overhead].*
- *Algorithm_Degree_of_complexity: [High complexity].*
- *Algorithm_Strength_points :*
  - *it is chosen the best VM by comparing the cost of executing a task on one VM with the other available VMs.*
  - *Takes into account excessive resources and under used ones*
  - *Number of assigned tasks less than number of handled by other VMs to reduce execution time*
- *Algorithm_Limitations:*
  - *Complexity*
  - *Increase the machine size does not increase throughput*
- *Algorithm_Performance: [High Performance].*
- *Algorithm_Throughput: [Low Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.8 AMLB

- *Algorithm_Name: [AMLB].*
- *Algorithm_Type: [Static].*
- *Algorithm_Overhead: [High Overhead].*
- *Algorithm_Degree_of_complexity: [medium complexity].*
- *Algorithm_Strength_points :*
  - *It identifies the least loaded VM*
  - *Increase performance*
  - *Update index table every time task executed*
- *Algorithm_Limitations:*
  - *Time consuming in index table calculation*
  - *algorithm gives better results when there is low variation in workload*
- *Algorithm_Performance: [High Performance].*
- *Algorithm_Throughput: [High Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.9 Genetic Algorithm

- *Algorithm_Name: [Genetic Algorithm].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [High Overhead].*
- *Algorithm_Degree_of_complexity: [High complexity].*
- *Algorithm_Strength_points :*
  - *Executed the cloudlets in less time*
  - *Improve load distribution in the cloud*
  - *Better resource usage*
  - *Better approve the best fitted population indexing in VMs*
- *Algorithm_Limitations:*
  - *Complexity*
  - *High resources required for population generation and calculation*
  - *Crossover mutation is highly effected by probability parameters*
- *Algorithm_Performance: [High Performance].*
- *Algorithm_Throughput: [High Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.10 First come first serve

- *Algorithm_Name: [First come first serve].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [There is Overhead].*
- *Algorithm_Degree_of_complexity: [Simple complexity].*
- *Algorithm_Strength_points :*
  - *Distribute tasks to all available VMs*
- *Algorithm_Limitations:*
  - *Not preventive*
  - *assigning tasks to all virtual machines without paying attention to the specifications of the available virtual machine*
  - *not paying attention to the time that the tasks in the queue will take when distributing the new tasks*
- *Algorithm_Performance: [Medium Performance].*
- *Algorithm_Throughput: [Medium Throughput].*
- *Algorithm_Fault_tolerance: [No FT].*

5.11 Generalized Priority Algorithm

- *Algorithm_Name: [Generalized Priority Algorithm].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [No Overhead].*
- *Algorithm_Degree_of_complexity: [medium complexity].*
- *Algorithm_Strength_points :*
  - *give priority to the tasks according to their size and the machine processor power*
- *Algorithm_Limitations:*
  - *take more execution time*
  - *could lead to starvation*
- *Algorithm_Performance: [Increased Performance].*
- *Algorithm_Throughput: [High Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.12 NBST Algorithm

- *Algorithm_Name: [NBST Algorithm].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [Yes the is Overhead].*
- *Algorithm_Degree_of_complexity: [medium complexity].*
- *Algorithm_Strength_points :*
  - *Distribute all task on all available VM*
- *Algorithm_Limitations:*
  - *take more time in an assign VMs for cloudlets*
- *Algorithm_Performance: [increased Performance].*
- *Algorithm_Throughput: [High Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.13 Load Balancing based on CSF

- *Algorithm_Name: [Load Balancing based on CSF].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [medium Overhead].*
- *Algorithm_Degree_of_complexity: [medium complexity].*
- *Algorithm_Strength_points :*
  - *schedules the tasks and allocates the overloaded VMs tasks to the under-loaded VMs*
  - *avoids the task imbalanced situations in the entire system*
  - *high interchange rate and less number of tuning parameter*
- *Algorithm_Limitations:*
  - *multiple algorithms have to be followed*
- *Algorithm_Performance: [Improved Performance].*
- *Algorithm_Throughput: [High Throughput].*
- *Algorithm_Fault_tolerance: [FT].*

5.14 predicted load balancing algorithm based on the dynamic exponential smoothing

- *Algorithm_Name: [predicted load balancing algorithm based on the dynamic exponential smoothing].*
- *Algorithm_Type: [Dynamic].*
- *Algorithm_Overhead: [Yes there is Overhead].*
- *Algorithm_Degree_of_complexity: [High complexity].*

- *Algorithm_Strength_points :*
  - *provide more accurate service demand for users and improve the resource utilization of the VMs*
- *Algorithm_Limitations:*
  - *the accuracy of the model prediction can be affected, because of the selection standard of α exponential smoothing is not very clear*
- *Algorithm_Performance: [Improved Performance].*
- *Algorithm_Throughput: [Medium Throughput].*
- *Algorithm_Fault_tolerance: [No FT].*

6 Conclusion and Future Work

This paper presents an extensive Review for load balancing that carry out load distribution among the VM in a different way. considering the lack of excessive resource utilization and not keeping the VMs idle. As a service that is carried out over the network is known as cloud computing, where a lot of significance is given to load balancing issues. The performance will decrease with an overloaded system. Hence, smart load balancing algorithm is needed to maintain the position of QoS. a description for algorithm techniques has been explained and followed by a comparison between them.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The Authors Thanks The Alzaytoonah University of Jordan for their support.

References

[1] A. A. AlKhatib, T. Sawalha, S. AlZu'bi, "Load Balancing Techniques in Software-Defined Cloud Computing: an overview," in 2020 Seventh International Conference on Software Defined Systems (SDS), 240–244, IEEE, 2020,doi: 10.1109/SDS49854.2020.9143874.

[2] N. A. Sultan, "Reaching for the "cloud": How SMEs can manage," International journal of information management, **31**(3), 272–278, 2011,doi: https://doi.org/10.1016/j.ijinfomgt.2010.08.001.

[3] S. AlZu'bi, Y. Jararweh, H. Al-Zoubi, M. Elbes, T. Kanan, B. Gupta, "Multi-orientation geometric medical volumes segmentation using 3d multiresolution analysis," Multimedia Tools and Applications, **78**(17), 24223–24248, 2019, doi:https://doi.org/10.1007/s11042-018-7003-4.

[4] A. Susarla, A. Barua, A. B. Whinston, "Understanding the 'service'component of application service provision: an empirical analysis of satisfaction with ASP services," in Information Systems Outsourcing, 481–521, Springer, 2006, doi: https://doi.org/10.1007/978-3-540-34877-1_17.

[5] G. Lu, W. H. Zeng, "Cloud computing survey," in Applied Mechanics and Materials, volume 530, 650–661, Trans Tech Publ, 2014, doi: DOI: 10.4028/www.scientific.net/AMM.530-531.650.

[6] A. Abusukhon, M. N. Anwar, Z. Mohammad, B. Alghannam, "A hybrid network security algorithm based on Diffie Hellman and Text-to-Image Encryption algorithm," Journal of Discrete Mathematical Sciences and Cryptography, **22**(1), 65–81, 2019, https://doi.org/10.1080/09720529.2019.1569821.

[7] S. Robison, "Everything-as-a-Service: A blue sky view of the cloud," 2009.

[8] L. Xin, C. Song, "Cloud-based innovation of Internet long tail," in 2011 International Conference on Product Innovation Management (ICPIM 2011), 603–607, IEEE, 2011, doi: https://doi.org/10.1007/s11042-019-7367-0.

[9] R. F. El-Gazzar, "A literature review on cloud computing adoption issues in enterprises," in International Working Conference on Transfer and Diffusion of IT, 214–242, Springer, 2014, doi: 10.1007/978-3-662-43459-8_14.

[10] V. Paul, S. Pandita, M. Randiva, "CLOUD COMPUTING REVIEW," 2018.

[11] P. Kumar, R. Kumar, "Issues and challenges of load balancing techniques in cloud computing: a survey," ACM Computing Surveys (CSUR), **51**(6), 1–35, 2019, doi:https://doi.org/10.1145/3281010.

[12] E. J. Ghomi, A. M. Rahmani, N. N. Qader, "Load-balancing algorithms in cloud computing: A survey," Journal of Network and Computer Applications, **88**, 50–71, 2017, doi: https://doi.org/10.1016/j.jnca.2017.04.007.

[13] M. R. Mesbahi, M. Hashemi, A. M. Rahmani, "Performance evaluation and analysis of load balancing algorithms in cloud computing environments," in 2016 Second International Conference on Web Research (ICWR), 145–151, IEEE, 2016, doi:10.1109/ICWR.2016.7498459.

[14] K. Al Nuaimi, N. Mohamed, M. Al Nuaimi, J. Al-Jaroodi, "A survey of load balancing in cloud computing: Challenges and algorithms," in 2012 second symposium on network cloud computing and applications, 137–142, IEEE, 2012, doi:10.1109/NCCA.2012.29.

[15] S. K. Upadhyay, A. Bhattacharya, S. Arya, T. Singh, "Load optimization in cloud computing using clustering: a survey," Int. Res. J. Eng. Technol, **5**(4), 2455–2459, 2018.

[16] L. Xu, M. Xu, J. Wang, R. Semmes, Q. Wang, H. Mu, S. Gui, H. Yu, W. Tian, R. Buyya, "A Data-driven Approach to Identify Resource Bottlenecks for Multiple Service Interactions in Cloud Computing Environments," .

[17] S. Afzal, G. Kavitha, "Load balancing in cloud computing–A hierarchical taxonomical classification," Journal of Cloud Computing, **8**(1), 22, 2019, doi: https://doi.org/10.1186/s13677-019-0146-7.

[18] M. A. Shahid, N. Islam, M. M. Alam, M. M. Su'ud, S. Musa, "A Comprehensive Study of Load Balancing Approaches in the Cloud Computing Environment and a Novel Fault Tolerance Approach," IEEE Access, **8**, 130500–130526, 2020, doi:10.1109/ACCESS.2020.3009184.

[19] R. Sajjan, B. R. Yashwantrao, "Load balancing and its algorithms in cloud computing: a survey," International Journal of Computer Sciences and Engineering, **5**(1), 95–100, 2017.

[20] M. M. D. Shah, M. Kariyani, M. Agrawal, "Allocation of virtual machines in cloud computing using load balancing algorithm," International Journal of Computer Science and Information Technology & Security (IJCSITS), **3**(1), 2249–9555, 2013.

[21] N. R. Tadapaneni, "A Survey Of Various Load Balancing Algorithms In Cloud Computing," 2020.

[22] G. Megharaj, K. Mohan, "A survey on load balancing techniques in cloud computing," IOSR Journal of Computer Engineering (IOSR-JCE), **18**(2), 55–61, 2016, doi: 10.9790/0661-18215561.

[23] V. Mathur, "A Comparative Study of Load Balancing Techniques in Distributed Systems," .

[24] J. James, B. Verma, "Efficient VM load balancing algorithm for a cloud computing environment," International Journal on Computer Science and Engineering, **4**(9), 1658, 2012.

[25] S. G. Domanal, G. R. M. Reddy, "Load balancing in cloud computing using modified throttled algorithm," in 2013 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), 1–5, IEEE, 2013, doi: 10.1109/CCEM.2013.6684434.

[26] S. G. Domanal, G. R. M. Reddy, "Load balancing in cloud environment using a novel hybrid scheduling algorithm," in 2015 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), 37–42, IEEE, 2015, doi : 10.1109/CCEM.2015.31.

[27] V. Arulkumar, N. Bhalaji, "Performance analysis of nature inspired load balancing algorithm in cloud environment," Journal of Ambient Intelligence and Humanized Computing, 1–8, 2020, doi: https://doi.org/10.1007/s12652-019-01655-x.

[28] S. Kaur, T. Sharma, "Efficient load balancing using improved central load balancing technique," in 2018 2nd International Conference on Inventive Systems and Control (ICISC), 1–5, IEEE, 2018, doi: 10.1109/ICISC.2018.8398857.

[29] D. Saranya, L. S. Maheswari, "Load balancing algorithms in cloud computing: a review," International Journal of Advanced Research in Computer Science and Software Engineering, **5**(7), 1107–1111, 2015.

[30] A. A. Prakash, V. Arul, A. Jagannathan, "A look at of efficient and more suitable load balancing algorithms in cloud computing," Int. J. Eng. Res. Comput. Sci. Eng., **5**(4), 7, 2018.

[31] P. Prajapati, A. K. Sariya, "A review: Methods of load balancing on cloud computing," Int. J. Res. Anal. Rev., **6**, 2019, doi: 10.6084/m9.doi.one.IJRAR19J2346.

[32] S.-C. Wang, K.-Q. Yan, W.-P. Liao, S.-S. Wang, "Towards a load balancing in a three-level cloud computing network," in 2010 3rd international conference on computer science and information technology, volume 1, 108–113, IEEE, 2010, doi:10.1109/ICCSIT.2010.5563889.

[33] W. Duan, X. Tang, J. Zhou, J. Wang, G. Zhou, "Load balancing opportunistic routing for cognitive radio ad hoc networks," Wireless Communications and Mobile Computing, **2018**, 2018, doi: https://doi.org/10.1155/2018/9412782.

[34] K. Nishant, P. Sharma, V. Krishna, C. Gupta, K. P. Singh, R. Rastogi, et al., "Load balancing of nodes in cloud using ant colony optimization," in 2012 UKSim 14th international conference on computer modelling and simulation, 3–8, IEEE, 2012, doi: 10.1109/UKSim.2012.11.

[35] W. Hashem, H. Nashaat, R. Rizk, "Honey bee based load balancing in cloud computing." KSII Transactions on Internet & Information Systems, **11**(12, doi: 10.3837/tiis.2017.12.001), 2017.

[36] J. Lim, D. Lee, "A Load Balancing Algorithm for Mobile Devices in Edge Cloud Computing Environments," Electronics, **9**(4), 686, 2020, doi:https://doi.org/10.3390/electronics9040686.

[37] P. Kumar, A. Verma, "Scheduling using improved genetic algorithm in cloud computing for independent tasks," in Proceedings of the international conference on advances in computing, communications and informatics, 137–142, 2012, doi: 10.1145/2345396.2345420.

[38] J. AnantKumar, S. PremChand, "Multi-Agent Genetic Algorithm for Efficient Load Balancing in Cloud Computing," in International Journal of Innovative Technology and Exploring Engineering (IJITEE), 45–51, 2020.

[39] R. Panwar, B. Mallick, "Load balancing in cloud computing using dynamic load management algorithm," in 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), 773–778, IEEE, 2015, doi:10.1109/ICGCIoT.2015.7380567.

[40] G. Joshi, S. K Verma, "Load balancing approach in cloud computing using improvised genetic algorithm: a soft computing approach," in IJCA, volume 122, 24–28, 2015, doi: 10.5120/21729-4894.

[41] D. Agarwal, S. Jain, et al., "Efficient optimal algorithm of task scheduling in cloud computing environment," arXiv preprint arXiv:1404.2076, 2014, doi:10.14445/22312803/IJCTT-V9P163.

[42] S. Sidana, N. Tiwari, A. Gupta, I. S. Kushwaha, "NBST algorithm: A load balancing algorithm in cloud computing," in 2016 International Conference on Computing, Communication and Automation (ICCCA), 1178–1181, IEEE, 2016. doi: 10.1109/CCAA.2016.7813914.

[43] M. Yakhchi, S. M. Ghafari, S. Yakhchi, M. Fazeli, A. Patooghi, "Proposing a load balancing method based on Cuckoo Optimization Algorithm for energy management in cloud computing infrastructures," in 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), 1–5, IEEE, 2015, doi: 10.1109/ICMSAO.2015.7152209.

[44] K. Kumar, T. Ragunathan, D. Vasumathi, P. Prasad, "An Efficient Load Balancing Technique based on Cuckoo Search and Firefly Algorithm in Cloud," Int. J. Intell. Eng. Syst, **13**(3), 422–432, 2020, doi: 10.22266/ijies2020.0630.38.

[45] L. Yan, X. Liu, "The predicted load balancing algorithm based on the dynamic exponential smoothing," Open Physics, **18**(1), 439–447, 2020, doi: 10.1515/phys-2020-0108.

[46] X. Ren, R. Lin, H. Zou, "A dynamic load balancing strategy for cloud computing platform based on exponential smoothing forecast," in 2011 IEEE international conference on cloud computing and intelligence systems, 220–224, IEEE, 2011, doi: 10.1109/CCIS.2011.6045063.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 862-870 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Accounting Software in Modern Business

Lesia Marushchak[1,*], Olha Pavlykivska[1], Galyna Liakhovych[2], Oksana Vakun[2], Nataliia Shveda[3]

[1]*Department of Accounting and Audit, Ternopil Ivan Puluj National Technical University, Ternopil, 46001, Ukraine*

[2]*Department of Accounting and Finance, Ivano-Frankivsk Education and Research Institute of Management Ternopil National Economic University, 76000, Ukraine*

[3]*Department Management and Administration Department, Ternopil Ivan Puluj National Technical University, 46001, Ukraine*

ARTICLE INFO

Article history:
Received: 06 November, 2020
Accepted: 19 January, 2021
Online: 05 February, 2021

Keywords:
Accounting software
Leading companies
Accounting system
Business performance

ABSTRACT

The purpose of the research is an investigation of different accounting software products, their functions, and specific features to make easier choice among variety of similar products and analysis of their pros and cons that can influence on companies' performance. Authors classified accounting software according to its capabilities to serve the different managerial purposes. Because accounting software contains hundreds, some of them even thousands of features, the grouping method gave a possibility to assort similar models that might suit the company's specific requirements – size, cost, customizing, formats, appointments, models, and providers. Observation and comparing of data showed that the cost of accounting programs is critical to making the right choice. As the global accounting software market has a tendency to abrupt change to e-accounting, so that makes it impossible to predict the future behavior of accounting software users. To determine the objectives of this research statistical procedures are conducted. Received results can help potential users of accounting software products to choose the appropriate one based on listed advantages and disadvantages among the best sellers – customization tools, foreign currencies handling, financial and managerial reporting system and analytical capabilities. Lack of prior research studies on the topic and lack of available data have caused significant limitation of the analysis scope. The obtained results gave possibility to identified the main elements in formation the list of features necessary for making right choice of accounting software products. Facts showed managers, who don't consider specific needs and features of accounting software, encounter with problem of discrepancy to company's requirements.
The research is based on theoretical and empirical data. To collect the necessary data for research there was used a quantitative approach. Analytical method helped to analyze and evaluate the ponderable factors which must be considered in selecting process the most appropriate accounting software for companies. The research is dedicated to problems connected with an uncertainty that appears in the accounting software market. This research adds new knowledge to the accounting field as there was disproving theoretical and practical knowledge about accounting software.

1. Introduction

Accounting software has become an integral part of all types of business. Nowadays accounting and financial management became the scientific tools in running the business where software programs help to manage finance more effectively. Companies also spread their activities on international markets to get their shares and as a result, there is an increased demand for new and more sophisticated accounting software packages capable for handling international accounting issues [1]. There are more than 150 well-identified software products around the world and most of them are targeted on large business. But the main stress must be made especially on small business as companies this size are most unprotected and cannot finance much in accounting programs. There are a number of vendors which serve this share

*Corresponding Author: Lesia Marushchak, lesyamar@ukr.net

of market [2], but there is a very short list of scientific researchers conducted on this theme which would help business owners make the right and cost suitable choice.

As accounting systems are connected with other management platforms such as inventory management, warehouse management, order management, customer relationship management, enterprise resource planning, sales and production planning, enterprise quality management, supply chain and distribution systems, accounting software must be more than the simple programs for record and storage accounting information. Except this point, when it comes to select the right accounting product, the choices can be contradictory how to choose necessary speed, accuracy and reliability. The relationship between accounting software and company performance might be distinguished only in case when the accounting software program is suitable for some special company features. It also should be stressed that many specialists are involved in the accounting process. Many conducted researches showed that accounting information system adoption really influenced effectively on companies' performance, profitability and operations efficiency in some countries: Malaysia, Finland, Spain, Iran and Pakistan [3]-[6]. But there is also limitation in similar researches in the EU countries.

So, an accounting software program together with specialists and methods used to gather financial information about business events create the accounting information system. Accounting software without other components of AIS is the only device for recording and storage of information. Owners often buy simple accounting software with limited functional specification and do not get all possible results when it connects with other managerial systems. Choosing among cloud accounting and traditional accounting software is the main dilemma and problematic challenge for a lot of small and medium sized businesses and the important issue of current research.

Problems related to the right choice of accounting software programs could be solved only in complex: with selection of appropriate interconnected technical components – input and output devices, information storage and processors; and after deep examination of all accounting software pros and cons.

Current research has shown that it is important for companies to make special investigation to find what type of accounting software is necessary and suitable for satisfaction of accounting and managerial needs. There was made authors' contribution to develop knowledge in the theoretical and practical acquisition of accounting sphere. There have no researches been conducted which would show statistical analyses of companies produce accounting software according to different types, comparison of price policies and only few researches were conducted to evaluate the impacts of using accounting software on the companies' performance (only in separate countries).

2. Literature Review

Professional technologies have changed the way professionals conduct accounting. For past several decades accounting software has been used only to conduct monetary transactions between companies and with individuals but with time it became more complex. Nowadays accounting software is an effective tool in managing business processes. There are thousands of firms producing different kinds of accounting software. Managers collide with problems in choosing the most suitable accounting software to fit all companies' requirements.

Many scientists tried to investigate the influence of accounting information system (AIS) on the performance of all types of business. But most of them are connected to separate countries. Research [7] showed the influence of accounting information systems (AIS) on the performance of small and medium enterprises (SMEs) in Iraq and proved that AIS is one of the most important indicators for SMEs sustainability. Another author [8] investigated the impact of accounting software on the business performance of Malaysian firms. The research was based on participants' attitudes towards the importance of AIS. Results were positive and the author affirmed that accounting software has the significant importance and a great value to businesses, organizations and the economy. Some scientists [9] indicated that accounting software plays a critical role in the creation of quality accounting information and in storage necessary information to the decision-makers. In scientific paper was [10] investigated the impact of accounting information systems on organizational performance in the context of Saudi's SMEs and there was defined that it is a tool that helps to improve control of company's managers about business performance. It also can reduce the cost of accounting spending. Accounting software has a direct influence on the quality of financial statements submitted to the different state departments. Conducted investigation [11] showed the sufficient impact of accounting IS on the quality of financial statements, time of their preparation in Jordan and showed the positive result in case of its usage. Several scientists conducted research connected with defining the impact of AIS on performance measures in a case of Spanish SMEs, [12] and they found that it has a positive influence on outcome indicators and productivity. Other scientists went ahead and proved that with rapid technological development the transactions of a business are becoming more complex than it was ever that course more demand for control. Good accounting software can put a business in efficient control.

But in the literature, there are almost no scientific approaches to the classification of features and specifications of accounting software which can meet separate requirements of companies according to their sizes. Only a few accounting software types were analyzed according to their advantages and disadvantages.

3. Results: Theoretical Background

Accounting tasks are spreading each year; it is much than a booking, so many specialists have been thought about simplification of accountants' work and making storage of information more reliable. A computerized accounting system could be defined as an effective means of keeping accounting records in the electronic version. Accounting software is a special class of computer programs that helps to manage a business. The most suitable definition of 'accounting software' was given by Barron's Accounting Dictionary: 'Accounting software are programs used to maintain books of account on computers. The software can be used to record transactions, maintain account balances, and prepare financial statements and reports'[13]. Many accounting software programs are designed to create simple

money transactions (bookkeeping) between companies and others to record entire financial comings and goings for simplifying analysis procedures and decisions making processes. It depends on their scopes and functions.

History of accounting software. Most accountants affirm that accounting software is an invaluable tool for modern business but at first it was only appointed to simple operations storage. And the essential meaning of accounting software was recording financial data and later its function was extended to turning it into useful financial information [14]Error! Reference source not found.. With time professionals titled the accounting software as a 'magic wand', since before the computerizing of accounting transactions many tasks were conducted by hands and companies kept large transaction journals. The change happened only in the 20th century when computers transformed accounting significantly. The first computer for accounting purposes was sold in 1955 [15]. Later, **Peachtree Software** introduced the first accounting software in 1978 – a package for the early personal compute [16]. The Internet era has caused a significant increase in demand for accounting software. A list of companies and competitiveness among them has been increasing each year. Such innovation gives a real picture of the company's performance or as it is modishly to talk 'fair view' and 'health status of company'.

Functions, features and requirements to users. But even though the automated program **makes mach manual accounting and bookkeeping easier processes,** all software should meet the necessary requirements for legal and ethical characteristics of each country. The general functions of accounting software in IAS are analyzed in Figure 1.

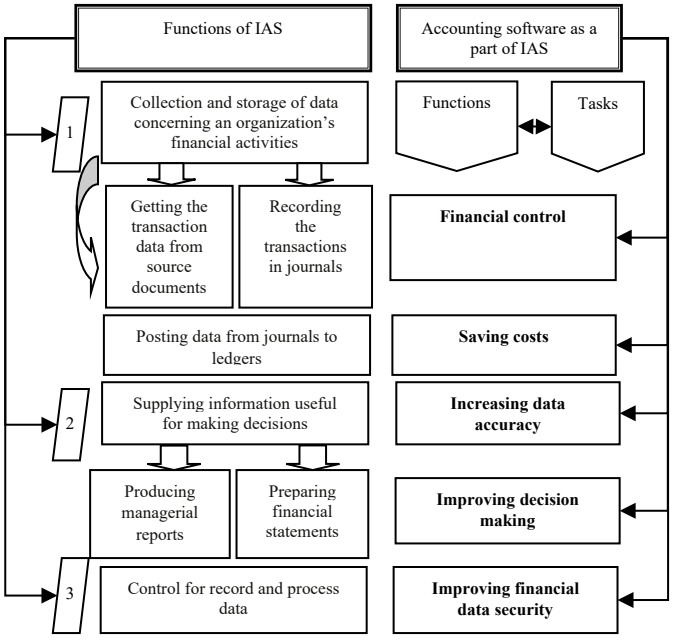


Figure 1: Managerial functions and tasks of accounting software in complex with IAS [17]; [14]

Accounting software has standard functions in the accounting field reflected in Figure 2.

The speed and accuracy of operations are the main characters of accounting software. Scientists also emphasize two other features of such computer programs; one of them is the managing of resources more efficiently in accounting departments and other reducing of costly bookkeeping mistakes [18]. Calculation of math figures and sums in the accounting process became far easier with new computer technology.

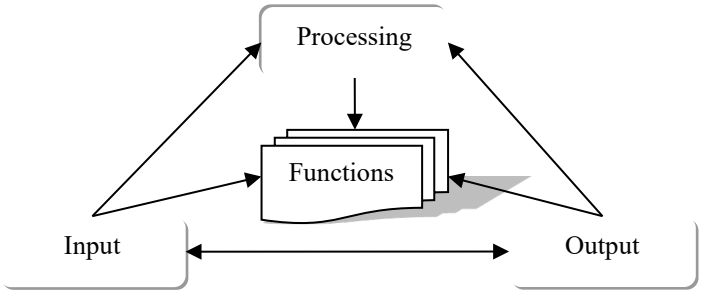


Figure 2: Productive (basic) functions of accounting software

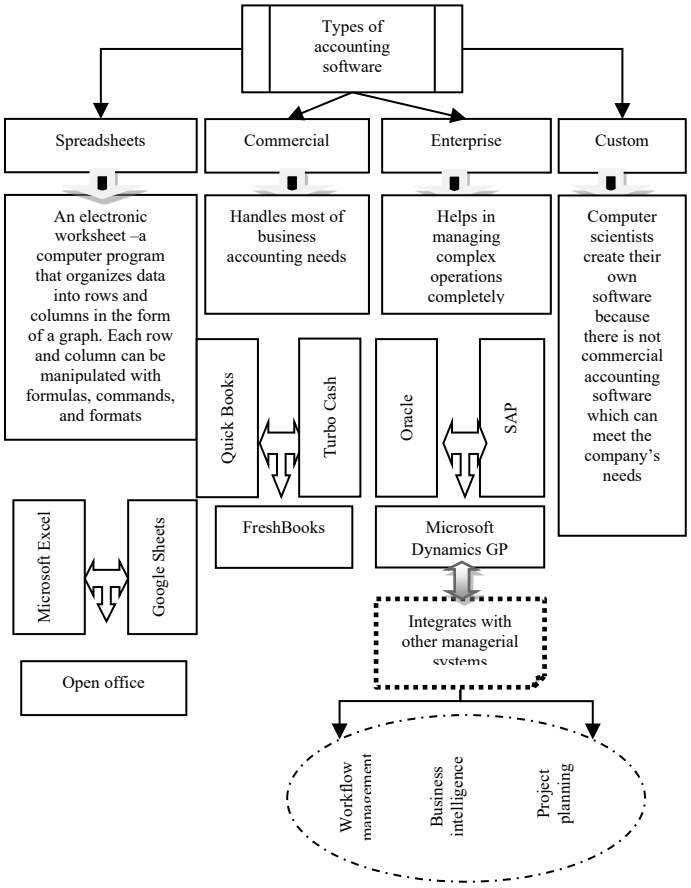


Figure 3: Types of accounting software [20]

But there was not proved how much time was reduced in an accounting cycle after computer programs appeared. Overall work from all departments is faster because a computer can keep and calculate thousands of indicators simultaneously. Most of managers and business owners separate another controversial dilemma – the requirement of qualified accountants on each stage of the accounting process. In past years accountants had to obtain necessary education, knowledge and experience in the accounting sphere and then they could work with accounting procedures for a long period, only changes in Government Standards on the state level required of Certification training or additional educational programs (courses). Accounting computer programs work faster

but only after clearly defined tasks. Computerized accounting procedures require accountants who can use specific software and therefore special computer skill. Only few accountants have possibilities to obtain the necessary qualification and experience themselves. Others need special education or training. But it is additional expense for many business owners. This problem is also connected with the high cost of training courses. But, the future benefit is higher than temporary expenses. Moreover, AIS can easily integrate with other management systems of enterprises.

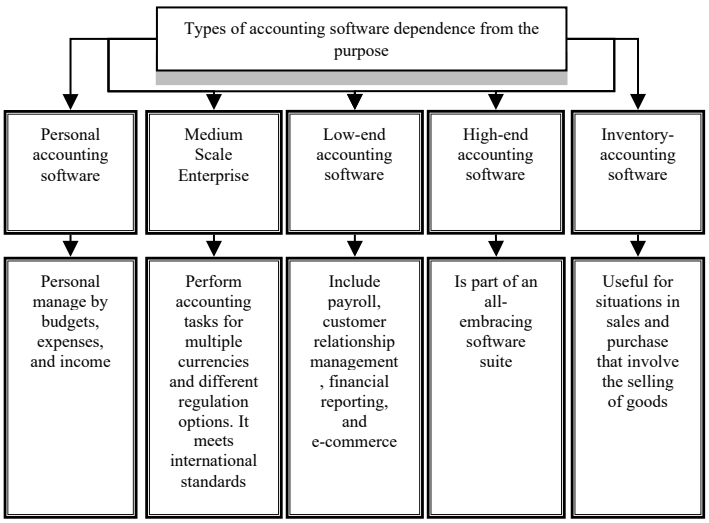


Figure 4: Types of accounting software serving the different purposes [24]

Accounting software has a row of advantages and disadvantages. But most practitioners do not take into account them; this is evidenced by fact that almost all organizations and companies have used computer programs for accounting purposes in everyday business [19]. Scientists predict that computerization of accounting work will continue in the future as managers have value benefit from it – a stable source of information expressed in three forms – a document query, report and results of calculation (mostly expenses and income).

The effectiveness level of accounting software units depends on many factors such as types, appointments and their functions (Figure 3).

It is also necessary to stress on other classification of accounting software [22], where the author excreted three types of accounting software – database, installed and cloud. Large companies and corporations need to secure own data so owners install database software that has a higher level of security, fulfills more complex tasks, integrates with large network and complicated accounting functions (ex. Oracle). Installed accounting software is useful for businesses where internet connection is limited or absent. Meanwhile, cloud accounting is becoming more popular among startups and small-sized companies (ex. Freshbooks, Quickbooks, Xero and Sage 50 Premium Accounting in different versions). Currently, 6% of SMEs use a 'cloud-based' accounting system worldwide [23]. There is also another classification of software which fits accounting purposes [24], and it is reflected in Figure 4.

Table 1: Advantages and disadvantages of accounting software

Advantages and disadvantage of separate accounting software					
Cloud accounting software		Commercial accounting software (Peachtree)		Enterprise accounting software	
Pros	Cons	Pros	Cons	Pros	Cons
Lower operational costs (investment, maintenance is absent), this type of accounting software is cheaper then purchasing software	Stable demand in access to the internet connection	Offer a set of the most widespread features to perform accounting tasks	Comparable high cost	Easy to install on a system and deploy to end-users quickly	Comparable high cost
Higher reliability	There is the risk of data confidentiality (hacking – there is no guarantee that data is 100% safe)	Affordability for companies' owners – the average cost	There is a small but real possibility to lose information if the system is attacked by computer viruses in case it is not enough protected	Easy to customize	There is a small but real possibility to lose information if the system is attacked by computer viruses in case it is not enough protected
Higher accuracy (fewer mistakes). Much easier tools then other types propose	Technical problems	There is no demand in well-experienced skills	Increasing of productivity	Meet almost all user requirements	The necessity of high qualified personals
Pay only for subscription	Vendor lack-in/ lack-of control	High level of security	Flexibility and timeliness	Increasing productivity	On-going support
24/7 access and possibilities to recover data (restore)	In some cases limits to the data that can be freely store	Low risk	The necessity of high qualified personals	Management information	Mistakes in data entry can throw off a whole set of data
Collaboration	A lack of specialized tools	Flexibility and timeliness	Mistakes in data entry can throw off a whole set of data	Flexibility and timeliness	Can be inflexible

Table 2: Cost comparison of 12 top accounting software programs and description of their capabilities [27]**Error! Reference source not found.**; [28]; [29]; [30]; [31]; [32]; [33]

№	Accounting software	Form	Manufacturer	Cost, $ per month/y	Appointment
For start-ups, small and medium-sized businesses					
1	QuickBooks	Online Desktop	Intuit United States	20-150 299, 95	Easy for working accounting software with features that can satisfy complex accounting needs: reconciliation, taxes, reporting and inventory management. 4,5 million customers globally use QuickBooks
2	FreshBooks		Designed by Mike Mc Derment in Toronto, Ontario	6-20	FreshBooks is a cloud-based accounting designed software. It is ideal for freelancers (accountants) and small businesses that look to fast track their sales cycle and keep their processes in line with standards and regulations. Over 12 million users in 120 countries
3	Accounting by Wave	Online Mobile version	Wave (Wave Apps)	Free (only several payroll services)	It is a very suitable and free accounting software for many owners of small companies. It is a powerful bookkeeping. Easily integrates with other management systems of companies
4	FreeAgent	Cloud-based	FreeAgent Central	Free	It is the accounting software for small business in the UK. There are over 80 000 small business users.
5	Manager	Online Desktop	NGSoftware	Free	For a small business, all industries, 70 languages
For mid-sized and large businesses					
6	QuickBooks Enterprise	–	Intuit United States	1419-3442	The end-to-end accounting software designed for growing business. Usage of the program helps to manage inventory, run payrolls, track sales and report. This type of software manages accounting transactions from start to finish. 135 000 companies chose QuickBooks Enterprise. There is another version QuickBooks Desktop Enterprise 20.0. Users: manufacturers, wholesalers, retailers and nonprofit (for different industries)
7	Xero	Online	Xero (New Zealand)	20-40	It is the smart online accounting for small business. Experts consider this type of accounting software one of the best for growing companies. It is also suited for mid and large businesses.
8	Finsync	Cloud-based	Finsync	From 40	It is a powerful accounting software (all-in-one) with different options. All types of business can use it.
9	NetSuite	Cloud-based	Oracle	–	It is suitable accounting software for companies that have grown out of their accounting software and are ready for full ERP[1] software
10	Zoho Books	Online	Zoho	9-29	The main features of accounting software are: end-to-end accounting, easy collaboration and integrated platform
11	FinancialForce accounting	Cloud-based	FinancialForce	From 175	This accounting software is designed for the companies with employees 100 +. It allows managing company's accounting and it works with a large amount of data. It also boasts powerful features. Experts prove that accounting software can satisfy complex enterprise requirements. The main characters of the product are multicurrency and global tax
12	*MYBOS Accounting & Billing* *MYBOS Accounting on Cloud*	Desktop Online	SBSCC Sri Lanka	249 199	Full integrated multi-user accounting software. In general, this accounting software is for medium-sized business. It is also accounting management that helps companies' owners to control sales and purchases, inventory, and journal entries. For companies, it is possible to obtain a hybrid cloud system

Some special industries order to design separate accounting software which belongs to Commercial off-the-shelf software. These are positive determinates of selecting accounting software. But other features might influence the appropriate choice of accounting software. Owners must choose among free, paid, or online accounting software. Most of the producers give a trial period for testing their products. In [25], the author stressed that selecting the most appropriate accounting software package has become one of the most critical decisions for most organizations in the fast-changing business world. A choice depends on the advantages and disadvantages of accounting software packages.

Vendors of accounting software try to advertise the main positive aspects of their products and it is very hard to exaggerate meaning of benefit from the accounting software. But many owners have a predisposition to make choice among types of software which give not only gain inside of companies but competitive advantages in the marketplace. In the question of

[1]ERS – enterprise resource planning.

choice among accounting software advantages must overweight completely disadvantages (Table 1).

The list of some of the best accounting software for three groups of business according to size is reflected in Figure 5.

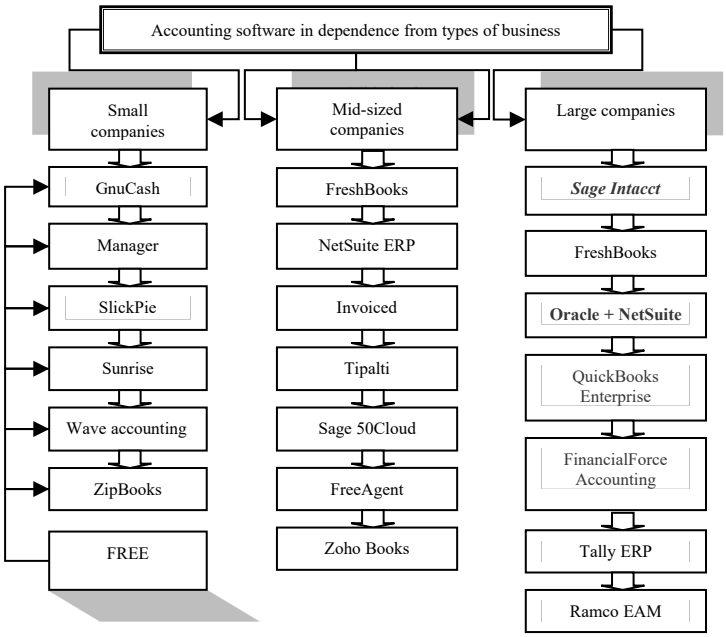


Figure 5: Different accounting software for three types of business

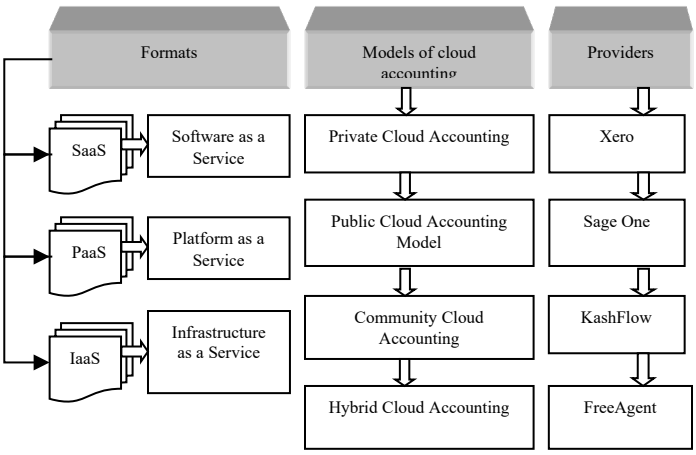


Figure 6: Brief characteristic of cloud accounting system [26]

There are other suitable accounting software programs for different types of business. More and more consumers need special software that contains features of effective management. Nowadays, true and fast information for external and internal users is only one of the main reasons for the purchase of accounting software. Cloud accounting software: This term appeared in far 2011. It gives the possibility to perform all accounting functions through internet-based devices. The main features of cloud accounting are shown in Figure 6.

Cloud accounting is a more flexible system with one strong requirement – connection to the internet. It is appointed to work with traditional accounting functions, updates financial information automatically and provides complex financial reports.

4. Practical aspects of accounting software choice

Cost of accounting software: The world market is full of different accounting programs with a variety of capabilities. Prices for accounting programs vary from several dollars to thousand. It depends on many factors and the number of users supported it. The cost of accounting software includes installation, implementation, training programs for users, customizing the software, and operational costs. Nowadays, many producers offer accounting programs on a cloud base and the number of venders who sell licenses has been decreasing each year. It is better for producers to get charge a monthly subscription fee from users. For small businesses and start-ups, the most appropriate accounting software is QuichBooks, Wave, Manager and FreshBooks, but some of their manufactures offer packages for growing business and thriving firms with competitive items. For large companies, there are several worldwide famous accounting software producers such as Intuit (USA), Xero (New Zealand) and Oracle (USA), (Table. 2).

The price is the second indicator which influences on the right choice. For small companies the price is affordable, it varies from $5 to $500 (Figure 7). But companies with multinational transactions and a large amount of staff need accounting software with additional opportunities and integration capabilities. The scientist [34] advised to choose optional enterprise-class accounting software that can meet the company's specific needs but the price is higher accordingly.

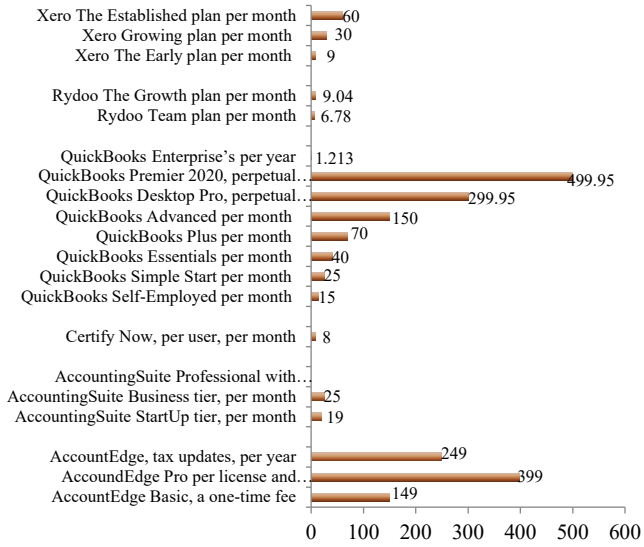


Figure 7: Compression of the prices for different accounting software products [35]

Many manufactures of accounting software offer additional payment options for some special requirements. Byers also desire to combine accounting software with human resources management (HRM) and POS[2] (software) systems; integration with the last one is limited in many accounting products. The right

[2]POS – point of sale.

accounting system must be adopted not only for information gathering but for synchronizing transactions, serving with oversight into the company's position, and proposing accounting solutions. The other point which must be accounted by users is a pricing policy for accounting software. Venders offer two popular pricing models: subscription-based on a cloud platform or 'pay-as-you-go' model[3] and perpetual licensing. Subscription price must be made monthly/year and it guarantees necessary support and maintenance. Some venders give buyers a choice to use all components, special modular components or features they need and pay for them. Many designers of accounting software include additional services and support in the base price of core products but others offer additional features, or modules for an extra payment. Before making the eventual choice for any type of accounting software it is very important to learn whether venders have professionals who are specializing in installation, supporting, and integrated with other company's systems processes or they involve the third party. Third-party consultants sometimes have divergences. It is worldwide accepted to provide the basic training services for free but there are some exceptions.

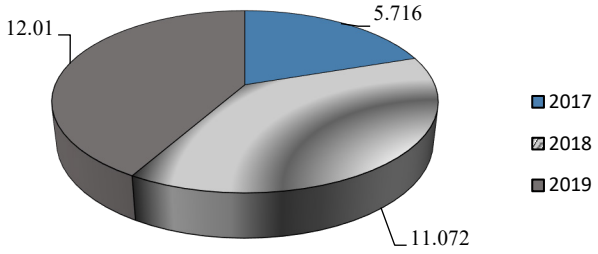


Figure 8: Global accounting software market in 2017-2019, Bn, USD dollars [36]-[38]

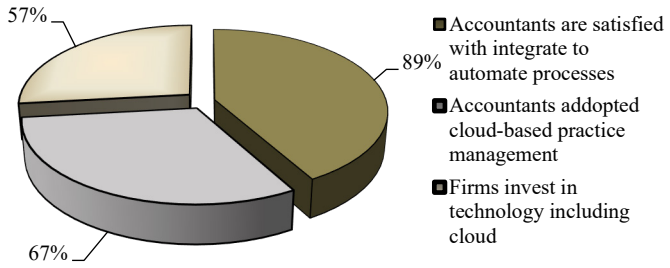


Figure 9: Perception of cloud technologies by accountants, 2017 [39]

The global market of accounting software: The market size of accounting software has been increasing each year and scientists predict the continuation of this tendency for the next few years (Figure 8). Most of the rising is observing in the cloud-based sector. Now, cloud accounting replaced the old desktop versions. Many business owners invest in the cloud accounting system that gives the future long-term benefit in a case when they plan to expand their activity. It is always useful to exploit such type of software when a lot of employees work on the same data. Large companies which are represented by foreign branches perceive cloud-based accounting software more flexible. Most of the users are worried by one main question – is it enough safe? It's protected by the individual password. Accounting software providers also create backup servers in different locations for reinsurance if one server would go down, users can access to necessary data. Excess to cloud accounting software is more protected than the traditional one. The team of 'Finances Online' investigated that more than 50 % of owners and accountants are satisfied by the cloud accounting platform (Figure 9).

Other useful statistical data shows that companies have increased their profits after adopted online accounting (Figure 10).

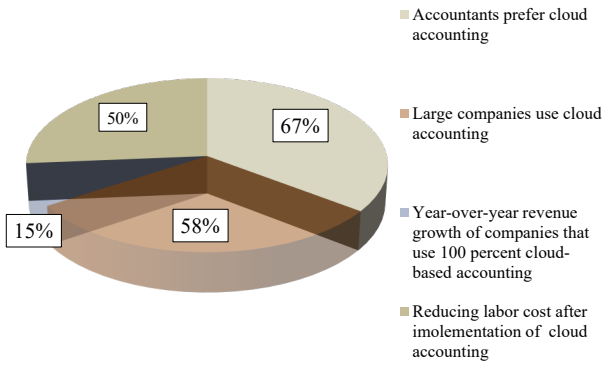


Figure 10: Cloud accounting statistics in 2018 [40]

Any type of accounting software is a great solution for each business as its advantage's overweight disadvantages. There is a huge variety of accounting software, producers, and venders. Many specialists try to model the best one or make the most suitable for all businesses. But the best accounting product for one company might not be the best for another even with similar features. The main character which must fit all customers is trust in it.

5. Accounting Software Influence on Companies Performance

Accounting software provides companies' owners and managers with quick and easy reports for decision making [41]. Accounting system can effect on all spheres of activity [42] and has positive influence on companies' performance. Appropriate accounting software gives some benefits, the most important among them are affordability, time savings, higher accuracy and one place for all financial tasks; they are all available to each form of business.

6. Generalization of Results

Accounting software in modern business plays very important role and it is a great progress made in the accounting sphere. The main goal of accounting software is tracking financial transactions and generating various financial reports. It is a part of whole accounting systems which might include different types of software from simple transaction-entry programs to advanced accounting systems integrated with other management

[3]Pay-as-you-go system is one in which you pay for a service before you use it and you cannot use more than you have paid for. Cambridge Dictionary.

programs of the company. Accounting software makes accountants work easier and gives information about the 'true health' of business, companies' owners agree to pay a substantial sum for complicated and advanced accounting technologies. Companies with foreign branches have always troublesome issues with coordination among them (mainly a time of reports' delivery that could cause a delay in decision making process) but after cloud technologies appeared, most problems have solved. Cloud accounting or e-accounting is rather a new type of service than IT. The main advantage of cloud accounting software is no time and place limitation as well as putting the whole accounting system to the cloud-based platform.

Analyzed scientific researches proved the positive interrelationship between accounting software and business performance. But in this relationship, many factors have an indirect influence. There are several benefits for companies' owners who use such software: the establishment of financial control, saving costs, increasing data accuracy, and improving data security. With a good accounting system, accountants and business owners can make financial forecasting and improve the decision-making process.

Nowadays, the global accounting market is represented by a lot of producers who offer from traditional old desktop accounting software to the modern cloud-based platforms on license subscription conditions. The tendency of the global market has a positive and stable increasing character. Research showed that more than 60% of accounting specialists are satisfied with their current accounting programs and most of them are ready to invest in accounting technologies for getting future benefits. Most specialists affirm that cloud-based accounting software together with other intelligent online programs is the future of accounting.

7. Conclusion

Research showed a significant increasing of accounting software popularity and that fact has caused sufficient diversification in computer accounting programs. And it's becoming difficult to decide which one to choose. There was grouping of accounting software programs of the most famous world producers using affordable and easy-to-use approaches, listed their advantages and disadvantages, analyzed price policy parameters. There were indicated items to keep in mind when selecting accounting software for companies: special features (accounts receivable and accounts payable tools, track inventory, time tracking, project management, payroll or advanced reporting capabilities), usability (desktop software, cloud software or mobile version and number of users) and costs (inexpensive, average price or expensive with extra features). Also research reports about some risks which connect with storage accounting date and information. As recommendation there are some offers for each type of accounting software program – to give detail explanation on its security level, use multi-business support, find providers with additional services and possible training programs.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M.I. Ajay Adhikari, "Firm characteristics and selection of international accounting software," Journal of International Accounting and Taxation, **13(1)**, 59–63, 2004.

[2] M. Burns, "Customer survey roundup," CA Magazine, Retrieved June 15, 2013.

[3] B. Gullkvist, "Towards Paperless Accounting and Auditing," E-Business Research Centre, 87–98, 2002.

[4] H. Sajady, M. Dastgir, H. Hashem Nejad, "Evaluation of the effectiveness of accounting information systems," International Journal of Information Science and Management, **6**(2), 49–59, 2008.

[5] K. Saira, M.A. Zariyawati, M.N. Annuar, "Information System and Firms," International Business Research, **3**(4), P28, 2010.

[6] R. Kouser, A. Awan, Gul-e-Rana, F.A. Shahzad, "Firm size, leverage and profitability: Overriding impact of Accounting Information System," Business & Management Review, **1**(10), 58–64, 2011.

[7] E. Harash, "Accounting performance of SMEs and effect of of accounting information system: A conceptual model," Global Journal of Management and Business Research (D), **17**(3), 21–26, 2017.

[8] I.N. Ch.Yvonne, "The impact of accounting software on business performance," International Journal of Information System and Engineering, **6 (1)**(November 2017), 2018, doi:10.24924/ijise/2018.04/v6.iss1/01.26.

[9] D.M.J. Wickramsainghe, R. Pemarathna, N. Cooray, T. Dissanayake, "Impact of accounting software for Business Performance," Imperial Journal of Interdisciplinary Research (IJIR, **3**(5), 1–6, 2017.

[10] Trabulsi, "The Impact of Accounting Information Systems on Organizational Performance : The Context of Saudi ' s SMEs," International Review of Management and Marketing, **8**(2), 69–73, 2018.

[11] A.A.J. Abdallah, "The impact of using accounting information systems on the quality of financial statements submitted to the income and sales tax department in Jordan," European Scientific Journal, **1**(December), 41–48, 2013.

[12] E.U. Grande, R.P. Estébanez, C.M. Colomina, "The impact of accounting information systems (AIS) on performance measures: Empirical evidence in spanish SMEs," International Journal of Digital Accounting Research, **11**(June 2010), 25–43, 2011, doi:10.4192/1577-8517-v11_2.

[13] J.K.S. J.G. Siegel, Dictionary of Accounting Terms, Third Edit, 2000.

[14] Accounting Information System, Aziroff, 2019.

[15] B. Scott, The Accounting Journal: 60 years of accounting software, MYOB, 2015.

[16] F. Sherman, The History of Computerized Accounting, Career Trend, 2019.

[17] Kanya, The Importance of Accounting Software for Businesses, Businesstech, 2019.

[18] U.D. Dr Mahesh, "Role of accounting software in today scenario," International Journal of Research in Finance and Marketing, **6**(6), 25–34, 2016.

[19] V.N.T. İ. Dalci, Benefits of Computerized Accounting Information Systems on the JIT Production Systems, Eastern Mediterranean University, 21–36.

[20] T. James, Types of Accounting Software, Chrone, 2019.

[21] What is a Spreadsheet?, My Accounting Course, 2020.

[22] K. Coloso, Best Accounting Software for Startups, Founders Guide, 2015.

[23] P. Griffiths, Accounting in the Cloud – Pros and Cons, Accru Chartered Accountants, 2020.

[24] What are the types of accounting software?, Shoebooks Team, 2012.

[25] A.A. Abu-Musa, "The Determinates Of Selecting Accounting Software: A Proposed Model," Review of Business Information Systems (RBIS), **9**(3), 85–110, 2005, doi:10.19030/rbis.v9i3.4456.

[26] S.A. Tarboush, "Cloud Accounting As a New Business Model and Its Influence on Accounting Process," (2), 1–14, 2017.

[27] Ch. Krause, QuickBooks Online VS QuickBooks Desktop: Which Is Better?, Merchant Maverick, 2019.

[28] The way you work is changing. Our dedication to small business has not., QuickBooks, 2020.

[29] Accounting software that's free and powerful, Wave Accounting, 2020.

[30] Xero is online accounting software for your small business, Xero, 2020.

[31] Looking for accounting software?, Manager, 2020.

[32] E. Seppala, Best Accounting Software For Large Businesses, Merchant Maverick, 2020.

[33] Accounting software world, FinancialForce, 2020.

[34] J. Martinez, 5 Mistakes to Avoid When Choosing Accounting Software, PC, 2017.

[35] M. Pardo-Bunte, How Much Does Accounting Software Cost? 2020 Pricing Guide, BetterBuys, 2020.

[36] Global Accounting Software Market to be worth US$ 11,771.6 Mn by 2026: Transparency Market Research, Cision, 2018.

[37] Hardware and software IT services / Accounting software market, Fortune

Business Insights, 2019.
[38] Accounting software market – growth, trends, and forecast (2020-2025), Mordor Intelligence, 2020.
[39] D. Epstein, What is Accounting Software? Analysis of Features, Types, Benefits and Pricing, FinancesOnline, 2020.
[40] M. Gigante, 30+ Important Accounting Statistics You Need to Know in 2019, 2018.
[41] M.M. Thottoli, K. V. Thomas, R.A. Essia, "Adoption of Audit Software by Audit Firms: A Qualitative Study," Journal of Information and Computational Science, **9**(9), 768–776, 2019.
[42] H. Rkein, "Saudi Journal of Business and Management Studies Impact of Automation on Accounting Profession and Employability: A Qualitative Assessment from Lebanon," Journal of Business and Management Studie, **4**(4), 372–385, 2019, doi:10.21276/sjbms.2019.4.4.10.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 871-886 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Detailed Assessment of Dissaving Risk Against Life Expectancy for Elderly People using Anonymous Data and/or Random Data: A Review

Yuya Yokoyama*, Yasunari Yoshitomi

Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Kyoto, 606-8522, Japan

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 27 March, 2021
Online: 10 April, 2021

Keywords:
Life expectancy
Savings
Dissaving risk
Anonymous data
Random data

ABSTRACT

With a view to detecting whether economic activity deterioration for elderly people at age of sixty-five or over could be observed, anonymous data (AD) were used as analysis data, which were obtained from the National Survey of Family Income and Expenditure (NSFIE) conducted by the Ministry of Internal Affairs and Communications (MIC). We have developed a method to detect dissaving risk among elderly people. In our previous analysis, AD were divided into test data and training data. Three kinds of methods were performed on the basis of income and savings. Then two-step methods were processed to determine dissaving risk. Nevertheless, in utilizing AD as it is, the security of anonymity could be questionable. Therefore, in order to enhance the anonymity of the data, random data (RD) were generated based on AD in this paper. Then RD were compared with the case of analyzing mere AD as it is, for the purpose of performance evaluation. Further analysis results suggest that using both RD and AD would be as effective as using only AD in evaluating the performance of the proposed method.

1. Introduction

This paper is an extension of work originally presented in SNPD2019 [1] by the further detail data analysis and algorithm explanation. It has been widely known that the decrease of nerve cells and the deterioration of cognitive function in accordance with aging can cause the deterioration of memory, thinking, and motivation [2]. As the symptoms of dementia, cerebral disorder is also known, e.g. memory impairment, judgment disorder, execution function disorder, dropping off of cognitive ability, etc. It has also been shown that the older people get, the higher the possibility of the occurrence of disease. In the case of dementia patients, a depressing symptom deprives him/her of vitality and even interest in what he/she used to be interested in before.

Doctors mainly determine whether dementia occurs on the basis of medical diagnostic criterion. Dementia can be classified into several types such as Alzheimer dementia, dementia with Lewy bodies (DLB), etc. Early diagnosis enables the specification of disease and appropriate prognosis treatment. However, there are not a few cases where temporal, economic, and geographical difficulties might obstruct early diagnosis by doctors.

It has been known that the deterioration of brain function can cause the impairment of consumption activity, such as being unable to withdraw money from banking facilities or to do shopping in spite of ample money. The case of increasing savings without using money because of abnormal adherence to money is also one form of declining consumption activity as well.

The relationships have been revealed between the deterioration of cerebral function and the reduction of consumption activity of elderly people. The malfunction of cerebral function advances if not treated properly or immediately, and can disrupt even daily life owing to dementia, etc.

For the single elderly person, existing techniques enable us to determine dementia by their behaviors in their daily lives. For example, a system has been developed that predicts dementia by using input information in order to inform and input their safety [3]. This system can observe the safety under the protection of the elderly's privacy as well as foresee dementia without regular medical check-ups. Another system is a household electric appliance that determines if the operator is possibly experiencing dementia by obtaining his/her operation history, sleeping hours, commodity purchase history, etc. [4]. If the possibility of dementia is determined, the information is sent to an external device owned by the family of the operator. However, it would be indispensable to monitor the behavior of elderly people to determine if he/she has dementia. Therefore, it would be necessary to purchase the devices and install them to monitor him/her.

*Corresponding Author: Yuya Yokoyama, Email: y_yokoyama@mei.kpu.ac.jp

The jeopardies of economic activity ability are classified into the following two types: (1) "Dissaving risk," including the possibility of being victimized by fraudulent means, which results from the impairment of judgement, etc., (2) "Under-consumption risk" that is caused by the lowering of motivations. Therefore, the objective of our work is to develop a system that detects those jeopardies early on and issues a warning. In this paper, "dissaving risk" is focused on.

So far we have employed anonymous data (AD) on the basis of the National Survey of Family Income and Expenditure (NSFIE) [5] conducted by Ministry of Internal Affairs and Communications (MIC) in 2004. As a basis of economy activity, for the data of single households at the age of sixty-five or older, clustering was performed with five clusters based on income and saving [6]. As a result of analysis, five clusters were obtained and characteristically classified. Then we developed a method to determine if savings are enough against life expectancy in terms of income and savings [6]. The analysis results revealed that correction detection ratio of the data determined as dissaving risk ranged from 52.6 to 94.1%.

Nevertheless, our first method showed the limitations with the additions of AD years 1994 and 1999. Therefore, our second method has been proposed with a view to determining immediately and easily cerebral function that has started to decline by focusing on economical ability of single-person and two-person households at the age of sixty-five or older [7]. Here, the analysis data were divided into test data and training data [7]. Then two-step methods were then taken to determine dissaving risk [7]. Three kinds of methods were performed in terms of income and savings. However, the security of anonymity for AD could be controversy [1]. In addition, AD can be available and must be returned within the designated durations after application is approved [8]. In order to reinforce the anonymity of AD, random data (RD) is generated along with AD and then compared with the previous analysis [7], for the purpose of performance evaluation. The analysis result conveys the conclusion that using the RD might be as effective as using mere AD [1] in evaluating the performance of the proposed method [7].

In our previous paper [1], the analysis methods and results were superficially and briefly summarized due to space limitations. Therefore, this paper provides more detailed analysis methods and results so that our proposed method [7] would be understood more profoundly.

The remainder of the paper is organized as follows. Related works are referred in Section 2. Preparation of the analysis is summarized in Section 3. The outline of our proposed method [7] using AD is explained in Section 4. Analysis results using RD to evaluate the proposed method [7] is explained in Section 5. Finally, Section 6 concludes the paper.

2. Related Works

In [9], authors pointed out that the shrink of economic activity of rapidly increasing dementia patients along with aging could influence actual economy. For example, when an elderly person develops dementia, he/she might repeat the following experiences; he/she is unable to withdraw money from his/her bank account because of not remembering password, forgets to pay bills, repeatedly purchases what he/she already had before. These experiences could result in considerable reduction of economic activities. On 2012 fiscal year, among the elderly people at the age of sixty-five or older, the number of dementia patients is approximately 4.62 million, while that of mild cognitive impairment (MCI) attains roughly 4 million [9]. The sum of these patients account for close to 7% of Japanese. Since the elderly people occupy most portions of dementia patients, their figure is expected to go on increasing on a global scale in the upcoming several decades [9]. Thus the shrink of economic activities will extremely impact on economy worldwide. From these viewpoints, our work focuses on the possible decline of economic activities of the elderly people.

In [10], authors conducted an experiment with a view to investigation on annual change of economical ability related to variation from MCI to dementia [10]. They managed the subjects on financial capacity instrument (FCI) under one-year follow-up. The subjects are composed of the following three types: 76 Controls, 25 MCI Converters, and 62 MCI Non-Converters; Control is healthy elderly people, whereas MCI Converter (MCI Non-Converter) is a subject on whom the inflection from MCI to Alzheimer-type dementia was (was not) observed. Using multivariate analysis of variance (MANOVA), the performance of FCI domain and global scores are compared within/among groups. As a result of analysis, MCI Converters fell short of MCI Non-Converters, from the viewpoints of economical concepts, money withdraw management of bank deal specification, payment on invoice, and total FCI scores. Although this research indicates the degradation of economic activities themselves of MCI patients, it is not based on actual consumption behaviors.

In [11], the authors researched on money management issues of the elderly people households along with aging and cognitive ability deterioration. In [12], the authors investigated if the knowledge of basic concepts necessary for effective and economical choice would be degraded after sixty years old, for the purposed of detecting the evidence of intelligence decline in accordance with aging that influences the ability necessary for money management. In [13], the authors analyzed the impact of information against cognitive ability decline of decision maker of finance by using time-series data related to elderly married couples. Authors made an analysis on cognitive ability by employing the actual examples of almost best behaviors and ample economic decisions of consumers. Analysis results conveyed the conclusion that consumers with overall higher scores on analysis items and mathematics are unlikely to end up with economic collapse. These researches [10-13] have been analyzed on the basis of anecdotal reports and multiple manifold methods such as regression analysis. Nevertheless, there have been no researches that focus on the relationships between cognitive ability and purchase motivation.

In [14], the authors evaluated default risks through comparing transition of representative financial ratio considered as the financial safety of companies among bankrupt firms and non-bankrupt firms. For financial ratio, fourteen feature values are extracted; four for profitability and ten for security. These fourteen time-series variation of financial ratio were described to investigate the validity of financial ratio analysis in default risk evaluation. This research analyzes bankrupt risk for enterprises.

However, no indications for individual bankrupt risk were referred.

In [15], the authors used AD to study on relationships between living expenses and fortunes for medical care risks by classifying family budgets, e.g. higher income, lower income, and so on. It has been suggested that appropriate household plans are possible that foresee life cycle of "how much savings will be able to keep covering anticipated household expenditure."

3. Preparation for Analysis

3.1. Anonymous Data (AD)

The analysis data employed for our previous works [1,6,7] is an AD obtained from National Statistics Center (Japan) [16], based on NSFIE carried out by MIC in 1994, 1999 and 2004. This data contains 1,919 (1,752 and 1,780, respectively) detailed family income and expenditure survey items for 1994 (1999 and 2004), e.g. the amount of income and expenditure, expenditure amount of food such as crop, fish, meat, etc. Among the AD, the present paper uses the data of single elderly people and two-person households (Twos) at the age of sixty-five or over; 717 single males, 3,475 single females, and 11,528 two-person households. There are two types of Twos; head of household is either male or female. However, the Twos whose head is male (1994: 2,607, 1999: 3,840, 2004: 5,081, total: 11,528) considerably outnumbers those whose head is female (1994: 20, 1999: 39, 2004: 44, total: 103). Thus, the Twos whose head is male are used for the analysis data.

3.2. Revised Disbursement

Disbursement consists of the following three components: (1) Expenditure, (2) Disbursements other than expenditure, and (3) Carry-over to next. Among the disbursement, the items that should be exempted from disbursement were considered [6]. As a result, five items were determined to exclude from disbursement [6]; saving, insurance premium payments, security purchase, property purchase and carry-over to next.

In AD, income is provided by ten thousand yen unit per year, whereas each household item related to disbursement is provided by yen per month. Therefore, the value of disbursement exempting the five items above is multiplied by 12 to calculate annual money. This value is defined as "revised disbursement." The amount of revised disbursement is denoted as D_{rev}. Using Formula (1), D_{rev} is calculated [6].

$$D_{rev} = \frac{12\{D_1+(D_2-v_1-v_2-v_3-v_4)\}}{10000} \quad (1)$$

,where D_1 is the amount of expenditure, D_2 is the total amount of disbursements other than expenditure, v_1 is the saving, v_2 is the amount of insurance premium payments, v_3 is the amount of security purchase, and v_4 is the amount of property purchase.

The terms v_1, v_2, v_3, v_4 are the feature values excluded from disbursement. The unit of D_{rev} is ten thousand yen, while those of D_1, D_2, v_1, v_2, v_3, v_4 are yen. Therefore, in order to unify the unit yen, the right side of Formula (1) is divided by 10,000.

3.3. Methodology of a Method to Determine Dissaving Risk

(1) Selection of Analysis Data

The analysis data utilized were selected from the following two viewpoints

- The data whose revised disbursement are not more than income were excluded, because their savings are unlikely to be exhausted.
- The data whose savings are 0 yen were also exempted, since their savings are apparently in dissaving risk.

From these two reasons, the data whose revised disbursement are larger than income with savings were extracted; 171 single males, 1,203 single females, and 2,367 two-person households. These data are used as analysis data in our works. These data can be divided in accordance with six groups of attributes made up of house types ("Own" and "Rent") and year ("1994," "1999" and "2004"). These amounts are shown in Table 1.

Table 1: Amounts of Analysis Data (Revised Disbursement > Income) [1]

Year	Single						Two-person Household		
	Male			Female					
	Rent	Own	Total	Rent	Own	Total	Rent	Own	Total
1994	11	30	41	104	234	338	78	445	523
1999	24	31	55	105	275	380	96	613	709
2004	26	49	75	122	363	485	132	1003	1135
Total	61	110	171	331	872	1203	306	2061	2367

(2) Condition of Dissaving Risk Determination

In order to determine the condition of dissaving risk, the term in which savings will be exhausted needs to be calculated through dividing savings by the difference between revised disbursement and annual income. The term is named as "saving exhaustion term (SET)." The SET is described as T. T is obtained through Formula (2) [6].

$$T = \frac{S}{D_{rev}-I} \quad (2)$$

, where S is the amount of savings, I is the amount of annual income, and D_{rev} is the revised disbursement.

Provided that P is the ratio of T to L_{exp} corresponding to its age group for each data, P is calculated through Formula (3).

$$P = \frac{T}{L_{exp}} \quad (3)$$

Provided that P_{th} is the threshold to determine dissaving risk, if P is less than P_{th}, the data is determined as having dissaving risk. In our previous analyses, P_{th} was set to 0.1 (0.3, respectively) for single males and females (Twos). In order to avoid confusions, the cases of single males and females ($P_{th} = 0.1$) will be focused on to explain the outlines of analysis method in the subsequent sections.

(3) Classification into Training Data and Test Data

The analysis data selected in Section 3.3-(1) are divided into training data and test data according to the six groups of attributes composed of house types ("Own" and "Rent") and year ("1994," "1999" and "2004"). Here, for the respective groups, the analysis data is sequentially classified into training data and test data

according to the ascending order of the amounts of savings. The amounts of training data and those of test data divided through these procedures are shown in Table 2. The columns entitled "R" indicate the training/test data determined as having dissaving risk. How to determine if those data are determined as dissaving risk is explained in Section 3.3-(2). Meanwhile, the columns entitled "N" show those data determined as not having dissaving risk. The columns entitled "T" indicates the sums of the amounts of "R" and those of "N."

Table 2: Amounts of Training Data and Those of Test Data
(a) Male, Rent

Age Group	Training			Test			Total		
	R	N	T	R	N	T	R	N	T
65 - 69	2	9	11	1	10	11	3	19	22
70 - 74	1	9	10	1	9	10	2	18	20
75 - 79	2	4	6	3	2	5	5	6	11
80 - 84	1	2	3	0	3	3	1	5	6
85 -	0	1	1	0	1	1	0	2	2

(b) Male, Own

Age Group	Training			Test			Total		
	R	N	T	R	N	T	R	N	T
65 - 69	3	16	19	2	16	18	5	32	37
70 - 74	4	9	13	2	11	13	6	20	26
75 - 79	0	16	16	1	15	16	1	31	32
80 - 84	1	4	5	0	5	5	1	9	10
85 -	0	1	1	0	2	2	0	3	3

(c) Female, Rent

Age Group	Training			Test			Total		
	R	N	T	R	N	T	R	N	T
65 - 69	10	41	51	8	42	50	18	83	101
70 - 74	13	47	60	11	48	59	24	95	119
75 - 79	4	32	36	5	31	36	9	63	72
80 - 84	1	14	15	1	13	14	2	27	29
85 -	1	4	5	0	5	5	1	9	10

(d) Female, Own

Age Group	Training			Test			Total		
	R	N	T	R	N	T	R	N	T
65 - 69	22	134	156	17	139	156	39	273	312
70 - 74	17	126	143	12	131	143	29	257	286
75 - 79	7	79	86	4	82	86	11	161	172
80 - 84	3	39	42	3	39	42	6	78	84
85 -	1	8	9	0	9	9	1	17	18

(e) Twos, Rent

Age Group	Training			Test			Total		
	R	N	T	R	N	T	R	N	T
65 - 69	25	52	77	20	57	77	45	109	154
70 - 74	17	35	52	11	41	52	28	76	104
75 - 79	1	18	19	3	16	19	4	34	38
80 - 84	0	5	5	0	4	4	0	9	9
85 -	0	1	1	0	0	0	0	1	1

(f) Twos, Own

Age Group	Training			Test			Total		
	R	N	T	R	N	T	R	N	T
65 - 69	117	422	539	129	409	538	246	831	1077
70 - 74	56	281	337	70	266	336	126	547	673
75 - 79	14	106	120	15	105	120	29	211	240
80 - 84	1	33	34	3	30	33	4	63	67
85 -	1	1	2	1	1	2	2	2	4

(4) Representative Age and Life Expectancy

Dissaving risk is evaluated along with life expectancy (L_{exp}) in 1994, 1999 and 2004 published by Ministry of Health, Labor and Welfare (MHLW) [17]. These values are published for all the ages every year by MHLW. In the AD, age is provided as a form of age groups, e.g. 65 to 69 that ranges from age 65 to 69. Actual age for each data is completely confident. In order to determine L_{exp}, the age of each data is assumed as the youngest of the age group so that dissaving risk must not be underestimated by assuming the unknown age older than actual age. The age determined in this manner is defined as "representative age (A_{rep})." L_{exp} is used as the form of rounded off to the nearest integer. L_{exp} for each age group is shown in Table 3.

For A_{rep} and L_{exp} of Twos, these values are averaged. Taking an example of a household composed of a male at the age of 70 to 74 whose L_{exp} is 13 years, and a female at the age of 65 to 69 whose L_{exp} is 22 years old, each A_{rep} is 70 and 65, respectively. Their average A_{rep} is 67.5, whose A_{rep} is regarded as 65 since the number is between 65 and 70. Their average L_{exp} is 17.5 years. In this fashion, L_{exp} for Twos are calculated and listed in Table 4.

Table 3: Amounts of Analysis Data (Revised Disbursement > Income) [1]

Age Group	A_{rep}	Male			Female		
		1994	1999	2004	1994	1999	2004
65 - 69	65	17	17	18	21	22	23
70 - 74	70	13	13	15	17	18	19
75 - 79	75	10	10	11	13	14	15
80 - 84	80	7	8	8	9	10	11
85 -	85	5	5	6	7	7	8

Table 4: Life Expectancy for Two-person Households [1]

		Female				
	Age (A_{rep})	65 - 69 (65)	70 - 74 (70)	75 - 79 (75)	80 - 84 (80)	85 - (85)
Male	65 - 69 (65)	65	65	70	70	75
	70 - 74 (70)	65	70	70	75	75
	75 - 79 (75)	70	70	75	75	80

	80 - 84 (80)	70	75	75	80	80
	85 - (85)	75	75	80	80	85

4. Proposed Method [7] Using Anonymous Data (AD)

4.1. Regression Line to Determine Training Data

Two thresholds of savings (S_{th0} and S_{th}) among the training data are explained by taking an example of single females with own house. The threshold S_{th0} is used to determine if the test data is processed through either primary determination or secondary determination subsequently explained in Section 4.3. On the other hand, threshold S_{th} is utilized to make the training data stated in this section. For each age group, representative age A_{rep} and maximum saving S_{MAX} are shown in Table 5. The column "Rep-Age" indicates the representative age. Here, each value of S_{MAX} is obtained as the maximum savings for each A_{rep} for the data determined as having dissaving risk among the analysis data selected in Section 3.3-(1). In determining each value of S_{MAX} for each A_{rep}, there are several data with greatly high savings despite being determined as dissaving risk. In such cases, their consumption items were investigated to confirm if they were compelled to pay extremely high amounts of money for specific reasons, e.g. renovations, purchasing new furniture, etc. Those data were then discarded for setting S_{MAX}. In Figure 1, S_{MAX} for each A_{rep} is shown as brown square. The horizontal axis indicates A_{rep}, while the vertical axis indicates S_{MAX}. Regression line R_l is plotted as dashed blue line in Figure 1. R_l is shown as Formula (4).

Table 5: Representative Age and Savings for Each Age Group (AD: Female, Own) [1]

Age Group	Rep-Age	Savings			
		Max	Predicted Value	Predicted Error	Threshold
	A_{rep}	S_{MAX}	S_{th0}	$S_{MAX} - S_{th0}$	S_{th}
65 - 69	65	506	939.2	-433.2	1403.7
70 - 74	70	1194	729.5	464.5	1194.0
75 - 79	75	800	519.8	280.2	984.3
80 - 84	80	89	310.1	-221.1	774.6
85 -	85	10	100.4	-90.4	564.9

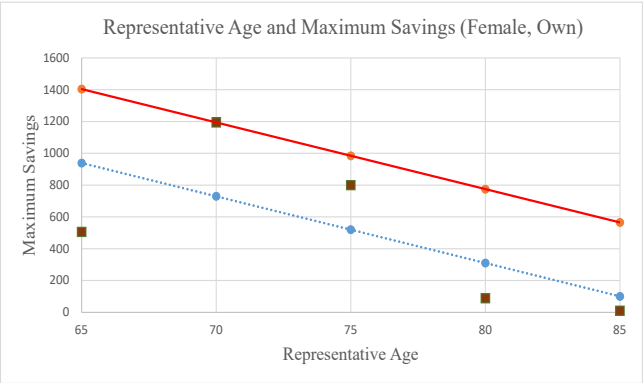


Figure 1: Example of Representative Age and Maximum Savings [1]

$$R_l: S_{MAX} = -41.9A_{rep} + 3665.3 \quad (4)$$

The vertical value obtained through R_l at each A_{rep} is defined as predictive value S_{th0}.

Here, the whole set of training data A is defined as "the set whose savings are less than or equal to those of the line R_l' defined as the vertically upward shift of R_l by the maximum of predicted error at the A_{rep} of its range focused on" [7]. Predicted error is calculated through subtracting S_{th0} from S_{MAX}. The line R_l' is plotted as solid red line in Figure 1 and expressed as Formula (5).

$$R_l': S_{MAX} = -41.9A_{rep} + 4129.8 \quad (5)$$

The maximum of predicted error is 464.5, which is shaded in Table 5. The vertical value obtained through line R_l' at each A_{rep} is defined as threshold S_{th}. On top of the maximum (S_{MAX}), the following three values are also shown in Table 5; predicted value (S_{th0}), predicted error ($S_{MAX} - S_{th0}$), and threshold (S_{th}).

4.2. Training Data and Partial Sets

Provided that D_{rev} is the revised disbursement, I is the amount of annual income, and S is the amount of savings, in order to determine the partial sets of training data, the subsequent two conditions are set [7].

1. Revised disbursement is 1.2 times or more than income. ($D_{rev} \geq 1.2I$)
2. Saving is 0.1 times or less than income. ($S \leq 0.1I$)

These coefficients, 1.2 and 0.1, are determined through trial and error. As defined in Section 4.1, the training data A is the set that satisfies the amounts of savings are not more than S_{th} at A_{rep}. Among the training data A, the set that meets the condition 1 (2, respectively) is denoted as B (C). The partial sets D_1, D_2 of the training data are then defined as Formulas (6) and (7) [7].

$$D_1 = A \cap (B \cup C) \quad (6)$$

$$D_2 = A \cap (\bar{B} \cap \bar{C}) \quad (7)$$

4.3. Two-Step Methods to Determine Dissaving Risk

The data divided into test data in Section 3.3-(2) is used as test data, while the training data classified in Section 3.3-(2) are used as training data in the following two-step methods. According to the value of S_{th0} set in the fashion explained in Section 4.1, the dissaving risk against each test data is evaluated through either (1) primary determination or (2) secondary determination stated as follows:

(1) Primary Determination

The data whose savings are not less than S_{th0} are determined as "no dissaving risk." The results of primary determination are shown in Table 6. Each column is explained as follows:

- "P" provides either "$< P_{th}$" or "$\geq P_{th}$." As stated in Section 3.3-(2), the threshold P_{th} to determine dissaving risk was set to 0.1 (0.3, respectively) for single males and females (Twos) in the previous analysis [7]. The abbreviation "Act"

listed in the second row stands for "Actual," indicating that P is an actual value.

- "Det" stands for determination. This column provides the result of determining dissaving risk, either "Dissaving Risk" or "No Dissaving Risk." For Table 6, mere "No Dissaving Risk" is provided, as explained above.
- "T/F" provides whether determination result is correct or not. If the determination result is correct, this column provides "TRUE," otherwise "FALSE."
- "Result" is classified as two results in Table 6. The results are "Leak Detection 1 (LD1)" and "No Risk 1 (NR1)."
- "Abb" indicates the abbreviations of results as stated above.
- "Data" indicates the amounts of the data.

Table 6: Primary Determination Results

No.	P	Primary Determination Results				
	Act	Det	T/F	Result	Abb	Data
1	$< P_{th}$	No Dissaving Risk	FALSE	Leak Detection 1	LD1	N_{L1}
2	$\geq P_{th}$		TRUE	No Risk 1	NR1	N_{N1}

(2) Secondary Determination

Secondary determination proceeds through each of three methods, against the test data that have not been determined in primary determination yet.

(i) Method 1

For a test data, estimated saving exhaustion term (*ESET*) is calculated as follows. Provided that D_{rev} is the revised disbursement, I is the amount of annual income, and S is the amount of savings, similar as Formula (2), *ESET* can be calculated as using Formula (8).

$$ESET = \frac{S}{D_{rev} - I} \tag{8}$$

A further procedure of method 1 will be explained using Figure 2-(a). Here, three training data whose Euclidean distances to the test data are the first to third shortest in terms of savings are extracted from training data belonging to D_1. The criterion of the training data with first to third shortest determined through trial and error. Among the *ESETs* of the three training data extracted, the shortest one is set as T_{est}. In Figure 2-(a), the training data with the shortest is described as a red circle, while the other two are shown as orange circles. The red arrow is then derived from the red circle to a blue circle, the test data, to set the shortest ESET as T_{est}.

The T_{est} of the test data will then be divided by the life expectancy L_{exp} corresponding to the age group and the year (1994, 1999 or 2004) of the training data. Provided that P_{est} is the ratio of T_{est} to L_{exp} for each data, similar as Formula (3), P_{est} is calculated through Formula (9).

$$P_{est} = \frac{T_{est}}{L_{exp}} \tag{9}$$

If P_{est} is less than P_{th}, the threshold to determine dissaving risk, the test data is determined as having "Dissaving risk." Otherwise as "No dissaving risk." The results of secondary determination are summarized in Table 7. The columns are basically the same as explained in Section 4.3-(1). Minor differences from Table 6 are explained as follows:

- In the column entitled "P_{est}", the abbreviation "Est" is listed, indicating that P_{est} is an estimated value.
- "T/F" provides "TRUE" if the column entitled "Estimated" and "Actual" correspond, otherwise "FALSE."
- "Result" is classified as four categories in Table 7. The categories are expressed as "Correct Detection (CD)," "Incorrect Detection (ID)," "Leak Detection 2 (LD2)" and "No Risk 2 (NR2).

(ii) Method 2

For a test data, three training data whose Euclidean distances to the test data are the first to third shortest in terms of savings are extracted from training data belonging to D_1 or D_2. If training data belonging to D_2 account for at least two out of those three, the test data is determined as having "no dissaving risk." On the other hand, if training data belonging to D_1 account for at least two out of those three, then similar as method 1, three training data whose Euclidean distances to the test data are the first to third shortest in terms of savings are extracted from the training data belonging to D_1. The procedure of method 2 is described in Figure 2-(b), with the case when two training data belonging to D_1. Similar to method 1 and Figure 2-(a), among the training data belonging to D_1, the smallest is set as T_{est}. Then the test data is determined as having "dissaving risk" if P_{est} is less than P_{th}.

(iii) Method 3

For a test data, three training data whose Euclidean distances to the test data are the first to third shortest in terms of the two-dimensional vector space composed of income and savings are extracted from the training data belonging to D_1 or D_2. Similar to method 2, if training data belonging to D_2 account for at least two out of those three, the test data is determined as having "no dissaving risk." Meanwhile, if training data belonging to D_1 account for at least two out of those three, T_{est} is then obtained as explained in methods 1 and 2. The procedure of method 3 is described in Figure 2-(b). Then the test data are determined as having "dissaving risk" if P_{est} is less than P_{th}.

4.4. Measures for Performance Evaluation

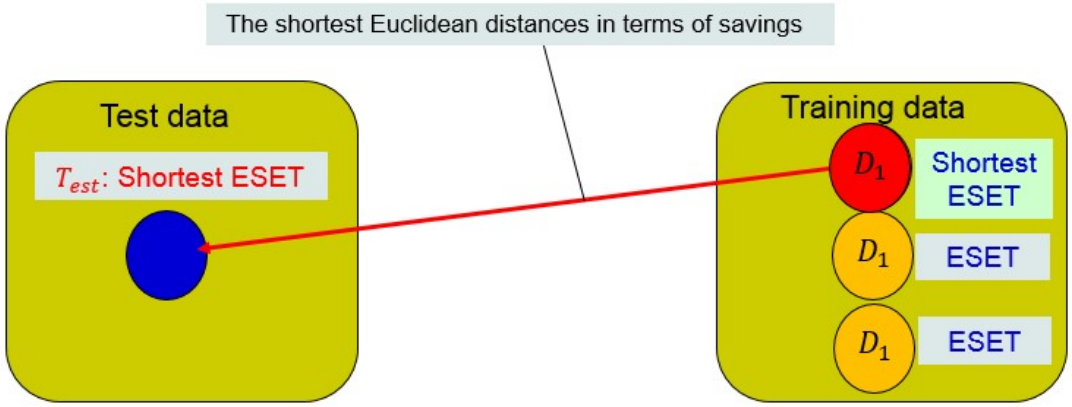
The data determined as the total six types of the results are summarized in Tables 6 and 7. For convenience, LD1 and LD2 (NR1 and NR2, respectively) are collectively named as LD (NR). Here, correct detection rate (CDR), leak detection rate (LDR) and correct judgment rate (CJR) are calculated as follows.

- CDR means the ratio of the amount of CD to the sum of the data determined as having dissaving risk in the secondary determination. For the denominator, these data are determined as having "Dissaving Risk" in the column entitled "Det" in Table 7 corresponding to the sum of the amount of CD and that of ID. CDR is calculated with Formula (10):

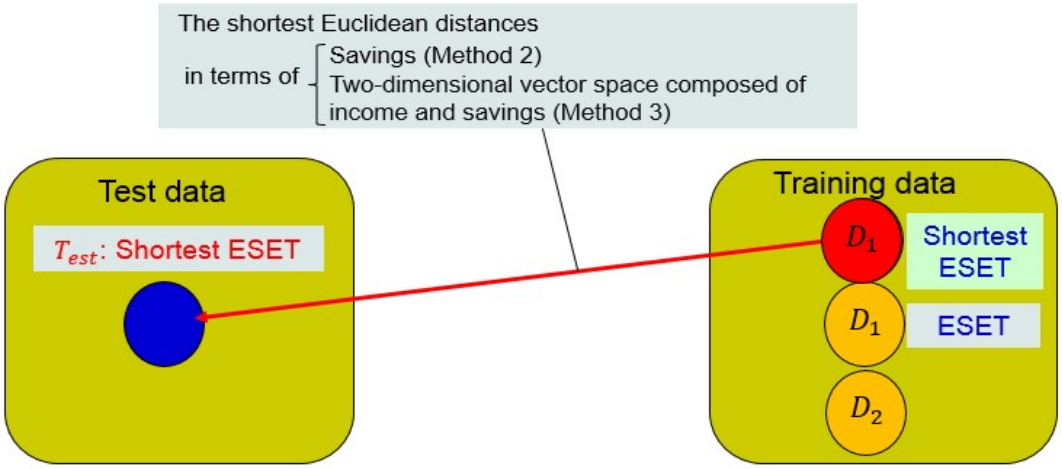
$$CDR = \frac{100N_C}{N_C + N_I} \tag{10}$$

Table 7: Secondary Determination Results

No.	P_{est}	P	Secondary Determination Results				
	Est	Act	Det	T/F	Result	Abb	Data
1	$< P_{th}$	$< P_{th}$	Dissaving Risk	TRUE	Correct Detection	CD	N_C
2	$< P_{th}$	$\geq P_{th}$		FALSE	Incorrect Detection	ID	N_I
3	$\geq P_{th}$	$< P_{th}$	No Dissaving Risk	FALSE	Leak Detection 2	LD2	N_{L2}
4	$\geq P_{th}$	$\geq P_{th}$		TRUE	No Risk 2	RD2	N_{N2}



(a) Method 1



(b) Method 2 and 3

Figure 2: The Procedures of Three Methods for Secondary Determination

- LDR means the ratio of the amount of LD among the data actually determined as having dissaving risk. For the denominator, these data are expressed as "$< P_{th}$" in the column entitled "P_{est}" and "Act" in Tables 6 and 7 corresponding to the sum of the amount of CD and that of LD. LDR is calculated with Formula (11):

$$LDR = \frac{100(N_{L1}+N_{L2})}{N_R+N_{L1}+N_{L2}} \quad (11)$$

- CJR means the ratio of the sum of amount of CD and the amount of NR to the number of all the analysis data. For the numerator, these data are expressed as "TRUE" in the column entitled "T/F" in Tables 6 and 7. CJR is calculated with Formula (12):

$$CJR = \frac{100(N_C+N_{N1}+N_{N2})}{N_C+N_I+N_{L1}+N_{L2}+N_{N1}+N_{N2}} \quad (12)$$

5. Evaluation of Proposed Method through Random Data (RD)

5.1. Division of Anonymous Data (AD)

The use of AD could be controversial in terms of the safety of its anonymity. In order to enforce the anonymity of AD, random data (RD) is generated along with the AD. For the purpose of performance evaluation, RD is then compared with the case of analyzing mere AD as it is.

Determination method is performed by setting each AD as test data and RD as training data. As a preparation of generating RD, AD (single: 171 males, 1,203 females, Twos: 2,367 households) is divided through the following procedures regarding savings and revised disbursements.

(a) The analysis data is just divided into three datasets according to the amount of saving; high-savings (H-s), medium-savings (M-s), and low-savings (L-s). The thresholds between the datasets are determined as follows:

- In AD, the amount of maximum savings is provided with 5.5 million yen. The threshold between M-s and H-s is set as the average of 0 yen and 5.5 million yen: 2.75 million yen.
- Similarly, the threshold between L-s and M-s is set as the average of 0 yen and 2.75 million yen: 1.375 million yen.

(b) Each of the three datasets divided in (a) is respectively divided into two datasets; high-disbursement (H-d), medium-disbursement (M-d), and low-disbursement (L-d) in accordance with the amount of revised disbursement. The threshold D_{rev_th} is set as double of the average of respective datasets. From this viewpoint, for example of single females with own houses, D_{rev_th} for L-s (M-s and H-s, respectively) is set as 483.7 (718.4 and 926.7). These dividing methods are determined thorough trial and error. The amounts of analysis data through the procedure are shown in Table 8. Nevertheless, there could be the case where it would be impossible to divide the test data into any smaller groups. Such cases result from insufficient amount of data after they are divided in accordance with savings and revised disbursements. Taking an example of the single males with rent house, there are only nine data whose savings are over 1.375 million yen (M-s and H-s). In such cases, these groups are not divided into smaller ones.

5.2. Generation of Random Data (RD)

For the analysis data divided into the six attribute datasets as stated in Section 3.3-(3), in accordance with respective savings and revised disbursement, RD are generated along with the following three variables; income, revised disbursement, and savings. Through *mvnrnd* function implemented in MatLAB [18], k sets ($k > 5$) of preliminary RD is generated for each of the respective subdivided datasets in order to avoid failing to extract five sets of RD (amount of data: N) owing to the shortage of preliminary RD.

Table 8: The Amounts of Datasets Divided in Accordance with Saving and Revised Disbursement [1]

(a) Male, Rent

S\D	L-d	H-d	Total
L-s			52
M-s			9
H-s			
Total			61

(b) Male, Own

S\D	L-d	H-d	Total
L-s	58	5	63
M-s			26
H-s	18	3	21
Total			110

(c) Female, Rent

S\D	L-d	H-d	Total
L-s	269	10	279
M-s			37
H-s			15
Total			331

(d) Female, Own

S\D	L-d	H-d	Total
L-s	529	43	572
M-s	171	9	180
H-s	110	10	120
Total			872

(e) Twos, Rent

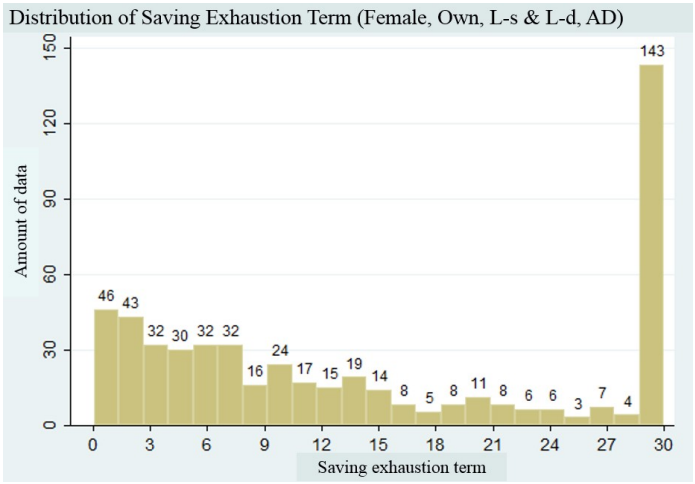
S\D	L-d	H-d	Total
L-s	255	10	265
M-s			32
H-s	5	4	9
Total			306

(f) Twos, Own

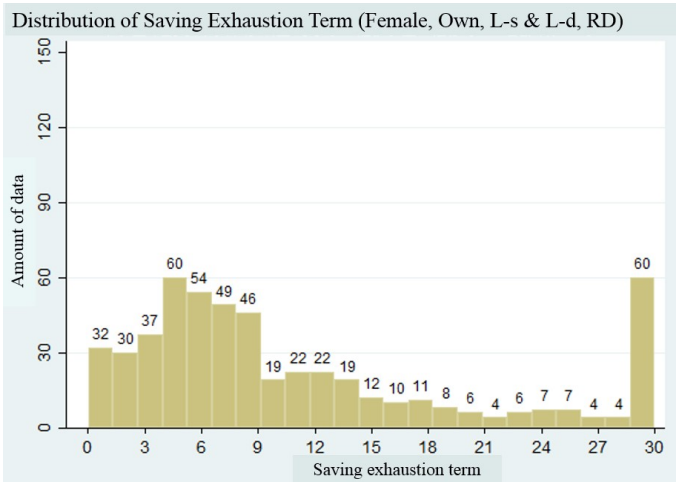
S\D	L-d	H-d	Total
L-s	1388	48	1436
M-s	373	46	419
H-s	172	34	206
Total			2061

Given that income is I, revised disbursement is D_{rev} and amount of savings is S, negative values can be automatically and randomly generated for I, D_{rev} and S. Therefore, the data that absolutely meets $I \geq 0,\ D_{rev} \geq 0,\ S \geq 0$ are sequentially extracted from the preliminary RD generated through *mvnrnd* function [18]. Here, the five sets of RD extracted through these procedures for the respective divided datasets are defined as candidate RD.

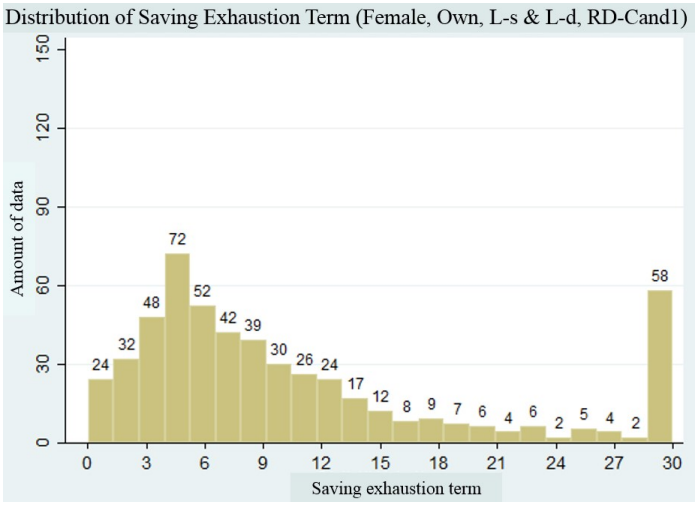
Next, one of the five sets of the candidate RD must be chosen to use for the subsequent analysis. In terms of savings exhaustion term (SET), T_{est} for five sets of RD are compared with T for AD by plotting them. Through visual judgment, one set of RD is chosen whose outline is the most similar to AD. How to determine RD out of the five candidate RD are explained by using the actual example of datasets for female with own house divided into L-s and L-D. SET (T for AD and T_{est} for candidate RD) are shown in Figure 3. Here, T and T_{est} shown in Figure 3 range from 0 to 30 years. If these terms are beyond 30 years, T or T_{est} for those data are regarded as just 30 years. Among these values, T for AD is



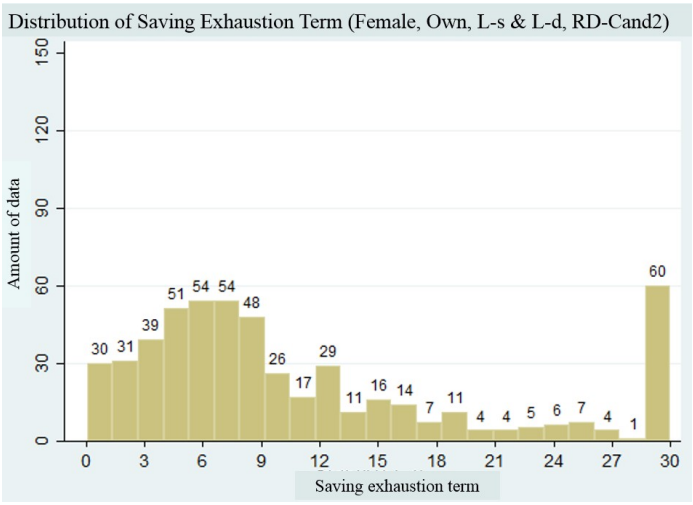
(a) AD



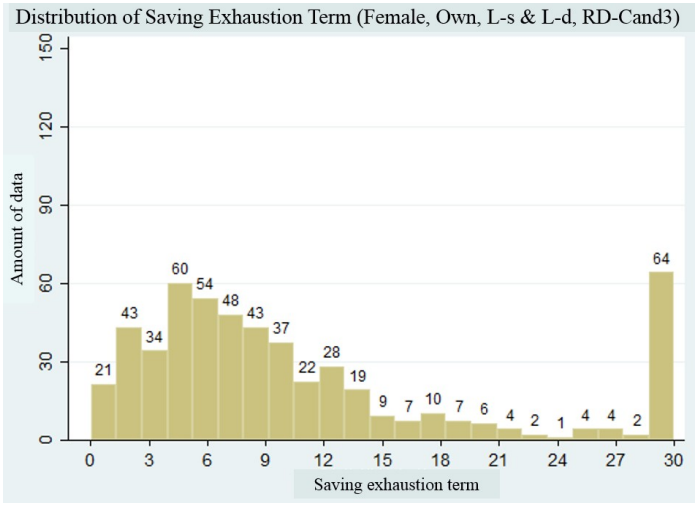
(b) RD



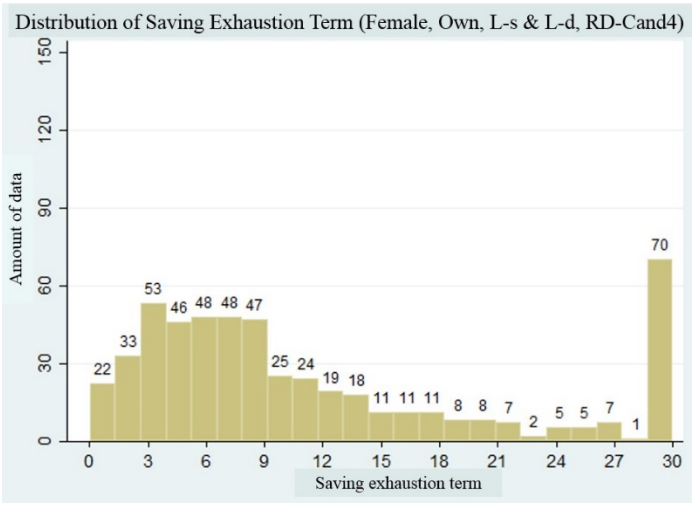
(c)-1 RD-Cand1



(c)-2 RD-Cand2



(c)-3 RD-Cand3



(c)-4 RD-Cand4

Figure 3: Distribution of Saving Exhaustion Term (Female, Own, L-s and L-d)

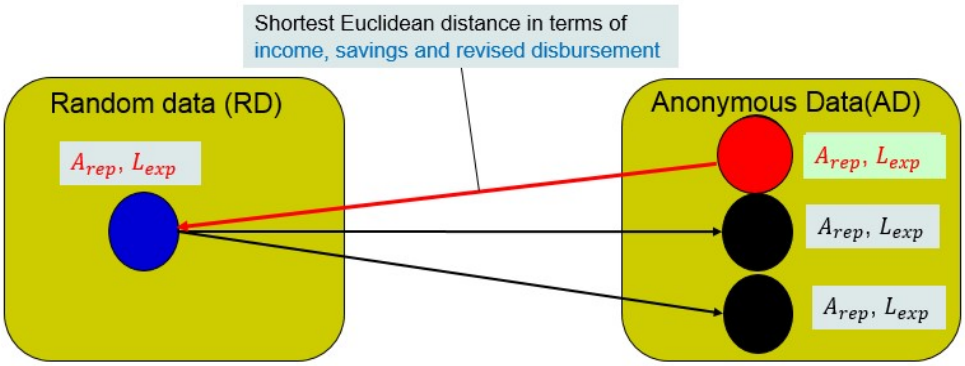


Figure 4: The Procedure of Setting Age Group and Life Expectancy on Random Data (RD) obtained from Anonymous Data (AD)

depicted in Figure 3-(a), while (b) T_{est} for the selected RD is shown in Figure 3-(b). On the other hand, T_{est} for the candidate RD that are not selected for the subsequent analysis are shown in Figure 3-(c). These candidate RD are named as RD-Cand1, Rd-Cand2, Rd-Cand3 and Rd-Cand4, respectively. Comparing these values, (b)RD resembles (a)AD more than (c)the rest of candidate RD do. Through these procedures, the selected RD for each of the six attribute datasets is utilized for the following analysis.

5.3. Setting Age Group and Life Expectancy to RD

RD generated and chosen through the procedures described in Section 5.2 already have three attributes; income I, revised disbursement D_{rev} and savings S. Nevertheless, age group A_{rep} and life expectancy L_{exp} of RD have not been assigned yet. Therefore, it would be necessary to determine these attributes. Here, in the three-dimensional space composed of I, D_{rev}, S, for one RD, one among the AD is extracted whose Euclidean distance is shortest. The age group and life expectancy of the AD are then assigned to the RD. The outline of this procedure is shown in Figure 4, similar as that of secondary determination explained in Section 4.3-(2) and shown in Figure 2. One AD with shortest Euclidean distance is denoted as red circle. Then a red arrow starts from the red circle to a blue circle (RD) to set A_{rep} and L_{exp}. The amounts of AD and those of RD determined through this procedure are shown in Table 9. The column entitled "AD (Test)" indicates the amount of AD data used as test data, whereas the column entitled "RD (Train)" means that of RD used as training data.

Table 9: Amounts of AD and RD [1]

(a) Male, Rent

Age Group	AD (Test)	RD (Train)
65 - 69	22	24
70 - 74	20	20
75 - 79	11	6
80 - 84	6	9
85 -	2	2

(b) Male, Own

Age Group	AD (Test)	RD (Train)
65 - 69	37	29
70 - 74	26	38
75 - 79	32	31
80 - 84	10	9
85 -	5	3

(c) Female, Rent

Age Group	AD (Test)	RD (Train)
65 - 69	101	84
70 - 74	119	121
75 - 79	72	100
80 - 84	29	20
85 -	10	6

(d) Female, Own

Age Group	AD (Test)	RD (Train)
65 - 69	312	324
70 - 74	286	304
75 - 79	172	153
80 - 84	84	78
85 -	18	13

(e) Twos, Rent

Age Group	AD (Test)	RD (Train)
65 - 69	154	146
70 - 74	104	114
75 - 79	38	40
80 - 84	9	5
85 -	1	1

(f) Twos, Own

Age Group	AD (Test)	RD (Train)
65 - 69	1077	1129
70 - 74	673	645
75 - 79	240	227
80 - 84	67	51
85 -	4	9

5.4. Analysis Results

The validity of AD used along with RD as training data is analyzed in the similar fashion as explained in Section 4.4. Here, RD generated through the method described in Sections 5.1 and 5.2 is utilized as training data, while AD is used as test data. The amounts of partial sets of D_1, D_2 are shown in Table 10.

There are two slight differences from the previous analysis explained in Section 4.4. Firstly, the threshold S_{MAX} is determined in a subtle different fashion. Similar to the AD, there are several RD with substantially high savings in spite of being determined as dissaving risk. Unlike AD, however, RD have no attributions of consumption items. Here, the distributions of savings for each age group are investigated and plotted. Then higher savings that apparently seem isolated from the other lower ones are then discarded for setting S_{MAX}. The thresholds determined are shown in Table 11. Another modification is the threshold P_{th} to determinate dissaving risk for single males and females. In the previous analysis, as stated in Section 3.3-(2), the value of P_{th} was set to 0.1 for single-person households, as explained in Section 3.3-(2). In this analysis, through trial and error, the threshold is set to 0.15 for single households.

Table 10: Amounts of Partial Sets of D_1 and D_2 [1]

(a) Male, Rent

Age Group	Partial Sets		
	D_1	D_2	Total
65 - 69	5	0	5
70 - 74	3	1	4
75 - 79	1	0	1
80 - 84	0	0	0
85 -	0	0	0

(b) Male, Own

Age Group	Partial Sets		
	D_1	D_2	Total
65 - 69	13	3	16
70 - 74	17	0	17
75 - 79	13	2	15
80 - 84	5	1	6
85 -	2	0	2

(c) Female, Rent

Age Group	Partial Sets		
	D_1	D_2	Total
65 - 69	56	12	68
70 - 74	49	20	69
75 - 79	34	4	38
80 - 84	9	2	11
85 -	2	0	2

(d) Female, Own

Age Group	Partial Sets		
	D_1	D_2	Total
65 - 69	152	33	185
70 - 74	147	33	180
75 - 79	74	14	88
80 - 84	34	8	42
85 -	8	1	9

(e) Twos, Rent

Age Group	Partial Sets		
	D_1	D_2	Total
65 - 69	88	39	127
70 - 74	60	18	78
75 - 79	14	5	19
80 - 84	0	0	0
85 -	1	0	1

(f) Twos, Own

Age Group	Partial Sets		
	D_1	D_2	Total
65 - 69	824	212	1036
70 - 74	454	119	573
75 - 79	124	30	154
80 - 84	25	4	29
85 -	2	0	2

Table 11: Threshold for Analysis Using RD [1]

(a) Male, Rent

Age Group	Savings		
	S_{MAX}	S_{th0}	S_{th}
65 - 69	200.0	168.7	200.0
70 - 74	98.6	161.0	192.2
75 - 79	184.5	153.3	184.5
80 - 84			184.5
85 -			184.5

(b) Male, Own

Age Group	Savings		
	S_{MAX}	S_{th0}	S_{th}
65 - 69	423.8	547.3	978.9
70 - 74	991.3	559.7	991.3
75 - 79	79.8	572.2	1003.7
80 - 84	769.0	584.6	1016.1
85 -			1028.6

(c) Female, Rent

Age Group	Savings		
	S_{MAX}	S_{th0}	S_{th}
65 - 69	491.3	593.6	749.7
70 - 74	563.0	487.7	643.8
75 - 79	537.9	381.8	537.9
80 - 84	146.6	275.9	432.0
85 -			432.0

(d) Female, Own

Age Group	Savings		
	S_{MAX}	S_{th0}	S_{th}
65 - 69	1034.2	1056.7	1184.5
70 - 74	1052.0	924.2	1052.0
75 - 79	603.3	791.7	919.4
80 - 84	742.0	659.1	786.9
85 -			786.9

(e) Twos, Rent

Age Group	Savings		
	S_{MAX}	S_{th0}	S_{th}
65 - 69	1647.8	1414.0	1647.8
70 - 74	951.4	1040.7	1274.4
75 - 79	333.5	667.3	901.1
80 - 84			527.8
85 -	109.9	-79.3	154.5

(f) Twos, Own

Age Group	Savings		
	S_{MAX}	S_{th0}	S_{th}
65 - 69	4757.9	4923.1	5236.3
70 - 74	3988.0	3674.8	3988.0
75 - 79	2578.8	2426.5	2739.7
80 - 84	593.1	1178.2	1491.4
85 -	213.8	-70.1	243.1

In order to compare the analysis results for this section with the previous ones [7], both of the analysis results are summarized in Tables 12 and 13, respectively. See Appendix on the detailed analysis results; the amounts of data determined as LD1, NR1, CD, ID, LD2 and NR2, and the detailed results for each age group.

In order to compare the differences of results between the analysis using RD (Table 12) and that utilizing mere AD (Table 13), their differences are summarized in Table 14. Here, these values are calculated by subtracting those with AD from those with RD. For CDR and CJR, the bigger they are, the better the result is. Therefore, if these values are positive, they are colored blue. Otherwise, if the values are negative, they are colored red.

For LDR, meanwhile, the smaller it is, the better the result is. Thus, on the contrary to CDR and CJR, if these values are negative, they are colored blue. Otherwise, if the values are positive, they are colored red.

Table 12: Analysis Results with RD [1]

(a) Rent

Methods [%]	Single						Two-person Household		
	Male			Female					
	CDR	LDR	CJR	CDR	LDR	CJR	CDR	LDR	CJR
Method 1	60.0	0.0	83.6	46.1	11.9	76.7	41.4	16.7	68.3
Method 2	60.0	0.0	83.6	47.2	11.9	77.6	38.3	43.1	68.3
Method 3	65.2	0.0	86.9	47.2	11.9	77.6	38.2	41.7	68.0

(b) Own

Methods [%]	Single						Two-person Household		
	Male			Female					
	CDR	LDR	CJR	CDR	LDR	CJR	CDR	LDR	CJR
Method 1	57.1	36.8	85.5	38.2	30.3	80.5	31.8	21.1	63.8
Method 2	57.1	36.8	85.5	39.1	35.3	81.4	32.3	29.0	66.0
Method 3	60.0	36.8	86.4	42.4	31.9	83.0	36.0	26.2	70.0

Table 13: Analysis Results with only AD [7]

(a) Rent

Methods [%]	Single						Two-person Household		
	Male			Female					
	CDR	LDR	CJR	CDR	LDR	CJR	CDR	LDR	CJR
Method 1	100.0	60.0	90.0	56.1	8.0	87.8	32.1	12.9	59.9
Method 2	100.0	60.0	90.0	60.6	20.0	89.0	39.5	50.0	71.7
Method 3	100.0	60.0	90.0	60.0	16.0	89.0	33.3	58.8	68.4

(b) Own

Methods [%]	Single						Two-person Household		
	Male			Female					
	CDR	LDR	CJR	CDR	LDR	CJR	CDR	LDR	CJR
Method 1	62.5	0.0	94.4	30.4	22.2	83.5	36.5	25.9	68.1
Method 2	50.0	40.0	90.7	32.4	38.9	86.2	36.2	47.7	70.4
Method 3	50.0	40.0	90.7	38.6	38.9	88.8	37.3	49.5	71.5

Table 14: Differences of Results between RD (Table 12) and only AD (Table 13)

(a) Rent

Methods [%]	Single						Two-person Household		
	Male			Female					
	CDR	LDR	CJR	CDR	LDR	CJR	CDR	LDR	CJR
Method 1	-40.0	-60.0	-6.4	-10.0	3.9	-11.1	9.3	3.8	8.4
Method 2	-40.0	-60.0	-6.4	-13.4	-8.1	-11.4	-1.2	-6.9	-3.4
Method 3	-34.8	-60.0	-3.1	-12.8	-4.1	-11.4	4.9	-17.1	-0.4

(b) Own

Methods [%]	Single						Two-person Household		
	Male			Female					
	CDR	LDR	CJR	CDR	LDR	CJR	CDR	LDR	CJR
Method 1	-5.4	36.8	-8.9	7.8	8.1	-3.0	-4.7	-4.8	-4.3
Method 2	7.1	-3.2	-5.2	6.7	-3.6	-4.8	-3.9	-18.7	-4.4
Method 3	10.0	-3.2	-4.3	3.8	-7.0	-5.8	-1.3	-23.3	-1.5

From the viewpoints of CDR and CJR, for the data with rent house, deterioration is observed in most cases except some for Twos: CDR and CJR for method 1, and CDR for method 3. For the data with own house, CDR is improved for single households except for males for method 1, while it becomes degraded for Twos. CJR becomes deteriorated for both single-person and two-person households. On the other hand, LDR is improved for most of the cases except those with method 1 for Twos with rent and males with own.

5.5. Considerations

Comparing the three methods using RD (Table 12), from the viewpoints of CDR and CJR, as a whole, CDR and CJR become best with method 3. However, for the data with rent house, CDR and CJR equally show similar performance for methods 2 and 3. For Twos, method 1 outperforms the others. From the viewpoint of LDR, on the other hand, method 1 works best since LDR is the lowest value among the three methods for all the cases. Therefore, method 3 is best in light of CDR and CJR, whereas there is room for improvement for LDR. Meanwhile, method 1 is better than the others considering LDR, while it leaves much to be desired for CDR and CJR.

In the light of the differences of analysis results between RD and AD (Table 14), as a whole, LDR is improved, while CDR and CJR are deteriorated. However, these analysis results are dependent on RD and can be variable. As limitations of analysis, AD are required to be returned within the designated periods [8]. Moreover, there could be insufficient analysis data with certain attributes, especially males. How to cope with the situation where there is not enough data is our requirement for the further analysis.

Nevertheless, these two types of analysis results show the similar results for all the three methods. Therefore, it could be concluded that using RD along with the AD for evaluating the performance of the proposed method [7] could be effective.

6. Conclusion

In the reported study [7], we proposed a method to detect dissaving risk of people at the age of sixty-five or older for AD. In order to strengthen the anonymity of the data, RD used along with the AD is generated and then compared with the previous analysis, with a view to performance evaluation [1]. Here, AD is set as test data, whereas RD is used as training data. As a result of analysis, from the viewpoints of CDR and CJR, for the data with rent house, degradation was observed with most cases. On the other hand, for the data with own house, CDR was improved with most cases for single-person households, while it became deteriorated for Twos. CJR dropped for both single-person and two-person households. Meanwhile, LDR was improved for most of the cases. As a whole, it could be concluded that using RD might be as effective as using AD for evaluating the performance of the proposed method [7].

For future work, while CJR exceeded 80% for single-person households, it attained as high as approximately 70% for two-person households. Therefore, CDR and CJR especially for the data of two-person households must be improved. How to deal with the cases with insufficient analysis data for further analysis is included in our future works.

Acknowledgment

This research was supported by the Research Institute of Science and Technology for Society within the Japan Science and Technology Agency.

References

[1] Y. Yokoyama and Y. Yoshitomi, “Assessment of Dissaving Risk against Life Expectancy for Elderly People through Anonymous Data and Random Data,” Proc. of the 20th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019), 274-279, 2019, doi: 10.1109/SNPD.2019.8935718.

[2] H. Ohba, Y. Kadoya, and J. Narumoto, “Influence on the household change Caused by Dementia Decline -Two-year Longitudinal Analysis Using JSTAR Data-” (in Japanese), Proc. of the 20th Japan Society of Geriatric & Gerontological Behavioral Sciences, 2017.

[3] Aomori Prefectural Industrial Technology Research Center, etc., “Dementia Prediciton Program and Method” (in Japanese), Japan patent JP2015-138488A, 2015-07-31.

[4] Toshiba Corporation, "Dementia Determination Device, Sysyem, Method, and Program" (in Japanese), Japan patent JP2017-104289A, 2017-06-15.
[5] e-Stat: Search for statistics (URL): https://www.e-stat.go.jp/en/stat-search/database?page=1, 2021-4-2.
[6] Y. Yokoyama and Y. Yoshitomi, "A Method to Assess Dissaving Risk against Life Expectancy for Elderly People," Proc. of the 20th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019), 262-267, 2019, doi: 10.1109/SNPD.2019.8935809.
[7] Y. Yokoyama and Y. Yoshitomi, "A Method to Determine Dissaving Risk of Elderly People" (in Japanese), Abstracts of the 2019 Spring National Conference of Operations Research Society of Japan, 240-241, 2-F-11, 2019.
[8] National Science Center: Use of Anonymous Data (URL, in Japanese), https://www.nstac.go.jp/services/anonymity.html, 2021-4-2.
[9] T. Fujita, S. Ogano, and J. Narumoto, "Dementia and Information" (in Japanese), ISBN978-4-326-44976-7, Keisou Shobou, 150, 2019.
[10] K. L. Triebel, R. Martin, H. R. Griffith, J. Marceaux, O. C. Okonkwo, L. Harrell, D. Clark, J. Brockington, A. Bartolucci and D. C. Marson, "Declining Financial Capacity in Mild Cognitive Impairment: A one-year Longitudinal Study", Neurology, **73**, 928-934, 2009, doi: 10.1212/WNL.0b013e3181b87971
[11] D. Bisdee, D. Price and T. Daly, "Coping with Age-related Threats to Role Identity: Older Couples and the Management of Household Money," Journal of Community & Applied Social Psychology, **23**, 505-518, 2013, doi: 10.1002/casp.2149
[12] M. S. Finke, J. S. Howe and S. J. Huston, "Old Age and the Decline in Financial Literacy," Management Science, **63**(1), 213-230, 2017, doi: 10.2139/ssrn.1948627
[13] J. W. Hsu and R. Willies, "Dementia Risk and Financial Decision Making by Older Households: The Impact of Information," Journal of Human Capital, **7**(4), 340-377, 2013, doi: 10.2139/ssrn.2339225
[14] H. Sakurai and T. Moriwaki, "The Effectiveness of the Default Risk Assessment Based on the Financial Ratios" (in Japanese), Journal of Economics & Business Administration, **214**(2), 1-17, 2016.
[15] T. Izumi, "Analysis on Consumption by Elder Family Type of Household based on the National Survey of Family Income and Expenditure" (in Japanese), Kaetsu University Research Review, **59**(2), 55-67, 2016.
[16] National Statistic Center (URL), https://www.nstac.go.jp/en/, 2021-4-2.
[17] Life expectancies at specified ages (URL), https://www.mhlw.go.jp/english/database/db-hw/lifetb04/1.html, 2021-4-2.
[18] MatLAB: mvnrnd (URL), https://jp.mathworks.com/help/stats/mvnrnd.html?lang=en/index.html, 2021-4-2.

Appendix

The detailed analysis results of Table 12 for rent for males, own for males, rent for females, own for females, rent for Twos, and own for Twos are shown in Table A.1 (a), (b), (c), A.2 (a), (b), (c), A.3 (a), (b), (c), A.4 (a), (b), (c) , A.5 (a), (b), (c) , and A.6 (a), (b), (c), respectively.

Table A.1 (a): The Detailed Analysis Results (Male, Rent: Method 1)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	0	15	15	5	2	0	0	7	22	71.4	0.0	90.9
70-74	0	13	13	3	4	0	0	7	20	42.9	0.0	80.0
75-79	0	3	3	6	2	0	0	8	11	75.0	0.0	81.8
80-84	0	3	3	1	2	0	0	3	6	33.3	0.0	66.7
85-	0	1	1	0	0	0	1	1	2	-	-	100.0
T	0	35	35	15	10	0	1	26	61	60.0	0.0	83.6

Table A.1 (b): The Detailed Analysis Results (Male, Rent: Method 2)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	0	15	15	5	2	0	0	7	22	71.4	0.0	90.9
70-74	0	13	13	3	4	0	0	7	20	42.9	0.0	80.0
75-79	0	3	3	6	2	0	0	8	11	75.0	0.0	81.8
80-84	0	3	3	1	2	0	0	3	6	33.3	0.0	66.7
85-	0	1	1	0	0	0	1	1	2	-	-	100.0
T	0	35	35	15	10	0	1	26	61	60.0	0.0	83.6

Table A.1 (c): The Detailed Analysis Results (Male, Rent: Method 3)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	0	15	15	5	2	0	0	7	22	71.4	0.0	90.9
70-74	0	13	13	3	3	0	1	7	20	50.0	0.0	85.0
75-79	0	3	3	6	1	0	1	8	11	85.7	0.0	90.9
80-84	0	3	3	1	2	0	0	3	6	33.3	0.0	66.7
85-	0	1	1	0	0	0	1	1	2	-	-	100.0
T	0	35	35	15	8	0	3	26	61	65.2	0.0	86.9

Table A.2 (a): The Detailed Analysis Results (Male, Own: Method 1)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	2	24	26	5	3	0	3	11	37	62.5	28.6	86.5
70-74	3	12	15	5	3	1	2	11	26	62.5	44.4	73.1
75-79	0	26	26	1	2	0	3	6	32	33.3	0.0	93.8
80-84	1	7	8	1	1	0	0	2	10	50.0	50.0	80.0
85-	0	3	3	0	0	0	2	2	5	-	-	100.0
T	6	72	78	12	9	1	10	32	110	57.1	36.8	85.5

Table A.2 (b): The Detailed Analysis Results (Male, Own: Method 2)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	2	24	26	5	3	0	3	11	37	62.5	28.6	86.5
70-74	3	12	15	5	3	1	2	11	26	62.5	44.4	73.1
75-79	0	26	26	1	2	0	3	6	32	33.3	0.0	93.8
80-84	1	7	8	1	1	0	0	2	10	50.0	50.0	80.0
85-	0	3	3	0	0	0	2	2	5	-	-	100.0
T	6	72	78	12	9	1	10	32	110	57.1	36.8	85.5

Table A.2 (c): The Detailed Analysis Results (Male, Own: Method 3)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	2	24	26	5	3	0	3	11	37	62.5	28.6	86.5
70-74	3	12	15	5	3	1	2	11	26	62.5	44.4	73.1
75-79	0	26	26	1	1	0	4	6	32	50.0	0.0	96.9
80-84	1	7	8	1	1	0	0	2	10	50.0	50.0	80.0
85-	0	3	3	0	0	0	2	2	5	-	-	100.0
T	6	72	78	12	8	1	11	32	110	60.0	36.8	86.4

Table A.3 (a): The Detailed Analysis Results (Female, Rent: Method 1)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	1	31	32	20	19	1	29	69	101	51.3	9.1	79.2
70-74	2	50	52	26	23	3	15	67	119	53.1	16.1	76.5
75-79	0	35	35	10	14	1	12	37	72	41.7	9.1	79.2
80-84	0	15	15	2	12	0	0	14	29	14.3	0.0	58.6
85-	0	6	6	1	1	0	2	4	10	50.0	0.0	90.0
T	3	137	140	59	69	5	58	191	331	46.1	11.9	76.7

Table A.3 (b): The Detailed Analysis Results (Female, Rent: Method 2)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	1	31	32	20	16	1	32	69	101	55.6	9.1	82.2
70-74	2	50	52	26	23	3	15	67	119	53.1	16.1	76.5
75-79	0	35	35	10	14	1	12	37	72	41.7	9.1	79.2
80-84	0	15	15	2	12	0	0	14	29	14.3	0.0	58.6
85-	0	6	6	1	1	0	2	4	10	50.0	0.0	90.0
T	3	137	140	59	66	5	61	191	331	47.2	11.9	77.6

Table A.3 (c): The Detailed Analysis Results (Female, Rent: Method 3)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	1	31	32	20	16	1	32	69	101	55.6	9.1	82.2
70-74	2	50	52	26	23	3	15	67	119	53.1	16.1	76.5
75-79	0	35	35	10	14	1	12	37	72	41.7	9.1	79.2
80-84	0	15	15	2	11	0	1	14	29	15.4	0.0	62.1
85-	0	6	6	1	2	0	1	4	10	33.3	0.0	80.0
T	3	137	140	59	66	5	61	191	331	47.2	11.9	77.6

Table A.4 (a): The Detailed Analysis Results (Female, Own: Method 1)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	148	152	37	45	11	67	160	312	45.1	28.8	80.8
70-74	5	127	132	26	48	9	71	154	286	35.1	35.0	78.3
75-79	1	77	78	12	22	3	57	94	172	35.3	25.0	84.9
80-84	0	40	40	8	19	1	16	44	84	29.6	11.1	76.2
85-	0	7	7	0	0	2	9	11	18	-	100.0	88.9
T	10	399	409	83	134	26	220	463	872	38.2	30.3	80.5

Table A.4 (b): The Detailed Analysis Results (Female, Own: Method 2)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	148	152	36	41	12	71	160	312	46.8	30.8	81.7
70-74	5	127	132	24	47	11	72	154	286	33.8	40.0	78.0
75-79	1	77	78	11	16	4	63	94	172	40.7	31.3	87.8
80-84	0	40	40	6	16	3	19	44	84	27.3	33.3	77.4
85-	0	7	7	0	0	2	9	11	18	-	100.0	88.9
T	10	399	409	77	120	32	234	463	872	39.1	35.3	81.4

Table A.4 (c): The Detailed Analysis Results (Female, Own: Method 3)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	148	152	39	38	9	74	160	312	50.6	25.0	83.7
70-74	5	127	132	24	47	11	72	154	286	33.8	40.0	78.0
75-79	1	77	78	12	15	3	64	94	172	44.4	25.0	89.0
80-84	0	40	40	6	10	3	25	44	84	37.5	33.3	84.5
85-	0	7	7	0	0	2	9	11	18	-	100.0	88.9
T	10	399	409	81	110	28	244	463	872	42.4	31.9	83.0

Table A.5 (a): The Detailed Analysis Results (Twos, Rent: Method 1)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	35	39	34	43	4	34	115	154	44.2	19.0	66.9
70-74	1	34	35	23	31	2	13	69	104	42.6	11.5	67.3
75-79	1	18	19	3	9	0	7	19	38	25.0	25.0	73.7
80-84	0	6	6	0	2	0	1	3	9	0.0	-	77.8
85-	0	1	1	0	0	0	0	0	1	-	-	100.0
T	6	94	100	60	85	6	55	206	306	41.4	16.7	68.3

Table A.5 (b): The Detailed Analysis Results (Twos, Rent: Method 2)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	35	39	25	31	13	46	115	154	44.6	40.5	68.8
70-74	1	34	35	14	25	11	19	69	104	35.9	46.2	64.4
75-79	1	18	19	2	8	1	8	19	38	20.0	50.0	73.7
80-84	0	6	6	0	2	0	1	3	9	0.0	-	77.8
85-	0	1	1	0	0	0	0	0	1	-	-	100.0
T	6	94	100	41	66	25	74	206	306	38.3	43.1	68.3

Table A.5 (c): The Detailed Analysis Results (Twos, Rent: Method 3)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	35	39	23	37	15	40	115	154	38.3	45.2	63.6
70-74	1	34	35	16	24	9	20	69	104	40.0	38.5	67.3
75-79	1	18	19	3	6	0	10	19	38	33.3	25.0	81.6
80-84	0	6	6	0	1	0	2	3	9	0.0	-	88.9
85-	0	1	1	0	0	0	0	0	1	-	-	100.0
T	6	94	100	42	68	24	72	206	306	38.2	41.7	68.0

Table A.6 (a): The Detailed Analysis Results (Twos, Own: Method 1)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	98	102	190	392	47	346	975	1077	32.6	21.2	58.9
70-74	1	111	112	98	194	23	246	561	673	33.6	19.7	67.6
75-79	1	74	75	19	69	5	72	165	240	21.6	24.0	68.8
80-84	0	43	43	3	9	1	11	24	67	25.0	25.0	85.1
85-	1	3	4	0	0	0	0	0	4	-	100.0	75.0
T	7	329	336	310	664	76	675	1725	2061	31.8	21.1	63.8

Table A.6 (b): The Detailed Analysis Results (Twos, Own: Method 2)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	98	102	171	334	66	404	975	1077	33.9	29.0	62.5
70-74	1	111	112	95	186	26	254	561	673	33.8	22.1	68.4
75-79	1	74	75	10	57	14	84	165	240	14.9	60.0	70.0
80-84	0	43	43	3	9	1	11	24	67	25.0	25.0	85.1
85-	1	3	4	0	0	0	0	0	4	-	100.0	75.0
T	7	329	336	279	586	107	753	1725	2061	32.3	29.0	66.0

Table A.6 (c): The Detailed Analysis Results (Twos, Own: Method 3)

Age Group	Preliminary			Secondary					T	Results		
	LD1	NR1	T	CD	ID	LD2	NR2	T		CDR	LDR	CJR
65-69	4	98	102	179	326	58	412	975	1077	35.4	25.7	64.0
70-74	1	111	112	93	136	28	304	561	673	40.6	23.8	75.5
75-79	1	74	75	15	45	9	96	165	240	25.0	40.0	77.1
80-84	0	43	43	3	9	1	11	24	67	25.0	25.0	85.1
85-	1	3	4	0	0	0	0	0	4	-	100.0	75.0
T	7	329	336	290	516	96	823	1725	2061	36.0	26.2	70.0

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 871-878 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Gene Selection for Cancer Classification: A New Hybrid Filter-C5.0 Approach for Breast Cancer Risk Prediction

Mohammed Hamim[*,1], Ismail El Moudden[2], Hicham Moutachaouik[1], Mustapha Hain[1]

[1]*ENSAM-Casablanca Université Hassan II, Casablanca, 20070, Morocco*

[2]*EVMS-Sentara Healthcare Analytics and Delivery Science Institute, Norfolk, 23324, United States*

ARTICLE INFO

Article history:
Received: 29 November, 2020
Accepted: 31 January, 2021
Online: 12 February, 2021

Keywords:
Breast Cancer
Gene Selection
C5.0 Decision Tree
Gene Expression
Classification

ABSTRACT

Despite the significant progress made in data mining technologies in recent years, breast cancer risk prediction and diagnosis at an early stage using DNA microarray technology still a real challenging task. This challenge comes especially from the high-dimensionality in gene expression data, i.e., an enormous number of genes versus a few tens of subjects (samples). To overcome this problem of data imbalance, a gene selection phase becomes a crucial step for gene expression data analysis. This study proposes a new Decision Tree model-based attributes (genes) selection strategy, which incorporates two stages: fisher-score-based filter technique and the gene selection ability of the C5.0 algorithm. Our proposed strategy is assessed using an ensemble of machine learning algorithms to classify each subject (patients). Comparing our approach with recent previous works, the experiment results demonstrate that our new gene selection strategy achieved the highest prediction performance of breast cancer by involving only five genes as predictors among 24481 genes.

1 Introduction

Ranked second among 36 kinds of cancers, breast cancer has the greatest number of incidences and moralities among females worldwide. As per a recent publication of the WHO (World Health Organization), in 2018, 627,000 females were estimated to lose their lives from this disease, representing approximately 15% of all cancer deaths among females worldwide. According to the same source, early detection is becoming a critical tool to reduce breast cancer morbidity and mortality [1]. At present, breast mammogram and ultrasound images are the most traditionally used breast cancer detection techniques. However, besides their high computational cost, these techniques may lead to an overdiagnosis (false positives), which can cause serious clinical consequences, including death from the side effects of a potential Overtreatment [2]. An alternative to this technique is to take the advantages of both data mining and gene expression technology to predict breast cancer. However, this alternative is not without challenges since the significant number of non-informative features in gene expression data may increase the search space size, which makes gene analysis an impossible task [3]. In the context of this new alternative, a gene selection phase is mandatory to deal with the problem of high-dimensionality in gene expression data. The gene selection process aims at getting rid of any irrelevant and redundant genes in gene expression data, which can simplify the learning model by using a strict minimum number of predictors, and thus improving breast cancer risk prediction performance [4]. In terms of contribution to this research field, this study proposes a new Decision Tree model-based gene selection approach, which incorporates two stages: fisher-score-based ranked technique and the gene selection capability of the C5.0 algorithm. The ranked method consists in reducing the computational cost by removing non-informative genes from the original gene expression data, while the C5.0 algorithm consists in identifying the optimal subset of genes, which are quite informative to classify patients. Finally, the genes contained in the obtained gene subset are used as predictors to construct an ensemble of predictive models using 6 learning algorithms: Random Forests (RF), Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN) , and C5.0 Decision Tree algorithm. In the next section, we briefly outline existing approaches. Methodology and Materials are described in section 3. Section 4 is devoted to discussing all experiment results. The final section concludes the proposed work.

[*]Corresponding Author: Mohammed Hamim, ENSAM-CASABLANCA, & mohamed.hamim@gmail.com

2 Existing literature

From different approaches, a variety of studies have been proposed on breast cancer risk prediction.

Through the implementation of Random Forest(RF) and Extreme Gradient Boosting (XGBoost) , S. Kabiraj et al. predicted breast cancer with an accuracy of 74.73% and 73.63%, respectively[5].

Using the breast cancer Coimbra dataset provided by the University of California Irvine (UCI), Naveen et al., proposed 6 machine learning-based prediction models. The learning process was initialized by a Z-Normalization and cross-validation technique. Experimental results show that the Decision tree and KNN algorithms achieved the highest prediction performance [6]. Agarap proposed a deep learning-based approach on normalized features implementing rectified linear units (ReLU) as the activation function, while the Softmax function as the classifier function. The average cross-validation performance using 10-fold does not exceed 87.96% in terms of accuracy on WBCD (Wisconsin Breast Cancer Data) dataset[7]. S. S. Prakash and K. Visakha also proposed a deep neural network-based approach to predict breast cancer. The authors optimized the neural network model using the early stopping mechanism and dropout layers to avoid overfitting problems. The selected predictors were obtained from the WBCD provided by UCI [8].

Using gene expression signature, Turgut et al., in their work, first proposed 8 machine learning-based predictive models without using any gene selection algorithms. Then they combined the proposed machine learning algorithms with two feature selection methods: RLR (Randomized Logistic Regression) and RFE (Recursive Feature Elimination). The most striking results concern the SVM algorithm since it achieved the best accuracy of 88.8% after using the selection process [9]. In the same context of gene expression data, Al-Quraishi et al. presented a breast cancer prediction model based on an ensemble classifier using FCBF (Correlation-based filter) algorithm. Compared to the existing works, their framework achieved an accuracy of 96.11% by involving 112 genes [10]. Aldryan et al. developed a prediction framework using MBP (Modified Backpropagation) with Conjugate Gradient Polak-Ribiere and the standard ACO (Ant Colony Optimization). The MBP was used as a classifier, while the ACO as a gene selection technique. For breast cancer classification using MBP without gene selection, they get the average F-Measure score of 0.2328. After combined with the feature selection using the ACO algorithm, it obtains an average of 0.6412 in terms of performance by involving 2448 genes [11]. Jain et al. presented a two phases-based hybrid feature selection strategy. Their approach combines the Correlation-based Feature Selection (CFS) with the improved-Binary Particle Swarm Optimization (iBPSO) algorithm. The proposed strategy was assessed using Naive–Bayes (NB) algorithm, and the experiments results showed an accuracy of 92.75% for breast cancer classification using an average of 32 genes [12]. To overcome class imbalance issue and time consumption of their old gene selection approach, Li et al. introduced a more efficient implementation of linear support vector machines based recursive feature elimination system (SVM-RFE). Their proposed approach was assessed on 6 public gene expression datasets, and results demonstrated a slight enhancement in terms of time consumption and prediction performance [13].

3 Research methodology

3.1 Description

Using gene expression data, the present paper proposes a new framework that improves breast cancer prediction performance. Figure1 summarizes the main steps of the proposed framework, where fisher score combined with C5.0 are used to select a small number of informative features, and ANN, SVM, KNN, LR, RF , and C5.0 algorithms are applied to the new gene subset to assess the effectiveness of the proposed gene selection approach. The overall pseudo-code of the whole prediction system is illustrated in Algorithm 1 . -discussed in the following sections.

Algorithm 1: Our proposed prediction system

Function *Classification(Training_set, Test_set)*:
  $List_Classifiers = [SVM, LR, C5.0, ANN, KNN, RF]$
  /* iterate over all proposed machine learning algorithms */
  for *each Classifiers in List_Classifiers* **do**
    $Model \leftarrow$Train *classifier* on the Training set
    Test the *Model* on the Test set
    Calculate average performance (Accuracy , F1-score, and the AUC).
  Return All obtained *Models* with their average performances

/* Main program */
Input :-A p-diemsional DNA microarray dataset $D = [y, x_1, x_2, ..., x_p]_{n\times 1}$, with n is the number of samples, x is the gene vector , y is the target vector and p is the number of genes
Output : List of generated prediction models with their average performance and running time for each one.

Split data D using the stratified K-fold cross-validation technique.;
for *each fold in D* **do**
  $D_{test} \leftarrow fold$;
  $D_{train} \leftarrow$ remaining $(K-1)folds$;
  /* Standardization of D_{train} and D_{test} using (12) */
  $SD_{train}, SD_{test} \leftarrow Standardization(D_{train}, D_{test})$;
  /* Filtering data using Fisher ratio using (1) */
  $F_{train}, F_{test} \leftarrow Filter(SD_{train}, SD_{test})$;
  /* Selection of the optimal gene subset using C5.0 */
  $Sub_{train}, Sub_{test} \leftarrow C5.0(F_{train}, F_{test})$;
  /* Classification using the optimal gene subset */
  $Models \leftarrow Classification(Sub_{train}, Sub_{test})$;
Return all generated prediction models "*Models*"

3.2 Genes selection strategy

Feature selection is an important stage in intelligent systems modeling, especially in prediction systems [14, 15]. The feature selection or gene selection in the context of microarray data analysis

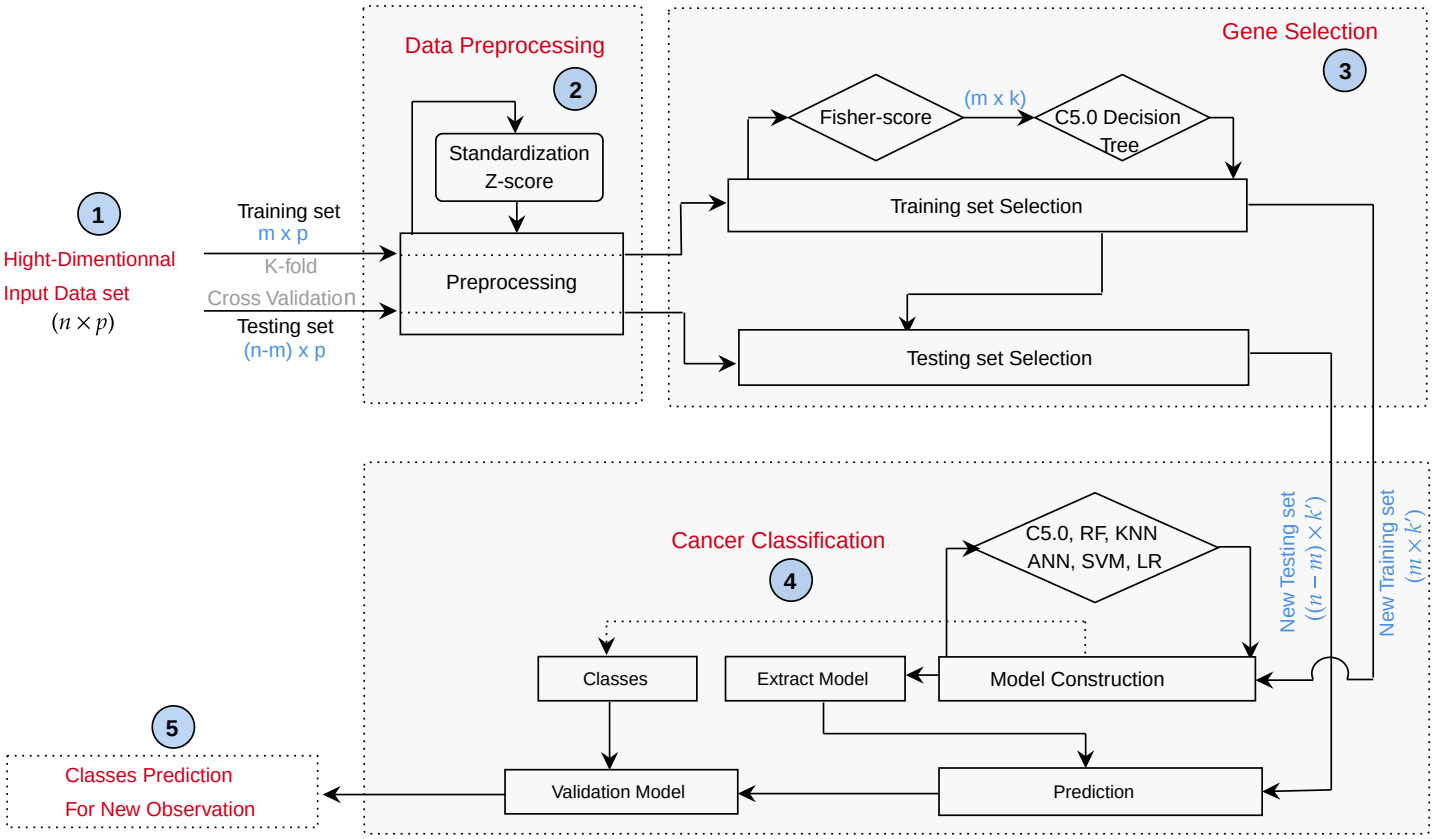


Figure 1: The proposed Framework

is a useful technique that can reduce dimensionality by removing any redundant, irrelevant, or noisy genes, which can lead to improve the classification performance and reduce the cost of computation [4], [16]–[18]. As shown in Figure 2, gene selection process can be reformulated as follows: given an original set of p genes, $X = (X_1, X_2, .., X_p)$, find a subset of genes $X_{sub} = (X_{i1}, X_{i2}, .., X_{ik<<p})$ such that the most informative genes are selected.

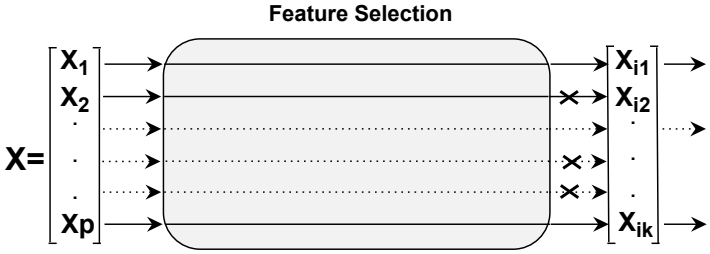


Figure 2: Feature Selection process

Here we present a novel gene selection strategy composed of two steps -Fisher score combined with the C5.0 algorithm- presented into the sequel.

3.2.1 Filter method using Fisher score

Because it acts independently of any classification process, Fisher score-based filter technique is considered as a fast supervised feature selection technique, the reason why it is frequently used when it comes to working with a large number of features (genes) [19]. It can be reformulated as follows: given an input data matrix G of genes, then the Fisher score of each gene (denoted by F) j can be represented by (1) :

$$F\left(G^j\right) = \frac{\sum_{k=1}^{c} \eta_k \left(\mu_k^j - \mu^j\right)^2}{\sum_{k=1}^{c} \eta_k \left(\sigma_k^j\right)^2} \quad (1)$$

With μ_k^j, σ_k^j are the mean and standard deviation of k-th class, corresponding to the j-th feature. μ^j denotes the mean of the whole j-th gene in the G matrix.

As the F of each gene is calculated independently from original genes, only 10% of the highest-ranked genes were selected to achieve the second round of our proposed gene selection strategy.

3.2.2 C5.0 Decision Tree

C5.0 is a new decision tree algorithm developed from C4.5, which has proved its high detection accuracy in many research fields. Compared to C4.5, C5.0 can handle different types of data, deal with missing values, and support boosting to improve classification accuracy [20]. Besides its ability in classification tasks, C5.0 was was used as an efficient feature selection technique in many research Fields [21]–[24]. In the present work, we take both advantages of C5.0, its ability as a powerful feature selection method combined

with the Fisher score-based filter technique, and as a classifier to achieve the classification task in our whole prediction framework.

In the context of gene selection using C5.0 Decision Tree, all genes were initially compared by using the following process: first of all, we set the pruning degree (Pruning is an inherent technique that consists in reducing the size of decision trees by getting rid of branches of the tree that provide little information which can reduce the complexity of the classifier, thus improving classification performance) to 75% as a default value to prevent overfitting. Then, the information gain ratio of each gene is calculated using the formula (2).

$$GainRatio(G) = \frac{Info(S) - Info(S/G)}{S\,plit(G)} \tag{2}$$

Where :

-$Info(S)$ is the Information Entropy which calculated as in (3) :

$$Info(S) = -\sum_{i=1}^{m} p_i log_2(p_i) \tag{3}$$

With m the number of classes in the training set (in our case m=2), S a given set of n samples ($|S| = n$) , and $p_i = \frac{n_i}{|S|}$ the probability that an object in S belongs to the class C_i (with n_i the number of samples that belong to the class C_i)

-$Info(S/G)$ denotes the Conditional Information Entropy which is defined as in (4):

$$Info(S/G) = -\sum_{j=1}^{v} \frac{|S_j|}{|S|} \sum_{i=1}^{m} \frac{n_{ij}}{|S_j|} log_2(\frac{n_{ij}}{|S_j|}) \tag{4}$$

Assuming that G divide the set S into v subsets ($S_1, S_2, ..., S_v$), then n_{ij} denotes the number of classes C_i samples in the subset S_j with $|S_j| = \sum_{i=1}^{m} n_{ij}$ and $|S| = n$.

-$S\,plit(G)$ is defined as in (5) :

$$S\,plit(G) = -\sum_{j=1}^{v} \frac{|S_j|}{|S|} log_2(\frac{|S_j|}{|S|}) \tag{5}$$

The most informative gene (best predictor) with the maximum information gain is selected to be the root node of the whole tree. Then, the root node in each level of the tree is selected from remained genes using the same principle. The gene selection process continues until a maximum depth is meet. As the tree was pruned, the optimal gene subset is determined [20, 25].

4 Classification Algorithms

4.1 Artificial Neural Network

The ANN is a form of distributed computation inspired by networks of human biological neurons. Typically, an ANN consists of a set of interconnected artificial neurons that are organized in several layers called, input, output, and one or several hidden layers. In a typical 3-layer network as shown in Figure fig :ANN, each neuron in a layer is connected to the all next layer neurons with no interconnectivity among the neurons of the same layer and no connection back. All the connections are defined by weight values denoted by w. In the input layer, all nodes get information from the outside and pass it to the nodes of the next layer. Each node computes the weighting sum of all the N neurons of the previous layer and passes it through an activation function (usually logistic) [26]

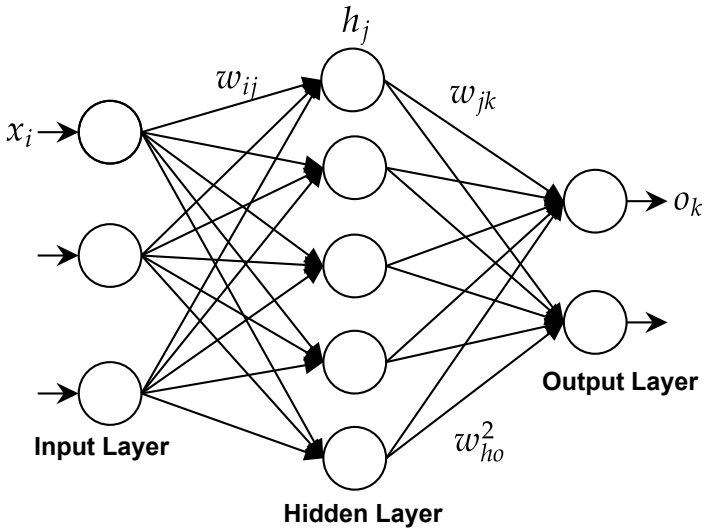


Figure 3: Artificial Neural Network diagram

4.2 Random Forest

As a powerful supervised pattern classification algorithm, RF is used in many intelligent systems. Random Forest is viewed as a combination of several independent decision tree classifiers. Each decision tree is constructed using a randomly selected subset of features. In RF, a majority voting process is used to affect samples to one of the classes by taking the most popular class among all predicted classes by all the tree predictors in the RF [27]. Many processes are used to construct a decision tree classifier; the most commonly used are the Information Gain (IG) and the Gini Index (GI) [28]. In the present paper, we used the GI for the random feature selection measure.

4.3 K-Nearest Neighbors

K-Nearest Neighbors is an easy and simple machine learning algorithm that is widely used in many domains of pattern classification. As the KNN is a non-parametric classifier, new samples are affected to the class represented by a majority of its K-nearest neighbors using a feature similarity rule; since there is no mathematical model to predict labels for new samples. The similarity process is defined using different distance metrics such as Hamming Distance, Euclidean Distance, and Minkowski distance [29] . In the present work, the similarity is measured between K-nearest neighbors using the Euclidean distance metric as in(6), and the number of neighbors K is set to k=4.

$$d(S_1, S_2) = \sqrt{\sum_{i=1}^{p} (s_{1i} - s_{2i})^2} \tag{6}$$

Where S_1 and S_2 are two given samples and p is the number of features .

4.4 Logistic Regression

As an extension of the linear regression algorithm for classification problems, Logistic Regression aims to find the best fitting model, which squeezes the output of a linear equation between 0 and 1 using the logistic function. In linear regression, the relationship between output and features is modeled by using a linear equation. However, in a classification problem, it is strongly recommended to have probabilities between 0 and 1, which can force the outcome to be only between 0 and 1 using (7).

$$L(x) = \frac{1}{1 + e^{-(a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n)}} \tag{7}$$

4.5 Support Vector Machine

The SVM is a binary classifier algorithm that has been successfully applied in many pattern recognition areas. In linear classification, SVM constructs a classification hyper-plane that separates the data into two sets by maximizing the margins and minimizing the classification error. The hyper-plane is constructed in the middle of the maximum margin. Thus, samples above the hyper-plane are classified positives. Otherwise, they are classified as negatives (Figure 4). SVM is a linear classifier algorithm, meaning that it uses a linear separation to classifier samples. However, in real intelligent systems, datasets are often linearly non-separable. To overcome this problem of non-linearity, a nonlinear transformation of the input vectors into a new feature space is performed, and then a linear separation is performed using the new feature space. In this work, the Gaussian kernel using (8), was used to overcome the non-linearity issue.

$$k(x_i, x_j) = \exp(-\frac{\| x_i - x_j \|^2}{2\sigma^2}) \tag{8}$$

Where: $\| x_i - x_j \|^2$ denotes the Euclidean distance and σ a positive parameter which denotes the smoothness of the kernel.

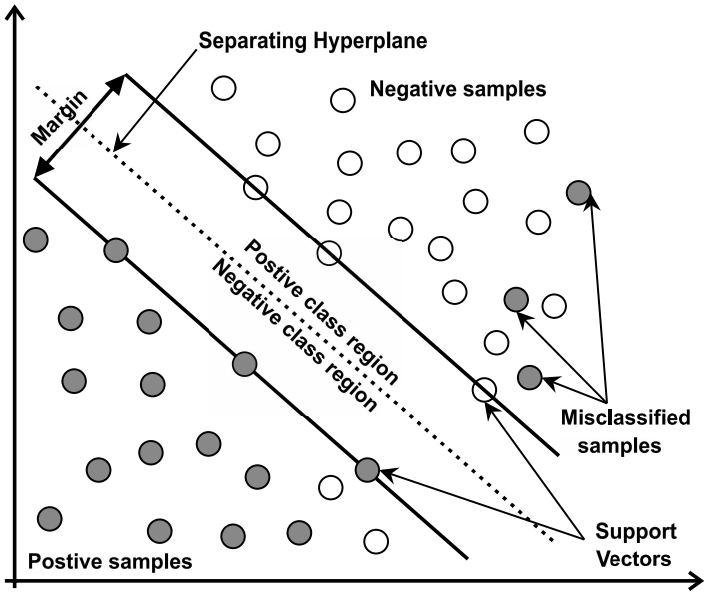


Figure 4: SVM diagram

4.6 Evaluation Metrics

The data mining process has several ways to check the performance of any classification model. The quality of any classification model is built from the confusion matrix (see Figure fig :CM), which summarizes the comparison between predicted and observed classes for all samples. Different types of evaluation measure are calculated from the confusion matrix, and the most commonly used in practice are the classification accuracy as in (9), and F1-score as in (10).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{9}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \tag{10}$$

		Predicted classes	
		Positives	**Negatives**
Observed classes	**Positives**	**TP** (True Positive)	**TN** (True Negative)
	Negatives	**FP** (False Positive)	**FN** (False Negative)

Figure 5: Confusion Matrix for binary classification

Another common evaluation metric used in Machine Learning is the receiver operating characteristic (ROC) curve, which is created by plotting the True Positive Rate (TPR=TP/(TP+FN)) against the False Positive Rate (FPR=FP/(FP+TN)). The Area Under the ROC Curve (AUC) provides a good idea about model performance. The model that gives 100% of corrects predictions has an AUC of 1, while the model that gives 100% of wrong predictions has an AUC of 0.

In the present paper, the average of each evaluation metric (described above) in training and test sets is calculated to evaluate the quality of each prediction model using (11).

$$metric = \frac{metric_{train} + metric_{test}}{2} \tag{11}$$

5 Experimental results and discussion

Before presenting our experimental results in the next section, we describe the gene expression data set used in this paper, K-Fold cross-validation, data preprocessing, and the system configuration.

5.1 Breast cancer dataset details

The proposed prediction framework was conducted on the public available microarray breast cancer dataset [30], which includes 24,481 features with 97 samples, 51 (52.58%) of which are healthy and 46 (47.42%) are diagnosed with breast cancer. Details of our used microarray dataset are given in Table 1.

Table 1: Microarray datasets details

Dataset	Genes	Samples	Classes	Ref
Breast cancer	24481	97	Relapse/Non-relapse	[30]

5.2 Stratified 10-fold Cross-validation

In order to avoid overestimating prediction, a stratified 10-fold cross-validation technique was employed. By using this technique, samples are split into five equal folds (subset) of samples. One of the 10- fold is used for the testing step, and the remaining four folds are put together to form the training data. This process is repeated ten times. The stratification process was used to ensure that all folds are made by preserving the same percentage of samples for each class.

5.3 Preprocessing

Before supplying the datasets to our analysis system (Figure 1), it was necessary to perform a data preprocessing as it is an important step to overcome data imbalance issue. In the present paper, Gene expression levels for each gene were standardized using (12). The result is that expression levels for each feature have a mean 0 and variance 1.

$$X = \frac{X - \overline{X}}{\sigma} \tag{12}$$

Where: $\overline{X}$ the overall mean of the feature X and σ its standard deviation.

5.4 System configuration

By using parallel processing, our proposed framework was implemented in Python 3.7 language under Ubuntu 18.04.3 with v5.0 based Linux kernel 64bits operating system. All of the experiments were carried out using an Intel Xeon E5-2637 v2 3.5 GHz PC with 64 GB of RAM.

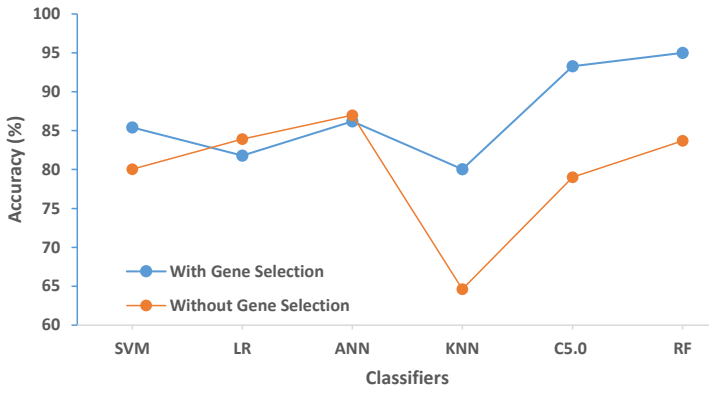


Figure 6: Classification performance with and without our proposed gene selection approach

5.5 Results and discussion

In order to see how good our new approach behaves in situations where there is a large number of genes versus few observations, this section aims at analyzing the results of our proposed framework in terms of classification accuracy, degree of dimensionality reduction, and the running time.

To prove the power of our gene selection approach in terms of prediction performance, we first applied our six proposed machine learning algorithms on the whole breast cancer gene expression data set (without using any feature selection process). Table 2 presents the experimental results and shows the classification performance in terms of accuracy, F1-score, and AUC of each constructed classifier. As we can notice, the most striking performance was achieved by the ANN classifier, which does not exceed 86.99%, while the KNN algorithm shows the lowest performance rate of 64.62%. Almost the classification performances of all constructed classifiers do not exceed the eighties in terms of all evaluation metrics.

Table 2: Performance measurement without gene selection

Classification Model	Time (s)	Accuracy (%)	F1-score	AUC
ANN	42	86.99	0.84	0.85
KNN	1	64.62	0.65	0.65
LR	13	83.91	0.83	0.84
RF	1	83.71	0.86	0.84
SVM	1	80.05	0.79	0.8
C5.0	25	79.01	0.79	0.8

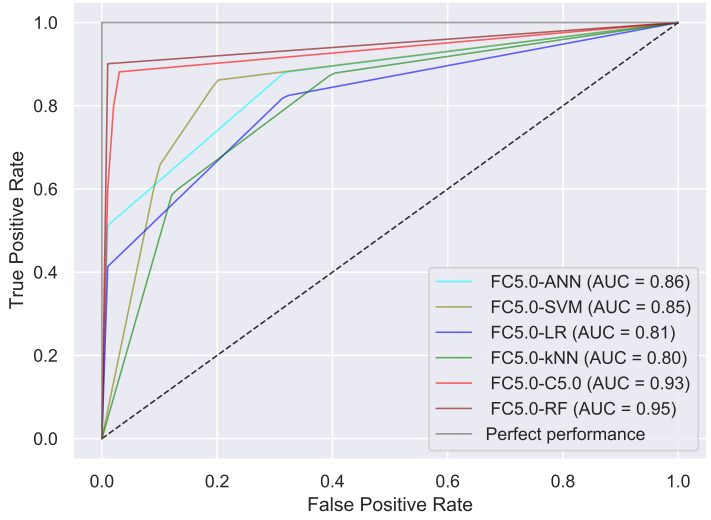


Figure 7: ROC curve of our shrinkage models using the proposed gene selection approach

Table 3 shows the experimental results of our gene selection method when used along with our proposed machine learning algorithms. The results are presented in terms of classification performance matched with the number of selected genes. As we can notice, using our approach, the dimensionality of our research space (the number of genes) was reduced in two phases. First, the number of genes passed from p=24481 (the original number of the gene as shown in Table 1) to k=2448 using the Fisher ratio-based filter method, the new k represents 10% of the original number of genes that have the highest score. In the second phase, the number of previously selected genes k was reduced for a second time using the inner feature selection ability of C5.0 Decision Tree algorithm; thus, the k passed from 2448 genes to five genes (k'=5). Thereafter, the new subspace of k'=5 predictors (genes) was used to construct the ensemble of classifiers (SVM, KNN, LR, ANN, C5.0, and RF). As

Table 3: Performance measurement using our framework

Input Data	Original Genes Number	Gene Selection approach		Selected Genes	Classification Model	Time (s)	Accuracy (%)	F1-score	AUC
		Fisher score (F)	C5.0						
Breast Cancer gene expression data	p=24481	k=2448	k'=5	NM_013438, AL137615, NM_003477, Contig26768_RC, Contig55662_RC	**FC5.0-RF**	**31**	**95.00**	**0.94**	**0.95**
					FC5.0-C5.0	31	93.28	0.94	0.97
					FC5.0-ANN	31	86.21	0.89	0.86
					FC5.0-SVM	31	85.4	0.86	0.85
					FC5.0-LR	31	81.78	0.82	0.81
					FC5.0-KNN	31	80.4	0.82	0.8

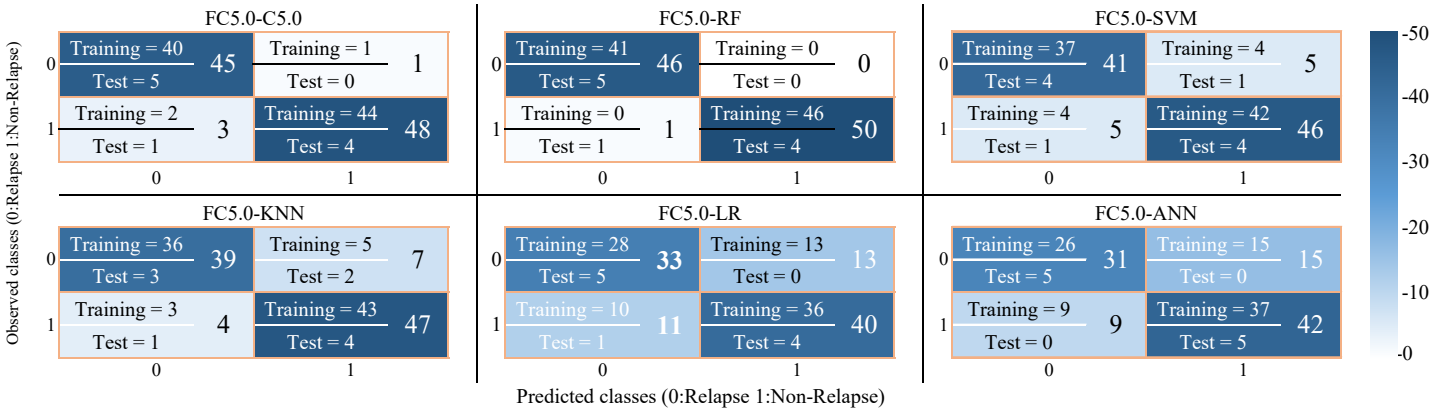


Figure 8: Confusion Matrix of our shrinkage models

we can see from experiment results (Table 3 and Figure 6), almost all constructed shrinkage models produced a classification performance that could exceed the eighties; and the most striking result concerns the FC5.0-RF and FC5.0-C5.0 models since they achieved the best prediction performance that could reach respectively to 95% and 93.28% in terms of all evaluation metrics, which is much better than what was reported in the first experiment (without applying gene selection Table 2). In contrast, the prediction performance has slightly deteriorated when it comes to LR and ANN classifiers when using our gene selection approach.

In Figure 7, the ROC curves are plotted to estimate the AUC of each constructed classifier model when using our gene selection method, which can help to compare their prediction performances. As we can see from this figure, ROC curves of FC5.0-RF and FC5.0-C5.0 models are the closest to the perfect performance curve, which can explain the quality of these models reported in Table 3. In the same context of the second experiment using our gene selection approach, Figure 8 gives a better overview of the relations between our generated prediction models outputs and the true labels. As we can notice, for our favorite model (FC5.0-RF) all samples in the training set are correctly classified, while in the test set, only one sample out of 10 is miss-classified, which confirms the achieved average classification accuracy of 95%.

According to all results reported above, it can be summed up that the power of our whole system (illustrated in Figure1)) comes from its ability to predict breast cancer risk with high performance by involving a bar minimum of genes predictors (only five genes).

Table 4 compares the performance (in terms of accuracy and the number of selected genes) of our gene selection-based breast cancer classification technique with existing gene selection methods. The comparison is made to demonstrate the capability of our approach over other techniques in predicting breast cancer. As it can be noticed, the proposed approach led to a higher prediction performance by involving only five genes, which is much better than was reported for the other techniques.

Table 4: Comparison of the proposed FC5.0 approach with other feature selection-based approaches

Paper	Appraoch	# Selected genes	Best accuracy (%)
[9]	RFE+SVM	50	88,82
	RLR+SVM		87,87
[10]	FCBF+DNN+SVM	112	96.11
[11]	ACO+MBP	2448 (10%)	64.12
[12]	CFS+iBPSO+NB	32	94.00
[13]	SVM-VSSRFE	-	90.03
Ours	FC5.0-RF	5	95.00

6 Conclusion

By using gene expression data to improve breast cancer risk prediction, this study proposed a new Decision Tree model-based gene selection strategy, which incorporates two stages: fisher score-based filter method and the feature selection capability of the C5.0 algorithm. The prefiltering phase using the Fisher score aims at reducing the dimensionality of breast cancer gene expression data by getting rid of any irrelevant or redundant genes in the predictive genes. Then in the second stage, we make use of the obtained low-dimensional research space to find the best gene subset that maximizes the prediction performance by using the feature selection capability of the C5.0 decision tree algorithm. The optimal genes subset was used as the input for cancer classification using six machine learning algorithms. To prove the impact of our approach on breast cancer

risk prediction, we compared the classification performance of our generates models between them and with the performance of classifiers without using any gene selection process. Experimental results show that our gene selection approach led to a higher prediction performance that reached 95% using FC5.0-RF model by taking fewer genes, which is better than what was reported in Table 2 and Table 4 .

This work can be enhanced in two different aspects: the classification models and the search engine. For the classification model, we intend to propose the use of other supervised machine learning algorithms in order to achieve more accurate results in terms of prediction. For the search engine, other gene selection approaches can be proposed or combined with our proposed one. We expect these improvements to predict breast cancer risk with high accuracy, which may guide further breast cancer researches.

References

[1] "WHO | Breast cancer," Publisher: World Health Organization.

[2] "Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: A comparative investigation in machine learning paradigm," Biocybernetics and Biomedical Engineering, **39**(2), 393–409, 2019, doi:10.1016/j.bbe.2019.03.001, number: 2 Publisher: Elsevier.

[3] J. Cao, L. Zhang, B. Wang, F. Li, J. Yang, "A fast gene selection method for multi-cancer classification using multiple support vector data description," Journal of Biomedical Informatics, **53**, 381–389, 2015, doi:10.1016/j.jbi.2014.12.009.

[4] H. Moutachaouik, I. El Moudden, "Mining Prostate Cancer Behavior Using Parsimonious Factors and Shrinkage Methods," SSRN Electronic Journal, 2018, doi:10.2139/ssrn.3180967.

[5] S. Kabiraj, M. Raihan, N. Alvi, M. Afrin, L. Akter, S. A. Sohagi, E. Podder, "Breast Cancer Risk Prediction using XGBoost and Random Forest Algorithm," in 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1–4, 2020, doi:10.1109/ICCCNT49239.2020.9225451.

[6] Naveen, R. K. Sharma, A. Ramachandran Nair, "Efficient Breast Cancer Prediction Using Ensemble Machine Learning Models," in 2019 4th International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT), 100–104, 2019, doi:10.1109/RTEICT46194.2019.9016968.

[7] A. F. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," CoRR, **abs/1803.08375**, 2018.

[8] S. S. Prakash, K. Visakha, "Breast Cancer Malignancy Prediction Using Deep Learning Neural Networks," in 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 88–92, 2020, doi:10.1109/ICIRCA48905.2020.9183378.

[9] S. Turgut, M. Dağtekin, T. Ensari, "Microarray breast cancer data classification using machine learning methods," in 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), 1–3, 2018, doi:10.1109/EBBT.2018.8391468.

[10] T. Al-Quraishi, J. H. Abawajy, N. Al-Quraishi, A. Abdalrada, L. Al-Omairi, "Predicting Breast Cancer Risk Using Subset of Genes," in 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT), 1379–1384, IEEE, Paris, France, 2019, doi:10.1109/CoDIT.2019.8820378.

[11] D. P. Aldryan, Adiwijaya, A. Annisa, "Cancer Detection Based on Microarray Data Classification with Ant Colony Optimization and Modified Backpropagation Conjugate Gradient Polak-Ribiére," in 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA), 13–16, IEEE, Tangerang, Indonesia, 2018, doi:10.1109/IC3INA.2018.8629506.

[12] I. Jain, V. K. Jain, R. Jain, "Correlation feature selection based improved-Binary Particle Swarm Optimization for gene selection and cancer classification," Applied Soft Computing, **62**, 203–215, 2018, doi:10.1016/j.asoc.2017.09.038.

[13] Z. Li, W. Xie, T. Liu, "Efficient feature selection and classification for microarray data," PLOS ONE, **13**(8), e0202167, 2018, doi:10.1371/journal.pone.0202167, number: 8.

[14] I. El Moudden, H. Jouhari, M. Ouzir, S. bernoussi, "Learned Model For Human Activity Recognition Based On Dimensionality Reduction," 2018.

[15] I. El Moudden, S. Lhazmir, A. Kobbane, "Feature Extraction based on Principal Component Analysis for Text Categorization," 2017, doi:10.23919/PEMWN.2017.8308030.

[16] I. El Moudden, M. Ouzir, S. ElBernoussi, "Feature selection and extraction for class prediction in dysphonia measures analysis: A case study on Parkinson's disease speech rehabilitation," Technology and health care: official journal of the European Society for Engineering and Medicine, **25**, 1–16, 2017, doi:10.3233/THC-170824.

[17] M. Hamim, I. E. Moudden, H. Moutachaouik, M. Hain, "Decision Tree Model Based Gene Selection and Classification for Breast Cancer Risk Prediction," in Smart Applications and Data Analysis, 165–177, Springer, Cham, 2020, doi:10.1007/978-3-030-45183-7_12.

[18] M. Hamim, I. El Mouden, M. Ouzir, H. Moutachaouik, M. Hain, "A NOVEL DIMENSIONALITY REDUCTION APPROACH TO IMPROVE MICROARRAY DATA CLASSIFICATION," IIUM Engineering Journal, **22**(1), 1–22, 2021, doi:10.31436/iiumej.v22i1.1447.

[19] Q. Gu, Z. Li, J. Han, "Generalized Fisher Score for Feature Selection," arXiv:1202.3725 [cs, stat], 2012.

[20] R. Rathinasamy, L. Raj, "Comparative Analysis of C4.5 and C5.0 Algorithms on Crop Pest Data," International Journal of Innovative Research in Computer and Communication Engineering, **5**, 2017, 2019.

[21] Y. Y. Wang, J. Li, "Feature-Selection Ability of the Decision-Tree Algorithm and the Impact of Feature-Selection/Extraction on Decision-Tree Results Based on Hyperspectral Data," Int. J. Remote Sens., **29**(10), 2993–3010, 2008, doi:10.1080/01431160701442070.

[22] D. McIver, M. Friedl, "Using Prior Probabilities in Decision-Tree Classification of Remotely Sensed Data," Remote Sensing of Environment, **81**, 253–261, 2002.

[23] Z. Qi, A. G.-O. Yeh, X. Li, Z. Lin, "A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data," Remote Sensing of Environment, **118**, 21 – 39, 2012, doi:https://doi.org/10.1016/j.rse.2011.11.001.

[24] L. Deng, Y.-n. Yan, C. Wang, "Improved POLSAR Image Classification by the Use of Multi-Feature Combination," Remote Sensing, **7**, 4157–4177, 2015.

[25] S.-l. PANG, J.-z. GONG, "C5.0 Classification Algorithm and Application on Individual Credit Evaluation of Banks," Systems Engineering - Theory & Practice, **29**(12), 94 – 104, 2009, doi:https://doi.org/10.1016/S1874-8651(10)60092-0.

[26] S. Rajasekaran, G. A. V. Pai, Neural networks, fuzzy logic, and genetic algorithms : synthesis and applications, New Delhi : Prentice-Hall of India, eastern economy ed edition, 2003, includes bibliographical references and index.

[27] A. Chowdhury, T. Chatterjee, S. Banerjee, "A Random Forest classifier-based approach in the detection of abnormalities in the retina," Medical & Biological Engineering & Computing, **57**, 2018, doi:10.1007/s11517-018-1878-0.

[28] M. Pal, "Random forest classifier for remote sensing classification," International Journal of Remote Sensing - INT J REMOTE SENS, **26**, 217–222, 2005, doi:10.1080/01431160412331269698.

[29] R. O. Duda, P. E. Hart, D. G. Stork, Pattern classification, Wiley, New York, 2nd edition, 2001.

[30] L. J. van 't Veer, H. Dai, M. J. van de Vijver, Y. D. He, A. A. M. Hart, M. Mao, H. L. Peterse, K. van der Kooy, M. J. Marton, A. T. Witteveen, G. J. Schreiber, R. M. Kerkhoven, C. Roberts, P. S. Linsley, R. Bernards, S. H. Friend, "Gene expression profiling predicts clinical outcome of breast cancer," Nature, **415**(6871), 530–536, 2002, doi:10.1038/415530a, number: 6871.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 879-885 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Cyclic Evaluation of Capacity of Recovered Traction Battery after Short-Circuit Damage

Matus Danko*, Marek Simcak

University of Zilina, Faculty of electrical engineering and information technologies, Department of mechatronics and autotronics, Zilina, 010 26, Slovakia

ARTICLE INFO

Article history:
Received: 09 December, 2020
Accepted: 24 January, 2021
Online: 12 February, 2021

Keywords:
Traction battery
$LiFePO_4$
Short-circuit
Damage recovery

ABSTRACT

Presented paper discusses possibilities related to the recovery of the damaged lithium batteries after the short-circuit. The recovery procedure was applied on the selected traction LiFePO4 40Ah cell which was initially short-circuited. After the short-circuit, the damaged cell has visible damage of the electro-mechanical properties. For the recovery of damaged traction cell as much as possible, the experimental recovery procedure has been proposed. For the realization of this recovery procedure, the automated workplace for the cell discharging and charging with the proposed algorithm was created. For verification of the proposed recovery algorithm, the traction cell was tested with a delivered ampere-hour test at the various discharging currents. Results of the delivered ampere-hour test of the recovered cell were compared to results of delivered ampere-hour tests of the new cell. From the final evaluation is seen that the proposed recovery algorithm can recover up to 90% of capacity within a wide range of discharge and charge current.

1. Introduction

Batteries (in general form) are increasingly applied in fully electric vehicles or hybrid electric vehicles, energy storage systems, and other consumer product. The increasing of usage lithium batteries requires their bigger production and minimization of their ecological impact like footprints of carbon dioxide. Moreover, the raw materials access must be ensured, and the cost of materials must be kept low, then production rates will dramatically rise in the next several years. To secure enough battery materials, the spent lithium for batteries, especially those for hybrid electric vehicles or electric vehicles, must be recycled or recovered and used in other energy storage systems [1-5].

In present, complete recycling of lithium batteries is very costly so recycling is unprofitable. Production of batteries from recycled materials is unattractive for manufacturers because is up to 5 times more expensive than the new battery production from newly extracted materials. One of the ways to increase number of reused batteries is to regenerate them. The batteries used in cars are unusable in a vehicle when its usable capacity falls below 60%, but can still be used, for example, like energy storage in the photovoltaic system. In this simpler application, battery can safely serve until the real end of life comes. It is assumed that up to 60% of li-ion batteries from vehicles can be used in this way [6-10].

Currently, the R&D institutions are not only involved into the research on recovery of raw materials, but also on reuse or second life of batteries which cannot fulfil the requirements of their original application area but are highly usable in less demanding application. Best example of battery second life is application of batteries from hybrid electric vehicles (HEV) or fully electric vehicles (EV) in small or medium battery storage systems (BSS) for example photovoltaic systems [11–13].

On the other side, there is also an issue that is common if operation of battery cells is considered. We can talk about improper storage or operation of cells which can resulting into the damage caused for example by deep discharge, over current discharging or short-circuit. Therefore, recovery or regeneration of such cells or harmed batteries is an important topic, due to the environment and the price of new batteries [14–18]. Nowadays, the recovery processes mainly used for NiMH and batteries lead-acid which was used in hybrid electric vehicles or large electricity storage systems. In present, lithium-based cells are largely used for e-mobility, so the topic of lithium battery recovery is more and more important, because of high lithium price [19–22].

Therefore, within the presented paper, the recovery procedure of damaged battery caused by application of long-term short-

*Corresponding Author: Matus Danko, University of Zilina, Faculty of electrical engineering and information technologies, Department of mechatronics and autotronics, Zilina, 010 26, Slovakia, +421 41 513 1610, matus.danko@feit.uniza.sk

circuit is being presented. The focus is given on the traction battery cell Sinopoly LiFePO4 - 40Ah/128Wh because LiFePO4 battery type are often used as traction batteries for electric or hybrid electric vehicles. Initially is being described the procedure of the application of long-term short-circuit, followed by the detection and evaluation of the damages caused by this improper operation. Consequently, the proposal procedure of regeneration is being described, while it is based on the pulsed charging with defined amplitude of charging current and durations of pluses, followed by adequate regeneration breaks. As an evaluation of suitability of proposed regeneration approach, the experimental test of the recovered battery capacity was realized, while received results have been compared to the similar tests of undamaged new cell.

2. Short-circuit damage of traction battery

During the operation of traction batteries, several hazardous conditions which are related to the behavior of the batteries can occur. These conditions can cause damage to the battery itself. Results of operation in inappropriate conditions primarily reflect into the open-circuit voltage (OCV) drop and significant loss of the capacity. If battery operation in hazardous condition is lasting, it can cause damage to the internal structure which can lead to damage of the package of the battery. In this paper is focus on operation at long-term short-circuit. Because unwanted hazardous operational conditions result in damage of battery structure, it is important to find the most suitable recovery procedure which can be applied for battery regeneration as near as possible to its nominal state.

In this paper, the attention is focused on the investigation of regeneration 40 Ah-128 Wh Sinopoly LiFePO4 3.2 V cell after short-circuit. In Table 1 are listed main parameters of tested cell.

Table 1: Parameters of selected cell

Electrical Parameter	Sinopoly $LiFePO_4$ 3.2 V 40 Ah	
Nominal voltage	3.2	(V)
Maximum charging voltage	3.65	(V)
Minimum voltage	2.5	(V)
Discharge current - maximum (continuous)	3	(C)
Discharging current - optimal	13	(A)
Charging current - maximum	60	(A)
Charging current - optimal	20	(A)
Operating temperature	−45 to +85	(°C)
Capacity	40	(Ah)
Package material	plastic	(-)

2.1. Experimental application of the short-circuit

For obtain a reference cell which can be used for recovery with proposal algorithm, selected $LiFePO_4$ 3.2 V, 40 Ah, 128 Wh battery cell was selected. Therefore, selected cell was short-circuited for obtaining reference short-circuited cell. Short-circuit, which is worst scenario of operation from safety point of view. On Figure 4 shown test set-up used for experiment of the short-circuit. This configuration uses mechanical switch rated for very high current for star of short-circuiting test. Thermo-vision camera FLIR E5 and thermo-vision camera FLIR SC 660 used for thermal analysis during short-circuit.

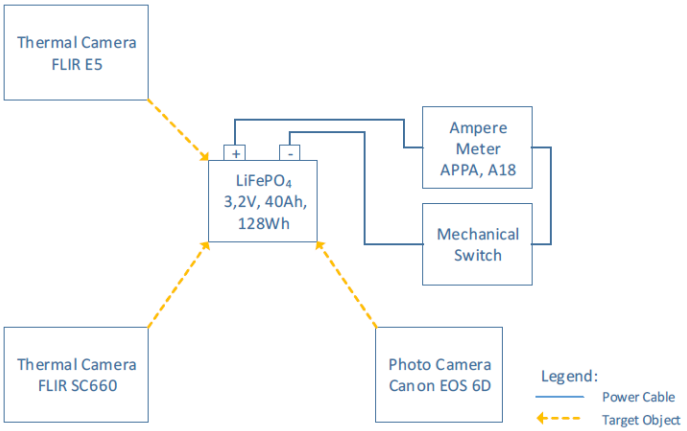


Figure 1: Block diagram of set-up for short-circuit test

The measurement of current during short circuit test was provided by current meter APPA A18, while measured values were stored on NI PXI with measurement cards NI PXI 1031 using LabVIEW. For the evaluation of battery dimensions and shape changes from pictures, camera Canon EOS 6D was used.

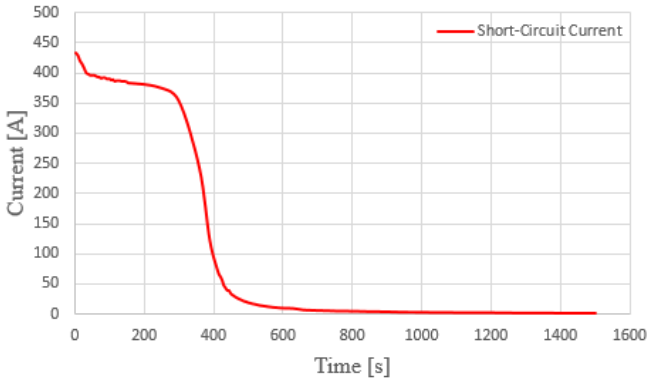


Figure 2: Current profile during the short-circuit experiment

On new battery cell which was formatted during manufacturing process was used for short-circuiting test. Secondary formatting was realized before short-circuiting test within laboratory conditions. Before short-circuiting test, battery open circuit voltage was 3,24 Vdc.

Figure 2 shows the current waveform during the short-circuit experiment. From this figure we can see that after short-circuiting, discharging current of the battery reached up to 430 A. Following more than 300 s, discharging current of the battery was over 350 A, than in next 100s is visible rapid fall of discharging current. After more than 600 s of test, current was reduced to the value of 11.5 A and finally to 2.7 A. The duration of this test was 25 min.

During short-circuited test, the battery discharging current reached more than 10C what is 3-times higher than maximal continuous discharging current determined from manufacturer. Even due to this critical operational condition, the maximum surface temperature recorded during experiment reached 49.2 °C (Figure 3), while during short-circuit test the ambient temperature was 18 °C. This maximum temperature was measured at the end

of the short-circuit test and the safety pressure valve was inactive during whole duration of short-circuit.

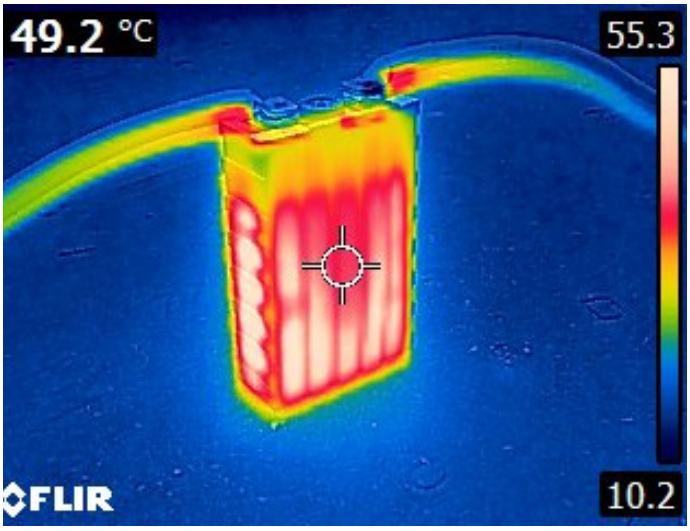


Figure 3: Maximal temperature during short-circuit test

The result of the short-circuit was visible mechanical change of the battery internal structure which cause change dimension of package. Figure 4 indicates structural damage after the end of the experiment, whereby depth of the package raised from 46 mm up to 54 mm, height of the package has risen from 186 mm up to 186.5 mm, and width had changed by 0.5 mm, i.e. from 116 mm down to 115.5 mm.

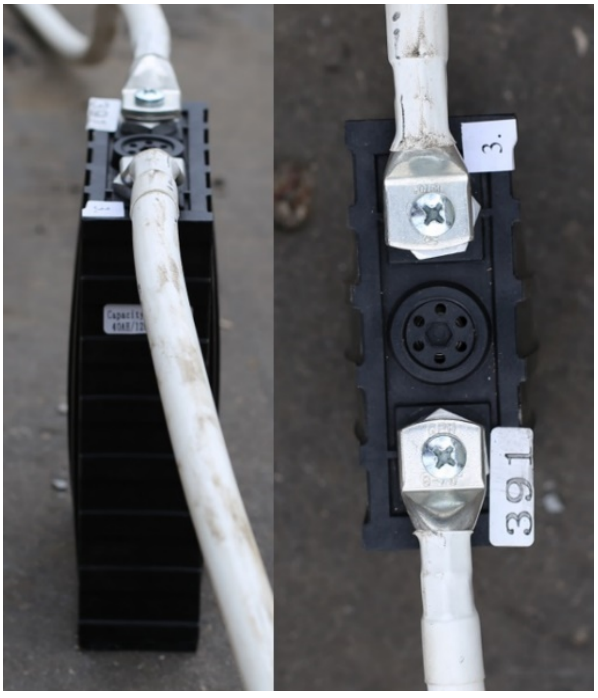


Figure 4: Changes of dimensions of the tested cell after experiment of long-term short-circuit

3. Regeneration procedure application after long-term short-circuit

After the short-circuit test, the tested cell was left resting for 22 days. After this rest period, geometrical dimension of cell was checked again while no dimensional or shape change was observed compared to measured dimensions immediately after the short-circuit test. Consequently, the regeneration procedure which is explained by table 2. Was applied. The proposal procedure is based on the sequential charging with pulse charging current. This charging sequence is split into six subsequences. After finish of each subsequence, approximately 16 h resting period is applied. Each pulse subsequence duration is 100 min is divided into four cycles which duration is 25 min. The main difference between the cycles is the charging current amplitude (Table 2).

Table 2: Parameters of the regeneration procedure after long-term short-circuit.

Subsequence	Duration	Charging current amplitude	Maximal charging voltage
1	25 min × 4 cycle = 100 min	Increasing after each cycle = 0,5 – 2A	3.65 V
2	25 min × 4 cycle = 100 min	Increasing after each cycle = 2,5 – 4A	3.65 V
3	25 min × 4 cycle = 100 min	Increasing after each cycle = 4,5 – 6A	3.65 V
4	25 min × 4 cycle = 100 min	Increasing after each cycle = 6,5 – 8A	3.65 V
5	25 min × 4 cycle = 100 min	Increasing after each cycle = 8,5 – 10A	3.65 V
6	25 min × 4 cycle = 100 min	Increasing after each cycle = 10,5 – 12A	3.65 V

Capacity of cell was tested with delivered ampere-hour test after regeneration and after formatting procedure, while the results have been presented within [23].

During initial capacity testing the cell showed similar performance to reference, unused new cell up to 80 A of discharge current. However, for maximal 3C discharging current, the recovered cell showed just less than half of the capacity (Figure 5).

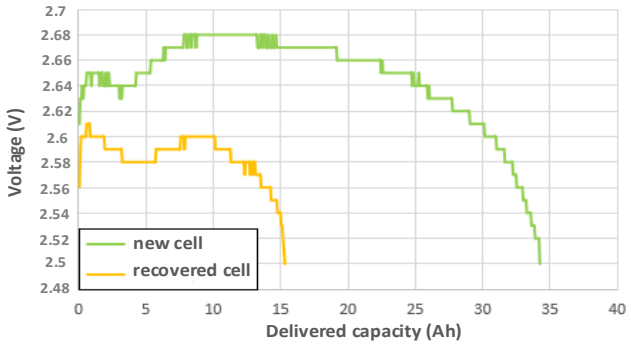


Figure 5: Voltage waveform during discharge with 3C discharging current for regenerated (yellow) and referenced cell (green)

After 28 days of storage, recovered battery was tested ten times in a row with delivered ampere-hour test. Differences in capacity values between repeated delivered ampere-hour tests were under

2%. The main issue, which was not satisfactory for the recovered cell, was the fact that for higher currents (above 2C) the battery lost its performance by more than 50 %. Therefore, second regeneration was applied once more, consequently resting period of 22 days was applied followed by formatting procedure. After this second regeneration procedure, once the cyclic testing of the capacity of regenerated battery was realized.

4. Cyclic testing of the capacity of recovered $LiFePO_4$ traction cell

Before this cycling testing, it is necessary to charge recovered batteries as well as reference one (non-damaged, unused and formatted cell). CC&CV charging (Constant Current and Constant Voltage) is recommended for selected types of batteries. Capacity of both batteries was verified by test of delivered ampere-hours.

For the battery capacity test, five discharging and charging strategies were chosen. Each strategy is specific by a different value of discharging and charging current, while the values of discharging currents was chosen based on manufacturer recommendations of selected cell 13A-120A (0.3-3C). Cyclic experiments have been made, i.e. recovered battery had continuously undertaken verification for each of charging and discharging test for 5 times to prove robustness of proposed regeneration procedure [24-25].

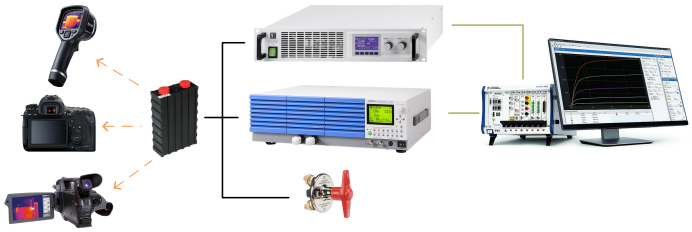


Figure 5: Block diagram of set-up for capacity testing with delivered ampere-hour test

4.1. Test of the discharged capacity

In figures 6 to Figure 10 can be seen voltage waveforms of battery cell during constant current discharge for 5 different current values (0.3C-3C). Each situation compares the situation for recovered cell and for reference (new) cell. From individual results is visible, that recovered battery exhibits the same performance as undamaged new cell.

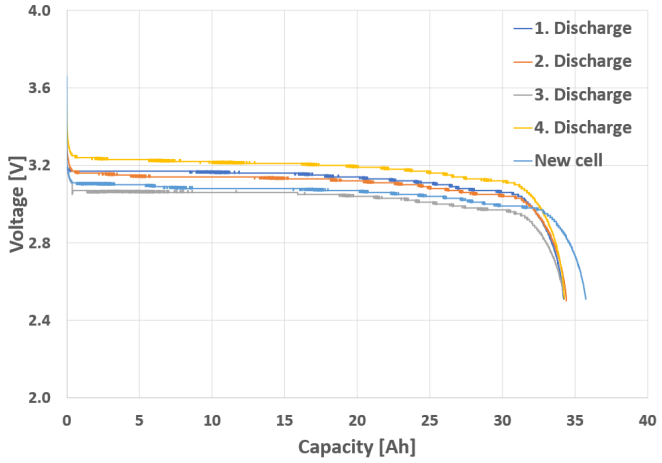


Figure 6: Cyclic discharging (I = 13A)

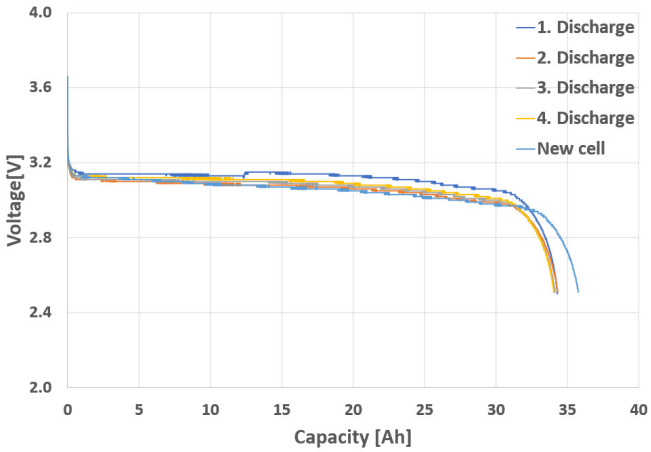


Figure 7: Cyclic discharging (I = 20A)

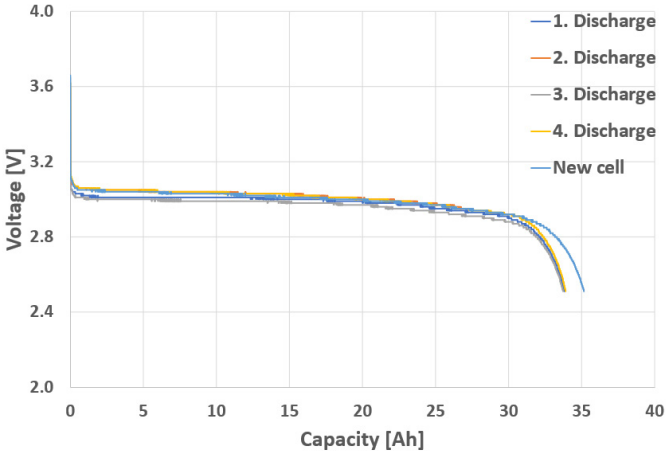


Figure 8: Cyclic discharging (I = 40A)

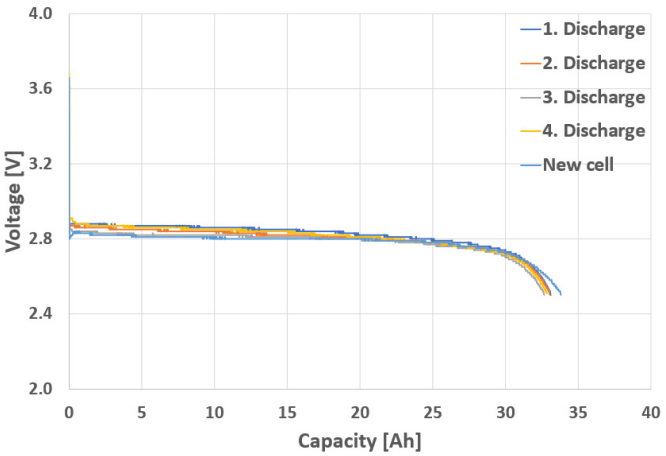


Figure 9: Cyclic discharging (I = 80A)

Each discharging test of recovered cell, i.e. for each value of discharge current, the measurement was validated for 4 times in order to prove stable value of recovered capacity. Simultaneously, there is a small difference between new and recovered cell. When comparing capacity for individual discharging currents, the highest difference of app. 2 Ah is visible for low discharging currents (Figure 6 – Figure 8). For higher values difference is reduced /Figure 9 – Figure 10).

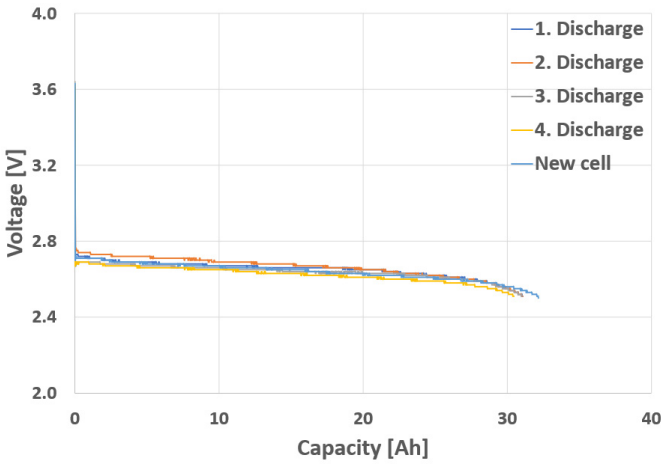


Figure 10: Cyclic discharging (I = 120A)

Within the datasheet of tested battery type it states that it is necessary to monitor the value of the temperature when discharging the cell with a current of 120 A. Temperature profile during whole discharging process by current of 120 A of recovered cell can be seen in Figure 11. From this figure is can be seen that temperature is reaching approximately 37.3 °C at the end of the discharging process. The value is much lower than critical discharging temperature listed in datasheet, which value is 85°C. This temperature profile was recorded for each discharging sequence.

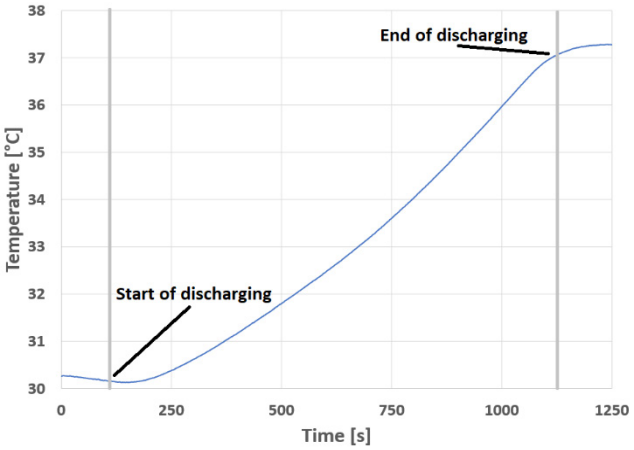


Figure 11: Temperature profile of recovered cell during discharge with 120A

The other two figures (Figure 12 and Figure 13) are representing reference results for individual situations, while independently the results for recovered and undamaged, new formatted cells are interpreted. Here the voltage waveforms are being showed during discharge with different discharging current values. At the beginning of discharging process there is a voltage drop. Voltage drop magnitude depends on the magnitude of the discharge current. The higher value of the discharge current cause the greater the voltage drop. This is caused by the effect of internal resistance of the cell. The difference between Figure 12 and Figure 13 is not markedly great, i.e. recovered cell exhibits similar behavior to undamaged. Therefore, it can be said that application of proposed capacity recovery algorithm has positively influenced the internal resistance of the cell as well.

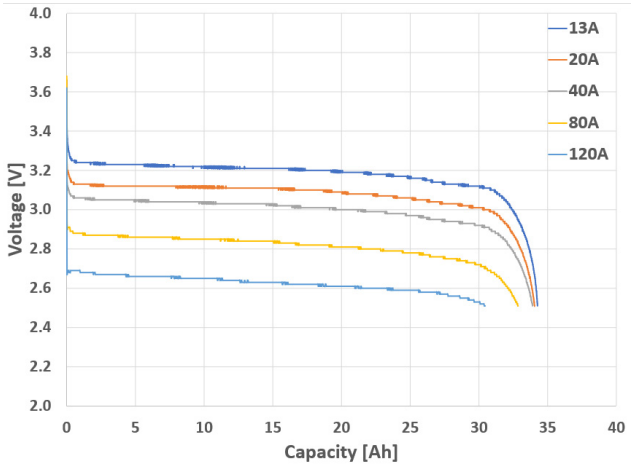


Figure 12: Voltage profile waveforms of recovered cell for various discharging currents

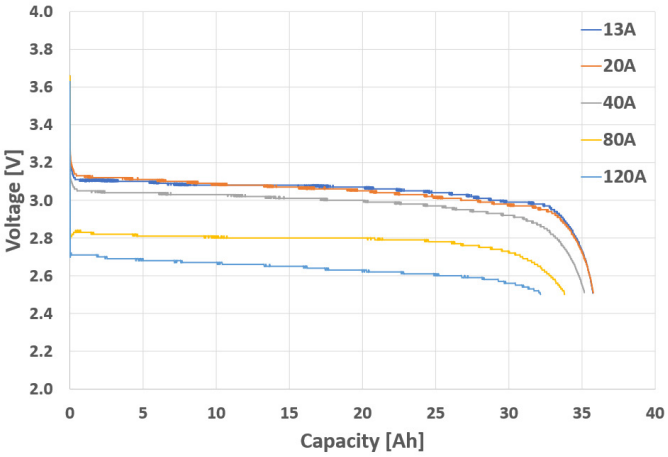


Figure 13: Voltage profile waveforms of new cell for various discharging currents

4.2. Test of charged capacity

Similary to the test of the discharging process, the test of the charging process was realized for recovered and new cell.

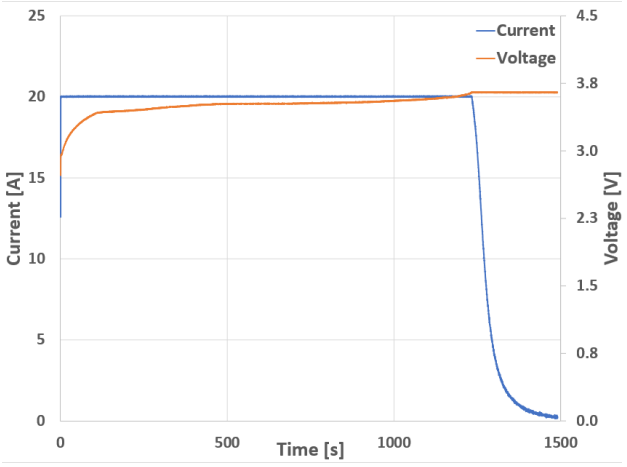


Figure 14: Charging waveforms of voltage and current (I_{CC} = 20A)

The recovered cell was tested several times to prove stability of the recovered capacity once again. Figure 14 shows the cell voltage and current waveforms during charging profile. Here we

can see the value of the constant current (in this case 20A) and the value of the charging voltage (3.65V). The charging profile is reflecting constant current, constant voltage procedure (CC/CV), what is an optimal charging method of tested type of bateries.

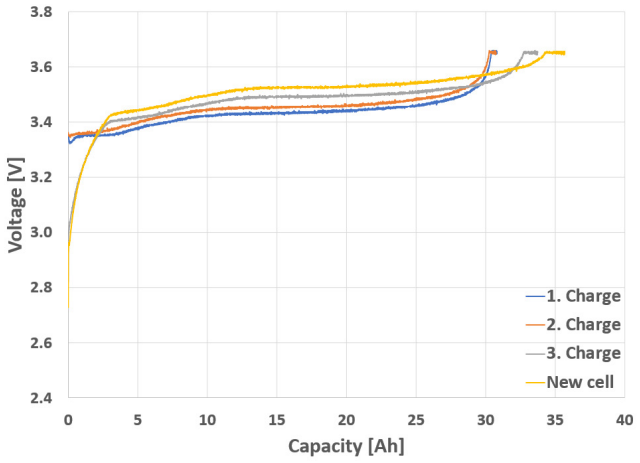


Figure 15: Cyclic charging (I = 20A)

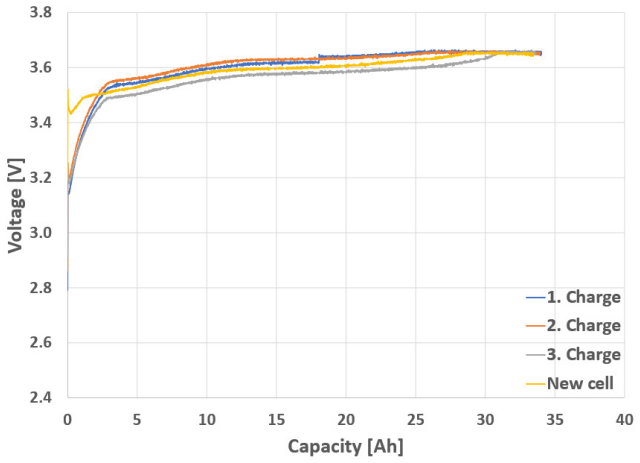


Figure 16: Cyclic charging (I = 40A)

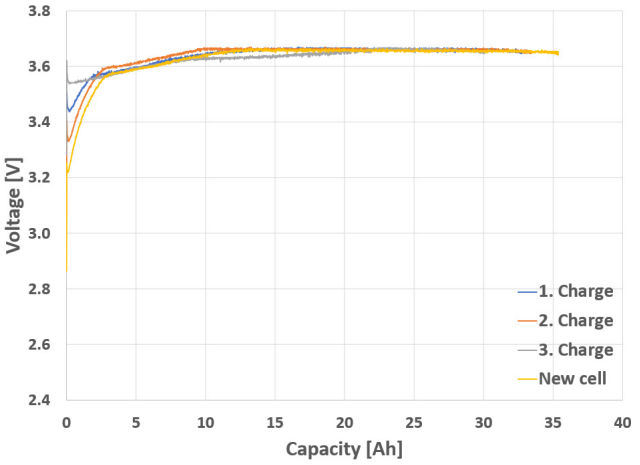


Figure 17: Cyclic charging (I = 60A)

In Figure 15 to Figure 17 are the voltage waveforms during charging proces. Individual figures are specific by the value of charging current. Each graph shows 3 measurements for recovered cell and one measurement for new cell. The graphs show that charging process of new cell exhibits higher capacity if low value of charging current is applied (Figure 15). This process was also visible during discharge process by low values of current. The higher the charging current, the smaller the capacity differences between the recovered and new cell.

Figure 18 and Figure 19 are representing reference voltage profiles during charging process. Figure 18 is shown that maximum capacity of recovered cell is achieved at 33 Ah, while voltage is at the level of 3,65 V. From figure 19 can be seen that the capacity of new cell achieves approximately 37 Ah at maximum voltage level 3,65 V. Based on this results we can state, that recovered cell lost about 10% of its nominal capacity. This result is also confirmed by the voltage waveform from Figure 12 and Figure 13 where is seen, that recovered cell provides less capacity than new. However, even due to reduction of available capacity, the recovery algorithm enables to recover more than 90% of damaged cell capacity.

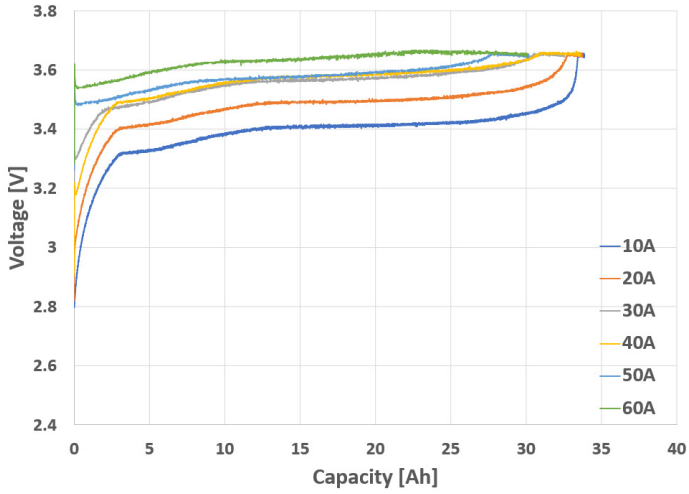


Figure 18: Reference waveforms of shorted cell based on the current

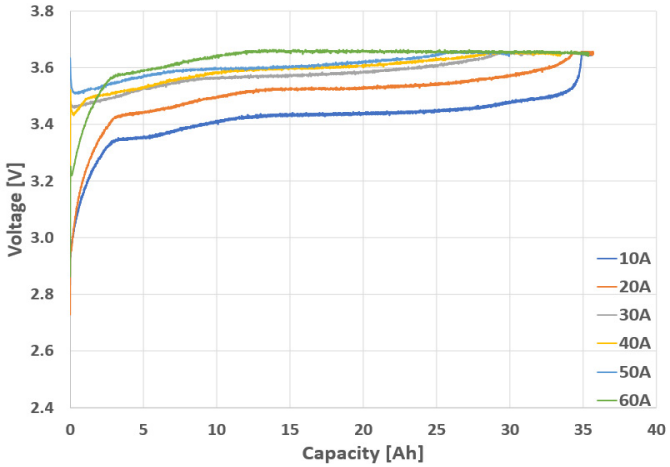


Figure 19: Reference waveforms of new cell based on the current

5. Conclusion

This paper present charging algorithm for recovery of lithium phosphate traction batteries after long-term short circuit. This recovery algorithm was applied to short-circuited Sinopoly LiFePO4 3.2V 40Ah cell. Proposal algorithm was verified with

delivered ampere-hour tests. Result of these tests was compared to delivered ampere-hour test of same type battery which was, new and formatted. From comparison of tests of new and recovered battery was found that proposed algorithm can recover up to 90% within wide range of discharge and charge currents.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This publication was realized with support of Operational Program Integrated Infrastructure 2014 - 2020 of the project: Innovative Solutions for Propulsion, Power and Safety Components of Transport Vehicles, code ITMS 313011V334, co-financed by the European Regional Development Fund.

References

[1] M. Frivaldsky, J. Cuntala, P. Spanik, A.Kanovsky, "Investigation of thermal effects and lifetime estimation of electrolytic double layer capacitors during repeated charge and discharge cycles in dedicated application," ELECTRICAL ENGINEERING, **100**(1), 11-25, 2018

[2] S. Chamnan-arsa, K. Uthaichana, B.Kaewkham-ai, "Modeling of $LiFePO_4$ battery state of charge with recovery effect as a three-mode switched system," Proceedings of the 2014 13th International Conference on Control, Automation Robotics & Vision (ICARCV), Singapore, 1712–1717, 2014, doi:10.1109/ICARCV.2014.7064574.

[3] J. Koscelnik, M. Frivaldsky, M. Prazenica, R. Mazgut, "A Review of Multi-elements Resonant Converters Topologies," ELEKTRO 10th International Conference, Rajecke Teplice, SLOVAKIA, 2014

[4] H.D. Yoo, E. Markevich, G. Salta, D. Sharon, D. Aurbach, "On the challenge of developing advanced technologies for electrochemical energy storage and conversion," Materials Today, **17**(3), 110-121, ISSN 1369-7021, 2014, https://doi.org/10.1016/j.mattod.2014.02.014

[5] P. Spanik, J. Cuntala, M. Frivaldsky, P. Drgona, "Investigation of Heat Transfer of Electronic System through Utilization of Novel Computation Algorithms," ELEKTRONIKA IR ELEKTROTECHNIKA, **123**(7) , 31-36, 2012

[6] V. D'Angelo, S. Cannavacciuolo, S. Lecce, V. Bendotti, O. Pennisi, "Enhanced hotplug protection in BMS applications. Part I: Theoretical Aspects and Practical Issues," 2019 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE), Torino, Italy, 1-5, 2019. doi: 10.23919/EETA.2019.8804513.

[7] M. A. Boukhal, I. Lagrat, O. Elbannay, "Implementation of a lithium-ion battery state of charge estimation algorithm for BMS on the real time target NI myRio," 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), Fez, Morocco, 1-5, 2019. doi: 10.1109/WITS.2019.8723849.

[8] M. Frivaldsky, B. Dobrucky, G. Scelba, P. Spanik, P. Drgona, "Bidirectional Step-Up/Step-Down DC-DC Converter with Magnetically Coupled Coils," Communications - Scientific Letters of the University of Zilina, 15(3), 21-25, 2013

[9] C. Huang-Jen, L. Li-Wei, P. Ping-Lung, T. Ming-Hsiang, "A novel rapid charger for lead-acid batteries with energy recovery," IEEE Trans. Power Electron. **21**(3), 640–647, 2006, doi:10.1109/TPEL.2006.872386.

[10] C. Sen, N. C. Kar, "Battery pack modeling for the analysis of battery management system of a hybrid electric vehicle," Vehicle Power and Propulsion Conference, 207– 212, 2009, doi: 10.1109/VPPC.2009.5289848

[11] P. Spanik, R. Sul, M. Frivaldsky, "Performance Investigation of Dynamic Characteristics of Power Semiconductor Diodes," ELEKTRONIKA IR ELEKTROTECHNIKA, 3, 3-6, 2010

[12] A. F. Moghaddam, A. Van Den Bossche, "An active cell equalization technique for lithium ion batteries based on inductor balancing," in 2018 9th International Conference on Mechanical and Aerospace Engineering (ICMAE). IEEE, 274–278, 2018

[13] M. Cacciato, G. Nobile, G. Scarcella, G. Scelba, A. G. Sciacca, "Energy management optimization in stand-alone power supplies using online estimation of battery SOC," 18th European Conference on Power Electronics and Applications (EPE'16 ECCE Europe), Karlsruhe, 1-10, 2016 doi: 10.1109/EPE.2016.7695559

[14] M. Galad, P. Spanik, M. Cacciato, G. Nobile, "Comparison of common and combined state of charge estimation methods for VRLA batteries, " 2016 ELEKTRO, Strbske Pleso, 220-225, 2016, doi: 10.1109/ELEKTRO.2016.7512069

[15] Y. Cai, Z. Zhang, Y. Zhang, Y. Liu, "A self-reconfiguration control regarding recovery effect to improve the discharge efficiency in the distributed battery energy storage system," Proceedings of the 2015 IEEE Applied Power Electronics Conference and Exposition (APEC), Charlotte, NC, USA,, 1774–1778, 2015. doi:10.1109/APEC.2015.7104587.

[16] P. Spanik, R. Sul, M. Frivaldsky, P. Drgona, "Performance Investigation of Dynamic Characteristics of Power Semiconductor Diodes," ELEKTRONIKA IR ELEKTROTECHNIKA, **3**, 3-6, 2010

[17] M. E. Orchard, M. Lacalle, B. E. Olivares, J. F. Silva, R. Palma-Behnke, P. Estevez, B. Severino, W. Calderon-Muoz, M. Cortez, "Information-Theoretic Measures and Sequential Monte Carlo Methods for Detection of Regeneration Phenomena in the Degradation of Lithium-Ion Battery Cells," IEEE Transactions on Reliability. **64**(2), 701–709, 2015, doi:10.1109/TR.2015.2394356.

[18] M. Cacciato, G. Nobile, G. Scarcella, G. Scelba, "Real-Time Model-Based Estimation of SOC and SOH for Energy Storage Systems," IEEE Transactions on Power Electronics, **32**(1), 794-803, 2017. doi: 10.1109/TPEL.2016.2535321

[19] P. Drgona, A. Prikopova, M. Frivaldsky, M. Priecinsky, "Simulation Based Method for Design and Application of Digital Control System, "Communications - Scientific Letters of the University of Zilina, **13**(2A), 32-37, 2011

[20] J. Koscelnik, M. Prazenica, M. Frivaldsky, S. Ondirko, "Design and Simulation of Multi-element Resonant LCTLC Converter with HF Transformer," ELEKTRO 10th International Conference, Rajecke Teplice, SLOVAKIA, 2014

[21] B. Dobrucky, M. Frivaldsky, P. Drgona, I. Kurytnik, "Measurement of swtitching losses in power transistor structure," ELEKTRONIKA IR ELEKTROTECHNIKA, **2**, 75-78, 2008

[22] M. Frivaldsky, P. Drgona, P. Spanik, "Experimental analysis and optimization of key parameters of ZVS mode and its application in the proposed LLC converter designed for distributed power system application," INTERNATIONAL JOURNAL OF ELECTRICAL POWER & ENERGY SYSTEMS, **47**, 448-456, 2013

[23] M. Frivaldsky, J. Adamec, M. Danko, P. Drgona, "Traction battery (40 Ah LiFePO4) recovery after long-term short circuit," 2020 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), Sorrento, Italy, 299-304, 2020. doi: 10.1109/SPEEDAM48782.2020.9161868.

[24] M. Frivaldsky, B. Dobrucky, M. Prazenica, J. Koscelnik, "Multi-tank resonant topologies as key design factors for reliability improvement of power converter for power energy applications," ELECTRICAL ENGINEERING, **97**(4), 287-302, 2015

[25] P. Spanik, M. Frivaldsky, P. Drgona, J. Kandrac, "Efficiency Increase of Switched Mode Power Supply through Optimization of Transistor's Commutation Mode," In: ELEKTRONIKA IR ELEKTROTECHNIKA, **9**, 45-52, 2010

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 886-897 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Method of Technological Forecasting of Market Behaviour of R&D Products

Vasyl Kozyk[1], Oleksandra Mrykhina[1], Lidiya Lisovska[2,*], Anna Panchenko [1], Mykhailo Honchar [3]

[1]*Department of Business Economics and Investment, Lviv Polytechnic National University, Lviv, 79013, Ukraine*

[2]*Management of Organizations Department, Lviv Polytechnic National University, Lviv, 79013, Ukraine*

[3]*Department of Management and International Business, Lviv Polytechnic National University, Lviv, 79013, Ukraine*

ARTICLE INFO

Article history:
Received: 09 December, 2020
Accepted: 24 January, 2021
Online: 12 February, 2021

Keywords:
R & D product
Commercialization
Technological forecasting
Market behaviour
Technological readiness
Strategy

ABSTRACT

The current concept of open innovation corresponds to the R&D products transfer model – "role changes". One of the fundamental provisions of the model is that R&D products are considered for commercialization not only at the final stage of technological readiness, but at any of them. In today's changing market environment, special attention is paid to the transfer and commercialization of R&D products at the early stages of readiness, but this process is characterized by significant problems from the point of view of technological forecasting. To solve the problems, the article substantiates the method of technological forecasting of market behaviour of R&D products at the early stages of technological readiness, which is based on taking into account the strengths and weaknesses, development factors and limiting factors of R&D product. The method allows you to predict indicators of product behaviour relative to the market where its commercialization is planned.

As a component of the above method and in order to increase the level of reliability of calculations and validity of results, a method for determining the correction factor of indicators of market behaviour of R&D product has been developed. The method was developed on the basis of fuzzy set theory algorithms using the fuzzy logic toolbox (MATLAB), which made it possible to integrate a set of different types of forecast data on the market behaviour of an R&D product, taking into account the relationships and interdependencies between them, into one correction factor. This coefficient contains the characteristics of signs of the impact of R&D product on the market (in particular, market effects, types of market changes) and the impact of market effects on R&D product (effects generated by R&D products, organizational and technological changes in R&D products). To justify the correction factor, a knowledge base of responses from subject area experts has been formed. In order to further select a commercialization strategy for R&D product, a system with normative indicators has been developed that interpret the following types of strategies: zero-level commercialization of R&D product; first-level commercialization of R&D product; commercialization of the second level of R&D product.

The author's method of technological forecasting of market behaviour of R&D products and choosing a commercialization strategy for a product is universal, can be applied to R&D product of any type of economic activity, transfer method, etc. Testing of the method on the example of a number of R&D products presented by the developers of the Lviv Polytechnic National University (Lviv, Ukraine) showed the validity of the author's method and its relevance in modern conditions of market singularity.

1. Introduction

1.1. Framework of theme relevance

This work is a continuation of a number of our scientific papers in the framework of research on the transfer of R&D products, in particular, presented at the 14th International Scientific and Technical Conference on Computer Sciences and Information Technologies (2019) [1].

In recent years, there has been a change in a number of models

*Corresponding Author: Lidiya Lisovska, Email: lida_lissovcka@ukr.net

of R&D products transfer, due to changes in the models of the innovation process, globalization and digitalization. The current concept of open innovation corresponds to the R&D products transfer model – "role changes". One of the fundamental provisions of the model is that R&D products are considered for commercialization not only at the final stage of technological readiness, but at any of them. This position is important from several positions. First, depending on the stage of R&D product readiness for commercialization, both its original purpose and the level of consumer value may change (for example, when R&D products converge at the stage of idea justification).

Secondly, in the context of the spread of digitalization, the commercialization of R&D product is much faster, compared to the similar situation that was inherent in the global economy several decades ago.

In general, the level of technological readiness of an R&D product indicates its presence at a certain stage (research, development, etc.), which is compared with the planned (or final) stage, and, taking into account the features of R&D product development, predict the duration, cost and nature of market behaviour of the product. In other words, the market behaviour of a commercialized R&D product, which is evaluated at different stages of its technological readiness, will be different.

The above actualizes the need to pay attention to the study of approaches, methods and models of technological forecasting of market behaviour of R&D products. The success of implementing the R&D products commercialization strategy will depend on the reasonable choice of methodological tools for technological forecasting. The development of such methodological tools is relevant both from the point of view of the needs of the modern economy, and from the point of view of predicting the market behaviour of R&D products in the conditions of a market singularity.

1.2. Statistical Background

The relevance of the above problem is confirmed by numerous statistical data. In particular, among the problems that will befall R&D products in the form of startups, there is insufficient funding in the early stages of product development – 29 %, which is the second significant reason in the list of their failures (the first reason is the lack of market demand – 42 %) [2]. At the same time, startups are characterized by a growth trend. For example, for two dozen top universities in the United States, the number of startups has more than tripled in 20 years-from 306 units in 1998, up to 1098 units in 2018 [3], which explain the focus of universities to generate R&D products in the early stages with the help of startups. At the same time, 20 % of startups fail in the first year, 50% of startups fail in the first five years [4].

Traditionally, among small business owners, about 77 % of early-stage product financing is done by raising their own funds [4], which requires careful technological forecasting of the development and payback of R&D products. However, only 2 – 3 % of development funding belongs to alternative sources (Venture Investors, Business Angels, crowd funding, etc. [4]).

The importance of paying attention to technological forecasting of market behaviour of R&D products in the early stages of technological readiness is confirmed by the results of the NanoCom Consortium study [5], in which 214 people were interviewed about barriers to commercialization of R&D products in the field of nanotechnology. Among them, 63 people noted significant barriers at 3 – 5 levels of TRL (technology readiness level), where R&D is conducted. Other significant barriers included commercially available at TRL levels 8 – 9 (93 responses). The participants of this organization also argue that the most significant impact of barriers occurs in the early stages of TRL – domain technological, and in subsequent stages TRL gives way to marketing and strategic domain.

R&D is one of the defining components of the structure of a number of global country competitiveness indices (Bloomberg Innovation Index, Global Competitiveness Index, Creativity Index, etc.), so technological forecasting of market behaviour of R&D products can be a strong argument for explaining other components of such indices (market perception, productivity, added value, etc.).

1.3. Formulation of the problem

Without a sound approach to technological forecasting of market behaviour of R&D products, there is a high probability that they will not forget or partially acquire the utility function that turns the product into a market innovation.

To develop a method of technological forecasting of market behaviour of R&D product in modern economic conditions, and, based on this, to justify the commercialization strategy for it, it is necessary to conduct a comprehensive study of both the product and the market. We should learn trends and patterns of the market, the readiness of R&D product for commercialization, and so on.

In this paper, we consider R&D product from the point of view of customer-oriented marketing, that is, the focus is on the consumer value of the product. Product R&D marketing is directly related to the consumer's consciousness. Therefore, in order to assess the level of technological readiness of an R&D product, it is necessary to keep in mind the readiness of the consumer (or market) to accept this technology. Starting from the first stage of evaluating the technological readiness of an R&D product it is necessary to formulate evaluation tasks in such a way that in the end we reach such technological, economic and other indicators that will best satisfy the business structures to which this technology will be transferred.

2. Theoretical Background

2.1. Approaches to substantiating R&D product development strategies

The development and commercialization of an R&D product requires considerable effort and resources from all participants in innovation activities. At the same time, one of the most important tasks is technological forecasting of changes that can occur with an R&D product and affect its commercialization and market distribution.

In modern scientific works on the problems of technological forecasting of R&D product development, various approaches, methods and models of research of R&D product and their market development are presented. In particular, to truly manifest innovation and reap its benefits, one must recognize that innovation is three different things: innovation is an outcome, innovation is a process, and innovation is a mindset [6].

A number of models of market development of R&D product

have been developed. The new technological innovations are changing the ways businesses are being operated. The sharing economy – based new business models (SEBMs) using technology have many benefits, it is paramount to understand these SEBMs models and the behaviour of the market, particularly on how to influence the market's attitude [7] this approach actualizes the importance of tasks to justify and choose a strategy for market development of R&D products.

The formation of a market development strategy should include forecasting factors and the dynamics of their impact on the market success of R&D products, sources of change and inertia that issue from the introduction of innovations in the market [8], business logic as consequence of far-reaching technological developments [9] using technological forecasting methods. Such factors should be systemic and complex, because focusing solely on the effectiveness of the technological innovation is detrimental to long-term operational benefits [10].

The above-mentioned problem is quite complex, since, as scientists note, chain reactions from one innovation can have effects so significant that the entire world alters its way of life to incorporate the new technology. Often times, the effects of a new type of technology build upon existing technology with future innovations building on the inventions of today. This constant evolution and growth forces businesses to adapt or expire [11].

In addition, scientists distinguish the interdependence between the introduction of different types of innovations, which determine the market prospects for each other's development, as in the case of the introduction of technological and organizational innovations [12]. Sustainability, adoption of various types of innovation [13], accounting for the environment in the agenda of companies [14]. The success of innovations depends in part on the business models used to distribute them [15]. Also important is the impact of specific types of information on commercialization rates of technologies [16].

Scientific papers note that the innovation adoption is a process, rather than a dichotomic choice [17]. Therefore, for a rational and effective process of adapting R&D products to the market, it is necessary to develop and implement strategic solutions.

For the formation of innovation development strategies, many approaches justify strategic management decisions from different positions. Marketing strategies and particularly brand associations with brand awareness, perceived quality, and brand loyalty [18], co-branding strategies become important [19].

In addition, market opportunities for innovation are associated with the implementation of competitive strategies and market-based assets (customer orientation, competitor orientation, and marketing creativity) were assessed for their ability to help an innovating firm deal with dysfunctional competition and improve the returns from innovation [20].

In connection with the development of the open innovation model, the following factors are becoming important: intellectual property strategy, the shaping of complementary and substitute appropriability regimes is central when strategizing in dynamic and systemic innovation contexts [21].

2.2. Tools for justifying R&D product development strategies

The study of tools for the formation of strategies for the development of technological innovations has led to a revision of methods of technological forecasting, that can be useful for bringing a new perspective to explain and generalize properties of the evolution of technology and predict which innovations are likely to evolve rapidly in society [22]. The tasks of modern technological forecasting are strategic foresight, sustainability, technology transfer, entrepreneurship, and absorptive capacity can be explored further [23].

When justifying the R&D products development strategy, scientists highlight two generative cognitive processes- analogical reasoning and conceptual computation [24].

More and more scientists are using matrix approach to technological innovations based on design science research, entrepreneurship, and innovation theories, especially in digital innovation [25].

Combined methods of forming strategic decisions based on MULTIMOORA (Multi-Objective Optimization on the basis of Ratio Analysis plus full multiplicative form) is a somewhat new multi criteria decision-making (MCDM) method which provides high efficiency and effectiveness in problem solving [26].

In order to avoid the subjectivism of expert assessments, the authors suggest novel approach for forecasting method selection and a recommendation-based ensemble forecasting approach [27].

Determination of the weighting coefficients of factors influencing the development of technological innovations can be carried out based on exponential smoothing; autoregressive integrated moving average and neural network are chosen to form the combined approach [28].

2.3. Theoretical generalization

The considered approaches to development strategies allow us to approach the justification of strategic decisions on the market development of R&D products from the standpoint of priority of a specific factor. However, according to successful practices, the market perception of R&D products is interdisciplinary in nature. In addition, the scientific and practical developments discussed above are based on the consideration of mostly completed innovations as objects, that is, the results of the completed innovation process. However, the object of this research is the intermediate result of the innovation process, namely, R&D product at the early stages of technological readiness.

These objects have different parameters of technological readiness [29] and features of market application. Therefore, it is necessary to take a differentiated approach to the development of methods for technological forecasting of market development of R&D products, choosing a strategy for their commercialization, and so on.

The author's method must be defined:

- the ability to differentiate in the projected period of market development of R&D products their functional advantages and weaknesses, factors that will restrain or stimulate the development of the product;
- the ability to take into account the features and characteristics of R&D product of various types of economic activity and business areas;
- with the ability to objectively predict market drivers and anti-drivers of R&D product development;
- it is an opportunity to take into account the dynamic nature

of the market development of R&D product in relation to the satisfaction of consumer needs;
— flexibility in changing input data, managerial maneuverability, and so on.

3. Research design

3.1. Methodology approach

The aim of the work is to develop a method for technological forecasting of market behaviour of R&D products at the early stages of their technological readiness.

The basis for conducting such a study is a number of hypotheses.

Hypothesis 1. The method of technological forecasting of market behaviour of R&D products is universal and can be applied to products of any type of economic activity, business sector, etc.

Hypothesis 2. Fuzzy set theory algorithms allow us to take into account changes in the market behaviour of R&D products in the short, medium and long term.

Hypothesis 3. Justification of the R&D product market behaviour indicator based on the author's method should correlate with the early stages of technological readiness of the product.

To prove / refute hypotheses and develop this method, it is advisable to conduct a study of the problems of technological forecasting of R&D products development at the empirical, theoretical and applied levels.

3.2. Methodology

3.2.1. Fundamentals of technological forecasting of market behaviour R&D product

When evaluating an R&D product for the purpose of market launch and developing strategies for its commercialization, analysts mainly use methods and models of technological forecasting of product *market development*.

Most of the known methodological tools for technological forecasting of market development of R&D products are based on the justification of its strengths and weaknesses, risks and chances, etc. [30–32]. However, the category "market development" is a general concept that implies a process that results in a change in the quality of a product, its transition from one qualitative state to another (higher). In the situation with technological forecasting of R&D product development, the category "market behaviour" will give a more accurate description of the above change. Within the framework of neobehaviourism, behaviour is considered visible manifestations that can be observed, internal states associated with external manifestations. Moreover, market behaviour is a set of states (characteristics) of an R&D product that a product reveals under the influence of certain environmental factors (market). However, the market behaviour of R&D products is not often taken into account by analysts in the context of technological forecasting. Therefore, traditional approaches to technological forecasting for R&D products do not always allow us to justify effective strategies for their commercialization.

If an R&D product is in the early stages of technological readiness, it is not always correct to assess the chances and risks of a product. Odds are the probability of such market behaviour of an R&D product that will contribute to its success in the market. Risks are a combination of the probability and consequences of adverse events based on the market behaviour of an R&D product. To determine the chances and risks, first of all, it is necessary to assess the future consumer value of an R&D product. However, when an R&D product is only in the process of development, significant errors may occur when determining the level of its consumer value. The assessment of chances and risks can be more reliable when evaluating R&D products at the stages of appropriability high technological readiness – then it is possible to predict the market behaviour of R&D product and its perception by the market environment with a significant degree of reliability. Therefore, in the case of evaluating an R&D product at the early stages of readiness, it is advisable to clarify the categories "*chances*" and "*risks*" in a meaningful way. It is proposed to replace them with "*development factors*" and "*limiting factors*". These factors determine the probability of events occurring in the R&D product market (events that will determine the chances or risks of R&D product spreading in the market).

When making technological forecasting of the market behaviour of R&D product, it is advisable to pay attention, on the one hand, to the stage of technological readiness of R&D product on the other – to the perception of R&D product by the market environment.

To determine the early stages of technological readiness of R&D products, we will use the NASA approach [33]:
— basic principles observed and reported;
— technology concept and / or application formulated;
— analytical and experimental critical function and / or characteristic proof-of-concept;
— component and / or breadboard validation in laboratory environment;
— component and / or breadboard validation in relevant environment.

In the early stages of R&D product readiness, it is advisable to conceptually determine the level of compatibility of the current market data set with the corresponding set of knowledge about the technological readiness of R&D product.

The set of data (characteristics) obtained based on studying the technological readiness and market perception of R&D product can be systematized into stimulants and destimulants of R&D product, highlighting from them – strengths and weaknesses, development factors or limitations of R&D product. To organize the characteristics of an R&D product, you can use the following list.

Strengths / weaknesses of R&D product
— stage of technological readiness;
— functional indicators of the main and additional purpose of R&D product;
— indicators of the value of technical / analytical / marketing support from the developer;
— the level of environmental safety of the product;
— the level of economic efficiency of investment in R&D product (commercialization);
— the level of costs for the use and disposal of R&D product;
— R&D product patent protection level;
— the level of uniqueness / standardization of the R&D product;

— the level of complexity of implementing an R&D product;
— the time level of expected use of the R&D product before replacement or improvement;
— the level of novelty / improvement of the main functional characteristics of the R&D product;
— scale of implementation and use of R&D products, etc.

Development factors / limitations for R&D product
— the nature and number of direct / indirect competitors of the R&D product;
— nature and number of substitutes R&D product;
— market capacity;
— stage of the R&D product lifecycle;
— dependence on R&D product participants / developers;
— the level of access to resources for using the R&D product;
— the level of availability of R&D product information, etc.

These characteristics can be adjusted according to the scope of R&D product, market specifics, comparison with market analogues, and so on.

In order to substantiate the method of technological forecasting of market behaviour of R&D product, we have proposed the following sequence of actions:
1) systematize the evaluation basis (reference matrix) of technological forecasting of market behaviour of R&D product;
2) develop a methodological approach to correcting indicators of market behaviour of R&D product;
3) substantiate the method of evaluating the results of technological forecasting of market behaviour of R&D product;
4) develop a system for selecting strategies for commercializing an R&D product based on technological forecasting of its market behaviour.

Systematization of existing methodological approaches to technological forecasting and the application of development factors and limitations of R&D product allowed us to develop a reference matrix for technological forecasting of market behaviour of R&D product – table 1.

Table 1: Reference matrix for technological forecasting of market behaviour of R&D product

Characteristics		Significance of the parties (Z)	Evaluation of parties (A)	Rank (F)
Strengths (S)	S_1	Z_{S_1}	A_{S_1}	$F_{S_1} = \frac{Z_{S_1} \times A_{S_1}}{\sum_{i=1}^{n} Z_S \times A_S}$
	S_2	Z_{S_2}	A_{S_2}	$F_{S_2} = \frac{Z_{S_2} \times A_{S_2}}{\sum_{i=1}^{n} Z_S \times A_S}$
	…	…	…	…
	S_n	Z_{S_n}	A_{S_n}	$F_{S_n} = \frac{Z_{S_n} \times A_{S_n}}{\sum_{i=1}^{n} Z_S \times A_S}$
	Σ	$\sum_{i=1}^{n} Z_S \times A_S$		1
Weaknesses (W)	W_1	Z_{W_1}	A_{W_1}	$F_{W_1} = \frac{Z_{W_1} \times A_{W_1}}{\sum_{i=1}^{n} Z_W \times A_W}$
	W_2	Z_{W_2}	A_{W_2}	$F_{W_2} = \frac{Z_{W_2} \times A_{W_2}}{\sum_{i=1}^{n} Z_W \times A_W}$
	…	…	…	…
	W_n	Z_{W_n}	A_{W_n}	$F_{W_n} = \frac{Z_{W_n} \times A_{W_n}}{\sum_{i=1}^{n} Z_W \times A_W}$
	Σ	$\sum_{i=1}^{n} Z_W \times A_W$		1
Development factors (D)	D_1	Z_{D_1}	A_{D_1}	$F_{D_1} = \frac{Z_{D_1} \times A_{D_1}}{\sum_{i=1}^{n} Z_D \times A_D}$
	D_2	Z_{D_2}	A_{D_2}	$F_{D_2} = \frac{Z_{D_2} \times A_{D_2}}{\sum_{i=1}^{n} Z_D \times A_D}$
	…	…	…	…
	D_n	Z_{D_n}	A_{D_n}	$F_{D_n} = \frac{Z_{D_n} \times A_{D_n}}{\sum_{i=1}^{n} Z_D \times A_D}$
	Σ	$\sum_{i=1}^{n} Z_D \times A_D$		1
Limiting factors (L)	F_{L_1}	Z_{L_1}	A_{L_1}	$F_{L_1} = \frac{Z_{L_1} \times A_{L_1}}{\sum_{i=1}^{n} Z_L \times A_L}$
	F_{L_2}	Z_{L_2}	A_{L_2}	$F_{L_2} = \frac{Z_{L_2} \times A_{L_2}}{\sum_{i=1}^{n} Z_L \times A_L}$
	…	…	…	…
	F_{L_n}	Z_{L_n}	A_{L_n}	$F_{L_n} = \frac{Z_{L_n} \times A_{L_n}}{\sum_{i=1}^{n} Z_L \times A_L}$
	Σ	$\sum_{i=1}^{n} Z_L \times A_L$		1

Table 2: Correlation matrix of strengths, weaknesses, development / limiting factors and indicators of R&D product market behaviour

Characteristics of R&D production			Development factors (D)				Limiting factors (L)				Market behaviour indicators (P)
			D_1	D_2	…	D_n	L_1	L_2	…	L_n	
		Ранг	F_{D_1}	F_{D_2}	…	F_{D_n}	F_{L_1}	F_{L_2}	…	F_{L_n}	
Strengths (S)	S_1	F_{S_1}	$K_{1,1}^{SD}$	$K_{1,2}^{SD}$	…	$K_{1,m}^{SD}$	$K_{1,1}^{SL}$	$K_{1,2}^{SL}$	…	$K_{1,m}^{SL}$	P_1^{SDL}
	S_2	F_{S_2}	$K_{2,1}^{SD}$	$K_{2,2}^{SD}$	…	$K_{2,m}^{SD}$	$K_{2,1}^{SL}$	$K_{2,2}^{SL}$	…	$K_{2,m}^{SL}$	P_2^{SDL}
	…	…	…	…	…	…	…	…	…	…	…
	S_n	F_{S_n}	$K_{n,1}^{SD}$	$K_{n,2}^{SD}$	…	$K_{n,m}^{SD}$	$K_{n,1}^{WL}$	$K_{n,2}^{WL}$	…	$K_{n,m}^{WL}$	P_n^{SDL}
Weaknesses (W)	W_1	F_{W_1}	$K_{1,1}^{WD}$	$K_{1,2}^{WD}$	…	$K_{1,m}^{WD}$	$K_{1,1}^{WL}$	$K_{1,2}^{WL}$	…	$K_{1,m}^{WL}$	P_1^{WDL}
	W_2	F_{W_2}	$K_{2,1}^{WD}$	$K_{2,2}^{WD}$	…	$K_{2,m}^{WD}$	$K_{2,1}^{WL}$	$K_{2,2}^{WL}$	…	$K_{2,m}^{WL}$	P_2^{WDL}
	…	…	…	…	…	…	…	…	…	…	…
	W_n	F_{W_n}	$K_{n,1}^{WD}$	$K_{n,2}^{WD}$	…	$K_{n,m}^{WD}$	$K_{n,1}^{WL}$	$K_{n,2}^{WL}$	…	$K_{n,m}^{WL}$	P_n^{WDL}

In order to assess the strengths and weaknesses, development factors and limitations of an R&D product, it is necessary to involve experts in the relevant subject area and skill level. Experts determine the significance of the parties to the R&D product (Z) and evaluate them (A) according to the expressions in table. 1.

Significance is a non – repeating indicator for each of the characteristics (according to the principle: highest value-highest significance). A rating can be assigned, for example, from the interval 1…10, where 1 is the weakest manifestation of the characteristic, and 10 is the strongest. After that, the rank of ratings is determining. It is the specific weight of the impact of each characteristic on their entire group.

When making calculations, the priority should be the value of the R&D product for the consumer, its market competitiveness, as well as the expected efficiency of commercialization.

After establishing the ranks, we determine the pairwise correlation coefficients (C) of factors influencing the R&D product and calculate the indicator of its market behaviour (P). The matrix of relationships between indicators is shown in table 2.

Therefore, K [x; y], where x and y are in the range 0…1, is the correlation coefficient of strengths and weaknesses, development factors, and limitations of the R&D product. The sum of "$x + y$" placed in the corresponding cells will reflect the strength of the influence of a particular pair of factors: *SD, SL, WD* and *WL,* respectively, on the indicator of market behaviour.

The market behaviour indicator shows the strength of the manifestation of R&D product characteristics, due to the corresponding development factors or restrictions on it. This indicator is relative and is measured in fractions of units. Calculating the market behaviours indicator helps determine the level of R&D ability of a product in relation to market development.

To determine the indicator of market behaviours of an R&D product, it is advisable to follow the following systematic procedure:

1) summarize the correlation coefficients (K) separately for the strengths, weaknesses, development factors, and limitations of the R&D product. Use expressions to determine the indicator of market behaviour of an R&D product to set its value. For this purpose, expressions (1) – (2):

$$P_{n,m}^{SDL} = \sum_{i=1}^{n} K_{n,m}^{SD} - \sum_{i=1}^{n} K_{n,m}^{SL}, \quad (1)$$

$$P_{n,m}^{WDL} = \sum_{i=1}^{n} K_{n,m}^{WD} - \sum_{i=1}^{n} K_{n,m}^{WL}; \quad (2)$$

2) interpret the obtained values of market behaviour indicators. In particular, $P_{n,m}^{SDL}$ shows how much the strengths of an R&D product supported by development factors can withstand limiting factors, $P_{n,m}^{WDL}$ reflects how much the weaknesses of an R&D product supported by developmental factors can overcome limiting factors. The "–" and "+" signs of the indicators indicate the dominance of the influence of factors of restriction or development, respectively. In general, the level of the indicator indicates the strength of influence of the characteristics used in the assessment.

In practice, when evaluating an R&D product, such important indicators of technological forecasting as trends and features of market development that will affect the market behaviour of an R&D product often fall out of sight. In particular: technology diffusion, convergence of products and markets, synergy of R&D products, multiplier and spillover effect of R&D product, etc. Such indicators are practically not taken into account by modern appraisers, which is due to the complexity of their determination. It is even more difficult to identify interdisciplinary factors influencing R&D products and their relationships. However, you cannot ignore them either. Market effects can determine the market behaviour of an R&D product and the choice of its commercialization strategy.

To take into account these aspects, it is advisable to introduce a correction factor for indicators of market behaviour of R&D product. In our opinion, the correction factor should consist of two interrelated components: I (the impact of R&D product on the market) and II (the impact of market effects on R&D product). In order to justify such a correction factor, it is advisable to apply an economic and mathematical apparatus that would be based on not only quantitative estimates, but also take into account qualitative assessments of the market behaviour of R&D product. To do this, you can apply the algorithms of fuzzy set theory.

The criterion for the effectiveness of correcting indicators of market behaviour of an R&D product based on fuzzy set theory algorithms can be expressed by maximizing the degree of suitability of the result in each specific case of technological forecasting.

The correction factor (M) of R&D product market behaviour indicators can be described by the function (3):

$$M = f\,(I - \textit{Impact of R\&D product on the market;}\ II - \textit{impact of market effects on R\&D product}), \quad (3)$$

which is a next-level system of functions (see table 3).

Table 3: Signs of adjustment of indicators of market behaviour of R&D products and their linguistic terms

Signs of adjustment of R&D products market behaviour indicators	Characteristics of signs of adjustment of indicators of market behaviour of R&D product	Meaning of linguistic terms of signs of adjustment of market behaviour indicators R&D product
I – Market impact of R&D product	*a) market effects:* market convergence, spillover effect, Diffusion, synergy, crowd effect, etc.	*Low* [– 100; – 80; – 60], *L; Admissible* [– 59; – 29.5; 0], *A; Middle* [0; 29.5; 59], *M; High* [60; 80; 100], *H*
	b) types of market changes: the level of market renewal; the number and types of market participants affected by R&D product; the level and depth of market changes that accompany the introduction of R&D product; barriers to market entry, etc.	*Low* [– 100; – 80; – 60], *L; Admissible* [– 59; – 29.5; 0], *A; Middle* [0; 29.5; 59], *M; High* [60; 80; 100], *H*
II – Impact of market effects on R&D product	*a) effects generated by R&D products:* convergence of R&D products; multiplier effect, etc.	*Low* [– 100; – 80; – 60], *L; Admissible* [– 59; – 29.5; 0], *A; Middle* [0; 29.5; 59], *M; High* [60; 80; 100], *H*
	b) organizational and technological changes of R&D products: flexibility (adaptation) of individual parameters of R&D product; the level of novelty of R&D product; the number and access to marginal resources that need to be attracted for the market development of R&D product and the level of costs for them; environmental friendliness of R&D product; scale of R&D product monetization; level of intellectual property protection, etc.	*Low* [– 100; – 80; – 60], *L; Admissible* [– 59; – 29.5; 0], *A; Middle* [0; 29.5; 59], *M; High* [60; 80; 100], *H*

To output the desired correction factor, use the Fuzzy Logic Toolbox (MATLAB). Within the framework of fuzzy set theory algorithms (based on the Mamdani model), the characteristics of signs of correction of market behaviour of R&D product are systematized by types (a, b, c, d), which are assigned the value of linguistic terms – table 3.

The correction factor of market behaviour is within the same limits as the characteristics of its features.

We have formed a knowledge base – a set of ratios of linguistic terms of adjustment features, taking into account their impact on the indicator of market behaviour of R&D product. In figure 1 shows a fragment of this knowledge base using the fuzzy logic toolbox (MATLAB).

1. If (a is Low) or (b is Admissible) or (c is Admissible) or (d is Low) then (Correction_factor is Admissible) (1)
2. If (a is Low) and (b is not Middle) and (c is Middle) and (d is Admissible) then (Correction_factor is Admissible) (1)
3. If (a is Low) and (b is Low) and (c is Low) and (d is Low) then (Correction_factor is Low) (1)
4. If (a is Low) or (b is High) or (c is High) or (d is Low) then (Correction_factor is Middle) (1)
5. If (a is Low) and (b is Low) and (c is Low) and (d is Low) then (Correction_factor is Low) (1)
6. If (a is Admissible) or (b is Admissible) or (c is Admissible) or (d is Admissible) then (Correction_factor is Admissible) (1)
7. If (a is Admissible) or (b is Admissible) or (c is Middle) or (d is Middle) then (Correction_factor is Middle) (0.7)
8. If (a is Admissible) and (b is Middle) and (c is Admissible) and (d is High) then (Correction_factor is Middle) (0.3)
9. If (a is Admissible) and (b is Low) and (c is High) and (d is Admissible) then (Correction_factor is Middle) (0.3)
10. If (a is Admissible) and (b is High) and (c is High) and (d is Middle) then (Correction_factor is Admissible) (0.7)
11. If (a is Middle) or (b is Middle) or (c is Admissible) or (d is Low) then (Correction_factor is Admissible) (0.7)
12. If (a is Middle) and (b is Admissible) and (c is Middle) and (d is High) then (Correction_factor is Middle) (0.7)
13. If (a is High) or (b is not High) or (c is High) or (d is High) then (Correction_factor is High) (0.5)
14. If (a is High) and (c is Middle) and (d is Admissible) then (Correction_factor is Middle) (0.8)
15. If (a is High) and (b is Middle) and (c is High) and (d is High) then (Correction_factor is Middle) (0.8)
16. If (a is High) and (b is Low) and (c is Middle) and (d is High) then (Correction_factor is Middle) (0.5)
17. If (a is High) and (b is not High) and (c is Admissible) and (d is Admissible) then (Correction_factor is Middle) (1)
18. If (a is Low) or (b is not Middle) or (c is Middle) or (d is Admissible) then (Correction_factor is Admissible) (1)
19. If (a is Low) or (b is Low) or (c is Low) or (d is Low) then (Correction_factor is Low) (1)

Figure 1: Knowledge base of relations of linguistic terms of adjustment features, considering their impact on the indicator of market behaviour of R&D product (fragment)

Visualization of the ratio of signs of adjustment of indicators of market behaviour of R&D product is shown in figure 2 – 3.

Implementing the described sequence, we obtain a relative correction factor, by the value of which we specify the indicators of market behaviour R&D product $P_{n,m}^{SDL}$ and $P_{n,m}^{WDL}$.

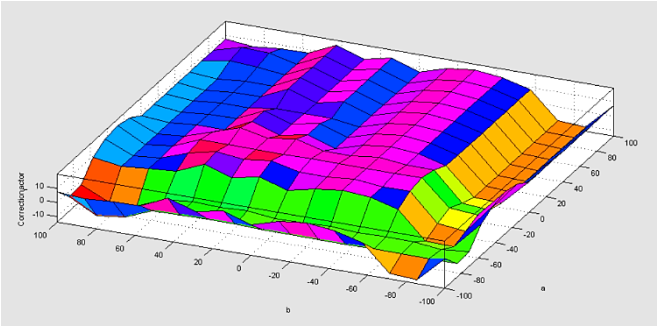


Figure 2: Visualization of the ratio of ***a*** and ***b*** signs of adjustment of R&D product market behaviour indicators

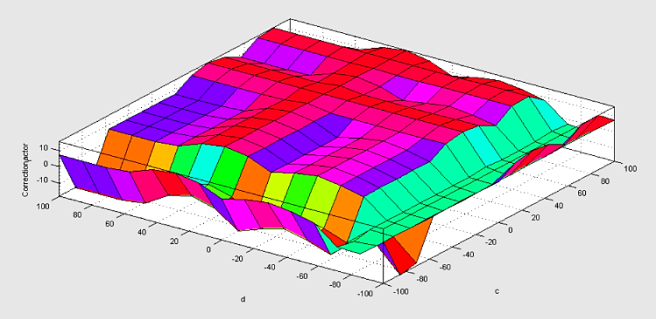


Figure 3: Visualization of the ratio of ***c*** and ***d*** signs of adjustment of indicators of market behaviour of R&D product

3.2.2. Evaluating results

There are many approaches to choosing an R&D product commercialization strategy, but none of the current ones takes into account technological forecasting of R&D product market behaviour. Taking into account the best practices of scientists [34 – 39], we have formed a system for choosing the strategy of

commercialization of R&D product, based on three methods of commercialization:

1) zero-level commercialization of R&D product;
2) commercialization of the first level of R&D product;
3) commercialization of the second level of R&D product.

If an R&D product is at the first or second stages of technological readiness (*basic principles observed and reported; technology concept and / or application formulated*), it is obvious that a zero-level commercialization strategy will be suitable for it. This strategy is the most efficient way to make a profit, since its object is the idea of an R&D product.

The object of the first – level commercialization strategy is an R&D product that is at the third or fourth stages of technological readiness: *analytical and experimental critical function and / or characteristic proof-of-concept; component and / or breadboard validation in the laboratory environment.*

The second – level commercialization strategy is possible for an R&D product at the fifth stage of technological readiness-*component and / or breadboard validation in a relevant environment.*

Calculated and adjusted values of market behaviour indicators (clause 3.2.1) should be noted in the system developed for this purpose (figure 4).

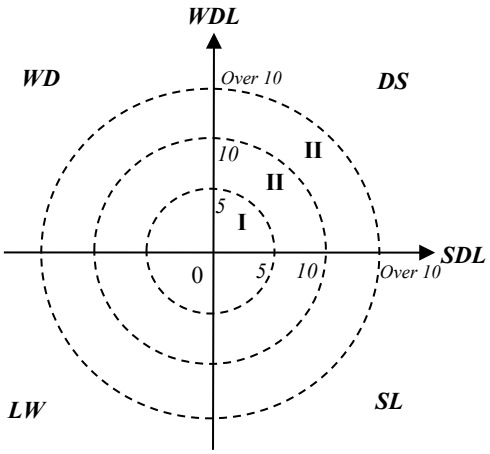


Figure 4: A system for selecting an R&D product commercialization strategy based on technological forecasting of market behaviour indicators*

** Symbols: I-zero-level commercialization strategy, II-first-level commercialization strategy; III-second-level commercialization strategy*

The system is based on a Cartesian coordinate system (on the abscissa – *SDL* axis, on the ordinate – *WDL* axis), described by three circles – the boundaries of the choice of a particular strategy for commercialization of an R&D product, according to technological forecasting of its market behaviour.

The explanation of the choice of this form of representation of the system is justified by us in [38]. All circles have a centre at the zero point. The radius of the smallest circle is 5 units, average – 10 units. For the largest circle of Radius, it is established that its radius must be more than 10 units.

According to Figure 4, the division of scenarios within certain commercialization strategies is as follows: *DS* sector – measures that are necessary to use the strengths of R&D product for its development; *WD* sector – measures that need to be carried out, overcoming the shortcomings of R&D product and using development factors in relation to it; *LW* sector – measures that minimize the shortcomings and take into account the limiting factors of R&D product; *SL* sector – measures that use the strengths of R&D product and take into account the limiting factors in relation to it.

3.3. Data analysis

The author's method has been tested on a number of developments of Ukrainian business entities. Here is an example of technological forecasting of market behaviour of R&D product – OSL dosimetry of ionizing radiation (IR) based on Yap:Mn crystals. The main purpose of R&D product is individual passive (cumulative) IR dosimetry of people working with IR sources or may be exposed to radiation from such sources. The technology was developed by scientists of the Department of semiconductor electronics of the Lviv Polytechnic National University (Ukraine) [40, 41]. The R&D product includes a dosimeter with several detectors in a cassette; detectors (based on single crystals of yttrium-aluminum oxide activated with manganese); a device for determining the radiation dose absorbed by the cassette with detectors (reader); a method for performing measurements that provides determination of dosimetric values with a given degree of accuracy and reliability.

Research of the global market of OSL dosimetry of IR in the field of security and defense and the study of this R&D product allowed us to form strengths and weaknesses, development factors and limitations in relation to it (table 4).

Table 4: Evaluation of the characteristics of OSL dosimetry of IR technology for technological forecasting of their market behaviour

	Characteristics	Z	A	$Z_n \times A_n$	F
S	Increased level of product functionality for rescuers	8	10	80	0.229
	Tissue equivalence of the dosimetric detector material and localization of dose measurements	6	9	54	0.154
	Ability to record the dose of different types of IR in a wide range of values	7	9	63	0.180
	Ability to identify the IR source type	3	8	24	0.068
	Dosimetric material with extremely high sensitivity	4	7	28	0.080
	The material of the dosimetric detector allows you to measure the absorbed dose of IR in a wide range (from units of micrograms to several tens of kg), is a more stable and stable material, compared to analogues	5	7	35	0.100
	Ability to perform a rapid assessment of radiation energy and determine the type of unknown source	2	6	12	0.034
	The ability to quickly determine the radiation dose (important for portable readers, allows you to save the dosimeter with the accumulated dose after its partial reading, makes it possible to re-read multiple times)	9	5	45	0.128

	Synthesizer of OSL dosimetry of IR technology. Package patenting	1	8	8	0.022
	Σ			**349**	**1**
W	Inflation, currency instability in the global IR dosimetry market	1	9	9	0.073
	Capital and knowledge intensity of product production	3	6	18	0.146
	Difficulty in finding a qualified workforce to start product production	5	10	50	0.406
	Significant need for investment to start production	4	8	32	0.260
	Undeveloped sales channels in the country of origin (Ukraine) [42]	2	7	14	0.113
	Σ			**123**	**1**
D	The North American region is promising from the point of view of the market launch (it is characterized by the largest global market share of manufacturers of IR dosimetry equipment and technology of 93.32 %, but the region's market is not saturated with these products) [38]	9	9	81	0.231
	Significant opportunities for product differentiation by Application Area	6	10	60	0.171
	In the country of producer (Ukraine), competitive products are presented in fragments; they are significantly more expensive than the analyzed product. The market of individual IR dosimetry in Ukraine is unsaturated today: up to 23,000 units. / year, or about $1.1 million (15.33 %) [43]	1	5	5	0.014
	Zoom effect	3	4	12	0.034
	The global dosimeters market will grow at a CAGR of 6.8% with a base value of $2.92 billion in 2020 to $4.36 billion in 2026 during the forecast period of 2020 – 2026 [44]	4	7	28	0.080
	There is a tendency to increase demand for IR dosimetry products and technology in the medical field (the radiotherapy market is projected to grow from an estimated $5.6 billion in 2018 to $6.8 billion by 2023, at a CAGR of 4.1% during the forecast period) [45]	8	8	64	0.182
	The market of manufacturers and consumers of IP dosimetry equipment is characterized by an increase in demand for OSL dosimetry of IR and a tendency to replace TLD dosimetry with OSL dosimetry, due to the expanded list of opportunities that OSL dosimetry provides	5	5	25	0.071
	The OSL dosimetry of IR market is 25% of the world market	2	6	12	0.034
	The OSL dosimetry of IR market is characterized by the expansion of the personal dosimetry segment, due to the improvement of the "performance / cost" ratio and wider application opportunities	7	9	63	0.180
	Σ			**350**	**1**
L	Low demand elasticity	1	4	4	0.024
	Insufficient communication between consumers and the manufacturer	3	5	15	0.093
	High level of specialization of competitive analogues presented in the markets	2	7	14	0.086
	Dependence on strategic relationships	4	7	28	0.173
	Product complexity, production	5	8	40	0.248
	The OSL dosimetry of IR market is monopolistic (the main market operator is LANDAUER, Inc. (USA)), which is due to the scale of production of this enterprise and legal barriers to other market participants (patent protection of technology, etc.)	6	10	60	0.372
	Σ			**161**	**1**

Table 5: Indicators of market behaviour of OSL dosimetry of IR technology

		D										*L*							P
		0.231	0.171	0.014	0.034	0.080	0.182	0.071	0.034	0.180	**Σ_1**	0.024	0.093	0.086	0.173	0.248	0.372	**Σ_2**	**Σ_1-Σ_2**
S	0.229	0.46	0.4	0.243	0.263	0.309	0.411	0.3	0.263	0.409	**3.058**	0.253	0.322	0.315	0.402	0.477	0.601	**2.37**	0.688
	0.154	0.385	0.325	0.168	0.188	0.234	0.336	0.225	0.188	0.334	**2.383**	0.178	0.247	0.24	0.327	0.402	0.526	**1.920**	0.463
	0.180	0.411	0.351	0.194	0.214	0.260	0.362	0.251	0.214	0.36	**2.617**	0.204	0.273	0.266	0.353	0.428	0.552	**2.076**	0.541
	0.068	0.299	0.239	0.082	0.102	0.118	0.250	0.139	0.102	0.248	**1.579**	0.092	0.161	0.154	0.241	0.316	0.440	**1.404**	0.175
	0.080	0.311	0.251	0.094	0.114	0.130	0.262	0.151	0.114	0.260	**1.687**	0.104	0.173	0.166	0.253	0.328	0.452	**1.470**	0.211
	0.100	0.331	0.271	0.114	0.134	0.150	0.282	0.171	0.134	0.280	**1.867**	0.124	0.193	0.186	0.273	0.348	0.472	**1.596**	0.271
	0.034	0.265	0.205	0.048	0.068	0.084	0.216	0.105	0.068	0.214	**1.273**	0.058	0.127	0.120	0.207	0.282	0.406	**1.200**	0.073
	0.128	0.359	0.299	0.142	0.162	0.178	0.31	0.199	0.162	0.308	**2.119**	0.158	0.221	0.214	0.301	0.376	0.500	**1.320**	0.799
	0.022	0.253	0.193	0.036	0.056	0.072	0.204	0.093	0.056	0.202	**1.165**	0.046	0.115	0.108	0.195	0.270	0.394	**1.128**	0.037
	Σ																		**3.258**
W	0.073	0.304	0.244	0.087	0.107	0.153	0.255	0.144	0.107	0.253	**1.654**	0.097	0.166	0.159	0.246	0.321	0.445	**1.434**	0.220
	0.146	0.377	0.317	0.160	0.179	0.226	0.326	0.217	0.180	0.326	**2.308**	0.170	0.239	0.232	0.319	0.394	0.518	**1.872**	0.436
	0.406	0.637	0.577	0.520	0.440	0.486	0.588	0.477	0.440	0.586	**4.751**	0.430	0.499	0.492	0.479	0.654	0.778	**3.332**	1.419
	0.260	0.491	0.431	0.274	0.294	0.340	0.442	0.331	0.294	0.440	**3.337**	0.284	0.353	0.346	0.433	0.508	0.632	**2.556**	0.781
	0.113	0.144	0.284	0.127	0.147	0.193	0.295	0.184	0.147	0.293	**1.814**	0.137	0.206	0.199	0.286	0.361	0.485	**1.674**	0.166
	Σ																		**3.022**

The results obtained are used to determine pairwise correlation coefficients of factors influencing the R&D product, based on which indicators of its market behaviour (table 5) are calculated.

Using the author's method for determining the corrective coefficient of market behaviour of an R&D product, with the involvement of a group of expert appraisers from the subject area and based on the fuzzy logic toolbox (MATLAB), the desired coefficient is determined for the OSL dosimetry of IR technology (figure 5).

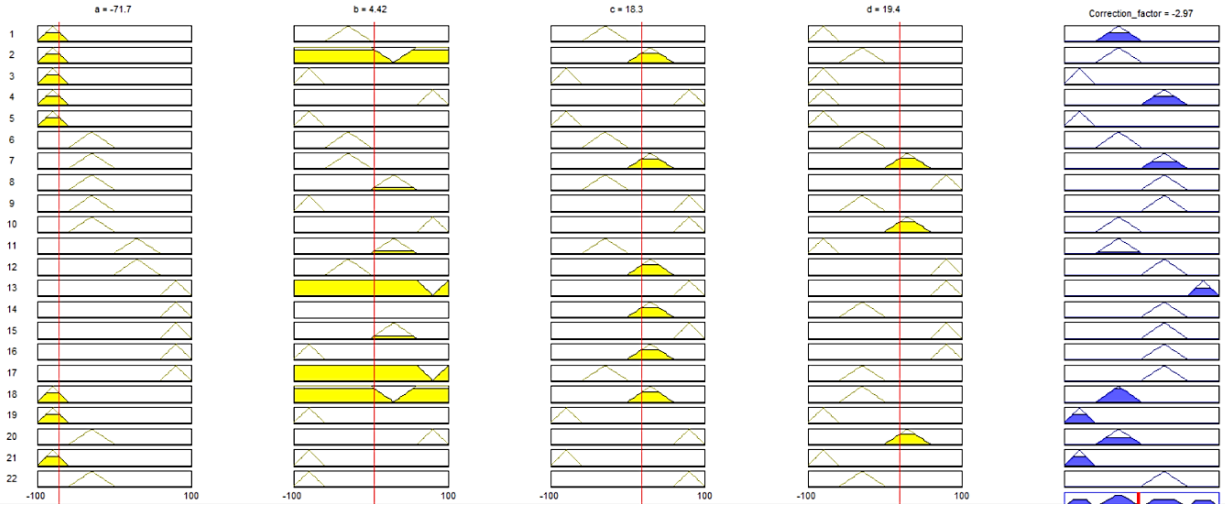


Figure 5: Results of determining the correction factor of indicators of market behaviour of OSL dosimetry of IR technology

The correction factor *M* for OSL dosimetry of IR technology will be – 2.97 %. The coordinates of the market behaviour indicator of this product, taking into account the correction factor, will be:

WDL: 3.258 * (– 2.97 %) = – 3.161;

SDL: 3.022 * (– 2.97 %) = – 2.93.

So, an indicator of the market behaviour of OSL dosimetry of IR technology in the R&D product commercialization strategy selection system (see figure 4) will have the coordinates – *K* [– 3.161; – 2.93], that is, it falls into the second matrix circle. In this case, a zero-level commercialization strategy is proposed for the product.

Since the product is located in the *LW* sector, as part of the chosen commercialization strategy for it, it is advisable to consider in detail for which positions of the correction factor the market development of the product may slow down. Based on this, appropriate corrective measures can be developed.

The result obtained correlates with the stage of technological readiness of OSL dosimetry of IR technology – *technology concept and / or application formulated* (second stage).

Testing of a reasonable method of technological forecasting of market behaviour of R&D product and the choice of commercialization strategy based on this, carried out on the example of a number of R&D products presented by the development teams of the Lviv Polytechnic National University (Lviv, Ukraine), showed the following results (table. 6).

So, in the case of technological forecasting of market behaviour for the 1st and 4th R&D products, a zero-level commercialization strategy is defined, for the 2nd, 6th and 7th – a first-level commercialization strategy, for the 3rd, 5th – a second-level commercialization strategy.

Table 6: Results of technological forecasting of market behaviour of OSL dosimetry of IR technology using the author's method

№	R&D products	Technological readiness stage of R&D product [1…5]*	Value of the correction factor, %	Coordinates (*K*) in the strategy system	Type of commercialization strategy
1	Technology of optically-stimulated luminescent dosimetry	2	– 2.97	[– 3.161; – 2.93]	Zero-level commercialization strategy
2	Continuation of the life of machine parts by the method of surfacing under a layer of flux	3	4,48	[5.45; 6.89]	First-level commercialization strategy
3	Autonomous smoke and leak detection system carbon monoscope	5	7.14	[10.38; 14.91]	Second-level commercialization strategy
4	Forecasting and providing a given value of the initial resistance of dissimilar and homogeneous materials of small thicknesses and cross sections during point capacitor contact micro welding	2	– 1.6	[– 2.85; – 4.03]	Zero-level commercialization strategy
5	Contact point capacitor micro welding equipment	5	6.31	[10.03; 11.82]	Second-level commercialization strategy
6	Technology of surface friction strengthening of working surfaces of machine parts	4	3.56	[7.21; 8.025]	First-level commercialization strategy
7	Technology of surface friction displacement of working surfaces of machine parts	4	– 1.56	[– 5.01; – 5.87]	First-level commercialization strategy

* Stages are defined according to the NASA approach [33].

Based on the R&D products that were taken to test the method, it can be applied to products of any type of economic activity, transfer method, etc. However, the use of the author's method is characterized by reservations:

1) in a situation where the chosen strategy has a marginal value (it is on the edge of the circles that determine the choice of strategies for commercialization of R&D product), decisions should be made based on the current market situation, the overall strategy of the business entity and the objective vision of the analyst;

2) to obtain reliable results when applying the correction factor of market behaviour based on the fuzzy logic toolbox (MATLAB), it is advisable to review and, if necessary, correct the set of data involved in the assessment in each evaluation situation.

4. Conclusions

The method of technological forecasting of market behaviour of R&D products at the early stages of their technological readiness is substantiated. The method is based on taking into account the strengths and weaknesses, development factors and limiting factors of the product R&D, which allows you to predict indicators of market behaviour of the product relative to the market where its commercialization is planned.

The market is characterized by an array of different types of factors that in one way or another influence the market behaviour of R&D products. Therefore, within the framework of the method of technological forecasting of market behaviour of R&D products and in order to increase the level of reliability of calculations and validity of results, a method for determining the correction factor of indicators of market behaviour of R&D product has been developed. The method is developed on the basis of fuzzy set theory algorithms using the fuzzy logic toolbox (MATLAB). This made it possible to integrate a set of data on the market behaviour of an R&D product, taking into account their relationships and interdependencies, into one correction factor. This coefficient contains the characteristics of signs of the impact of R&D product on the market (in particular, market effects, types of market changes) and the impact of market effects on R&D product (effects generated by R&D products, organizational and technological changes in R&D products). To justify the correction factor, a knowledge base was formed based on the answers of experts who have knowledge of trends, patterns and features of market development in the subject area.

To select the R&D product commercialization strategy, a system with normative indicators that interpret the types of strategies has been developed. In particular, this allowed us to offer the following types of strategies: commercialization of the zero level of R&D product; commercialization of the first level of R&D product; commercialization of the second level of R&D product.

Due to the clarification of content categories (development factors and limitations) and the introduction of a correction factor for market behaviour indicators, the results of choosing an R&D product commercialization strategy are more accurate and reliable.

The **hypothesis 1** put forward in this paper is justified. The method of justifying the indicator of market behaviour of R&D products for the purpose of technological forecasting is universal, can be applied to products of any type of economic activity, transfer method, etc.

The method of justifying the indicator of market behaviour of R&D products for the purpose of technological forecasting can only give errors if S, W, D, and L are poorly defined, compared, and weighed. Therefore, experts are recommended to pay thorough attention to studying the features of the product and market in each evaluation situation.

Hypothesis 2 is not fully justified. The proposed correction factor for indicators of market behaviour of R&D products based on fuzzy set theory algorithms makes it possible to take into account changes in the market behaviour of products only in the short term. The characteristics of the features that this coefficient is based on can only be relevant at the time of technological forecasting. To apply such methods in the medium and long term, it is necessary to take into account the volatility and singularity of the market.

Hypothesis 3 is true. Justification of the indicator of market behaviour of R&D product based on the principles of the author's method correlates with the early stages of technological readiness of the product, since: 1) the economic content of the stages of technological readiness of R&D products was laid in the author's method of technological forecasting of market behaviour of R&D products; 2) the characteristics of the signs of the correction factor of market behaviour indicators are based on phenomena, market effects, etc., which are characteristic of R&D product at a specific stage of technological readiness.

The economic interpretation of the obtained indicators of market behaviour of R&D products and the commercialization strategies selected based on this showed the validity of the author's method and its relevance in the changing conditions of the modern market. The study of the characteristics of the features of the correction factor of market behaviour indicators is the subject of the following scientific work.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We thank the developers of R&D products and scientists of Lviv Polytechnic National University (Ukraine), who shared the pearls of their knowledge and experience, which made it possible to conduct research and justify the results.

References

[1] N. Chukhray, N. Shakhovska, O. Mrykhina, M. Bublyk, L. Lisovska, “Consumer Aspects In Assessing The Suitability of Technologies for the Transfer”, in 2019 IEEE 14th International Scientific and Technical Conference on Computer Sciences and Information Technologies, 142–147, 2019, https://doi:10.1109/STC-CSIT.2019.8929879

[2] N. McCarthy, “The Top Reasons Startups Fail”, 2017

[3] D. Nag, A. Gupta, A. Turo, “The Evolution of University Technology Transfer: By the Numbers”, 2020

[4] T. Tetreault, “10 Surprising Startup Statistics”, 2019

[5] NanoCom Consortium, Best Practices to Lower the Barriers for Commercialisation of Nanotechnology Research. NanoCom Lowering Barriers for Nanotechnology Commercialisation. Barriers and Success Factors; Commercialisation Readiness Scale

[6] ”, Business Horizons, **61**(3), 453–460, 2018, https://doi.org/10.1016/j.bushor.2018.01.011

[7] S. S. Dadwal, A. Jamal, T. Harris, G. Brown, S. Raudhah, “Technology and Sharing Economy-Based Business Models for Marketing to Connected Consumers”, in Handbook of Research on Innovations in Technology and Marketing for the Connected Consumer, IGI Global, 62–93, 2020.

[8] S. Sarasini, M. Linder, "Integrating A Business Model Perspective into Transition Theory: The Example of New Mobility Services", Environmental Innovation and Societal Transitions, **27**, 16–31, 2018, https://doi.org/10.1016/j.eist.2017.09.004

[9] K. E. Bosbach, J. F. Tesch, U. C. M. Kirschner, "A Business Model Perspective on Innovation Susceptibility", in: Tesch J. (eds) Business Model Innovation in the Era of the Internet of Things. Progress in IS. Springer, Cham, 2019, https://doi.org/10.1007/978-3-319-98723-1_6

[10] R. Santa, M. Ferrer, L. M. Jørsfeldt, A. Scavarda, "The Impact of the Quality of the Service From IS/IT Departments on the Improvement of Operational Performance: the Point of View of Users of Technological Innovations", International Journal of Business Information Systems, **28**(2), 125–146, 2018, https://doi.org/10.1504/IJBIS.2018.10012923

[11] S. T. McKnight, B. Maniam, J. Robertson, "Effects of Technology Innovations on Business", Journal of Business and Behavioral Sciences, **31**(1), 84–101, 2019.

[12] G. Azar, F. Ciabuschi, "Organizational Innovation, Technological Innovation, and Export Performance: the Effects of Innovation Radicalness and Extensiveness", International Business Review, **26**(2), 324–336, 2017, https://doi.org/10.1016/j.ibusrev.2016.09.002

[13] T. Rantala, J. Ukko, M. Saunila, J. Havukainen, "The Effect of Sustainability in the Adoption of Technological, Service, and Business Model Innovations", Journal of Cleaner Production, **172**, 46–55, 2018, https://doi.org/10.1016/j.jclepro.2017.10.009

[14] E. Martínez-Ros, "Revisiting Product and Process Innovations", International Journal of Business Environment, **10**(3), 270–280, 2019, https://doi.org/10.1504/IJBE.2019.097983

[15] T. B. Long, V. Blok, K. Poldner, "Business Models for Maximising the Diffusion of Technological Innovations for Climate-smart Agriculture", International Food and Agribusiness Management Review, **20**(1030-2017-2134), 5–23, 2016, https://doi.org/10.22434/IFAMR2016.0081

[16] J. T. Eckhardt, M. P. Ciuchta, M. Carpenter, "Open Innovation, Information, and Entrepreneurship Within Platform Ecosystems", Strategic Entrepreneurship Journal, **12**(3), 369–391, 2018, https://doi.org/10.1002/sej.1298

[17] F. Damanpour, F. Sanchez-Henriquez, H. H. Chiu, "Internal and External Sources and the Adoption of Innovations in Organizations", British Journal of Management, **29**(4), 712–730, 2018, https://doi.org/10.1111/1467-8551.12296

[18] D. S. de Oliveira, M. Caetano, "Market Strategy Development and Innovation to Strengthen Consumer-based Equity: The case of Brazilian airlines", Journal of Air Transport Management, **75**, 103–110, 2019. https://doi.org/10.1016/j.jairtraman.2018.12.006

[19] M. Grębosz-Krawczyk, J. Pointet, "Co-branding Strategy as a Source of Innovation on International Market", Journal of Intercultural Management, **9**(3), 63–77, 2017, https://doi.org/10.1515/joim-2017-0014

[20] W. Liu, K. Atuahene-Gima, "Enhancing Product Innovation Performance in a Dysfunctional Competitive Environment: the Roles of Competitive Strategies and Market-based Assets", Industrial Marketing Management, **73**, 7–20, 2018, https://doi.org/10.1016/j.indmarman.2018.01.006

[21] M. Holgersson, O. Granstrand, M. Bogers, "The Evolution of Intellectual Property Strategy in Innovation Ecosystems: Uncovering Complementary and Substitute Appropriability Regimes", Long Range Planning, **51**(2), 303–319, 2018, https://doi.org/10.1016/j.lrp.2017.08.007

[22] M. Coccia, "The Theory of Technological Parasitism for the Measurement of the Evolution of Technology and Technological Forecasting", Technological Forecasting and Social Change, **141**, 289–304, 2019, https://doi.org/10.1016/j.techfore.2018.12.012

[23] S. Singh, S. Dhir, V. M. Das, A. Sharma, "Bibliometric Overview of the Technological Forecasting and Social Change journal: Analysis from 1970 to 2018", Technological Forecasting and Social Change, **154**, 119963, 2020, https://doi.org/10.1016/j.techfore.2020.119963

[24] R. Amit, C. Zott, Business Model Innovation Strategy: Transformational Concepts and Tools for Entrepreneurial Leaders, John Wiley & Sons, 2020.

[25] A. Hevner, S. Gregor, "Envisioning Entrepreneurship and Digital Innovation Through a Design Science Research Lens: A Matrix Approach", Information & Management, 103350, 2020, https://doi.org/10.1016/j.im.2020.103350

[26] J. H. Dahooie, E. K. Zavadskas, H. R. Firoozfar, A. S. Vanaki, N. Mohammadi, W. K. M. Brauers, "An Improved Fuzzy MULTIMOORA Approach for Multi-criteria Decision Making Based on Objective Weighting Method (CCSD) and its Application to Technological Forecasting Method Selection", Engineering Applications of Artificial Intelligence, **79**, 114–128, 2019, https://doi.org/10.1016/j.engappai.2018.12.008

[27] M. Züfle, A. Bauer, V. Lesch, C. Krupitzer, N. Herbst, S. Kounev, V. Curtef, "Autonomic Forecasting Method Selection: Examination and Ways Ahead", in 2019 IEEE International Conference on Autonomic Computing (ICAC), 167–176, 2019, https://doi.org/10.1109/ICAC.2019.00028

[28] L. Zhou, "Prediction of a Service Demand Using Combined Forecasting Approach", Journal of Physics: Conference Series, **887**(1), 012075), 2017, https://doi.org/10.1088/1742-6596/887/1/012075

[29] O. Mrykhina, L. Lisovska, I. Novakivskyj, A. Terebukh, V. Zhukovska, "Method of Modelling Prices for R&D Products in the Case of their Transfer from Engineering Universities to the Business", Advances in Science, Technology and Engineering Systems, **5**(5), 80–93, 2020, https://doi.org/10.25046/aj050512

[30] R. M. Elavarasan, S. Afridhis, R. R. Vijayaraghavan, U. Subramaniam, M. Nurunnabi, "SWOT Analysis: a Framework for Comprehensive Evaluation of Drivers and Barriers for Renewable Energy Development in Significant Countries", Energy Reports, **6**, 1838–1864, 2020. https://doi.org/10.1016/j.egyr.2020.07.007.

[31] N. I. Chukhrai, L. S. Lisovska, "Formation of Innovation's Consumer Utility", Actual Problems of Economics, **149**(11), 27–34, 2013.

[32] R. Puyt, F. B. Lie, F. J. De Graaf, C. P. M. Wilderom, Origins of SWOT Analysis. Academy of Management Proceedings (Ed: Atinc G.), 2020 [online] Available at: https://doi.org/10.5465/AMBPP.2020.132

[33] NASA, "Technology Readiness Levels (TRLs)", 2020

[34] H. Pylypenko, V. Prokhorova, O. Mrykhina, O. Koleshchuk, S. Mushnykova, "Cost Evaluation Models of R&D Products of Industrial Enterprises", Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, **5**, 163–170, 2020, https://doi.org/10.33271/nvngu/2020-5/163

[35] V. Dimitrova, "Commercialization of the Academic Scientific Product in Bulgaria, Economic Science, Education and the Real Economy: Development and Interactions in the Digital Age", Publishing House Science and Economics Varna, **1**, 343–356, 2020

[36] S. K. Daneshjoovash, P. Jafari, A. Khamseh, "Effective Commercialization of High-technology Entrepreneurial Ideas: a Meta-synthetic Exploration of the Literature", Journal of Small Business & Entrepreneurship. 2020

[37] W. Sutopo, R. W. Astuti, R. T. Suryandari, "Accelerating a Technology Commercialization; with a Discussion on the Relation Between Technology Transfer Efficiency and Open Innovation", Journal of Open Innovation: Technology, Market, and Complexity, **5**, 95, 2019, https://doi.org/10.3390/joitmc5040095

[38] O. Mrykhina, Technology Transfer from Universities to the Business Environment: Paradigm, Concept and Tools of Evaluation, Lviv Polytechnic Publishing House, 440, 2018.

[39] L. Halkiv, O. Karyy, I. Kulyniak, S. Ohinok, "Modeling and Forecasting Of Innovative, Scientific and Technical Activity Indicators under Unstable Economic Situation in the Country: Case Of Ukraine", Communications in Computer and Information Science, **1158**, 79-97, 2020. https://doi:10.1007/978-3-030-61656-4_5

[40] S. Ubizskii, D. Afanassyev, Y. Zhydachevskii, A. Luchechko, V. Rabyk, "Technique and Apparatus for Pulsed OSL Readout of YAP:Mn Dosimetric Detectors with Enhanced Dynamic Range," in 2019 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), Odessa, Ukraine, 1–6, 2019, https://doi.org/10.1109/UkrMiCo47782.2019.9165365

[41] D. Afanassyev, S. Ubizskii, Ya. Zhydachevskyy, A. Luchechko, A. I. Popov, A. Suchocki, "Time-resolved Pulsed OSL of Ceramic YAP:Mn phosphors", Integrated Ferroelectrics, 196:1, 24–31, 2019, https://doi.org/10.1080/10584587.2019.1591980

[42] L. Halkiv, O. Karyy, I. Kulyniak, S. Ohinok, "Innovative, Scientific and Technical Activities in Ukraine: Modern Trends and Forecasts" in 2020 IEEE 3rd International Conference on Data Stream Mining and Processing, 9204148, 321-324, DSMP 2020. https://doi:10.1109/DSMP47368.2020.9204148

[43] Ministry of Defense of Ukraine, 2020

[44] Global Dosimeter Market Forecast to 2026 by Technology, End User, Device and Geography. Dublin, Aug. 25, 2020 (GLOBE NEWSWIRE), 2020

[45] Radiotherapy Market by Type (External (IGRT, IMRT, 3D-CRT, Stereotactic), Brachytherapy (LDR, HDR)), Product (LINAC, CyberKnife, Gamma Knife, Tomotherapy, Particle Therapy, Cyclotron), Application (Prostate, Breast), End User (Hospital) – Forecasts to 2023, 2020

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 887-893 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Detection and Counting of Fruit Trees from RGB UAV Images by Convolutional Neural Networks Approach

Kenza Aitelkadi*,[1], Hicham Outmghoust[1], Salahddine laarab[1], Kaltoum Moumayiz[2], Imane Sebari[1]

[1]*Cartography-photogrammetry Department, Agronomic and veterinary Hassan II Institute, Rabat, 6202, Morocco*

[2]*GolobalEtudes Company, Rabat, 6202, Morocco*

ARTICLE INFO

Article history:
Received: 30 September, 2020
Accepted: 08 March, 2021
Online: 10 April, 2021

Keywords:
Unmanned Aerial Vehicle
RGB images
Convolutional Neural Networks
Fruit tree counting

ABSTRACT

The use of Unmanned Aerial Vehicle (UAV) can contribute to find solutions and add value to several agricultural problems, favoring thus productivity, better quality control processes and flexible farm management. In addition, the strategies that allow the acquisition and analysis of data from agricultural environments can help optimize current practices such as crop counting. The present research proposes a methodology based on the exploitation of deep learning approach, especially Convolutional Neural Networks (CNN) on UAV data for fruit tree detection and counting. We build models for the automatic extraction of fruit trees. This approach is divided into main phases: dataset pre-treatment, implementing a fruit trees detection model by exploiting several CNN architectures, validating and comparing the performances of different models. The exploitation of RGB UAV images as input information will allow the learning models to find a statistical structure, which will result in rules capable of automating the detection task. They can be applied to new images for automatically identify and count fruit trees. The application of the methodology on collected data has made it possible to reach estimates of detection and counting until 96 %.

1. Introduction

Currently, agriculture continues to modernize and follow the evolution of new technologies to improve production practices and crop management. That has become a need in many countries due to the increasing demand for food. The use of new technologies, such as Unmanned Aerial Vehicle (UAV), can help find solutions and add value to several agricultural problems, thus promoting productivity, better quality control processes and flexible tree management [1]. Additionally, strategies that allow data analysis from agricultural environments, including artificial vision systems, can help optimize current practices, such as crop counting, yield estimation, diseases detection and classification of crop maturity [2]. Information on the number of plants in a crop field is essential for farmers as it helps them estimate productivity, assess the density of their plantations and errors occurring during the planting process [3]. From a perspective of detection, delimitation and counting of trees and in particular fruit trees, in [4] the author developed and tested the performance of an approach, based on RGB UAV imagery, to extract information about individual trees in an orchard with a complicated background which includes apple and pear trees. In [5], the author proposed an efficient method for an individual trees segmentation and the measurement of the width and area of identified trees crowns, based on images acquired by RGB UAV camera. The collected images from a peach orchard in Okayama, Japan, were integrated into Pix4Dmapper software for processing and generation of derived products (Digital Surface Model DSM). Using the intersection of the polygons corresponding to the peach branch line with the summer season DSM as markers indicating the sources of flooding, authors were able to delineate the peach trees crown via watershed segmentation. In [6], the author developed a specialized approach for citrus detection based on the DSM extraction. In [7], the author applied the Canny filter for edge detection applied to the images followed by the "Circular Hough Transform CHM" algorithm thus achieving the extraction and delineation of citrus trees.

The use of UAV in various arboriculture applications has many advantages and benefits. However, this depends on the types of sensors, mission objectives and their platforms [8], [9]. Nevertheless, there are some problems that need to be considered when using UAV such as the reliability of the platform, sensor

*Corresponding Author: K. Aitelkadi, Agronomic and veterinary Hassan II Institute, k.aitelkadi@iav.ac.ma

capacity and adequate treatment of the images [10]-[12]. To solve the problem of image processing which presents a main source of the success of predictions or good estimation, researchers have in recent years turned to artificial intelligence algorithms, particularly deep learning and especially the convolutional neural network. In [13], the author evaluated the use of convolutional neural network (CNN) -based methods for the detection of legally protected tree species. Three detection methods were evaluated: Faster R-CNN, YOLOv3 and RetinaNet. In [14], the author proposed a deep learning method to accurately detect and count bananas. In [15], the author exploited convolutional neural networks for the detection and enumeration of citrus fruits in an orchard in Brazil characterized by the high density of its trees. In [16], the author developed an automatic strawberry flower detection and yield estimation system.

Following the state of art review, we have selected some of efficient CNN architectures that we have implemented and tested on our context and data. The second phase is the preparation of the data especially the collection, cleaning, analysis, visualization and necessary treatments. The third phase consist to conceive, implement and analysis of the models performance. One of our objectives is to understand the advantage of one architecture over another. We have started by testing different architectures and their hyperparameters. The knowledge acquired as a result of these experiments allowed us to understand the influence of several parameters on the expected performance and to subsequently build a model intended to detect fruit trees.

The rest of this paper is organized as follow. Section 2 gives the material and the used methods. The experimental results and setup are shown in section 3. Section 4 presents the result discussion and section 5 the conclusion followed by the most relevant references.

2. Material and Methods

2.1. Material

2.1.1. Data collection

Among the many challenges of deep learning algorithms is the data collection which is considered to be one of the most critical points in the processes of artificial intelligence in general and deep learning in particular. The required time to run a deep learning algorithm depends on data preparation including collection, cleaning, analysis, and visualization. To answer this problem, we consulted several sources, resulting in a fairly large repertoire of images serving as the basis for feeding our algorithm. In this work, we will not treat the acquisition step and orthophoto establishment. We are limited in their uses and treatment for a successful training operation.

We present in Table 1 the number and size of each of the data acquired as well as the source.

Table 1: The number, size and source of data tested in this study

Data	Image number	Size	Camera resolution MPx	Altitude	Type of flight	Source
D1	103	960 x 540	-	-	-	https://github.com/skygate/skyessays
D2	170	400 x 400	-	-	-	(http://data.neonscience.org/
D3	1	9649 x 4532	20	120	Ebee sensefly	www.etafat.ma
D4	17	5472 x 3846	18.6	80	Ebee	Globetude company
D5	13	4896 x 3672	16	40	DJI	www.terramodus.ma

2.1.2. Computer tools and deep learning libraries used

Keras

Keras is a high-level neural network Application Programming Interface (API) written in Python and interfaceable with TensorFlow, CNTK and Theano. It was developed with the aim of allowing rapid experiments [17]. Keras can allow rapid and easy prototyping (due to its user-friendliness, modularity and extensibility). It supports both convolutional networks and recurrent networks as well as a combination of the two. Also, it Works seamlessly on CPU and GPU.

Darknet

Darknet is an open source neural network framework written in C and CUDA. It is fast, easy to install, and supports CPU and GPU computing [18]. It was developed by Joseph Redmon. Unlike Keras which is well known as much as a deep learning library, Darknet on the other hand is the library where versions of the YOLO object detector are implemented.

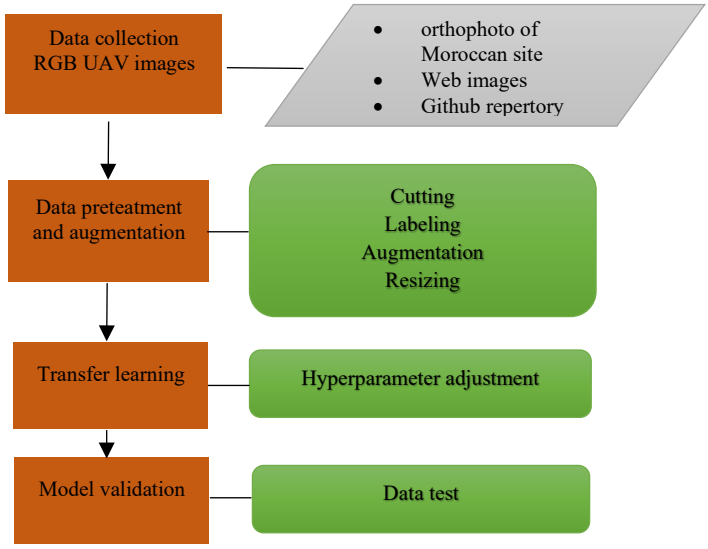


Figure 1. Methodology flowchart.

Google colaboratory

For any neural network, the training phase of the deep learning model is the most resource-intensive task. During training, a neural network receives input data, which is then processed in hidden layers using weights which are adjusted during training and the model then gives a prediction. The weights are adjusted to find patterns to make better predictions. Memory in neural networks is needed to store input data, weight parameters, and activations as an input propagates through the network [19]. Due to the memory and limited power of our computer, we used the Google

colaboratory to perform all of our deep learning operations. Google Colab or Colaboratory is a cloud service, offered by Google (free), based on Jupyter Notebook and intended for training and research in machine learning. This platform allows you to train machine learning models directly in the cloud (Google Colab: The Ultimate Guide). Colab provides a free Tesla K80 type GPU graphics processor and 13 GB of random-access memory (RAM) that work entirely in the cloud.

2.2. Methods

Our methodology is based on a following processes as presented in figure 1.

In the following sections we develop the methodology process.

2.2.1. Cutting operation

In order to increase the images number for the algorithm training, a cropping operation was performed on large images as well as the orthophoto. This operation consisted of splitting the original image to smaller images with a dimension of 816 x 816 pixels, the number of images resulting from the trimming operation depends on the initial dimensions of the image. This was done using the OpenCV image processing graphics library on Python. Figure 2 shows how an image of dimension 4 x 4 is divided into images of dimension 2 x 2.

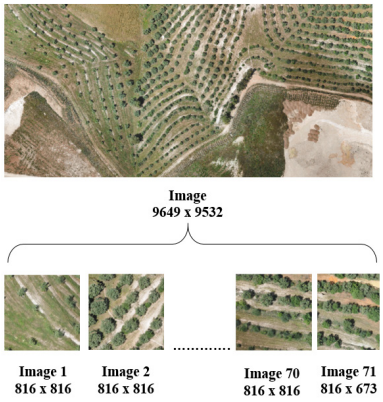


Figure 2: Cutting the orthophotographie using OpenCV algorithm

2.2.2. Labeling of images

Labeling images is a human task that involves annotating an image with labels. These labels are predetermined by the person and are chosen to give the computer vision model information about what is shown in the image. The tool used for this task is LabelImg. This is a graphical image annotation tool, it is written in Python and uses for its GUI. Annotations are saved as XML files in Pascal VOC format, the format used by ImageNet. Besides, it also supports YOLO format.

2.2.3. Image augmentation

One of the main difficulties in training a CNN model is that of overfitting. That is, the model produced fits too well on the training data. But, therefore the generalization error of the model is much too high, in other words the model modifies its predictions based on the size, angle and position of the image. To deal with the over-adjustment problem, several techniques are used and among them: Dropout, Transfer learning, Batch normalization and data augmentation. Unlike the techniques mentioned above, increasing the data resolves the problem from its origin, namely the training data set. This is done on the assumption that more information can be extracted from the original dataset through augmentation. Data augmentation is the technique of increasing the size of the data used to train a model. To obtain reliable detection, deep learning models often require a lot of training data, which is not always available. Therefore, the existing data is augmented in order to obtain a better generalized model. Some of the most common data augmentation techniques used and applied to our images are listed below: Scaling, Rotation, Translation, Shear, Brightness, Contrast, Saturation, Hue, Noise, and Blur the image. Figure 3 presents some criteria used in the augmentation phase.

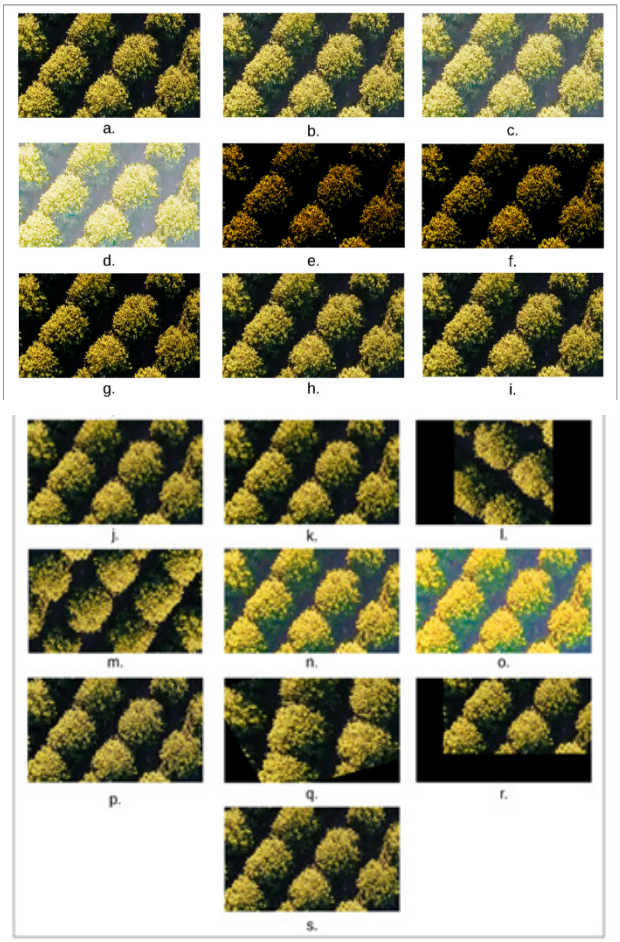


Figure 3: Augmentation algorithm by change of image pixel value : (a) initial image. (b) brightness γ = 1.5. (c) brightness γ = 2.5. (d) brightness γ = 5. (e) brightness γ = 0.3. (f) brightness γ = 0.4. (g) brightness γ = 0.7. (h) Gaussian Blur = 2. (i) Gaussian Blur = 1. (j) Average Blur. (k) Median Blur. (l) rotation θ = 90 °. (m) rotation θ = 180 °. (n) saturation = 50. (o) saturation = 100. (p) Sharpen. (q) Shear. (r) translation. (s) noise.

2.2.4. Resizing training images

We opted to resize the training images because there is a correlation between the dimensions of the images used for training the model and the RAM size needed for the model training. The dimensions chosen for the images training are 416 x 416 and 608 x 608. By resizing the images, all the annotation files must be adapted to the new dimensions of the images so that each enclosing rectangle drawn during the annotation phase remains placed around the object. For this task, the algorithm used is written in python using the OpenCV library.

2.2.5. CNN models used for fruit trees detection

Considering the huge set of applications for the object detection, a very large number of models suitable has been introduced. In what follows, we present the three models chosen for the realization of our work. To make our choice, we are based on the model's performance as well as on the models used for our similar application. The three models chosen are:

• YOLOv4

• YOLOv3

• RetinaNet-101

YOLO, You Only Look Once, is a real-time object detection system that recognizes various objects in a single enclosure. In addition, it identifies objects faster and more accurately than other recognition systems [20]. A modern detector is generally made up of two parts, a "backbone" part of the detector which is pre-trained on ImageNet and a Head which is used to predict the classes and limit boxes of objects. For detectors running on a GPU platform, their base network can be VGG, ResNet, ResNeXt, or DenseNet. For detectors running on a CPU platform, backbone can be SqueezeNet, MobileNet or ShuffleNet. As for the main part of the detector head, it is generally classified into two categories, namely the single stage object detector and the two-stage object detector. The most representative two-stage object detector is the R-CNN series, which includes Faster R-CNN, R-FCN and R-CNN Libra. For the one-step object detector, the representative models are YOLO, SSD and RetinaNet.

YOLOv4 network implements CSPDarknet53 as the backbone network. YOLOv4 is considering a few options for the neck, including: FPN, PANet, NAS-FPN, BiFPN, ASFF, SFAM. Neck components generally flow from top to bottom between layers and connect only a few layers at the end of the convolutional network. As part of this work, we chose PANet for the aggregation of network characteristics whose efficiency has been approved by the authors [21]. In addition, we have added an SPP block after CSPDarknet53 to increase the receive field and separate the most important features of the backbone.

YOLOv3 addresses object detection as a problem of direct regression from pixels to the coordinates of bounding rectangles and class probabilities. The input image is divided into S × S tiles. For each bounding rectangle, an objectivity or confidence score is predicted by logistic regression, which indicates the probability that the bounding rectangle in question has an object of interest. In addition, the probabilities of a class C are estimated for each bounding rectangle, which indicates the classes it may contain. In our case, each bounding rectangle can contain a fruit tree or the background (interesting object). Thus, each detection in YOLOv3 is composed of four parameters for the bounding rectangle (coordinates), an objectivity or confidence score and class C probabilities. To improve the precision of detection, YOLOv3 predicts these bounding rectangles at three different scales using the idea of setting up a network of pyramids. As a backbone network YOLOv3 uses Darknet-53.

RetinaNet is a single stage object detector similar to YOLOv4 and EfficientDet. However, unlike these two detectors presented previously and which both recent (2020), RetinaNet on the other hand was introduced in 2018. It is based on the ResNet network which makes it possible to add a connection connecting the input of a layer with its release. In order to reduce the number of parameters, ResNet does not have fully connected layers. GoogleNet and ResNet are much deeper but contain fewer parameters. This can make them more expensive in memory during training.

2.2.6. Hyperparameters selection

Several hyperparameters have been tested and iterated by empirical tests. We present in what section the hyperparameters whose variation has remarkably influenced the detection results rate.

The epochs number, stages per epoch and the size of the batch

Epoch's number is a hyperparameter that defines the number of times the training algorithm will work through the training data set. The size of the batch is a hyperparameter which defines the number of samples to be processed before updating the internal parameters of the model. The number of samples reviewed in a single epoch is set simultaneously based on the number of steps per epoch and the batch size. For example, taking a batch size of 32 and a number of steps of 1000 per epoch, then the number of the network samples worked on is 32000.

Learning rate

Considered to be the most important of all hyperparameters. If it is too low the convergence is slow, if it is too large the gradient descent algorithm may unintentionally increase the training error rather than decrease it.

Subdivisions Number

This is a specific hyperparameter to YOLO models and which allows the batch to be subdivided into mini batches which will be supplied to the machine during training. For example, a batch size of 64 and a number of subdivisions 16 means that 4 images will be loaded at the same time. It will take 16 of these mini batches to perform a full iteration. For all the hyperparameters, several values were tested for each of the models and on the basis of the performance obtained on the validation data that we validate the best hyperparameters for each of the models. Table 2 presents the hyperparameters values tested during training phase.

Table 2: The hyperparameters values tested during training phase

Hyperparameters	Tested value
Epochs number	Variable according to training time
Learning rate	0.01, 0.001, 0.0001, 0.00001
Subdivisions Number	16, 8

3. Results and experimentation

3.1. Data Preparation Results

The augmentation operation allowed us to have, from a limited number of images, a very large set. We distributed the images obtained respectively into training image, validation image and test image as presented in table 3.

Table 3: Distribution of images on the training, validation and test sets

Set	Training	Validation	Test
Rate	80%	10%	10%

Total number of images	2355	263	266
Tree number	140293	17605	24637

The size of the adopted images depends on the network size and the GPU capacity (graphics card) used for training. You must insert a reasonable size batch into the GPU memory. We did some tests with different images sizes. Then we compared the details obtained for the different sizes and the training time to finally choose the most suitable one. The dimensions of images used for training the models are 416 x 416 and 608 x 608. These dimensions are multiples of 32 since YOLO is designed in such a way that the input images for its training have this characteristic.

3.2. Model training and hyperparameters choice

The goal during the training phase is to minimize the loss function. The optimization algorithm used during this work is that of the gradient descent applied to batches (Mini-batch gradient descent). This method updates the weights for each batch consisting of N training samples. At each iteration, we therefore take a batch. It is said that an epoch has passed after the entire dataset is covered. The number N of samples in a batch represents another hyperparameter to be adjusted. During training, there are several choices available with respect to the sizes of the hyperparameters. These are the parameters whose value are set before the start of the learning process. They cannot be learned directly from data and must be predefined. Unfortunately, there is no generic way to determine the best hyperparameters. The optimization or modification of hyperparameters is the major problem for training our model. The same model with different choices of hyperparameters can generalize different data models. Additionally, the same type of machine learning model may require different constraints, weights, or learning rates to generalize different data models. So we experimented with different values of these parameters including the learning rate, the batch size and the network architecture. We train our model and at the end we choose the parameters that provide the best precision. Table 4 shows the values taken from the hyperparameters to reduce the loss function.

Table 4: the list of the different pre-trained models and their hyperparameters

Model	Basic network	Image size	learning rate
RetinaNet 101	ResNet 101	608	0.00001
YOLOv4	CSPDARKNET53	416	0.001
YOLOv3	DARKNET53	416	0.001

The loss function of training data is plotted over time. We start with the RetinaNet architecture by referring to the hyperparameters presented in Table 4. Figure 4 presents the evolution of loss functions by epoch

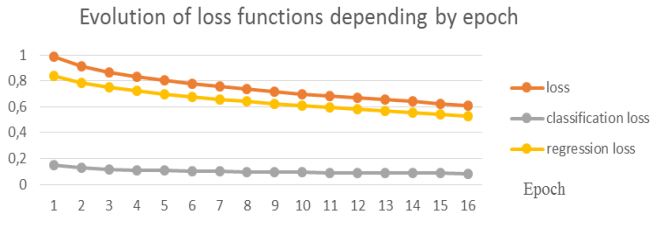


Figure 4: The evolution of loss functions by epoch.

We notice a decrease in the loss function. The experienced learning rate is 0.00001. We also observe that the loss function for the classification has a lower value compared to the other two. This is explained by the fact that we only have one class which is fruit trees.

Unlike RetinaNet which is executed using the Keras library and which allows to visualize the progress of the loss function during the training using the tenserboard module, with Darknet (base network of YOLOv3 and YOLOv4) the progress of the training is not recorded. Nevertheless, we were able to visualize his progress during the training.

For the YOLOv3 model, we could not complete training for a frame size of 608 due to the size of the RAM. We were satisfied with images of dimensions 416 x 416 for the training. Furthermore, it was possible to use a larger batch size for YOLOv3 such as 64. This is made possible thanks to a hyperparameter called subdivision. The smaller the number of subdivisions, the faster the training and the more memory it requires. For example, a number of subdivisions of 8 means that 8 images will be loaded at a time, so more memory will be consumed. For us, the number of subdivisions that can support our work environment is 16 that is mean to load 4 images at once for a batch size of 64. Table 5 shows the values of the loss function of the YOLOv3 model.

Table 5: Values of the loss function at the start and the end of learning- YOLOv3.

Model	Batch size	Subdivision Number	Loss at the start of learning	Loss at the end of learning for 5000 iterations
YOLOv3	64	16	2500	3.67

The results of the YOLOv4 model will be presented in a similar way to those of YOLOv3. Unlike YOLOv3, the training for YOLOv4 has been done up to 10000 iterations mainly because the training speed is much faster compared to YOLOv3. Similar to YOLOv3, YOLOv4 is run using DarkNet and not Keras. Thus, we cannot obtain at the end of the training the progress of the latter. However, we can get an idea of the behavior of the loss function. This function has been tested with a learning rate of 0.001 and 0.0001. This decrease is much faster with a learning rate of 0.001. Table 6 presents the values of the loss function of YOLOv4.

Table 6: The values of loss function at the start and the end of learning- YOLOv4

Model	Batch size	Subdivision Number	Loss at the start of learning	Loss at the end of learning for 5000 iterations
YOLOv4	64	16	2800	12.5624

3.3. Detection rate for the 3 architectures

As test data, we will use the 10% of the images reserved for this step. Figure 5 shows the detection results:

From the detection result, we can see that the two models YOLOv3 and YOLOv4 give better predictions. In addition, the RetinaNet model fails to give good detections of trees if the number of trees in the images is very small. Table 7 presents the predictions of tree number detection in UAV images:

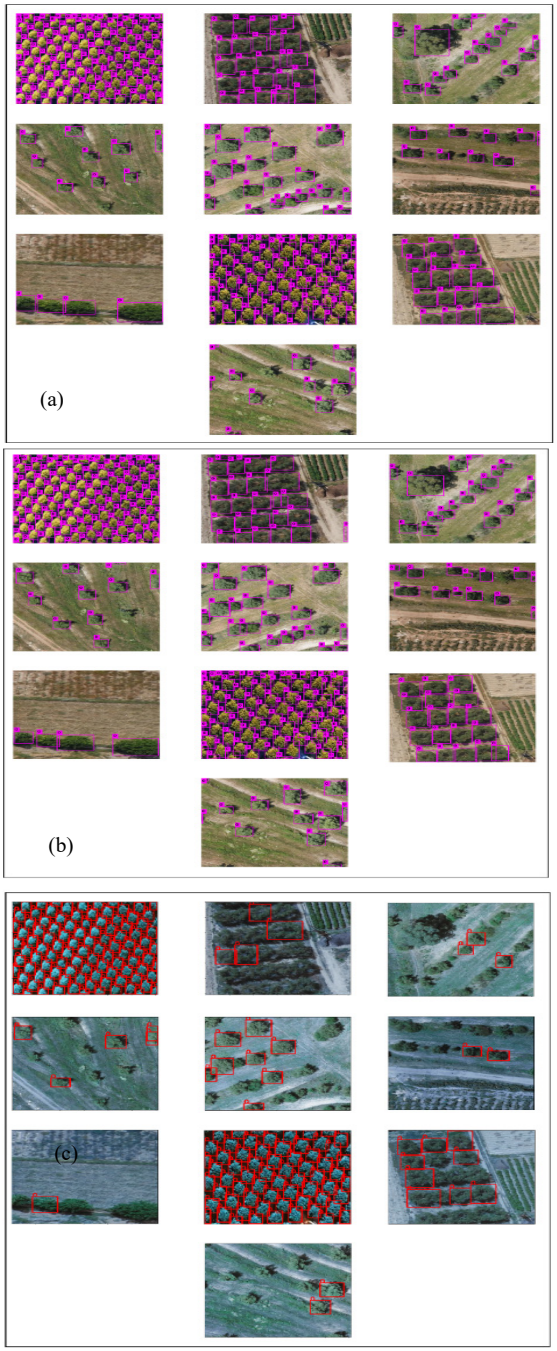


Figure 5: Example of detection results: (a) YOLOv4, (b) YOLOv3, (c) RetinaNet101.

Table 7: The detections obtained for the three tested models

Model	Detection rate %
RetinaNet-101	73%
YOLOv3	80%
YOLOv4	96%

The results obtained demonstrate that the YOLOv4 model is the model which gave the best validation and test results.

The reasons for this choice are as follows:

- ✓ Yolov4 allowed a faster improvement of the loss function compared to other architectures. For the same number of epochs, this reaches its lowest value. In addition, it does not require a lot of time for training.
- ✓ Yolov4 is the model which sacrifices the least precision to improve its Recall.
- ✓ Yolov4 provides very satisfactory results on the test set where it gives better detections with minimum time and better detection speed.

4. Discussion

Fruit trees have a very complex and heterogeneous form with vegetation classes that appear visually similar in orchards. Their categorizations require the detection model to be able to represent the spatial context of the image appropriately by learning a set of attributes adequate to distinguish between the different categories of the agricultural scene and the efficient extraction of fruit trees. Such an analysis is often carried out on images with very high spatial resolution and which have sufficient detail. Hence the considerable contribution of using RGB UAV images for the detection and enumeration of fruit trees.

The proposed approach based on the Yolov4 model that we implemented, succeeded in identifying and counting these trees. The results obtained on the validation and test sets are very satisfactory, it gave an accuracy of 96%. In addition, the results on various evaluation metrics undoubtedly confirm the contribution of convolutional neural networks algorithms. Despite their good performance, the use of convolutional neural networks presents challenges, the most important of which is the recurrent risk of overfitting. CNNs are big consumers of data. The general rule of learning algorithms is that a model trained on a large amount of data produces, in the majority of cases, much better results on new data than a model trained on small amounts. The amount of data required is not fixed and depends closely on the complexity and the mission objective.

Finally, despite the relevance of the obtained results, our work nonetheless presents some limitations which are most often due to the available computing resources which have proved insufficient for the conduct of some experiments. In recent years, the graphics processor (GPU) has established itself as an important and even indispensable player in heavy calculations that have long been done only by central processors (CPUs). Fortunately, the Google colaboratory cloud space offers excellent technical specifications and allowed us to run several experiments related to this work. Furthermore, we cannot present an experiment that takes place over a succession of stages without addressing the notion of time. Time is one of the deciding factors in choosing the best architecture. Of course, while weighting with the detection rate obtained. Table 5 presents a comparison of the overall training time for the three tested models.

Table 5: The overall training time for the three tested models

Model	Overall training time
RetinaNet-resnet101	11h
Yolov3	5h
Yolov4	8h08min

According to table 5, it can be seen that the Yolov3 and yolov4 model require less time than the Retinanet-101 model. Also the Yolov3 model consumes less time in the training phase. However, as we have mentioned before. Time cannot be the only parameter in the choice of the model. It will be necessary to weight according

to the precision of detection. In this sense, Yolov4 allowed us to obtain the best precision in a fairly satisfactory time.

5. Conclusion

We have tried to simplify the various experiments carried out either on the hyperparameters or on the architectures obtained. Our main goal is to obtain good details on the fruit tree detection and counting. This was accomplished with a detection success to 96%. These results are generally satisfactory given the difficulties encountered either in the availability of data or in the implementation of architectures. During this work, a considerable amount of time is spent for data preparation, collection, cleaning, analysis, visualization and the necessary treatments namely resizing, augmentation and annotation. Our primary concern is to find a sufficient number of images of fruit trees taken by UAV and which can be used subsequently to train, validate and test by the deep learning algorithm. In addition, the field of deep learning is constantly evolving and requires having solid basic knowledge and keeping up with all the news. Something that slowed down our work and took a long time for us to resolve implementation errors. Finally, the complexity of the calculations generated by the different models always makes their training requires more time and powerful machines. In conclusion, despite all these difficulties, it emerges that convolutional neural networks are very promising for the detection and enumeration of fruit trees and open a very interesting field of application for other uses such as yield estimation.

References

[1] K. R. Krishna, "Agricultural drones: A Peaceful Pursuit,"1e edition Editions Apple Academic Press, Inc. USA & Canada, 2018.

[2] J. P. Vasconez, J. Delpiano, S. Vougioukas, F. Auat Cheein, "Comparison of convolutional neural networks in fruit detection and counting : A comprehensive evaluation," Computers and Electronics in Agriculture, **173**, 105348, 2020, doi:10.1016/j.compag.2020.105348.

[3] Y. Ampatzidis, V. Partel, "UAV-Based High Throughput Phenotyping in Citrus Utilizing Multispectral Imaging and Artificial Intelligence, " Remote Sens., **11**(4), 410, 2019, doi:10.3390/rs11040410.

[4] X. Dong, Z. Zhang, R. Yu, Q. Tian, X. Zhu, "Extraction of information about individual trees from high-spatial-resolution UAV-acquired images of an orchard," Remote Sensing, **12**(1), 133, 2020, doi:10.3390/rs12010133.

[5] Y. Mu, Y. Fujii, D. Takata, B. Zheng, K. Noshita, K. Honda, S. Ninomiya, W. Guo, "Characterization of peach tree crown by using high-resolution images from an unmanned aerial vehicle," Horticulture Research, **5**(1), 1–10, 2018, doi:10.1038/s41438-018-0097-z.

[6] A.O. Ok, A. Ozdarici-Ok, "2-D delineation of individual citrus trees from UAV-based dense photogrammetric surface models," International Journal of Digital Earth, **11**(6), 583–608, 2018, doi:10.1080/17538947.2017.1337820.

[7] D. Koc-San, S. Selim, N. Aslan, B.T. San, "Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform," Computers and Electronics in Agriculture, **150**, 289–301, 2018, doi:10.1016/j.compag.2018.05.001.

[8] A.C. Watts, V.G. Ambrosia, E.A. Hinkley, "Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use," Remote Sensing, **4**(6), 1671–1692, 2012, doi:10.3390/rs4061671.

[9] S. Nebiker, A. Annen, M. Scherrer, D. Oesch, "A light-weight multispectral sensor for micro UAV—Opportunities for very high resolution airborne remote sensing," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, **37**(B1), 1193–1200, 2008.

[10] C. Zhang, J.M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: a review," Precision Agriculture, **13**(6), 693–712, 2012.

[11] A.S. Laliberte, A. Rango, J.E. Herrick, "Unmanned aerial vehicles for rangeland mapping and monitoring: a comparison of two systems," in ASPRS Annual Conference Proceedings, 2007.

[12] P.J. Hardin, T.J. Hardin, " Small‐scale remotely piloted vehicles in environmental research," Geography Compass, **4**(9), 1297‐1311, 2010, doi:10.1111/j.1749-8198.2010.00381.x.

[13] A.A. dos Santos, J. Marcato Junior, M.S. Araújo, D.R. Di Martini, E.C. Tetila, H.L. Siqueira, C. Aoki, A. Eltner, E.T. Matsubara, H. Pistori, "Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs," Sensors, **19**(16), 3595, 2019, doi:10.3390/s19163595.

[14] B. Neupane, T. Horanont, N.D. Hung, "Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV)," PloS One, **14**(10), e0223906, 2019, doi:10.1371/journal.pone.0223906.

[15] M. Zortea, M.M.G. Macedo, A.B. Mattos, B.C. Ruga, B.H. Gemignani, "Automatic citrus tree detection from uav images based on convolutional neural networks," in 2018 31th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018, doi:10.3390/s19245558.

[16] Y. Chen, C. Hou, Y. Tang, J. Zhuang, J. Lin, Y. He, Q. Guo, Z. Zhong, H. Lei, S. Luo, "Citrus tree segmentation from UAV images based on monocular machine vision in a natural orchard environment," Sensors, **19**(24), 5558, 2019, doi:10.3390/s19245558.

[17] A. Gulli, S. Pal, Deep learning with Keras, Packt Publishing Ltd, 2017.

[18] S.R. Masurkar, P.P. Rege, "Human Protein Subcellular Localization using Convolutional Neural Network as Feature Extractor," in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE: 1–7, 2019, doi:10.1109/ICCCNT45670.2019.8944812.

[19] T. Carneiro, R.V.M. Da Nóbrega, T. Nepomuceno, G.-B. Bian, V.H.C. De Albuquerque, P.P. Reboucas Filho, "Performance analysis of google colaboratory as a tool for accelerating deep learning applications," IEEE Access, **6**, 61677–61685, 2018, doi:10.1109/ACCESS.2018.2874767.

[20] A. Bochkovskiy, C.-Y. Wang, H.-Y.M. Liao, "Yolov4: Optimal speed and accuracy of object detection," ArXiv Preprint ArXiv:2004.10934, 2020.

[21] D. Wu, S. Lv, M. Jiang, H. Song, "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments," Computers and Electronics in Agriculture, **178**, 105742, 2020, doi:10.1016/j.compag.2020.105742.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 894-906 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Assessment of Electricity Industries in SADC Region Energy Diversification and Sustainability

Kakoma Chilala Bowa[1,*], Mabvuto Mwanza[2], Mbuyu Sumbwanyambe[3], Kolay Ulgen[4], Jan-Harm Pretorius[1]

[1]*Faculty of Engineering & Built Environment, University of Johannesburg, Johannesburg, 524, South Africa*

[2]*School of Engineering, University of Zambia, 32379, Zambia*

[3]*Department of Electrical and Mining Engineering, University of South Africa box 392 South Africa*

[4]*Solar Energy Institute, Department of Energy Technology, Ege University, Izmir, 35100, Turkey*

ARTICLE INFO

Article history:
Received: 24 October, 2020
Accepted: 11 March, 2021
Online: 10 April, 2021

Keywords:
Energy Sustainability
Social, Environmental & Economical (SEE)
Renewable Energy Technologies (RETs)
SADC
Zambia

ABSTRACT

Before the COVID-19 crisis, the Southern African Developing Countries (SADC) had a varied energy mix including renewable energy, fossil fuels, and military energy production. The use of fossil fuels in the energy mix is known to be the source of the growing levels of greenhouse gases in the atmosphere. However, there was a reduction in GHG emissions following the pandemic, which reduced travel and trade, and worldwide disruption in economic activities. The priority of priority B in the 2015-2020 Regional Indicative Strategic Development Plan, which is Energy, continues. As a result, the availability of affordable and renewable energy is still a priority for south of the equator countries and their growth agenda. This paper is aimed at exploring the sustainability of SADC countries' electricity sectors by using three sustainability pillars: Social, Environmental and Economic (SEE). SEE offers the main concepts of renewable energy, in a way that is socially, environmentally appropriate and economically viable. Study shows a gap in access rate in SADC countries with only Mauritius and Seychelles reaching 100% access to modern energy services (electricity) for both rural and urban areas. Currently all the member countries have set their RE goals for the year 2030. However, the subsidies by SADC member countries indicate that they are practiced as a way to make electricity affordable, and also to make electricity available to lower income households. In the period 2014-2017, big national budget deficits happened in various Southern African countries because of subsidies. Thus, this paper is of crucial importance to the foundational advancement of sustainable electricity sector growth in the country. The findings of this paper play a crucial role in helping and guiding politicians to better understand the existing and challenges future in the energy market and alternatives to address these problems. Additional research is given on how to arrive at sustainable decisions for the electricity sector in the region.

1. Introduction

This paper is an extension of work originally presented at the IEEE 2nd conference on renewable energy and power engineering [1]. In the past decade, the electricity industry has expanded rapidly with greenhouse gas emissions (see fig.1). As the economy is rising, carbon dioxide emissions are increasing in proportion to increasing energy demand, industry and access. Electricity is now the number one contributor to greenhouse gases around the globe due to reliance on hydrocarbon (fossil) fuels such as coal, diesel, and gas [2, 3]. There is currently a global concern about environmental impacts due to countless provocations emanating from rapid and extremely precarious demand for energy resulting from population increase, decreased reserves of fossil fuel,

*Corresponding Author: Kakoma Chilala Bowa, Email: chilalakakoma@gmail.com

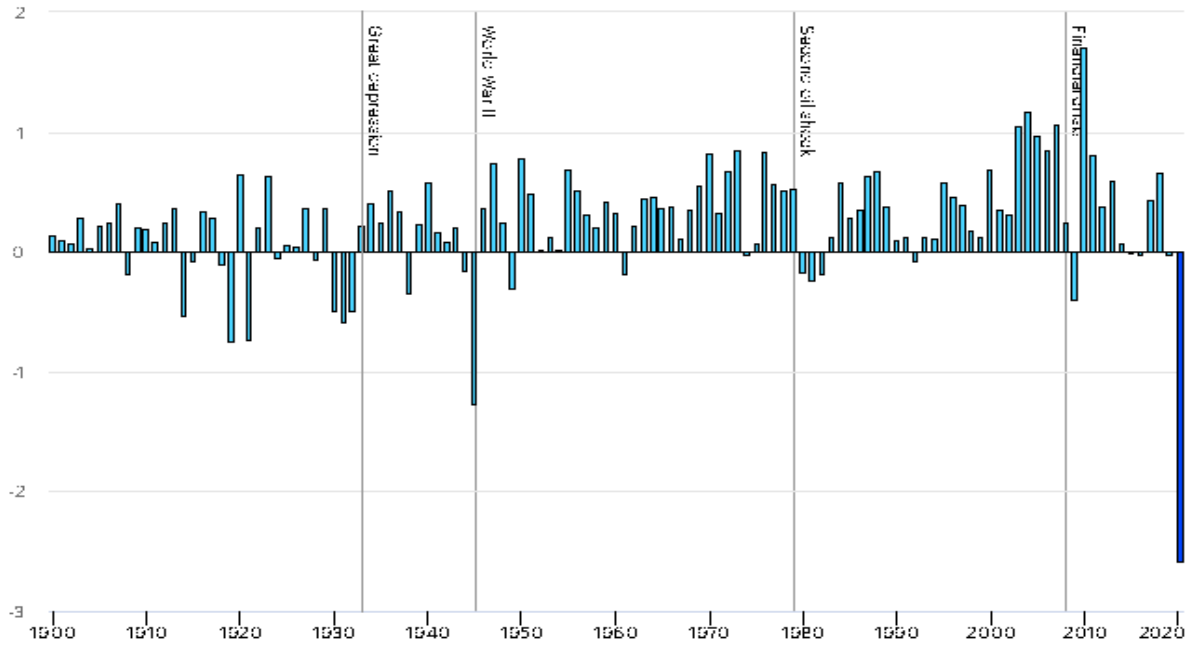


Figure 1: Globe Energy Related CO_2 Emission,1900-2020 [2, 7]

economic activities and high volatile fossil fuel [4, 5]. Following the COVID-19 pandemic, due to the new trend [6, 7], there will be a decrease in GHG emissions, foreign travel disruptions and economic activities. The electricity sector therefore faces difficulties, both politically and socially, in fostering sustainability.

The most pressing global challenge is the effort to guide the energy industry's future growth in a way that guarantees secure, sustainable, and sufficient access to electricity while being environmentally and socially appropriate as well. But the solutions presented must be commercially viable and must be combined with solutions from the real world. In Sub-Saharan Africa, this is needed in order to resolve numerous pressing energy challenges [8, 9]. Technologies for the generation of low-carbon electricity such as geothermal, solar photovoltaics (PV), wind turbines (WT), concentrated solar power (CSP), , biomass energy, and hydroelectric are rapidly recognized as crucial for a renewable electricity industry.

"The ultimate aim of SADC states (2016) is to "ensure that sufficient, effective, low-cost energy supplies are made available to help achieve economic efficiency and eliminate poverty while maintaining the efficient usage of energy resources [10]. The SADC member states have set up several subsidiary bodies to undertake energy related programs and initiatives in order to promote the harmonization of the energy sector. These numerous organizations include the SADC Center for Clean Energy and Energy Conservation, the Southern African Power Pool (SAPP) and the Southern African Regional Electricity Regulators Association (SARERA). These subsidiary organizations manage electricity planning, generation, delivery and marketing, as well as helping to harmonize regulatory policies [11].

The current RISDP 2015-2020 Priority B, provides guidelines on ensuring that regional and national strategies are compatible with the stability, sustainability, and efficiency of the region's energy market, with Member Countries working together to study and improve low-cost energy technology and energy sources for the region [10, 11]. With regard to the supply of electricity that is secure and affordable, Southern African countries Development Community considers energy access as a very important development priority.

This paper presents a comparative assessment of the SADC countries' electricity industry sustainability using the energy approach of Four (4) Aspects of Electricity (4EA's) which is based on sustainability criteria and indicators. The method was adopted on the basis of claims that it would stick to the four aspects belonging to 4EAs for any energy scheme to be viable. The four aspects of a comprehensive electricity industry assessment in the SADC region are grouped into sixteen (16) indicators. These indicators are linked to the performance of the electricity industry in terms of social and environmental including economical and technical performance. The 4EAs approach includes; Affordability, Accessibility, Availability and Acceptability of electricity sector in the SADC member countries [12]. The methodology of the 4EA discusses the three main sustainability pillars; societal, economic and environmental influences [13]. In particular, these pillars are major principles to achieve electricity sustainability in a way that is technically and economically viable plus a more socially and environmentally acceptable and appropriate for sustainable electricity sectors.

An appropriate quantitative and qualitative indicators and criteria were drawn up for the SADC electricity industry in the evaluation of the SADC Member State electricity sector. The findings of the electricity sector sustainability assessment in the SADC countries will therefore play an essential role to provide a basis to guide and assist policy-makers. The findings will help policy makers have a better understanding of the current and future challenges that arise in the SADC regional electricity industry and by providing alternatives to address them [14]. This will also assist

in the development of an informed investment strategy and decisions to achieve resilience in the region's electricity market.

2. Electricity Industries Sustainability Assessment in the SADC Region

The electricity sector in the SADC region, similarly to other parts of sub-Saharan African, faces numerous energy security related challenges due to rapid energy demand growth, reduced accessibility to energy because of grid lower coverage, high reliance on fossil fuels due to greenhouse gas emissions, leading to environmental and social impacts. Therefore, the SADC region is facing two (2) main energy policy challenges; lower access levels to electricity which are estimated at 48%, with 32% rural and 75% urban (see Table 5), and energy security issues because of dependence on fossil fuels or only single (one) source of energy in the national energy mix [11].

IEA defines energy security as "uninterrupted availability of energy sources at an affordable price" (2019). It can be categorized into two (2) major aspects, short and long-term. The long term aims at making timely investments in the energy system without compromising the environment in meeting economic development. However, the short-term primarily deals with the current energy supply system's ability to meet energy demand sudden changes [15]. Sustainability of energy is one of the key issues in the energy sector because of its strong links to social, economic, and environmental aspects [16]. The effect of access to affordable, reliable, and clean electricity is measured to improve the standard of human well-being. Economic aspects; represents the impact that the energy availability may have on economic growth and development prospects of any society or nation. The research conducted in 2013, highlighted that access to affordable, reliable and clean modern energy services, (e.g. electricity), is essential for the alleviation of poverty and economic development in developing countries [17]. The impact on energy systems that contribute to the environment as a whole is reflected in Environmental Sustainability. Thus, the current main drivers of environmental challenges such as greenhouse gas emission including climate change across the globe are energy systems based on fossil fuels. Therefore, the use of non-renewable energy sources in the energy mix poses a higher risk to natural environmental sustainability as compared to the use of renewable energy sources.

The 4EA sustainability indicators and criteria adopted for assessing the sustainability of the performance of the electricity industry in SADC member states are illustrated in table 1 below.

2.1. SADC Socio-Economic Development

The SADC consists of sixteen (16) member countries, including, as new Member States, Botswana Angola,, Lesotho, the DRC Congo (DRC), Malawi, Madagascar., Eswatini, Mauritius, Namibia, Mozambique, South Africa, Seychelles, Zimbabwe, Zambia and Tanzania. Approximately 33 percent of Sub-Saharan African population belongs to SADC region. The area has a projected 1.9 percent annual population growth. The area has seen an increase in urbanization in recent decades, for example in the period between 2013 and 2017, with the region seeing an increase in urban population share raising to 46% in 2017 from 35% in 2013. However, there exist a differences in the size and complexity of the gross domestic product (GDP) between SADC member states, with the lowest of US$1.498 billion (Seychelles) and highest of US$349.3 billion (South Africa) (check Table 2). This is the same with the GDP per capita, with ranges of US$800 (Malawi) to US$29300 per capita (Seychelles). However, on the overall, there is decline experienced by the SADC region in GDP per capita annual growth because of a drop in commodity prices on the international market (worldwide) including droughts that caused electricity deficit in industrial and agricultural production [16]. There are also large differences in member countries' economic and social growth profiles in the region. The SADC member countries are classified into three categories, lower, medium and high according to United Nations Human Development Index (HDI). The HDI in the region ranges between 0.418 (Mozambique) and 0.782 (Seychelles), with nine Member States below 0.50, four between 0.50 and 0.70 and only two above 0.70 (see Table 2 [11].

Table 1: Summary of 4EA's SADC Electricity Industry Sustainability Indicators and Criteria

Aspects of 4A's	Issues	Sustainability Criteria & Indicators
Acceptability	Global Warming	1) Policies on Renewable energy technologies (checklist) 2) Renewable energy share in electricity generation (%) 3) CO_2 emission per capital (t CO_2/capital) 4) CO_2 emission (ktCO_2/year)
Accessibility	Access to Modern Energy Services and Future Targets	5) Electrification rate (%) 6) Electricity consumption per capital(kWh) 7) Electricity Intensity (kWh/GDP) 8) Future Electrification Targets
Availability	Energy security	9) Diversity in Electricity generation mix (% share of each fuel type) 10) Diversity of fuel in Electricity generation (Shannon Wiener Index) 11) Dependence on imported electricity generation ▪ Self-Supply Sufficiency of Electricity (%) 12) Future Share of Renewable Energy Targets
Affordability	Energy Affordability	13) Electricity Prices ($/kWh) 14) GDP per Capita

Table 2: Socio-Economic Statistics for the SADC Member States [11,18,19]

Member States	Socio-economic indicators for the SADC Countries as of 2017					kWh per Capita	CO2 Emission (kilotons/year)	National Electrification Rates (%)
	Surface Area (km^2)	Population	GDP (US$ Billion)	GDP per capita ($)	HDI			
Angola	1,246,700	30,355,880	126,5	6,800	0.533	336.01	34,763	41
Botswana	581,730	2,249,104	17,38	17,000	0.698	1123.56	7,033	61
DRC	2,344,858	85,281,024	41,44	0,800	0.435	106.07	4,672	17
Eswatini	17,364	1,087,200	4,417	10,100	0.541	350.44	1,203	66
Lesotho	30,355	1,962,461	2,749	3,300	0.497	259.88	2,468	34
Madagascar	587,041	25,683,610	11,50	1,600	0.512	66.42	3,077	23
Malawi	118,484	19,842,560	6,240	1,200	0.476	71.56	1,276	11
Mauritius	2,040	1,364,283	13,33	22,300	0.781	2124.19	4,228	100
Mozambique	799,380	27,233,789	12,59	1,300	0.418	674.90	8,427	24
Namibia	823,290	2,533,224	13,24	11,200	0.640	553.84	3,755	56
Seychelles	455	94,633	1,498	29,300	0.782	3698.50	495	100
South Africa	1,219,090	55,380,210	349,3	13,600	0.666	4234.36	489,772	86
Tanzania	947,300	55,451,343	51,76	3,200	0.531	120.81	11,562	33
Zambia	752,614	16,445,079	25,71	4,000	0.579	702.34	4,503	31
Zimbabwe	390,757	14,030,368	17,64	2,300	0.516	484.66	12,020	38
SADC	**9,509,776**		**695,294**	**8,533**	**0.574**	**993,84**		**48**

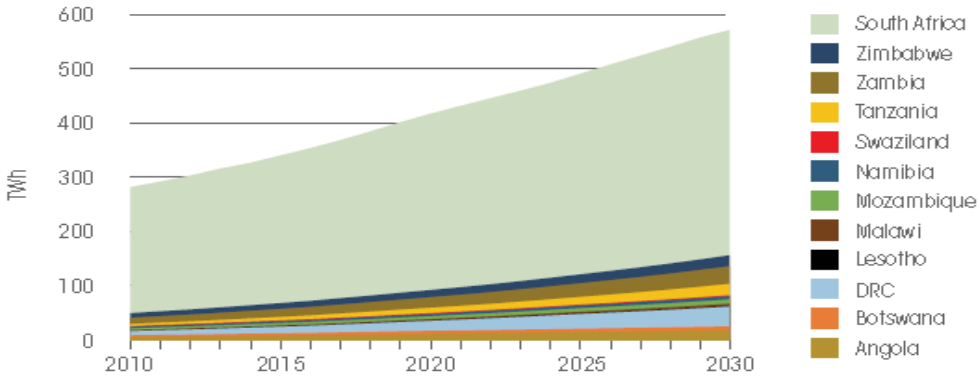


Figure 2: Projected Energy Demand in Southern Africa by 2030 [IRENA, 2015b]

2.2. SADC Share of Electricity Industry

As of 2018, the overall installed capacity of the SADC region was estimated at more than 61,859 GW, with 59,539 GW of installed capacity and 54,397 GW of operational capacity for SAPP Member States alone. Regional electricity demand growth is expected at an average rate of between 3 percent and 6 percent annually, in line with economic growth [20]. Complete electricity demand forecasts from the International Renewable Energy Agency (IRENA) (2015) indicate that demand in the SADC region is projected to double from 280 TWh (2010) to 570 TWh (2030) by 2030 (see Fig.2) owing to numerous activities such as strong economic growth, demographic growth, urbanization and industry production in the region [21, 22]. The demand for energy will continue to outpace production, with real costs, unless SADC countries raise cooperation with each other. Electricity will remain high and SADC's aspirations for wide-scale industrialization will continue to be suppressed by an inadequate supply of available electricity at fair prices.

In the region, however, generation capacity and installed power grids are limited; thus, electricity supply has lagged behind demand growth from 2007 until 2018, when the SADC region experienced excess energy supply as the result of the commissioning of the new energy projects [23]. In most SADC member states, this delay in the supply demand balance has resulted in a complex and persistent electricity shortage. Moreover, even in some highly hydroelectricity dependent SADC Member States due to climate change that has triggered droughts in recent years, the presence of a grid link has not guaranteed access to electricity for the end user, or even would be used by consumers due to prolonged load shedding, brownout and power cuts [24]. The SADC region's level of access to modern energy resources remained lower, with the total level of access to electricity as of 2017 being approximately 48% with 32% for rural and 75% urban areas (checkTable.5). The zone, however, aims to boost levels from the current level to 85.5 percent by 2030 [11].

2.3. SADC CO2 Emission

While Africa contributes little to global CO2 emissions overall, with many countries contributing almost zero, South Africa, by comparison, is the fourteenth largest CO2 emitter in the world. Tackling climate issues and related problems in the SADC region is obviously closely linked to tackling South Africa's and SADC's dependency on coal. At the 2016 Conference of the Parties (COP) to the United Nations Framework on Climate Change in Paris, countries committed to reducing emissions by individual pledges in a collective effort to minimize global emission levels. Such obligations are known as nationally specified contributions (NDCs). Although the SADC does not have a detailed shared strategy on how to achieve NDC targets at regional level,

organized discussions on how to define and formulate national targets were conducted in the lead-up to COP 21. And the SADC Regional Infrastructure Growth Master Plan 2012 sets out the target to increase renewable energy share in the grid to 39% by 2030 from the current 29% and reaching an off-grid share of 7.5% by 2030 [11, 20].Almost 85.7% of the electricity produced in the SADC region comes from coal-fired thermal power plants, posing a high health risk and rising the region's greenhouse gas emissions. Despite contributing fewer greenhouse gas emissions in the region, which was around 2.2 percent of global emissions in 2011, some SADC Member States, such as South Africa, have remained higher greenhouse gas emitters with 13th place in the world ranking. South Africa alone accounted for 48% of the carbon dioxide released in the SADC region in 2016 (see table 2) [11].

2.4. SADC Electricity Generation and Consumption

The only approach to solving existing electricity problems and promoting natural environmental resilience in the area is the transition to renewable energy in the electricity market of regional. In recent research in the field, the viability of renewable energy projects has been increasingly highlighted as a viable alternative. Renewable energy technologies such as biomass, wind, hydro, geothermal, and solar power is therefore rapidly gaining popularity in the SADC region, with total installed capacity reaching 18,069MW as of 2017 (see Figure 3 and Table 3), and capacity increased further to 21,760MW as of mid-2018 [11, 25]. The SADC region is projected to have an annual combined wind and solar capacity of 800 TWh, 219.5 TWh (solar PV) and 109.3 TWh (concentrated solar power) [21, 26]. However, the estimated electricity generation capacity from renewable energy sources for the period between 2010 and 2030 [26] in the region for consolidated energy systems is 62,781MW and 24,725MW for decentralized off-grid projects. A recent SACREE report published in 2018 showed that for the SADC area [11], wind and solar power could be an environmentally and economically competitive choice in the near future. Table 4 highlights advances in the generation of renewable energy (RE) for the period 2006-2017 in the SADC Member States.

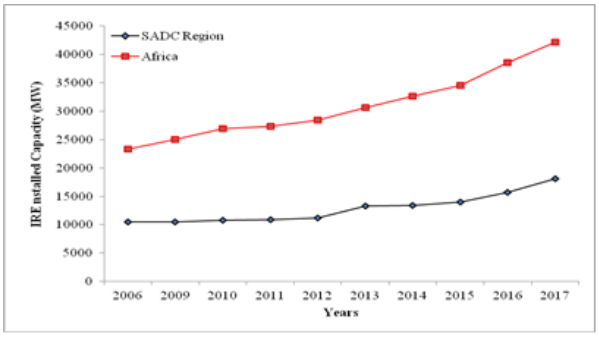


Figure 3: Renewable Energy Trend Comparison in Africa and SADC Region

Table 3: RE Development Trends for Eelectricity Generation in SADC Member States [IRENA 2018]

SADC Member States	RE Development Trends in SADC Member States (MW)										
	2006	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Angola	776	777	779	779	867	948	1016	1017	1418	2800	2428,0
Botswana	-	-	0	0	1	2	2	6	6	6	1,3
DRC	2514	2514	2514	2514	2514	2515	2515	2537	2687	2687	2601,0
Eswatini	124	124	124	186	186	186	186	186	186	186	167,6
Lesotho	73	73	73	73	75	75	75	75	75	75	77,3
Madagascar	122	122	131	131	165	167	170	173	175	177	177,3
Malawi	299	299	299	301	301	303	370	374	374	374	378,4
Mauritius	130	134	135	135	152	155	171	171	189	191	247,3
Mozambique	2198	2198	2198	2198	2201	2205	2207	2211	2214	2233	2198,9
Namibia	252	253	254	257	343	346	348	359	379	391	404,5
Seychelles	-	-	-	-	-	6	7	7	8	8	8,7
South Africa	838	848	853	897	903	1358	2524	3060	4069	4959	8679
Tanzania	607	609	620	636	642	645	649	656	657	659	612,9
Zambia	1715	1723	1937	1937	1944	2304	2304	2314	2434	2435	2597,8
Zimbabwe	835	835	837	856	856	862	863	879	884	888	1180,0
SADC Region	**10483**	**10509**	**10754**	**10900**	**11150**	**13339**	**13407**	**14025**	**15755**	**18069**	**21759,7**
Africa	**23381**	**24986**	**26940**	**27319**	**28485**	**30639**	**32666**	**34511**	**38603**	**42139**	

Table 4: Energy Imports, Export, Generation, and Consumption, for the SADC countries,2016/2017 [11,18,25]

Member States	Energy indicators for the SADC Countries (TWh/Year)				Primary Energy Generation Mix (%)				Diversification	ESE(%)	SSR(%)
	Generation	Consumption	Exports	Imports	Other RES	Hydro	Nuclear	Fossil			
Angola	10,20	9.036	0.00	0.00	2	64	0	34	0,67	88,59	100,0
Botswana	2,527	3,636	0,00	1,673	0	0	0	100	0,00	86,57	60,17
DRC	9,046	7,430	0,422	0,020	0	98	0	2	0,14	86,61	99,77
Eswatini	0,381	1,431	0,00	1,077	41	20	0	39	0,96	98,15	26,13
Lesotho	0,510	0,847	0,00	0,373	1	99	0	0	0,08	95,92	57,76
Madagascar	1,706	1,587	0,00	0,00	2	24	0	74	0,59	93,02	100,0
Malawi	1,420	1,321	0,00	0,00	6	93	0	1	0,28	93,03	100,0
Mauritius	2,898	2,726	0,00	0,00	14	7	0	79	0,59	94,06	100,0
Mozambique	18,39	11,57	12,88	9,928	1	83	0	16	0,45	86,34	35,69
Namibia	1,403	3,891	0,088	3,073	8	64	0	28	0,77	88,90	29,97
Seychelles	0,350	0,326	0,00	0,00	9	0	0	91	0,44	93,14	100,0
South Africa	234,5	207,1	16,55	10,56	10	1	4	85	0,39	91,26	95,38
Tanzania	6,699	5,682	0,00	0,102	6	40	0	55	0,79	83,55	98,50
Zambia	11,55	11,04	1,176	2,185	2	93	0	5	0,27	88,94	82,60
Zimbabwe	6,800	7,118	1,239	2,220	5	37	0	58	0,76	92,65	71,47

The primary energy generation combination of imports, exports and consumption as of 2016 by the SADC member states is shown in Table 4. It can be seen that South Africa had the highest generation, consumption, imports and exports in a region followed by Mozambique and, least of all, Seychelles, which has zero imports and exports.

3. Energy Sector Sustainability Evaluation of SADC Countries based on 4EAs

3.1. Electricity Acceptability

In terms of social and environmental acceptability, electricity generation technologies cause various and numerous environmental and social impacts such as visual impacts, land use and degradation, fugitive dust, noise, soil and air pollution including greenhouse gas emission which is the most serious threats to both the environment and society arising from present electricity industry. These impacts pose different policy challenges and there is no one straightforward solution. Hence, this dimension reflects environmental and social acceptability of the electricity sector by society and environmental organizations based on national energy mix. According to WEC(2007), the best alternative to attain environmental sustainability and increase social acceptability is by reducing greenhouse gas emission through adopting of low-carbon renewable energy technologies in the national energy mix [16]. Thus, an electricity generation mix with combination of various sources from low-carbon technologies is expected to increase social and environmental acceptability. Therefore, to assess the acceptability of SADC countries' electricity industry the following criteria have been considered; RE share in electricity production (%), emission of CO2 per capital (t CO2/capital), country's CO2 emission and Policies on Renewable energy technologies.

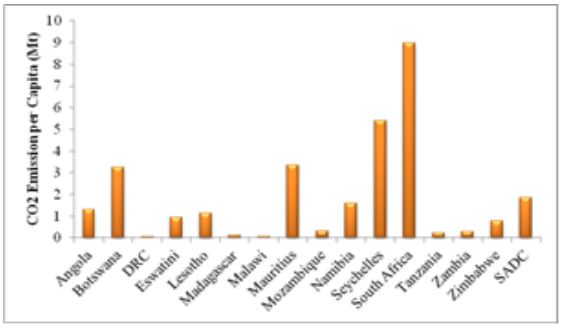


Figure 4: CO2 Emission per capita in SADC Countries [27]

1) CO2 Emission Per Capita

Figure 4 shows the emissions of CO2 per capita in the SADC countries as of 2015. The results indicate that, except for South Africa, the region's member countries have lower emissions of CO2 per capita, mainly because of higher dependence on fossil fuels (coal) in the electricity sector. Therefore, South Africa (SA) has the highest per capita emissions (8,98 tonnes of CO2) in the country, followed by Seychelles (5,42 tonnes of CO2) with the lowest value of 0,06 tonnes of CO2 in DRC [27]. This makes SA, as previously mentioned, the largest contributor and emitter of CO2 emissions in the SADC region. South Africa draws 85% of its energy from fossil fuels at the same time contributes more than 76% to the final electricity generation in the SADC region, hence contributing to nearly half of the region's CO2 emissions [27].

2) Policies on Renewable energy technologies

As seen in Table 8, it appears that all SADC member states have progressed towards the establishment of RE technology national and regional policies. These policies, however, differ according to the priorities and objectives of the country set in the national energy master plans.

3) Share of RE in Final Electricity Generation Mix

Figure 5 shows the renewable energy (RE) share in the electricity generation in member states of SADC. It can be seen that due to the high use of hydro for electricity generation, DRC shows a highest share of RE of 92.87 percent in the final national general electricity generation mix, with the lowest share of 1.03 percent shown by Seychelles because of its reliance on fossil fuels. In general, the shares of RE in the final energy generation mix have increased in the region and member countries (see table 4).

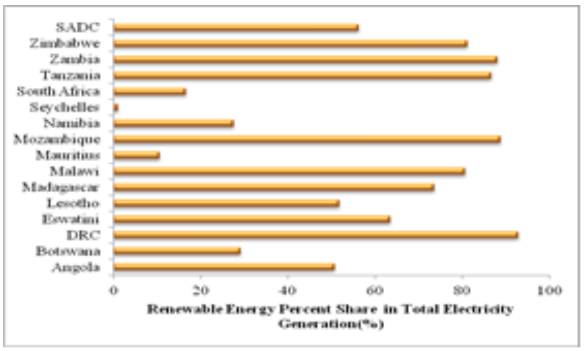


Figure 5: RE Share in SADC Countries 'Electricity Generation [11]

4) SADC Countries CO2 Emission Per Year

Most SADC member countries have low CO2 emissions because the overall generation of electricity is highly dependent on renewable energy. Nonetheless, few Member States such as South Africa, Botswana, Seychelles, Mauritius, and Namibia are heavily dependent on fossil fuels with less than 50% share of RE in the final electricity generation mix. Therefore, in terms of the total contribution of CO2 emissions to the region, South Africa is at the top of the group, followed by Angola, Zimbabwe, Tanzania and Seychelles at the very least (see table 2 above).

3.2. Electricity Accessibility

This dimension underlines society's accessibility to modern energy facilities, both now and in the future, in a sustainable way for everyone. Access to electricity can have multiple social impacts, including enhancing human well-being by meeting the basic needs of human being, such as entertainment, gender equality, food, education, clean water and health [14]. Moreover, accessibility to energy services will also help to reduce the effects on the environment, such as deforestation. Therefore, the idea of efficiency, quality and easy access to energy services to society poses a policy challenge. Societies must clearly have easy accessibility to critical energy resources, but services must be efficient and of high quality. A comparative study is also

conducted in this dimension to determine the accessibility of electricity in the SADC countries, as well as the possible access goals. In this dimension, three parameters are taken into account: electricity intensity (kWh/GDP) and electricity consumption per capital (kWh/capital), existing levels of electrification and potential access targets for connectivity and electricity Disruption Risk Index [28] to assess the efficiency and quality of services.

1) Consumption of Electricity per Capita

Figure 6 shows the energy consumption per capita in member states as of 2017. In general, electricity consumption per capita in most SADC countries is low below the average value of the region (993.84 kWh/ capita), with the exception of Seychelles, Mauritius and South Africa, with more than 2000 kWh per capita. Electricity consumption per capita varies significantly, with the smallest being 66.42kWh per capita for Madagascar and the highest of 4234.36kWh per capita for South Africa.

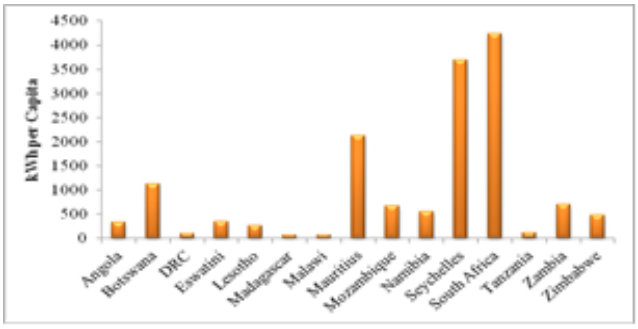


Figure 6: Consumptions of Electricity per Capita

2) Electricity Intensity

Figure 7 shows the energy intensity (kWh per US$) in member states. With Mauritius, Botswana, Seychelles and Namibia having the lowest because of higher GDP due to lower population with Zimbabwe, Madagascar and DRC having the highest energy intensity among the member countries. This means that growth in electricity demand is far below economic growth in most member countries. Hence, this shows that there is less modern energy services accessibility including lower penetration and utilization of electrical appliances and equipment in the region and in the member States.

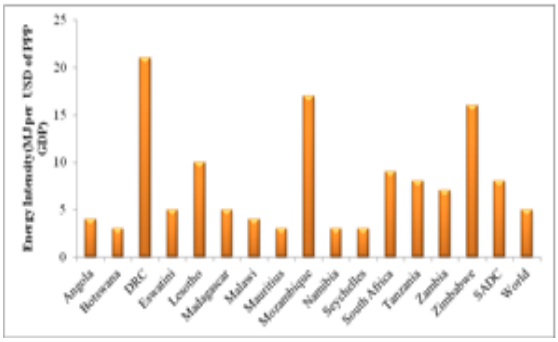


Figure 7: Energy Intensity in SADC Member States in 2017 [11]

3) Electrification Levels and Future Access Targets

Table 5 shows the electricity accessibility in SADC member states and the potential goals for access to electricity as set in the Member States' national energy master plans to be achieved by 2030. As of now, with the exception of a few countries like: Botswana, South Africa, Eswatini, and Namibia that are above 50%, the country and several Member States have lower access levels to electricity. The only two member countries that are now able to have 100% access to electricity are Mauritius and Seychelles. For most member countries, however, urban electrification levels are above 60 percent, with only Malawi having the lowest 42 percent. Nevertheless, with the exception of Mauritius and Seychelles, which have already crossed 100 percent, rural electrification levels are still lower in most member countries. Rural electrification should therefore be listed as a priority area for the SADC region and its member countries in order to achieve this aim.

Table 5: Electrification Rates in SADC Member, 2016 [11]

Member States	Electrification Rates (%)				Energy Access Targets (%)
	Rural	Urban	National	People without Access to Electricity (million)	
Angola	16	69	41	17	100% by 2030
Botswana	37	78	61	1	100% by 2030
DRC	.	78	17	68	60% by 2025
Eswatini	61	83	66	<1	100% by 2030; 75% by 2018; 85% by 2020
Lesotho	16	66	34	1	40% by 2020
Madagascar	17	67	23	19	-
Malawi	4	42	11	16	30% by 2030
Mauritius	100	100	100	-	-
Mozambique	5	65	24	21	100% by 2030; 30% Modern Cooking Fuel by 2030
Namibia	29	77	56	1	50% by 2020; 100% by 2030
Seychelles	100	100	100	-	-
South Africa	68	93	86	8	100% by 2025
Tanzania	17	65	33	36	75% by 2030
Zambia	4	67	31	11	66% by 2030; 90% by 2030 urban, 51% by 3030 rural
Zimbabwe	16	86	38	11	100% by 2030
SADC	**32**	**75**	**48**	**49**	**-**

3.3. Electricity Availability

Electricity reliability and supply continuity are very vital for the social and economic development of every community, society or country. In order to achieve the social and economic well-being of every community and country, electricity plays an important role. However, there are different issues that can cause short-term and long-term unreliability and electricity discontinuity. These vary from nation to nation, but it's very similar and popular in the SADC member states. The causes maybe due to insufficient electricity generating units, insufficient transmission lines to transport electricity generated to demand centers, inadequate demand-compliant generating units, poor maintenance of transmission networks or generating power station, weak electricity systems to withstand disruptions and contingencies. For example, severe weather conditions such as low precipitation (droughts) and potential availability of primary energy sources such as fossil fuels may be other factors [16]. However, it is argued, according to the [16] literature, that a well-diversified energy generation industry comprising of different types of local primary energy sources, including reliable generation technologies, is capable of helping to improve energy security and to reduce the risks that may arise from high and volatile fluctuations in fuel prices.

Therefore, the use of available local resources in the generation of electricity is important. Furthermore, timely and adequate investment in the electricity sector is essential to ensure the reliability and continuity of the supply of electricity [17]. This dimension examines the energy security of the SADC countries with regard to the short and long-term reliability and continuity of the supply of electricity by meeting four of the following availability indicators and criteria: the reserve margin of the supply of electricity, the efficiency of the supply of electricity, the diversification of the supply of electricity, the sufficiency of the self-supply of electricity (SSR).

1) *Efficiency of Electricity Supply (ESE) in SADC Countries*

The ESE is described as " the ratio of electricity not lost(ENL) to the total supply of electricity (TES)"[28]. This indicator shows the overall output of the country. The estimate was made using Equation 1 (1a to 1d). As given below, the general equation is expressed as.

$$ESE = \left(\frac{ENL}{TES}\right) \tag{1a}$$

In SADC countries that are involved in exporting electricity to other SADC countries within the region.

$$ESE = \left(\frac{EC + EXE}{EP}\right) \tag{1b}$$

On the other hand, for such SADC member countries that export and imports electricity from other SADC countries, the ESE was calculated using equation 1c given below

$$ESE = \left(\frac{EC}{EP + IE}\right) \tag{1c}$$

However, for those countries in SADC region that are not involved in either exports or import of electricity such Angola, Madagascar, Malawi, Seychelles and Mauritius. The ESE was calculated using the equation 1d.

$$ESE = \left(\frac{EC}{EP}\right) \tag{1d}$$

where, “EC” denotes consumption of electricity, IE-import of electricity, EP-output of electricity, EXE-export of electricity. Figure 8 below shows the output of SADC countries in terms of electricity supply reliability. The results show that among the member countries, the lowest ESE is shown by Tanzania whereas highest ESE is shown by Eswatini. In other words, with the exception of Eswatini, the majority of SADC countries have an ESE of less than 95%. This shows that the general electricity industry faces problems with electricity losses in most of the member countries. This may be due to theft of electricity or losses in the grid due to long grid distances or inadequate maintenance.

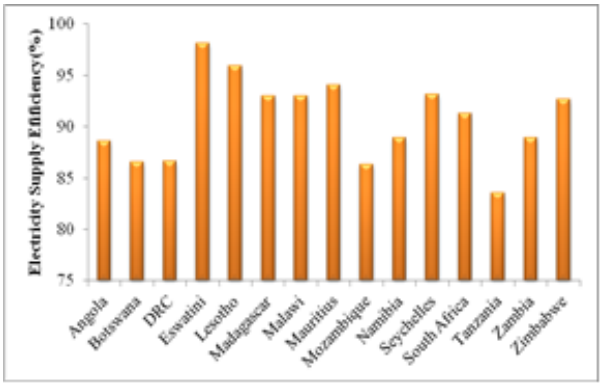


Figure 8: ESE in SADC Countries

2) *Electricity Supply Diversification of SADC Countries*

Using the Shannon Wiener Index given in equation 2 [5,29], the diversification of electricity supply is calculated.

$$SWI = -\frac{\sum_{l=1}^{n} s_l * Log(s_l)}{Log(n)} \tag{2}$$

In the equation “si” represents the portion of each energy source in the electricity generation mix, “I and n” is the number of main energy sources in the electricity generation mix. When the equal share (1/n) of all energy sources contributes to the final electricity generation mix, the maximum “Ln(n)” occurs. In terms of electricity supply diversification, the higher the value means higher energy resilience and security. Hence, lower risk for the electricity industry (see Fig.9). The findings show that diversification in the SADC countries’ electricity industry varies widely, with highest in Eswatini and lowest in Botswana. However, Table 7 below shows that all SADC member states have put in place plans for the diversification of the local mix of electricity production by 2030. Hence, it is expected that the future of SADC electricity industry of the Member States will be more diversified. However, the energy security and accessibility will be strengthened only if all Member States follow and meet their planned targets. Nonetheless, fossil fuels (natural gas and coal) which are readily available locally are expected to continue playing a major role in the region's energy generation mix.

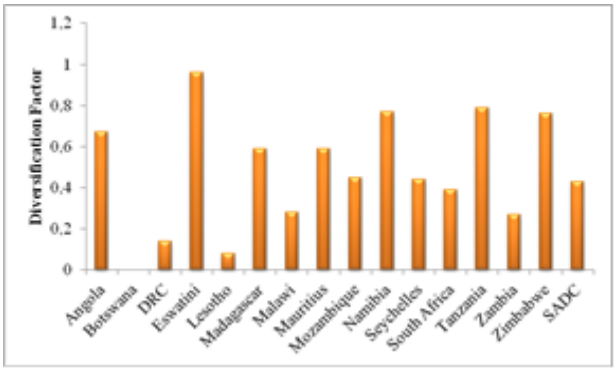


Figure 9: Energy Diversification in SADC Countries, 2016

3) Self-Supply Sufficiency Rate of SADC Countries

For each member state, the Self-Supply Sufficiency Rate (SSR) in the national electricity sector was assessed using equations 3a-d [28] to determine the energy diversification impact on the SADC member countries' electricity sector. This deminsulation tests the capability of the SADC countries' electricity sector to meet the current demand with local generation. Consequently, greater the value or equal to 100% represents that the nation is self-sufficient, while less than 100% means that the country is dependent on the energy it imports. General equation 3a.

$$SSR = \left(1 - \frac{NIE}{TDES}\right) \quad (3a)$$

For nations importing part of their electricity from other SADC members within the region, SSR was estimated using equation 3b.

$$SSR = \left(1 - \frac{NIE}{EP + IE}\right) \quad (3b)$$

In case of nations that are exporting their electricity surplus to other SADC member states, SSR was estimated using equation 3c..

$$SSR = \left(1 - \frac{NIE}{EP - EXE}\right) \quad (3c)$$

The SSR given in equation 3d represents countries that are not interconnected nor export or imports electricity from other SADC countries (Madagascar, Seychelles, and Mauritius).

$$SSR = \left(1 - \frac{\text{NIE}}{\text{EP}}\right) \quad (3d)$$

Where "NIE" denotes the net electricity imports whereas "TDES" is the total electricity generation. In terms of sustainability, the country with higher SSR value have higher energy security hence lower the risk due to lower dependency on the imports of electricity (see fig.10). The results shows that the majority of the SADC countries are self-sufficient in the electricity supply with only handful of countries facing challenges.

4) RE Share and Reserve Margin in Electricity Supply in SADC Countries

As of mid-2018, Table 6 highlights the share of RE in SADC member countries. In spite of the number of Member States producing electricity from RESs, the majority of electricity production in the region is highly dependent on fossil fuels. However, the region and Member States have increased their RE share in the overall mix of final electricity generation as of 2018, with South Africa leading the ladder (see table 4). As regards the reserve margin, it can be seen that, as of 2016, seven of the member countries had an energy deficit. Except for South Africa, which had more than 10.0GW, the other member countries had negligible reserve margins. These negligible reserve margins are mainly due to the region's rapid electricity demand, but also due to underinvestment in the electricity industry in many SADC nations.

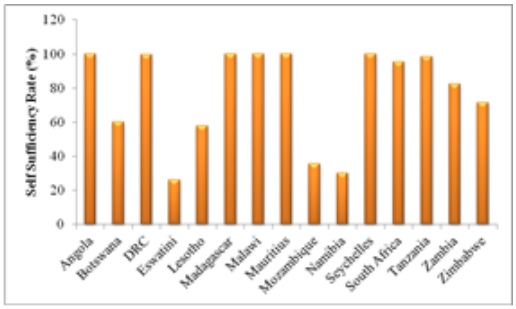


Figure 10: SSR for SADC Countries

Table 6: Share of Energy Sources in Total Final Energy Generation (TFEG)in SADC Countries, 2016[18,20]

Member States	Share of Energy Sources in TFEG (%)			
	Other RES	Hydro	Nuclear	Fossil
Angola	2	64	0	34
Botswana	0	0	0	100
DRC	0	98	0	2
Eswatini	41	20	0	39
Lesotho	1	99	0	0
Madagascar	2	24	0	74
Malawi	6	93	0	1
Mauritius	14	7	0	79
Mozambique	1	83	0	16
Namibia	8	64	0	28
Seychelles	9	0	0	91
South Africa	10	1	4	85
Tanzania	6	40	0	55
Zambia	2	93	0	5
Zimbabwe	5	37	0	58
SADC	**7,0**	**48,2**	**0,3**	**44,5**

5) RE Share Targets and Interconnection in SADC Countries by 2030

In order to promote in the regional and national electricity supply in the sustainable manner and promote the use of and increase access to clean modern energy services for the greater benefit of all people in the region and member states. The region and SADC states have set up targets plans to be achieved by 2030 for the renewable energy share in the electricity generation as illustrated in Table 7. In addition, to increase regional and national energy security and reliability in electricity sector almost all mainland member countries are interconnected and are involved in international electricity trade except for Malawi and Tanzania which are expected to be interconnected with other members by 2030. However, Madagascar, Mauritius and Seychelles due to their geographical location will always remain not connected to the region electricity grid. Furthermore, in terms of grid losses as shown in table 7, it indicates that Lesotho followed by Angola have higher losses in the SADC region with South Africa having the lowest among the member states. Hence, it can be concluded that the South African power grid has higher efficiency in the region.

Table 7: Regional and National Targets in SADC Member States by 2030 [11]

Member States	Renewable Energy Targets	Transmission and Distribution Losses (%)	Interconnections with other countries
Angola	Increase in RE capacity by 2025 as follows: ▪ Small Hydro:100MW with 60MW for municipalities ▪ Solar:100MW with 10MW off-grid ▪ Wind:100MW ▪ Biomass:500MW	10	✓
Botswana	100% access to modern energy services by 2030 Capacity increase expected from REFIT programme 15% RE share in final energy consumption by 2025 but may increase to 20%	3,7	✓
DRC	60% overall energy access by 2025	9	✓
Eswatini	50% RE share in Electricity Consumption by 2030 60MW intermittent resource by 2030 e.g solar	6	✓
Lesotho	Targets pnding completion of Sustainable Energy Strategy 2018	11	✓
Madagascar	85% RE Share in final energy by 2030	-	❖
Malawi	By 2025/2030 307% access to electricity 100% use of efficient cook stoves in off grid households 6% RE share in energy mix Biofuels mandate of 20% ethanol and 30% biodiesel	6	▪
Mauritius	35% RE Share in electricity generation by 2025: Bagasse:17%, Wind:8%, Waste to enrgy:4%, Hydro:2%, Solar:2%, Geothermal: 2%	6	❖
Mozambique	400MW increase in installed RE capacity by 2024 Wind:150MW Hydro:100MW large\scale & 100MW small scale	6,4	✓
Namibia	70% RE share in electricity generation by 2030	3,2	
Seychelles	5% RE share in electricity generation by 2020 & 20% by 2030	-	❖
South Africa	21% RE Share in electricity generation by 2030 17.6GW solar capacity by 2050 37.4GW wind capacity by 2050	0,1	✓
Tanzania	5% RE share in electricity generation by 2030	6,0	▪
Zambia	200MW increase in RE capacity by 2020	6,2	✓
Zimbabwe	16-5% RE Share increase in overall 1100MW increase in RE capacity by 2025 2100MW increase by 2030 2400GWh increase in RE generation by 2025 4600GWh increase in RE generation by 2030 (26.5%overall increase)	4	✓
SADC	71% by 2025, 85,5% by 2030 Regional Electricity Access targets 39% Regional Renewable Energy Mix target in the grid 7,5% Off-Grid share of RE as per total grid electricity capacity 15% Energy efficiency % savings achieved from grid consumption	5.97	✓

3.4. Electricity Affordability

According to literature [13], for electricity to be considered as being affordable the following conditions have to be fulfilled:

- The electricity tariffs should be low as compared to household income, that is, the household should be able to afford their electricity bills.
- The electricity tariffs should be low as compared to other alternative prices of competitor, that is, the price of electricity should be lower than other option such as wood fuel, traditional fuels, fossil fuels etc.
- The electricity tariffs should be low enough to ensure the energy import bill is small compared to export earnings [13],
- The electricity tariff should be high enough to ensure sufficient benefit and profitability for independent Power Producers and Utility Companies i.e., the tariff should be able to reflect the real value of the electricity supply. In short, the electricity price should take care of all the value chain cost of electricity supply (Production, Transmission/Distribution and other cost of the industry) [16].

This dimension investigates the electricity affordability of SADC countries using the following affordability indicators and criteria; Electricity Prices, GDP per Capita and Cost of Electricity Supply (i.e. the amount spent on electricity as compared to national GDP).

1) Electricity Tariffs

Figure 11 highlights the average tariffs for electricity as of 2018 in the SADC countries. The lower the price of electricity in terms of affordability, means the higher the government subsidies practiced, hence, a huge burden on expenditure of government. This suggests that the supply of energy is heavily subsidized and does not reflect the actual electricity cost. The results shows that the majority of SADC member countries have electricity tariff rates of less than US$c10/kWh. However, Namibia among member states has the highest electricity tariff. These lower tariffs

suggest that majority of SADC countries' energy is heavily government-subsidized, hence does not attract the private investors, which is the biggest obstacles to the sustainable growth of the electricity industry in the region. Nonetheless, with expectation to meet the cost-reflective rates in SADC member countries, it is expected that the sector will draw a lot of private investment by 2030.

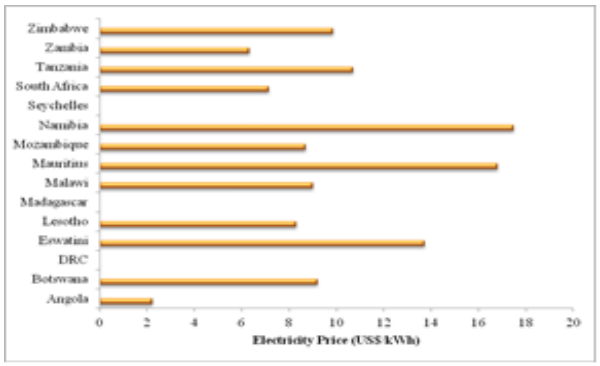


Figure 11: Electricity Tariffs in SADC Member, 2018 [24]

2) GDP per Capita

The per capita GDP reflects the citizen's purchasing power, thus nations with higher per capita GDP mean that people have more buying power and vice versa. The scenario represents electricity investment in terms household electricity affordability. Figure 12 shows that most SADC states face problems because the population has less purchasing power, so government subsidies are required to make electricity affordable.

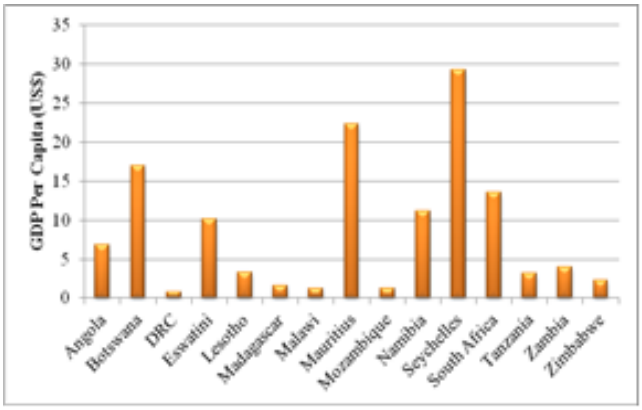


Figure 12: GDP per Capita in SADC Countries, 2016

3) Real Cost of Electricity Supply (RCES)

RCES is defined as "the ratio of the proportion of real GDP not dedicated to cover the net electricity supply expenditures (NESE) to real GDP (RGDP)or it is one minus the share of GDP dedicated to cover the cost of the electricity supply(ES)" [28]. The criteria measures the ability of the SADC countries to reduce the RCES and the vulnerability of member states to high RCES. The higher the RCES shows the country's ability to reduce its RCES whereas low values shows exposure of the country to high RCES. RCES is estimated using equation 4 given below.

$$RCES = \left(1 - \frac{ES}{RGDP}\right) = \left(\frac{NESE}{RGDP}\right) \quad (4)$$

Where ES is estimated by firstly converting the total electricity supply into barrel of oil equivalent (bbl) and multiplying it with the annual real average crude oil price (COP) given in US$ per bbl. Figure 13 below indicates the RCES for SADC member states as of 2019 at COP of US$59.06 per barrel (1kWh=0,006Barrels of Oil Equivalent) [30]. Apart from South Africa, it can be seen that the SADC member countries have RCES above 80%. As earlier mentioned this shows that the member countries have the ability to reduce their RCES. In short this indicates that the majority of SADC countries spend less on electricity from their GDP.

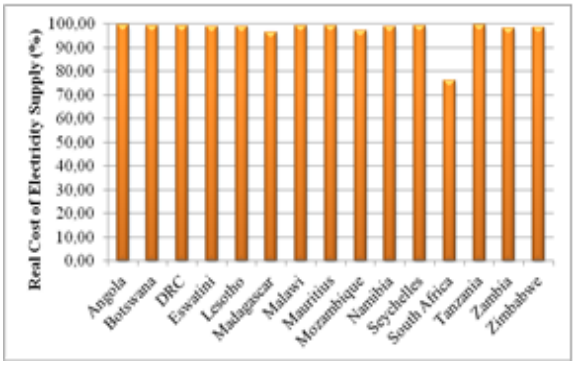


Figure 13: Real Cost of Electricity Supply of SADC Countries

4. RE Support Policies for improving the Sustainability of Energy in the SADC Countries

To ensure private sector participation in electricity Ensuring the participation of the private sector in the electricity industry and the successful development and dissemination of programs for renewable energy technologies (RETs) in the SADC region, as well as encouraging investment in RETs. A number of additional incentives and RE support policies have been/have been established by the SADC member countries, including (see table 8)[11,31,32]:

- Renewable Energy Targets
- Renewable Energy in NDC or INDC
- Regulatory Policies
  - Feed-in-Tariff/ Premium Payment
  - Electric Utility Quota Obligation
  - Net Metering
  - Grid Code Revisions
  - Tradable Renewable Energy Credit
  - Renewable Energy Project Tendering
  - Capital Subsidy' Grant or Rebate
- Fiscal Incentives & Public Financing
  - Investment /Production Tax Credits
  - Reductions in Sales, Energy' CO2, VAT or other taxes
  - Energy Production Payment
  - Public Investment Loans or Grants

As illustrated in Table 8, all SADC member countries have progressed to establishing policies on RE technologies. However, these policies vary according to the country targets and goals set in the national master plans.

Table 8: RE Support Policies in SADC Member States as of Mid-2018 [9,31]

Member States	RE Targets	RE in NDC or INDC	Regulatory Policies								Fiscal Incentives & Public Financing			
			FiT/ Premium Payment	Electric Utility Quota Obligation	Net Metering	Biofuels Obligation/ Mandates	Grid Code Revisions	Tradable RE Credit	Tendering	Capital Subsidy' Grant or Rebate	Investment /Production Tax Credits	Reductions in Sales, Energy' CO2,VAT or other taxes	Energy Production Payment	Public Investment Loans or Grants
Angola		▪	▪		▪	▪	▪		▪			▪	▪	
Botswana	▪		✓						▪					✓
DRC	▪													
Eswatini	✓	▪				▪	✓		✓					
Lesotho	▪	▪	✓		✓				✓					✓
Madagascar	▪	▪	✓				✓		✓					
Malawi	▪	▪	✓			▪	✓		✓			✓		✓
Mauritius	▪	▪	✓		✓		▪		▪	▪		✓		✓
Mozambique	▪	▪							✓					
Namibia	▪	▪	✓		▪				▪					
Seychelles	▪	▪			✓		✓		✓			✓		✓
South Africa	▪	▪		✓	▪	▪	▪		▪	▪	▪	▪		✓
Tanzania		▪	▪		✓	✓	✓		✓	▪		✓		✓
Zambia	▪	▪	▪			▪	▪		✓	▪		▪		✓
Zimbabwe	▪	▪	✓		✓	▪			▪	▪				✓

5. Key Findings and Conclusion

A total of 15 SADC Member States are analyzed in this paper, with a focus on the ability to provide reliable, effective and environmentally friendly energy systems. The analysis was carried out on the basis of performance measures of the electricity sector and factors such as availability, supply and demand for electricity, access to electricity and the environmental social effects of the use and development of electricity in the Member States. The research outcomes;

The economy of the SADC region is highly dependent on commodities. The world market saw a decline in oil prices over time and this culminated in a dramatic decrease in the SADC countries' economy. SADC Member States are currently facing a challenge with regard to electricity access levels, especially in rural areas, with only Mauritius and Seychelles achieving 100% access to electricity in both rural and urban areas. The region's population is around 341 million people, but just 48 percent of the total households have access to electricity. Studies by SACREE (2018) show that the majority of people live in rural areas and only 32% of the total population in rural areas has access to modern energy facilities, with electricity being available to only 75% of the total urban population.

– Low-carbon renewable energy technologies (RETs) are expected to contribute significantly to future electricity as all SADC member countries set goals are met by 2030, and the SADC region will be greener and cleaner relative to current trends.

SADC member countries have adopted subsidies as a affordability and accelerate access to electricity. Subsidies are proving to burden the national budget. In the case of Zambia, the country experienced enormous costs between the years 2014-2017 due to the electricity deficit resulting from electricity imports priced in neighboring countries at a cost-reflective tariff. This increased expenditure for the utility company (ZESCO) and the government.

– SADC Member States will benefit greatly from the diversification of RET and energy conservation policies as set out in the regional objectives. Adherence to the vast potential of RE sources in the area will negatively increase the degree of access to electricity in the region. In addition, the use of RE in the regional and national mix of electricity generation would lead to better health services and better living conditions for the population in both urban and rural areas, which will be significantly enhanced, resulting in safe and clean cities and communities.

It is therefore argued that policies addressing basic human needs, such as increased access to electricity and improved national or regional energy protection, are required in developing countries, and environmental and social acceptability should be taken into account. The cost of electricity is ideally representative of the cost and also represents the true cost of electricity services.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors gratefully acknowledge the support provided by University of Johannesburg, University of Zambia, University of South Africa and Ege University.

References

[1] K.C.Bowa, M.Mwanza, M.Sumbwanyambe, K.Ulgen, J.H.Pretorius, Comparative Energy Sustainability Assessment of Electricity Industries In SADC Region: The Role of Renewable Energy In Regional And National Energy Diversification, 2019 IEEE 2nd international conference on renewable and power engineering REPE, Toronto Canada,2019.

[2] IEA, Global Energy and CO2 Status Report 2017,International Energy Agency (IEA), 2018.

[3] WEC, World Energy Trılemma Index 2017; Monitoring the Sustainability of National Energy Systems, World Energy Council (WEC), Available at:www.worldenergy.org (Accessed 14/03/2019)

[4] IPCC, Special Report on Renewable Energy Sources and Climate Change Mitigation, Summary for Policymakers, Intergovernmental Panel on Climate

Change (IPCC), ISBN97-92-9169-131-9, 2012.
[5] P.Vithayasrichareon, Portfolio-Based Decision-Support Tool for Generation Investment and Planning in Uncertain and Low-Carbon Future Electricity ındustries, PhD Thesis, The University of New South Wales, Sydney, Australia, June 2012
[6] IEA, The impact of the Covid-19 crisis on clean energy progress, IEA,2020a, Paris, Available at: https://www.iea.org/articles/the-impact-of-the-covid-19-crisis-on-clean-energy-progress
[7] IEA, Annual change in global energy-related CO2 emissions, 1900-2020, IEA,2020b, Paris, Available at: https://www.iea.org/data-and-statistics/charts/annual-change-in-global-energy-related-co2-emissions-1900-2020
[8] A.Brew-Hammond, Energy Access in Africa:Challenges ahead, Energy Policy, **38**, 2291-2301, 2010.
[9] S.Karekezi, W.Kithyoma, Renewable Energy in Africa: Prospects and Limits, Renewable Energy Development, 2003.
[10] SADC, SADC Energy Monitor 2016; Baseline Study of the SADC Energy Sector. SADC, SARDC, Gaborone, Harare, 2016.
[11] SACREEE, SADC Renewable Energy and Energy Efficiency Status Report 2018, SADC Centre for Renewable Energy and Energy Efficiency(SACREEE), 2018
[12] WEC, Global Energy: Agenda, Challenges, Policies. presentation at 12th International Energy Forum, Cancun, Mexico, 29-31 March 2010.
[13] C.Aleh and J.Jessica, The concept of energy security.Beyond the four As, Energy Policy, **75**, 415-421, 2014.
[14] N.Wamukonya,. Power sector reform in developing countries: mismatched agendas, Energy Policy 31(12), 1273-1289, 2003.
[15] IEA, Energy Security, International Energy Agency (IEA), (2019). Available at: https://www.iea.org/topics/energysecurity/ (Accessed 09/03/2019).
[16] WEC, Deciding the Future: Energy Policy Scenarios to 2050. World Energy Council. London, 2007.
[17] J.Bengt, Security aspects of future renewable energy systems-A short overview, Energy 61, 598-605, 2013
[18] CIA, The World FactBook, Central Intelligence Agency (CIA), Available at: https://www.cia.gov/library/publications/resources/the-world-factbook/geos/sf.html (Accessed 20/03/2019)
[19] The world Bank, The World Energy Mix for Electricity Generation; Economic Indicators for Over 200 Countries,2015 Available at :https://www.theglobaleconomy.com/rankings/Carbon_dioxide_emissions_per_capita/ (Accesssed 18/03/2019)
[20] Oliver J, Cassilde M, Mbeo O, et al, Powering Africa, Unlocking Opportunities for Energy development in Southern Africa, Stockholm Environment Institution (SEI), 2018.
[21] IRENA, Africa 2030: Roadmap for a Renewable Energy Future, IRENA, 2015a, Abu Dhabi, www.irena.org/publications
[22] IRENA, African Power Sector: Planning and Prospects for Renewable Energy, Synthesis Report, International Renewable Energy Agency (IRENA),2015b, Abu Dhabi.
[23] K. Anme, SADC Sees growth in electricity generation capacity, CREAMER MEDIA, 2017.
[24] Energy Regulation Board (ERB). Energy Sector Report 2017. Lusaka: Zambia). Available at: www.erb.org.zm/downloads/esr2016.pdf (Accessed in October 2018)
[25] IRENA, Renewable capacity statistics 2018, International Renewable Energy Agency (IRENA),2018, Abu Dhabi
[26] IRENA, Southern African Power Pool: Planning and Prospects for Renewable Energy. International Renewable Energy Agency, 2013,Abu Dhabi. https://www.irena.org/ DocumentDownloads/Publications/ SAPP.pdf
[27] The world Bank, The World Energy Mix for Electricity Generation; Economic Indicators for Over 200 Countries, 2015, Available at :https://www.theglobaleconomy.com/rankings/Carbon_dioxide_emissions_per_capita/ (Accesssed 18/03/2019)
[28] A.Dakpogan, E.Smit, Measuring Electricity Security Risk, University of Stellenbosch Business School, 24 September 2018. Available at: https://mpra.ub.unimuenchen.de/89295/1/MPRA_paper_89295.pdf (Accessed 20/01/2019)
[29] M.Andre, J.Bengit, J.N.Lars, Assessing energy security:An overview of commonly used methodologies, Energy **73**, 1-14, 2014.
[30] OilPrice(2019), Oil Price Chart, Available at: https://oilprice.com/oil-price-charts/45 (Accessed 23/02/2019)
[31] Department of Minerals and Energy, South Africa, Renewable Energy Policy of the Republic of South Africa, 2003.
[32] Africa-EU Energy Partnership, Power Sector Market Brief: Zambia, 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 898-911 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Active Disturbance Rejection Control Design for a Haptic Machine Interface Platform

Syeda Nadiah Fatima Nahri[1], Shengzhi Du[1,*], Barend Jacobus van Wyk[2]

[1]*Department of Electrical Engineering, Tshwane University of Technology, Pretoria, 0001, South Africa*

[2]*Faculty of Engineering, the Built Environment and Technology, Nelson Mandela University, Port Elizabeth 6031, South Africa*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 24 January, 2021
Online: 12 February, 2021

Keywords:
Active disturbance rejection control (ADRC)
Human-machine interface (HMI)
Bilateral teleoperation systems
Mathematical modelling
Simulation

ABSTRACT

This paper proposes an active disturbance rejection control (ADRC) design for a haptic display platform structure. The motivation for the following scheme originates from the shortcomings faced by classical proportional integral derivative (PID) controllers in control theory. The ADRC is an unconventional model-independent approach, acknowledged as an effective controller in the existence of total plant uncertainties, and these uncertainties are inclusive of the total disturbances and unknown dynamics of the plant. The design and simulation for ADRC are established in MATLAB/ Simulink. The concerned electro-mechanical platform consists of dual ball screw driving system and DC motors. This overall physical system constitutes the haptic interface. Modelling of the two-dimensional physical platform is also explained in this article. Designing of ADRC controller and the human-machine interface (HMI) is followed by their integration, in order to obtain simulation results, thus proving the practicality and validity of the overall system. The results of the proposed controller are compared with the Proportional Integral (PI) controller, which suggests that the ADRC controller performs better as compared to the conventional PI controller.

1. Introduction

This paper is an extension of work originally presented in *2020 International SAUPEC/RobMech/PRASA Conference* [1]. Haptic feedback control systems containing human- in- the- loop (HITL) is relatively a young field of research, that display the interaction between human-machine systems. A typical bilateral teleoperation system consisting of the master device (human operator), communication channel and slave teleoperator, provide human operator's necessary interface and the experience of 'telepresence' [2, 3]. In simple text, the human perceives direct control and manipulation of the environment with their own hands, which is achieved by having their actions mediated physically by means of a robot, communication channel and control systems (i.e., controller), by providing the haptic sensing feedback signals (mostly interaction forces and position constraints) to the operators. Thus, the human operator can execute tasks remotely without physically being present there. Potential applications of teleoperation systems include highly specialized professions like telesurgery, mining, micro, and nanoparticles handling and artificial intelligence to name just a few [4]. These operations are performed globally, mainly to aid and assist the population with special needs.

The block diagram shown in Figure 1 gives a control system viewpoint of haptic teleoperation system. The main purpose of this system is the operator's ability to interact with the remote environment via haptic feedback. In this haptic teleoperation system, a human operator gets the feel of telepresence in the form of force operation and vibrations (sensations) from the haptic interface system, for instance, a simlulator. Thus, the movement of the human arm is based on the haptic feedback signals received from the haptic interface device. Accordingly, the human arm will send force and position signals to the controller (highlighted in the top red circle) which will forward the modified force and position signals to the robot system. The robot will then exert the received control signal to the concerned object present in the remote environment. The measured signals (force, position, and vibration) are then fed back to the haptic interface (highlighted in the bottom red circle) resulting in haptic feedback and HITL system. Thus, the human operator feels they are directly executing control operation of the remote side themselves.

*Corresponding Author: Shengzhi Du, Tshwane University of Technology, Pretoria, South Africa, Email: dushengzhi@gmail.com

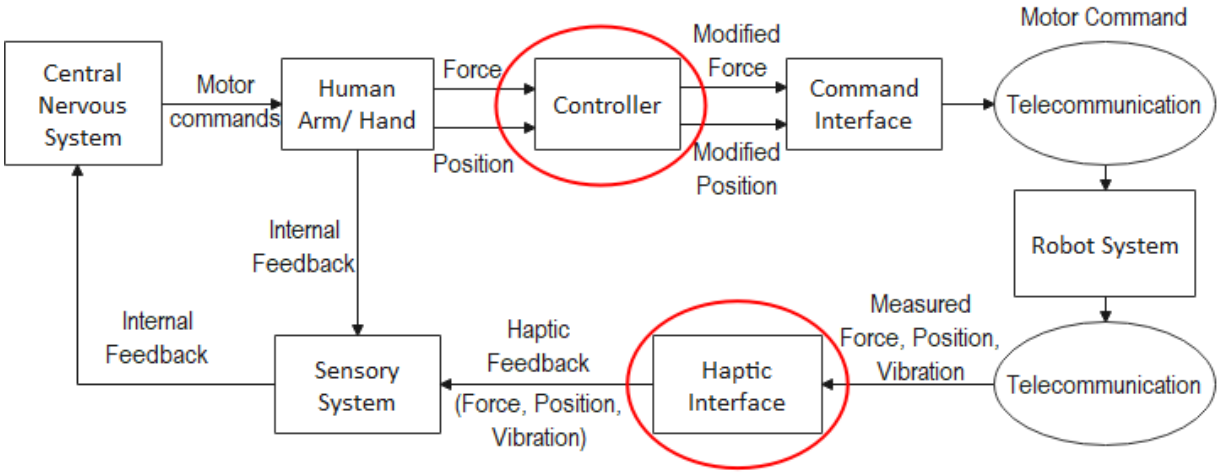


Figure 1: Control system viewpoint of haptic teleoperation system

In teleoperation systems, stability and transparency are both critical properties and conflicting to each other [5]. One of the main concerns relating to control systems design, is to achieve a decent tradeoff between these two contrasting goals. Stability can be achieved using Proportional integral derivative (PID) controls [6], wave-variables based passivity control [7] and time-domain passivity approach [8]. Whereas transparency is related to the human operator's potential to sense the slave side interaction with the environment. To enable this, force sensors are installed at the end effector providing force feedback control. Conversely, these force sensors impose measurement limitations due to their high cost, restricted area of contact and noise/ disturbances in the environment. Hence, disturbance observers were introduced [9]. On the contrary, these observers do not consider specific knowledge of the system like unmodelled dynamics and parameter uncertainties, resulting in narrowing the effectiveness of model-based controllers. Thus, these shortcomings have motivated the application of a model independent approach in the domain of bilateral teleoperation systems. Hence, an Active disturbance rejection control (ADRC) has been proven effective to tackle the limitations of both PID and model-based approaches [10,11]. ADRC's dynamic structure components, such as, the ESO (Extended state observer) and TD (Tracking differentiator), are used to compensate for internal and external disturbance(s) respectively. Therefore, the total system from the outside looks linear. The ADRC takes into account the time-varying nature of the system, unlike the PID controller, that considers particular time instants only. References [12–14] are survey papers that review the development of ADRC from its early beginning, along with, the controller's analysis being conducted from its methodology and theoretical perspective.

This paper aims to addresses and release the above mentioned constraints related to model- based approaches in control theory. This can be attained by introducing a model free approach (ADRC controller) for the physical interface platform (HMI) developed for this research. The objective of this paper includes the study of the disturbance rejection performances and control variable responses of the ADRC and PI controllers. The former provides a better understanding on the system response curves obtained, whereas the latter depicts more information on controller behaviour. Contributions of this paper include the algorithmic innovation of the ADRC controller and development of a controlled object that serves as a movable two- dimensional (2D) interface platform. Based on this platform, we propose an ADRC for the system shown in Figure 1. Furthermore, this paper focusses on the integration of the designed controller with the human machine interface platform, followed by their experimental analysis and comparison with the previously obtained PI controller results in [1]. This integration and analysis serves as the main focal point of the paper. Thus, by this comparison, the effectiveness of the total proposed system is measured. All the simulations will be conducted via MATLAB and Simulink in this study. Apart from MATLAB/ Simulink, experimental analysis for ADRC based applications have been conducted using other softwares, like LabView, Scilab/ Xcos and OpenPCS, to name a few.

The rest of the paper is structured in this way: Initially, the related work is briefly explained in Section 2, subsequently the proposed hardware platform is addressed in Section 3. Furthermore, Section 4 discusses the design and simulation of an ADRC controller, followed by results and discussions in Section 5. Lastly, Section 6 closes the paper.

2. Related Works

2.1. ADRC Controller

For almost a century, the PID controller has dominated the control technology to meet the ever-increasing demands of the automation industry. However, the PID controller design being system-dependent falls under the category of error-based empirical design paradigm (EDP) [11]. Tuning of such controllers is challenging due to model mismatch between the actual system and approximate model. Other challenges include computational error, noise deterioration caused by derivative control, performance degradation in the control law (linear weighted sum) and complications by integral control due to accumulation of

errors [10]. These conditions necessitated the introduction of a special type of nonlinear controller being system model-independent, called the ADRC. This controller possesses features of strong robustness, fast response speed, high accuracy control, minimum overshoot (smooth curve) and capability to estimate and compensate the overall effects of uncertainties and is thus minimized. These uncertainties include unknown plant dynamics, external disturbances and internal disturbances, in real-time. The capability to control uncertainties in a given system is one of the principal concerns in the field of modern control theory. The ADRC design was first proposed in [15], followed in 1998 [16], for plants with a large number of uncertainties. The same was introduced in [17], followed by an extension in [11]. An adequate transient response is attained using an ADRC design, which possesses a relatively simple framework.

2.1.1. Disturbance Rejection Mechanism

The existing paradigms in the field of control engineering, are the Modern Control Paradigm (MCP) (model-based approach) and error-based Empirical Design Paradigm (EDP) (classical trial and error approach) [11]. But there is a significant gap between theory and practice in the field of feedback control engineering. This calls for a paradigm shift introducing a new paradigm, namely, Disturbance Rejection Paradigm (DRC) [18].

The basic idea of ADRC can be put straight in the following train of thoughts. Firstly, cascade integral design of the canonical (ideal) configuration of the plant. Secondly, portion of the plant that is distinct from the ideal form can be called the 'total disturbance', and is rejected. This includes internal disturbance (i.e. variations in plant dynamics) and external disturbance (i.e. arises from the environment). Lastly, estimate and compensate the total disturbance using an extended state observer (ESO), so as to reduce the nonlinear time-varying system to the canonical form. In short, two main characteristics of modern control theory on which the fundamentals of ADRC are set: Firstly, the concept of canonical form; Secondly, the objective of a state observer [12]. Thus, with such a unique design concept, ADRC provides exceptional outstanding solutions to various pressing problems in the engineering field. Thus, practical applications of ADRC include robot motion control and aerial robotic systems [19], control tool for practitioners [20], control of humanoid robot [21], quadrotor helicopter [22], in several chemical processes, MEMS systems and power converters [23,24], to name a few.

2.2. Haptic Feedback Control

Over the last four decades, several bilateral teleoperation architectures were developed serving as the key technology for human interaction with the remote environment, by supplying the operator with haptic feedback. This haptic feedback control improves the quality of human-robot interaction, on the other hand, stability and transparency parameters are greatly affected due to time delay and packet loss in the communication channel, between the master side (operator) and slave side (remote robot interaction with the environment), as shown in Figure 2.

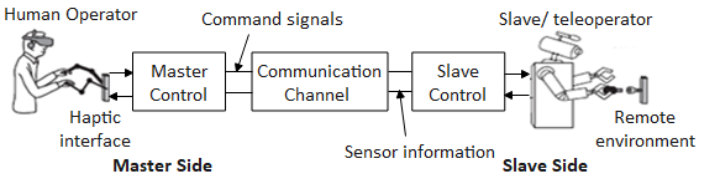


Figure 2: A basic haptic teleoperation system

Some of the important teleoperation architectures, along with their methods, studied and analysed in the past include Wave-variable (WV) approach [25,26], Time-domain passivity approach (TDPA) [27,28], and Model- mediated teleoperation (MMT) approach [29]. A detailed survey on the bilateral teleoperation algorithms and model-based control architectures for haptic interfaces is given in [3]. Such model-based methods represent input-output relationships using mathematical expressions, like the state-space model, Lagrange model, and transfer function model. They are also represented using various control laws, for example, the proportional integral (PI), proportional derivative (PD), PID, model predictive, sliding mode, and energy-based controllers [8,30].

The haptic teleoperation system control and design experience certain shortcomings, such as, uncertainties related to the communication channel and human operational dynamics, along with unsteady human operator decisions [31]. These uncertainties are time-varying and distinctive among individuals. Such inaccuracies make the model-based controller operation challenging and stimulate the use of model-free controllers, like the ADRC for smooth operations [10]. Thus, ADRC can contribute to the development of applications in connection with feedback contro1l systems in man- machine interface systems. Recent studies on ADRC controller for teleoperation systems based on distinct parameters, and for nonlinear systems, have been explained in [32] and [33,34] respectively.

3. Hardware design and modelling

3.1. Proposed Hardware Platform

The basic idea behind the platform shown in Figure 3, is to manoeuvre the operating handle in a rectangular coordinate system.

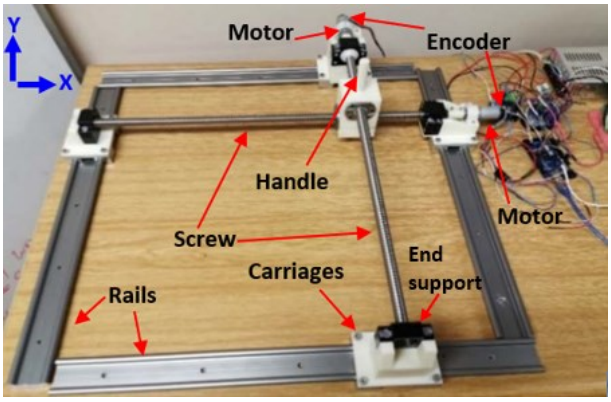


Figure 3: Experimental setup of the assembled hardware platform- 2 degree of freedom (2 DOF)

3.1.1. Mechanical components of the HMI design

Mechanical and electrical components constitute the experimental platform. The mechanical components used are stated in this subsection. The operating handle is a three-dimensional (3D) printed object, obtained using a designing software called Onshape. Both X-axis and Y-axis have a ball screw feed drive system (SFU1605) aligned parallel to each of the axis, with the operating handle assembled over the lead screw system. The ball-lead screws and nut dimensions are 600mm (length) x 16mm (diameter) x 5mm (pitch). Ball screw nuts provide the required linkage between the lead screws and the operating handle. Hence, ball lead-screw's rotational motion is transformed to operating handle's linear translation motion. As seen in Figure 3, two pairs of linear guide rails of dimensions 600mm (length) x 4mm (width) hold up the handle and screws. The guide rails used are called the Igus W series (WS-10-40-600), which possess benefits like low wear and tear, low friction, low noise system, and requires no maintenance being resistant to dust and dirt. The 3D printed carriages that firmly hold both ends of the lead screws, together with the sliding rails, are all mounted on a levelled surface. The end supports used for the ball screw (SFU16XX - BK 12 + BF 10) contains two parts, one for the front and the other for the rear of the lead screw. The fixed part is 10mm in diameter, whereas the floating part is 12mm in diameter. Also includes circlips and deep groove ball bearings. A DC motor mounted on one end-support component, and the motor shaft mechanically connected to the lead screw, help to drive the lead screw system.

3.1.2. Electrical components of the HMI design

The electrical components used in the experimental setup are explained in this subsection. As seen in Figure 3, a DC motor shaft equipped with an encoder is incorporated at one end of both the lead screws. The motors chosen are called the Metal Gearmotor 75:1 (gear ratio), with a 48 CPR quadrature encoder integrated on each motor, which provides 3592 counts per revolution of the gearbox's output shaft. Size of the motor is 25mm (diameter) x 66mm (length), with a shaft diameter of 4mm. These motors are intended for comfortable operation in the 3- 9V range. The encoders are connected to the Arduino MEGA 2560 board. Both motors are provided with a specific motor driver called the BTS7960 High Power H- Bridge motor driver. The motor driver has the following specifications, 43A of maximum current and input voltage range of 6V to 27V. According to the inout voltage received from the Arduino board, the motor drivers will adapt to these incoming signals, and then apply the appropriate voltage to the motor. The system contains two ACS714 current sensors, one for each motor driver, to aid in measuring the strength applied on the operating handle.

On accomplishing the hardware installation, signal transmission between MATLAB and the proposed interface system (HMI) was performed using Arduino 2560. This Arduino board was powered by an external power supply. Currently, only a single axis movement of the lead screw drive system is considered in this paper.

3.2. Mathematical Modelling of the Assembled Hardware Platform

Two of the single-axis electro-mechanical platform shown in Figure 3 are coupled together to implement the 2D movement. Torque (T_M) generated by the DC motor is transmitted to the ball screw shaft via coupling, which in turn causes the operating handle to move. Transmission ratio 'i' of ball screw system is defined as the distance travelled 'h' during one revolution of the shaft [35], also denotes, the transformation of rotational movement of screw shaft into linear motion of the operating handle. 'i' is given by the following equation,

$$i = {}^{h}/_{2\pi} \tag{1}$$

The lumped mass model (LMM) developed in [35,36] is adopted to derive the mathematical model for this structure. Such a model indicates several functions related to the system, by simplifying the simulation model to a reduced number of DOF system. As seen in Figure 4, all the rigidity parameters (k), inertial components (M), and damping parameters (c) are denoted in terms of axial and rotational components of the same parameter, i.e., total rotational and axial rigidity components, represented as (k_{rot}) and (k_{ax}) respectively.

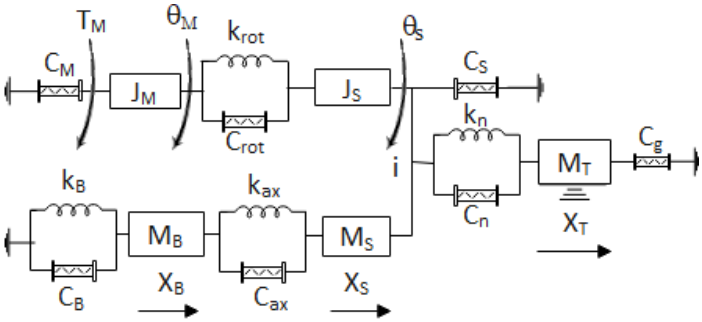


Figure 4: The LMM of the recommended HMI platform

The LMM parameters in Figure 4 consists of screw shaft rotary inertia J_S, base mass M_B, motor inertia J_M, screw shaft side equivalent mass M_S, handle mass M_T, equivalent torsional rigidity k_{rot}, ball screw nut rigidity k_n, axial base rigidity k_B, equivalent axial rigidity k_{ax}, DC motor torsional damping C_M, equivalent rotational damping C_{rot}, sides screw shaft damping C_S, ball screw nut damping C_n, axial guide damping C_g, axial base damping C_B, and equivalent axial damping C_{ax}.

Parameters (DOF) of the model shown in Figure 4 include,

θ_M- Angular rotation of the DC motor

θ_S- Angular rotation of screw shaft at handle position

X_B- Axial base displacement

X_S- Axial displacement of the screw shaft at handle position

X_T- Operating handle position

Equation 2 represents the mathematical model of a DC motor,

$$\dot{T}_M = \frac{-K_{bm}K_t}{L_a}\dot{\theta}_M - \frac{R_a}{K_t}T_M + \frac{K_t}{L_a}E_a \tag{2}$$

where, K_t and K_{bm} denote motor torque and back emf constants, L_a is winding leakage inductance, $\dot{\theta}_M$ is the angular velocity of the

rotor inside motor, E_a is the voltage source, and R_a is the armature resistance. References [1,36] provide a detailed step by step Lagrange and state-space mathematical modelling for the physical structure in Figure 4.

3.2.1. Lagrange Model

Second order Lagrange's equations are employed to develop the ball screw feed drive system's dynamic model. Lagrange function is defined by the variation between kinetic and potential energies of a physical system. Equation (3) gives the mathematical relations for total kinetic energy (T), potential energy (U) and dissipation function (D), associated with the total system.

$$\begin{cases} T = \frac{1}{2}J_M\dot{\theta}_M^2 + \frac{1}{2}J_S\dot{\theta}_S^2 + \frac{1}{2}M_B\dot{X}_B^2 + \frac{1}{2}M_S\dot{X}_S^2 + \frac{1}{2}M_T\dot{X}_T^2 \\ U = \frac{1}{2}k_{rot}(\theta_M - \theta_S)^2 + \frac{1}{2}k_B{X_B}^2 + \frac{1}{2}k_{ax}(X_S - X_B)^2 + \frac{1}{2}k_n(X_T - X_S - i\theta_s)^2 \\ D = \frac{1}{2}C_M\dot{\theta}_M^2 + \frac{1}{2}C_{rot}(\dot{\theta}_M - \dot{\theta}_S)^2 + \frac{1}{2}C_S\dot{\theta}_S^2 + \frac{1}{2}C_B\dot{X}_B^2 + \frac{1}{2}C_{ax}(\dot{X}_S - \dot{X}_B)^2 + \frac{1}{2}C_n(\dot{X}_T - \dot{X}_S - i\dot{\theta}_S)^2 + \frac{1}{2}C_g\dot{X}_T^2 \end{cases} \quad (3)$$

$$\text{Lagrange equation, } L = T - U = \frac{1}{2}J_M\dot{\theta}_M^2 + \frac{1}{2}J_S\dot{\theta}_S^2 + \frac{1}{2}M_B\dot{X}_B^2 + \frac{1}{2}M_S\dot{X}_S^2 + \frac{1}{2}M_T\dot{X}_T^2 - \frac{1}{2}k_{rot}(\theta_M - \theta_S)^2 - \frac{1}{2}k_B{X_B}^2 - \frac{1}{2}k_{ax}(X_S - X_B)^2 - \frac{1}{2}k_n(X_T - X_S - i\theta_s)^2 \quad (4)$$

Equation (5) represents the Lagrangian function of a system. This function is calculated about force 'Q' and the generalized coordinate 'q'.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} - \frac{\partial D}{\partial \dot{q}} = Q \quad (5)$$

Coordinates 'q' and 'Q' indicate independent coordinates and force inputs to the physical system respectively, given by the following relations in (6),

$$\begin{cases} q = (\theta_M \ \ \theta_S \ \ X_B \ \ X_s \ \ X_T)^T \\ Q = (T_M \ \ 0 \ \ 0 \ \ 0 \ \ 0)^T \end{cases} \quad (6)$$

On solving (4) using (5) and (6), the following set of equations were obtained,

$$\begin{cases} J_M\ddot{\theta}_M + k_{rot}(\theta_M - \theta_S) + C_M\dot{\theta}_M + C_{rot}(\dot{\theta}_M + \dot{\theta}_S) = T_M \\ J_S\ddot{\theta}_S - k_{rot}(\theta_M - \theta_S) - ik_n(X_T - X_S - i\theta_S) + C_S\dot{\theta}_S + C_{rot}(\dot{\theta}_M + \dot{\theta}_S) + iC_n(\dot{X}_T + \dot{X}_S + i\dot{\theta}_S) = 0 \\ M_B\ddot{X}_B + k_BX_B - k_{ax}(X_s - X_B) + C_B\dot{X}_B + C_{ax}(\dot{X}_S + \dot{X}_B) = 0 \\ M_S\ddot{X}_S + k_{ax}(X_s - X_B) - k_n(X_T - X_s - i\theta_S) + C_{ax}(\dot{X}_S + \dot{X}_B) + C_n(\dot{X}_T + \dot{X}_S + i\dot{\theta}_S) = 0 \\ M_T\ddot{X}_T + k_n(X_T - X_S - i\theta_S) + C_n(\dot{X}_T + \dot{X}_S + i\dot{\theta}_S) + C_g\dot{X}_T = 0 \end{cases} \quad (7)$$

After rearranging the terms in (7), a set of five equations are obtained, that is given in (8). Based on these equations, the Simulink model is built on the MATLAB software, this software-based mathematical model is displayed in Figure 5.

$$\begin{cases} \ddot{\theta}_M = \frac{1}{J_M}[T_M - k_{rot}(\theta_M - \theta_S) - \dot{\theta}_M(C_M + C_{rot}) - C_{rot}\dot{\theta}_S] \\ \ddot{\theta}_S = \frac{1}{J_S}[-\dot{\theta}_S\left(C_{rot} + C_S + i^2C_n\right) - \theta_S\left(k_{rot} + i^2k_n\right) - C_{rot}\dot{\theta}_M + k_{rot}\theta_M + ik_nX_T - ik_nX_S - iC_n\dot{X}_T - iC_n\dot{X}_S] \\ \ddot{X}_B = \frac{1}{M_B}[-X_B(k_{ax} + k_B) + k_{ax}X_S - C_{ax}\dot{X}_S - \dot{X}_B(C_B + C_{ax})] \\ \ddot{X}_S = \frac{1}{M_S}[k_{ax}X_B - X_S(k_{ax} + k_n) + k_nX_T - ik_n\theta_S - C_{ax}\dot{X}_B - \dot{X}_S(C_{ax} + C_n) - C_n\dot{X}_T - iC_n\dot{\theta}_S] \\ \ddot{X}_T = \frac{1}{M_T}[-k_nX_T - \dot{X}_T(C_n + C_g) + k_nX_S + ik_n\theta_S - C_n\dot{X}_S - iC_n\dot{\theta}_S] \end{cases} \quad (8)$$

LMM of the proposed design, is given in matrix form by the following relation in (9),

$$m\ddot{q} + c\dot{q} + kq = Q \quad (9)$$

Equations (10), (11) and (12) indicate the values for matrices *m*, *k* and *c* respectively. These matrix values are acquired from (7).

$$m = \begin{bmatrix} J_M & 0 & 0 & 0 & 0 \\ 0 & J_S & 0 & 0 & 0 \\ 0 & 0 & M_B & 0 & 0 \\ 0 & 0 & 0 & M_S & 0 \\ 0 & 0 & 0 & 0 & M_T \end{bmatrix} \quad (10)$$

$$k = \begin{bmatrix} k_{rot} & -k_{rot} & 0 & 0 & 0 \\ -k_{rot} & (k_{rot} + i^2 k_n) & 0 & i k_n & -i k_n \\ 0 & 0 & (k_B + k_{ax}) & -k_{ax} & 0 \\ 0 & i k_n & -k_{ax} & (k_{ax} + k_n) & -k_n \\ 0 & -i k_n & 0 & -k_n & k_n \end{bmatrix} \tag{11}$$

$$c = \begin{bmatrix} (C_M + C_{rot}) & C_{rot} & 0 & 0 & 0 \\ C_{rot} & (C_{rot} + C_S + i^2 C_n) & 0 & i C_n & i C_n \\ 0 & 0 & (C_B + C_{ax}) & C_{ax} & 0 \\ 0 & i C_n & C_{ax} & (C_{ax} + C_n) & C_n \\ 0 & iC_n & 0 & C_n & (C_n + C_g) \end{bmatrix} \tag{12}$$

3.2.2. State-Space Model and Transfer Function

Matric values given in (10-12), are substituted in (13), to provide the state-space representation of system model.

$$\ddot{q} = \frac{1}{m}(Q - c * \dot{q} - k * q) \tag{13}$$

For the state-space model, we first determine the number of states. In this experiment, each of the five equations in (7) are second-order differential equations, i.e, each equation will include at least two states.

State-state model is represented as, $\dot{x} = Ax + Bu$ and $y = Cx + Du$. In this system A is (11x11) matrix, B is (11x1) matrix, C is (1x11) matrix and D matrix is 0.

Equations (14) and (15) represent the State- space model and Transfer function of the system.

$$\dot{X} = \begin{bmatrix} \frac{-(C_M + C_{rot})}{J_M} & \frac{-k_{rot}}{J_M} & \frac{-C_{rot}}{J_M} & \frac{k_{rot}}{J_M} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-C_{rot}}{J_S} & \frac{k_{rot}}{J_S} & \frac{-(C_{rot} + C_S + i^2 C_n)}{J_S} & \frac{-(k_{rot} + i^2 k_n)}{J_S} & 0 & 0 & \frac{-i C_n}{J_S} & \frac{-i k_n}{J_S} & \frac{-i C_n}{J_S} & \frac{i k_n}{J_S} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-(C_B + C_{ax})}{M_B} & \frac{-(k_B + k_{ax})}{M_B} & \frac{-C_{ax}}{M_B} & \frac{k_{ax}}{M_B} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-i C_n}{M_S} & \frac{-i k_n}{M_S} & \frac{-C_{ax}}{M_S} & \frac{k_{ax}}{M_S} & \frac{-(C_{ax} + C_n)}{M_S} & \frac{-(k_{ax} + k_n)}{M_S} & \frac{-C_n}{M_S} & \frac{k_n}{M_S} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-i C_n}{M_T} & \frac{i k_n}{M_T} & 0 & 0 & \frac{-C_n}{M_T} & \frac{k_n}{M_T} & \frac{-(C_n + C_g)}{M_T} & \frac{-k_n}{M_T} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{-K_{bm} K_t}{L_a} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{-R_a}{L_a} \end{bmatrix} X + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{K_t}{L_a} \end{bmatrix} E_a \tag{14}$$

$$Y = [\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,]\mathrm{X} \tag{15}$$

where $X = [\,x_1\, x_2 \ldots x_{11}\,]$. The output is the operating handle position (X_T), which is assigned to the state variable x_{10}.

The transfer function (G) stated in (16) is obtained from the state-space model by using the following MATLAB codes,

```
Sys = ss(A,B,C,D);G= tf(Sys)
```

$$G = \frac{0.001122\,s^6 - 1.1 * 10^8 s^5 + 3.453 * 10^{12} s^4 - 7.294 * 10^{14} s^3 + 2.29 * 10^{19} s^2 - 1.896 * 10^{19} s + 5.954 * 10^{23}}{s^{11} + 407.7 s^{10} + 1.817 * 10^7 s^9 + 7.334 * 10^9 s^8 + 4.094 * 10^{13} s^7 + 1.552 * 10^{16} s^6 + 9.493 * 10^{18} s^5 + 3.353 * 10^{21} s^4 + 3.329 * 10^{23} s^3 + 7.323 * 10^{25} s^2 + 3.045 * 10^{27} s + 1.138 * 10^{16}} \tag{16}$$

The Controllability and Observability of the system can be obtained from its state- space model in (14). The Controllability matrix (C_M) and Observability matrix (O_M) were calculated using the following MATLAB commands, ' $ctrb(A, B)$ ' and '$obsv(A, C)$' respectively. Both the matrices were found to be invertible with their determinant values not equal to zero. Thus, this shows that the system is controllable and observable.

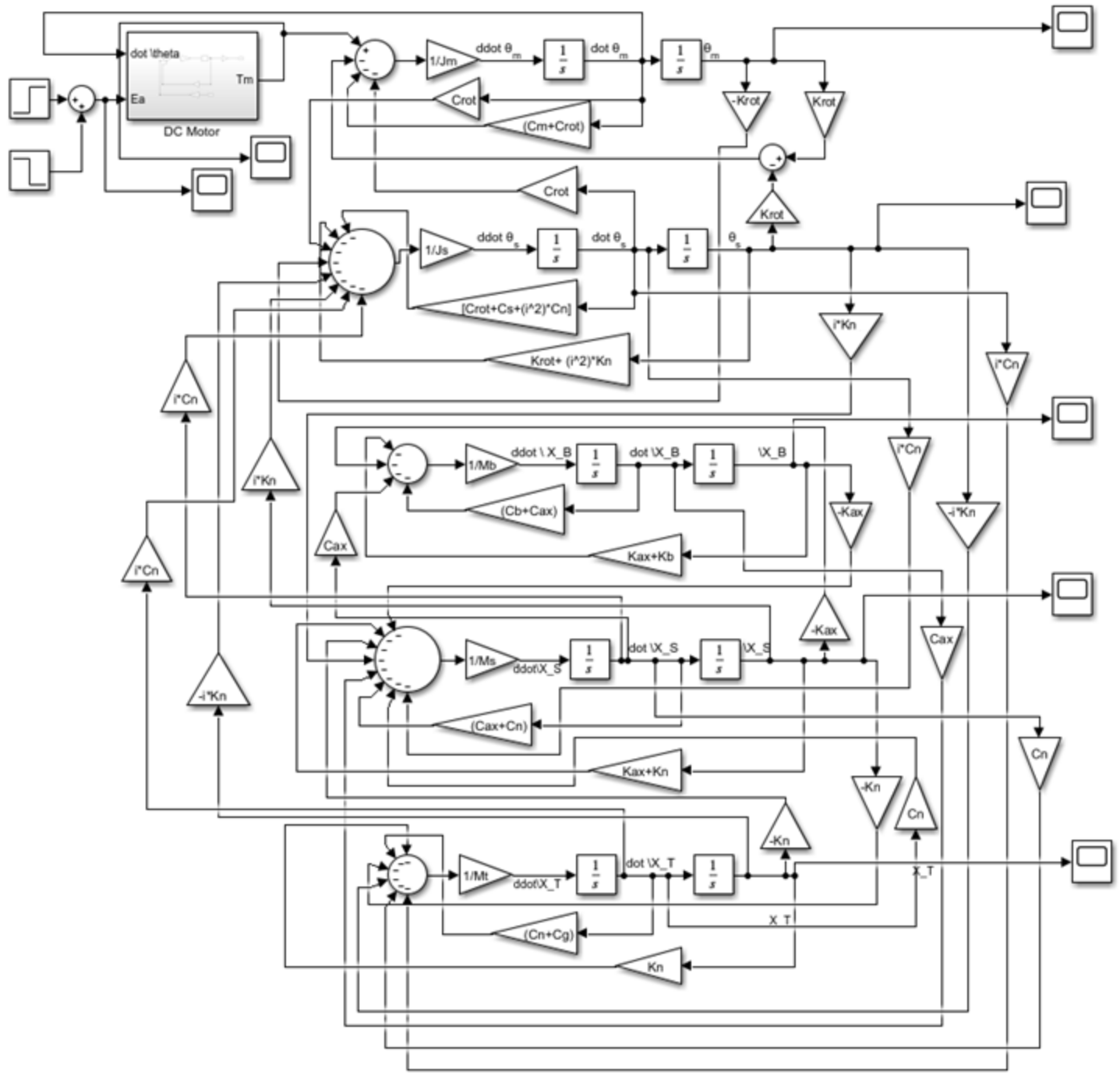


Figure 5: Simulink Block of the concerned Haptic Machine Interface (hardware platform)

4. Algorithm and Simulink Modelling of ADRC

ADRC originates from the classical PID controller, retaining the central idea of error feedback control. The drawbacks of the PID controller are overcome by the significant dynamic structure of ADRC. As seen in Figure 6, an ADRC architecture is composed of three parts, namely, the Tracking Differentiator (TD), Extended State Observer (ESO) and Nonlinear State Error Feedback (NLSEF). Each of the subsystems is explained here, followed by their Simulink model created on MATLAB.

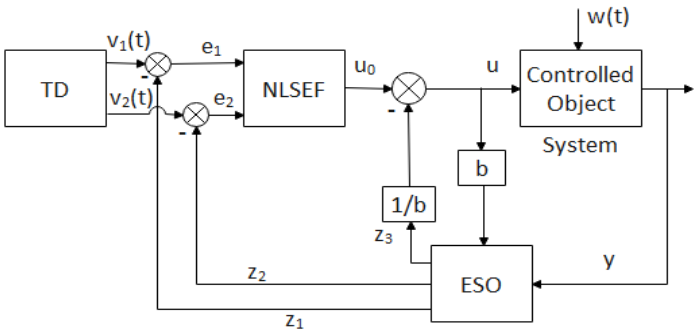


Figure 6: Basic ADRC structure of a two- order system

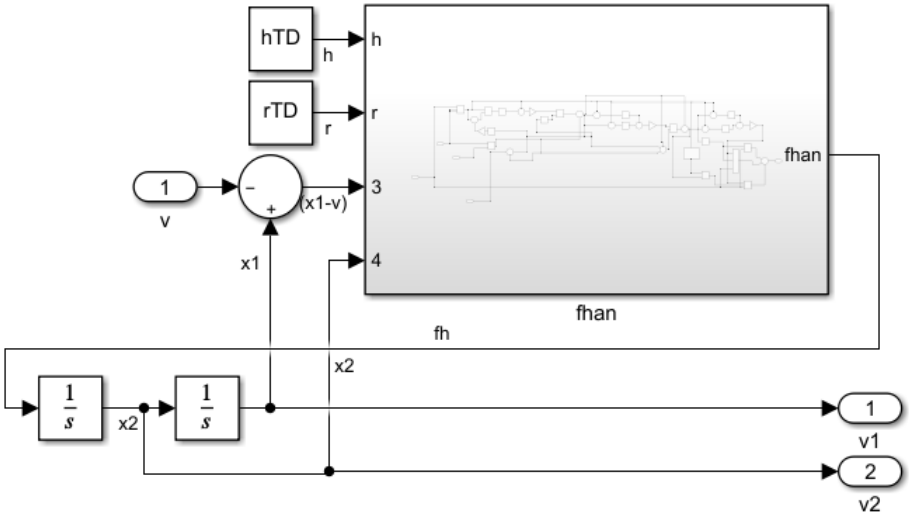


Figure 7: Simulink Model of TD subsystem

Definitions of parameters used for ADRC design: $v_1(t)$ or v_1 is the tracking signal, and $v_2(t)$ *or* $v_2$is the differential signal of the given input signal. The output signal of the system and input signal of ESO is given by 'y'. u is called the control volume, it is the input signal of the controlled object and the ESO. $z_1$and $z_2$are variables of the estimated state. $z_3$is the estimated signal of the total disturbance, i.e., both internal disturbance and the external interference present. $w(t)$ is the external disturbance acting on the concerned system. $e_1$and e_2 are the error signals fed to the NLSEF.

$$\begin{cases} e_1 = v_1 - z_1 \\ e_2 = v_2 - z_2 \end{cases} \tag{17}$$

4.1. Tracking Differentiator (TD)

TD determines the transition process for the system input, in order to obtain a smooth input signal $v_1(t)$ which is the desired trajectory and its differential signal $v_2(t)$. Consider the input signal $v(t)$ is to be differentiated, thus, TD provides the fastest tracking $v(t)$ given by (18). By solving the following differential equation, the required transient profile is obtained.

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -r\, sign(x_1 - v(t) + \frac{x_2 |x_2|}{2r}) \end{cases} \tag{18}$$

The parameter r is chosen according to the speed of the transient profile, i.e., to expedite or moderately slow down the transient profile. v is the desired value for x_1. The Simulink model of TD is shown in Figure 7, wherein $fhan$ is a nonlinear function denoted as $fhan(x_1 - v, x_2, r, h)$. h is the simulation step, also called the sampling period.

$$fh = fhan(v_1 - v, v_2, r_0, h_0) \tag{19}$$

The function $fhan$ is presented later in this paper. Equation (19) is a time-optimal solution provides no overshoot and fastest convergence from v_1 to v. Parameters r_0 and h_0 are equated to r and h respectively [10].

4.2. Nonlinear State Error Feedback (NLSEF)

The control law of PID controller implements a linear combination of the present, past and future predictive kinds of the tracking error. For an infinite time, the tracking error attains zero value for linear feedback systems. Whereas, for a nonlinear feedback function given in (20), the error can reach zero value much earlier in a finite time ($\alpha < 1$). Such a mechanism significantly alleviates the steady-state error in comparison to an integral control [10]. Thus, in an ADRC control framework, the nonlinear feedback functions fal and $fhan$ play a significant role during operation.

$$fal(x, \alpha, d) = \begin{cases} \frac{x}{d^{1-\alpha}}, |x| \le d \\ |x|^{\propto} sign(x), |x| > d \end{cases} \tag{20}$$

The Simulink model of NLSEF seen in Figure 8, has been developed using the mathematical relations from (20-23).

$$u = k_0 e_0 + fhan(e_1, c * e_2, r, h_0) \tag{21}$$

Parameter c denotes the damping coefficient. And function $fhan(x_1, x_2, r, h)$ is represented using the following set of equations [37],

$$fhan = -r\left(\frac{a}{d} - sign(a)\right) s_a - r sign(a) \tag{22}$$

where,

$$s_a = \frac{[sign(a + d) - sign(a - d)]}{2} \tag{23}$$

$$\begin{cases} a = (a_0 + y - a_2)s_y + a_2 \\ d = rh^2 \\ a_0 = hx_2 \\ y = x_1 + a_0 \\ a_2 = a_0 + [sign(y)(a_1 - d)/2] \\ s_y = [sign(y+d) - sign(y-d)]/2 \\ a_1 = \sqrt{d(d+8|y|)} \end{cases} \quad (24)$$

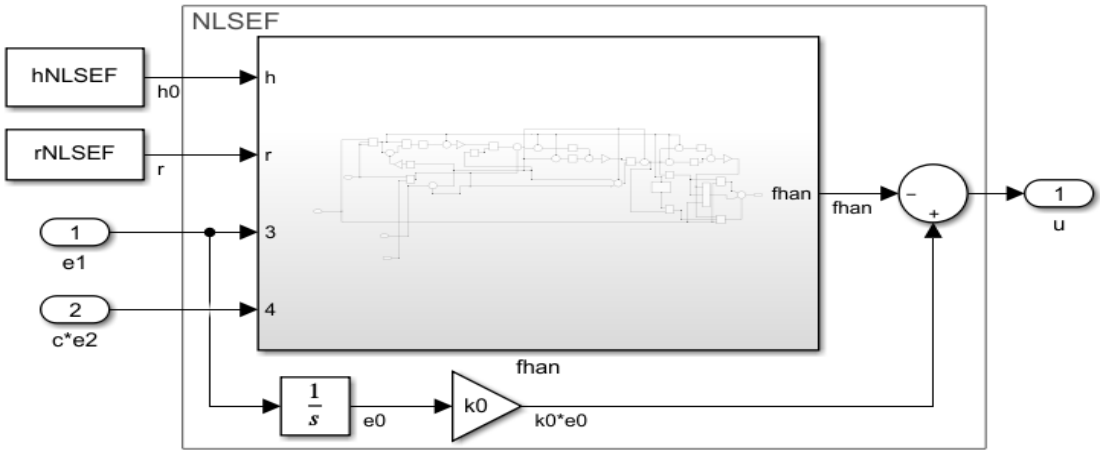


Figure 8: Simulink Model of NLSEF subsystem

4.3. Extended State Observer (ESO)

ESO estimates full system states and the total effect of disturbances (or extended states) in real-time. These disturbances may arise from unknown or nonlinear system dynamics of the manipulator, external disturbances, and mismatch of control parameters. Thus, in robot control, ESO present in the feedback loop is used to estimate and cancel the effects of these uncertainties. Figure 9 gives a detailed Simulink model of the ESO block. This subsystem constitutes of the nonlinear feedback function fal defined in (20). The algorithm for ESO is characterized as follows [37,38]:

$$\begin{cases} fe_1 = fal(e, 0.25, h) \\ fe = fal(e, 0.5, h) \end{cases} \quad (25)$$

$$\begin{cases} e = z_1 - y \\ \dot{z}_1 = z_1 + h(z_2 - \beta_{01}e) \\ \dot{z}_2 = z_2 + h(z_3 - \beta_{02}fe + b_0u) \\ \dot{z}_3 = z_3 + h(-\beta_{03}fe_1) \end{cases} \quad (26)$$

where, $\beta_{01}, \beta_{02}, \beta_{03}$ are parameters of ADRC that can be obtained from the system's sampling step. Thus, four main parameters to be controlled are, simulation step (h), control gain (r), damping coefficient (c) and compensation factor (b_0).

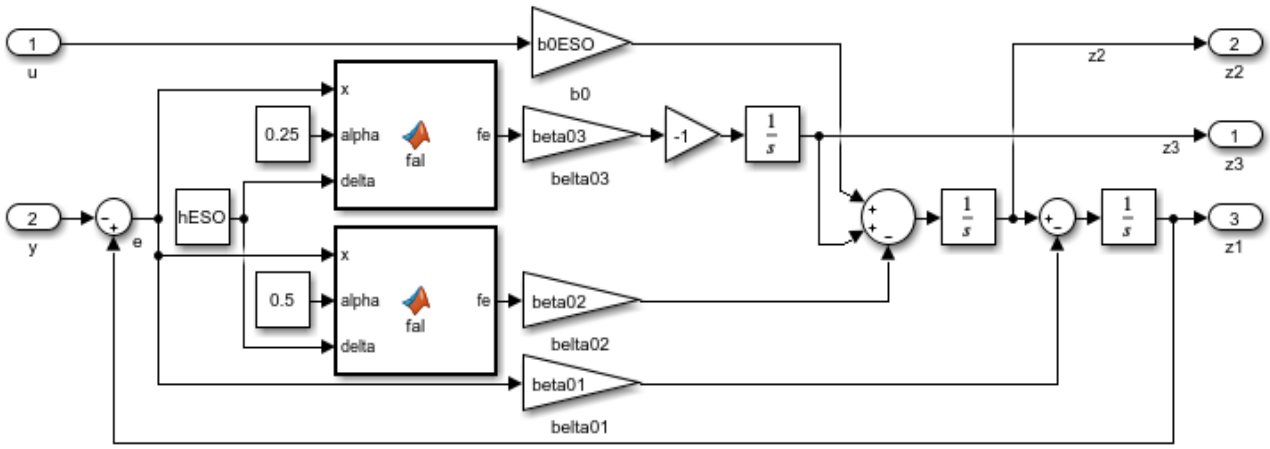


Figure 9: Simulation Model of ESO subsystem

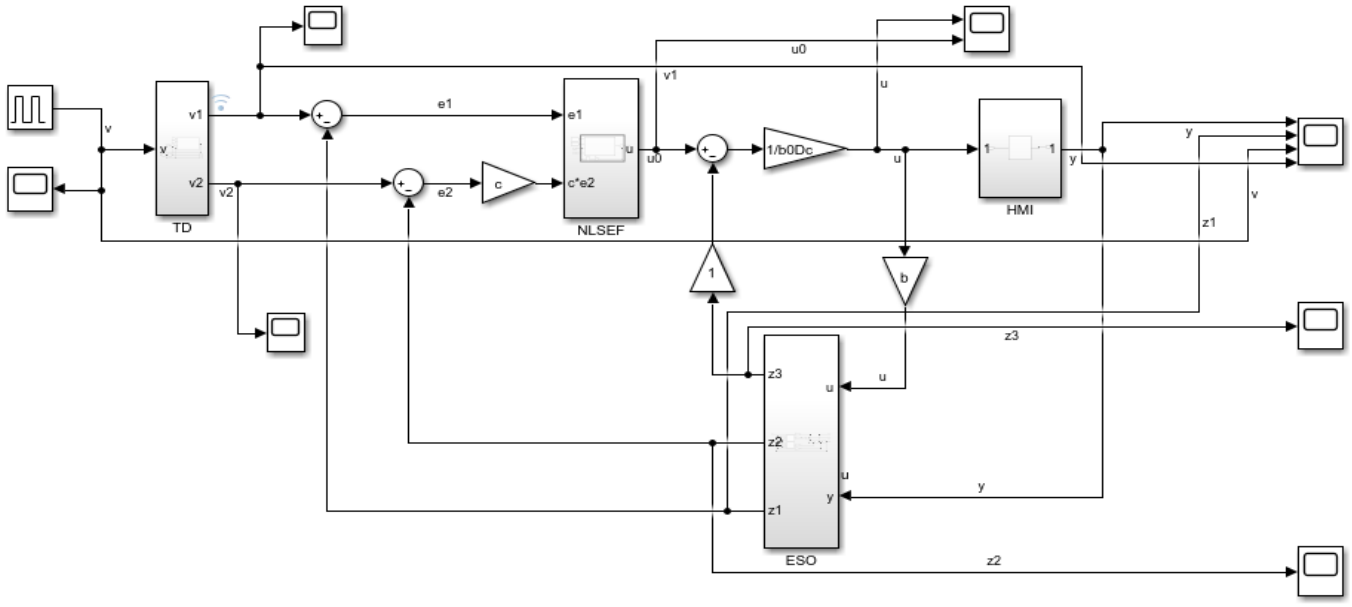


Figure 10: Simulink model of ADRC system including the HMI as the Controlled object

4.4. Controlled Object

The integration of TD structure, with the NLSEF structure, followed by the nonlinear ESO structure (which provides total disturbances estimation and rejection), altogether constitute the ADRC system. This controller is integrated with the HMI system. The Simulink model of HMI shown in Figure 5, is called the Controlled object in the ADRC system as seen in Figure 10.

5. Results and Discussions

5.1. Validation of the consistence between the physical interface system and the Simulink model

Figure 11 depicts the output response of the proposed physical system with PI controller, and Figure 12 depicts the LMM's response executed via MATLAB and Simulink.

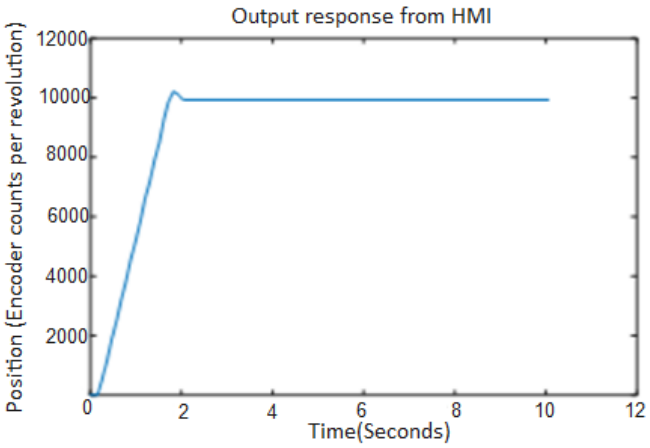


Figure 11: Output response of handle position from the physical system

As seen in Figures 11 and 12, except for the magnitude, output response curves obtained from the physical interface platform and from the Simulink model respectively, are found to be similar. Dissimilarity in magnitude is mainly because, firstly, the units of the Y-coordinates are distinct. Secondly, it is difficult to measure the lumped parameter values of the physical platform. Thus, the proposed platform's parameter values are slightly different from the simulation model values (i.e. designed via Simulink). Yet the simulation model substantiated the fundamental characteristic of the recommended HMI platform. The parameter values that were used in the Simulink model are listed in Table 1.

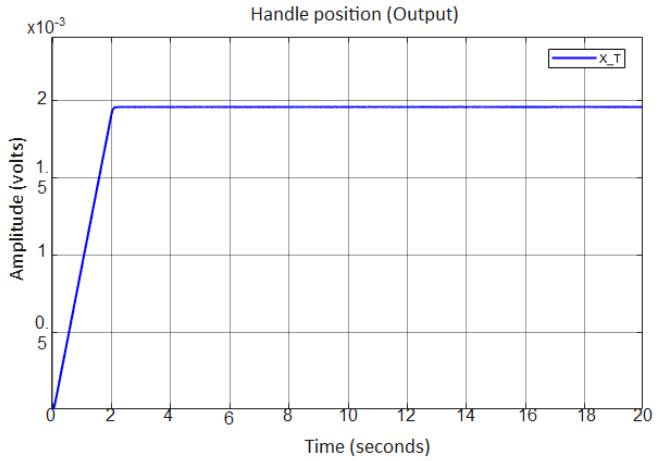


Figure 12: Output response of the Handle position (X_T) from Simulink model

Table 1: Parameter values of HMI used in Simulink Model

Parameters	Value
J_M	$6.75 * 10^{-3}(kg \cdot m^2)$
J_S	$1.7 * 10^{-3}(kg \cdot m^2)$
M_B	$3820\ (kg)$
M_S	$11.28\ \ (kg)$
M_T	$206\ (kg)$
k_{rot}	$3.14 * 10^{3}(N \cdot m \cdot rad^{-1})$
k_n	$9.8 * 10^{7}(N \cdot m^{-1})$
k_B	$1 * 10^{8}(N \cdot m^{-1})$
k_{ax}	$0.743 * 10^{8}(N \cdot m^{-1})$
C_M	$0.001\ (N \cdot s \cdot m^{-1})$
C_{rot}	$0.1\ (N \cdot s \cdot m^{-1})$

C_n	$0.001\ (N \cdot s \cdot m^{-1})$
C_S	$0.001\ (N \cdot s \cdot m^{-1})$
C_B	$0.1\ (N \cdot s \cdot m^{-1})$
C_{ax}	$0.001\ (N \cdot s \cdot m^{-1})$
C_g	$0.001\ (N \cdot s \cdot m^{-1})$
L_a	$1.5 * 10^{-3}(H)$
R_a	$0.5\ (\Omega)$
K_t	$0.05\ (N \cdot m \cdot A^{-1})$
K_{bm}	$0.05\ (V \cdot rad^{-1} \cdot s^{-1})$
T	$0.01\ (s)$
i	$7.96 * 10^{-4}(m \cdot rad^{-1})$

5.2. Disturbance rejection performance of the designed ADRC controller and comparison with PI controller

Experimental analysis of the comparison between ADRC and PI controllers is based on the following performance indices, that is, the system response obtained in the vicinity of disturbance present in the motor voltage, and controller variable values of both the controllers. The volatage unstability in the motor is one of the primary disturbance to the control system. In this experiment, the disturbance on the voltage supply to the motor is added to validate the disturbance rejection performance of the designed ADRC and PI controllers. The disturbance is added at the 5th second with an amplitude of 5000mV and lasting 0.2s, as shown in Figure 13.

Figure 14 shows the system response of the system controlled by ADRC and PI controllers. For the sake of fair comparison, the ADRC and PI controllers are tuned to have similar overshoot values. From Figure 14, one can find the system response with ADRC (Figure 14(a)) goes back to the reference value quickly after the overshot, however, the PI controller takes longer time to do so. Both system responses are affected by the disburance (starting from 5s in Figure 14), but the ADRC controlled system gets much smaller shift from the reference value than the PI controller.

Figure 15 demonstrates the control variable values of the two controllers, which depicts more details on controller behaviour. From Figure 15(a), one finds the ADRC controller starts with a mild control at the beginning due to the soften effect obtained in the tracking differentiator, which makes a smaller peak value on the control (Figure 15(a)). When the disturbance happens, the designed ADRC depresses the shift from reference value greatly, which leads to only a small change on the system response (refer to Figure 15(a)) from 5s. The PI controller is much more rigid when compared with the designed ADRC, which makes the controller start with an aggressive control (Figure 15(b)) at the beginning, and bigger shift from the reference value after the disturbance (refer to Figure 15(b)) from 5s.

These advantages of the system response controlled by ADRC can be explained in Figure 16, i.e. the total disturbances estimated by the ADRC. From Figure 16, one finds at the beginning of the control, the intial stage of the input step signal is considered as a disturbance (around 0s in Figure 16) which is compensated to obtain a soft rise of the system. This is obtained by the TD. When the disturbance occurs (around 5s in Figure 16), it is considered as disturbance (z_3 in (26)) as well and is compensated accordingly. Now the ESO observes the disturbance, and is compensated to the control variable generated by the NLSEF. This mechanism is shown in Figure 10. Thus, this disturbance rejection property of ADRC controller makes it unique in operation as compared to the PI controller.

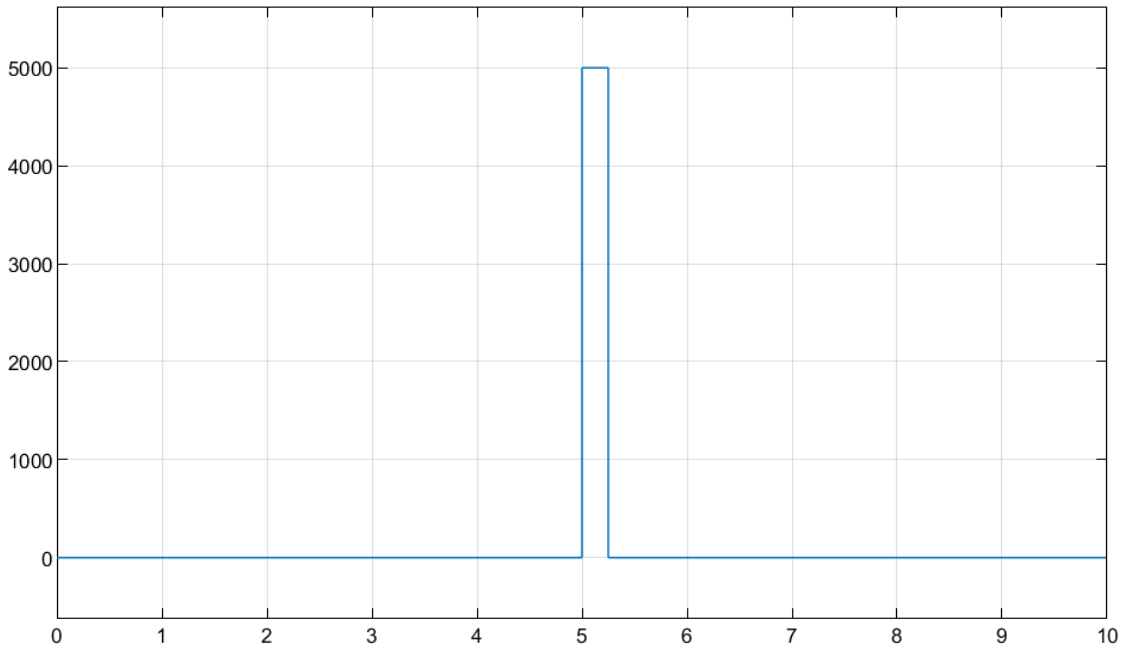
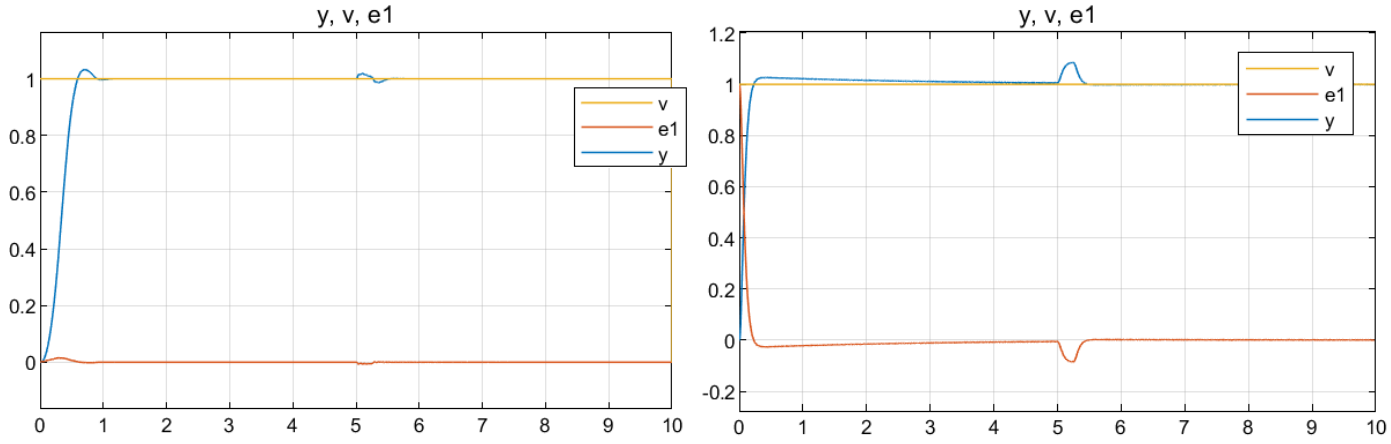
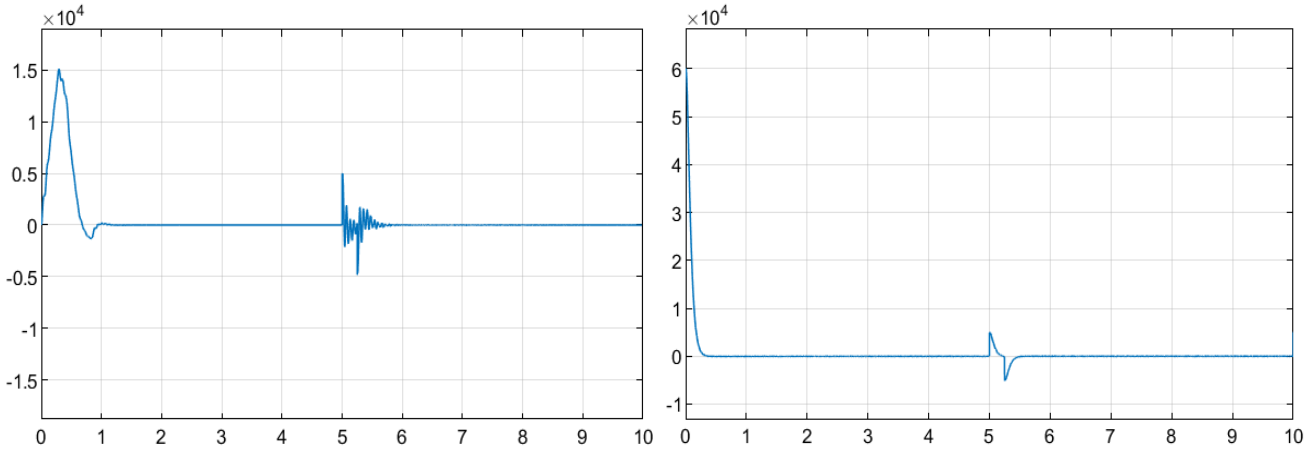


Figure 13: The added voltage disturbance in motor



(a) ADRC controller (b) PI controller

Figure 14: System response



(a) ADRC controller (b) PI controller

Figure 15: Control variable values

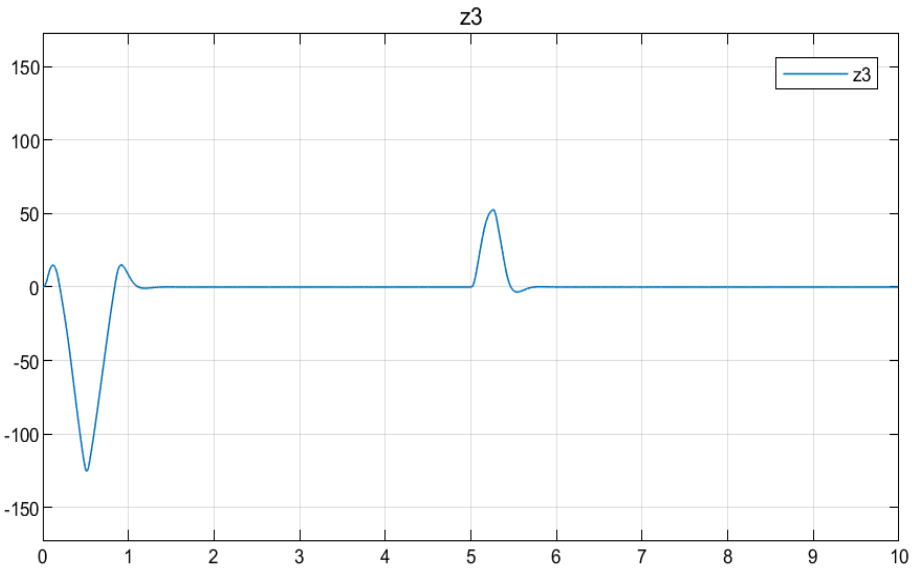


Figure 16: Estimated total disturbances by the ADRC

Table 2: Parameter values of the ADRC system

Parameters	Value
Tracking Differentiator (TD)	
hTD	0.03
rTD	10
Extended State Observer (ESO)	
beta01	100
beta02	300
beta03	1000
hESO	0.01
b0ESO	0.01
Nonlinear State Error Feedback (NLSEF)	
c	0.1
rNLSEF	20000
hNLSEF	0.001
k0	0
Disturbance Compensation	
b	1
b0Dc	1

6. Conclusions and Future work

The shift of feedback control paradigm design has been extensively studied over the past few decades. With a huge potential of disturbance rejection control method, this paradigm shift will play a critical role in the migration from the primitive PID controller to an ADRC. A PID controller passively responds to disturbances, causing zone oscillations. But on the other hand, ADRC actively rejects disturbances, delivering a smooth control system. In an ADRC, the unmeasurable states and total disturbances are actively compensated by the ESO present in the feedback loop. Whereas TD effectively acquire the continuous and differential form of signals from a measurement signal containing random noise, or from a discontinuous signal. It is easy to tune the parameters of TD due to its model independence property. The nonlinear feedback combination (NLSEF) of the two nonlinear functions play a significant role in this recently proposed control framework. Thus, considering the system parameters and model uncertainties, ADRC shows higher tolerance with merits of simple and intuitive model-free control design methods. Along with, it needs less energy in control in comparison to other control strategies.

This paper gives a detailed step by step design and mathematical modelling of the proposed HMI system developed for the research. After installation of the mechanical and electrical hardware components, Lagrange equations were derived based on the LMM of the system. An ADRC controller was designed and integrated with the haptic display platform Simulink model. Simulation studies were conducted to illustrate the efficiency and robustness of the ADRC controller.

However, up to date, debates and inconsistencies on the new paradigm concept still exist, and reports and validation on the successful integration of ADRC into the field of control theory are quite limited. This work pushes the realms of ADRC integration with actual physical plants for practical applications, and potentially opens up new realms constituting of new control laws formulated through experimentation.

Future research of this project will include the analysis of ADRC controller for a two-axis ball screw driving system dealing with time delay, and coupling between the two axes of the experiment. This analysis will be conducted with the modelling environment created on MATLAB/ Simulink.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

References

[1] S.N.F. Nahri, S. Du, B. Van Wyk, "Haptic System Interface Design and Modelling for Bilateral Teleoperation Systems," in 2020 International SAUPEC/RobMech/PRASA Conference, IEEE: 1–6, 2020, doi:10.1109/SAUPEC/RobMech/PRASA48453.2020.9041010.

[2] K. Abidi, Y. Yildiz, B.E. Korpe, "Explicit time-delay compensation in teleoperation: An adaptive control approach," International Journal of Robust and Nonlinear Control, **26**(15), 3388–3403, 2016, doi:10.1002/rnc.3513.

[3] P.F. Hokayem, M.W. Spong, "Bilateral teleoperation: An historical survey," Automatica, **42**(12), 2035–2057, 2006, doi:10.1016/j.automatica.2006.06.027.

[4] D.S. Nunes, P. Zhang, J. Sa Silva, "A Survey on Human-in-the-Loop Applications Towards an Internet of All," IEEE Communications Surveys & Tutorials, **17**(2), 944–965, 2015, doi:10.1109/COMST.2015.2398816.

[5] D.A. Lawrence, "Stability and transparency in bilateral teleoperation," IEEE Transactions on Robotics and Automation, **9**(5), 624–637, 1993, doi:10.1109/70.258054.

[6] L.Y. Ang K, Chong G, "PID control system analysis, design, and technology," IEEE Transactions on Control Systems Technology, **13**(4), 559–576, 2005, doi:10.1109/tcst.2005.847331.

[7] D. Sun, F. Naghdy, H. Du, "Wave-Variable-Based Passivity Control of Four-Channel Nonlinear Bilateral Teleoperation System Under Time Delays," IEEE/ASME Transactions on Mechatronics, **21**(1), 238–253, 2016, doi:10.1109/TMECH.2015.2442586.

[8] E. Nuño, L. Basañez, R. Ortega, "Passivity-based control for bilateral teleoperation: A tutorial," Automatica, **47**(3), 485–495, 2011, doi:10.1016/j.automatica.2011.01.004.

[9] A. Gupta, M. Kilchenman, O.' Malley, M.K. O'malley, "Disturbance-Observer-Based Force Estimation for Haptic Feedback MRI-compatible robotics View project Haptic Paddle View project Disturbance-Observer-Based Force Estimation for Haptic Feedback," Article in Journal of Dynamic Systems Measurement and Control, **133**(1), 2011, doi:10.1115/1.4001274.

[10] J. Han, "From PID to Active Disturbance Rejection Control," IEEE Transactions on Industrial Electronics, **56**(3), 900–906, 2009, doi:10.1109/TIE.2008.2011621.

[11] Z. Gao, "Active disturbance rejection control: a paradigm shift in feedback control system design," in 2006 American Control Conference, IEEE: **7**, 2006, doi:10.1109/ACC.2006.1656579.

[12] Y. Huang, W. Xue, "Active disturbance rejection control: Methodology and theoretical analysis," Elsevier, **53**(4), 963–976, 2010, doi:10.1016/j.isatra.2014.03.003.

[13] H. Feng, B.-Z. Guo, "Active disturbance rejection control: Old and new results," Annual Reviews in Control, **6**, 1–11, 2017, doi:10.1016/j.arcontrol.2017.05.003.

[14] H. Jin, W. Lan, Y. Chen, "Replacing PI Control With First-Order Linear ADRC," in 2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS), IEEE: 1097–1101, 2019, doi:10.1109/DDCLS.2019.8908981.

[15] J. Han, "A class of extended state observers for uncertain systems," Control and Decision, **10**, 85–88, 1995.
[16] J.Q. Han, "Auto Disturbance Rejection Controller and It's Applications," Control and Decision, **13**, 19–23, 1998, doi:NII Article ID (NAID) 10007201534.
[17] Z. Gao, Y. Huang, J. Han, "An alternative paradigm for control system design," in Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228), IEEE: 4578–4585, 2001, doi:10.1109/CDC.2001.980926.
[18] Z. Gao, "On the centrality of disturbance rejection in automatic control," ISA Trans, **53**(4), 850–857, 2014, doi:10.1016/j.isatra.2013.09.012.
[19] Q. Zheng and Z. Gao, "On practical applications of active disturbance rejection control," in Proceedings of the 29th Chinese control conference, 6095–6100, 2010, doi:INSPEC Accession Number: 11572423.
[20] G. Herbst, "A Simulative Study on Active Disturbance Rejection Control (ADRC) as a Control Tool for Practitioners," Mdpi.Com, **2**(4), 246–279, 2013, doi:10.3390/electronics2030246.
[21] A. Ortiz, S. Orozco, I. Zannatha, "ADRC controller for weightlifter Humanoid robot," in 2019 International Conference on Electronics, Communications and Computers (CONIELECOMP), IEEE: 41–46, 2019, doi:10.1109/CONIELECOMP.2019.8673147.
[22] Z.Q. W. Chenlu, C. Zengqiang, S. Qinglin, "Design of PID and ADRC based quadrotor helicopter control system," in 2016 Chinese Control and Decision Conference (CCDC), 5860–5865, 2016, doi:https://doi.org/10.1109/CCDC.2016.7532046.
[23] W.T. C. Fu, "Tuning of linear ADRC with known plant information," ISA Trans, **65**, 384–393, 2016, doi:doi:10.1109/access.2018.2805782.
[24] Z. Chen, Q. Zheng, Z. Gao, "Active Disturbance Rejection Control of Chemical Processes," in 2007 IEEE International Conference on Control Applications, IEEE: 855–861, 2007, doi:10.1109/CCA.2007.4389340.
[25] G. Niemeyer, J.-J.E. Slotine, "Stable adaptive teleoperation," IEEE Journal of Oceanic Engineering, **16**(1), 152–162, 1991, doi:10.1109/48.64895.
[26] Z. Chen, F. Huang, W. Sun, W. Song, "An Improved Wave-Variable Based Four-Channel Control Design in Bilateral Teleoperation System for Time-Delay Compensation," IEEE Access, **6**, 12848–12857, 2018, doi:10.1109/ACCESS.2018.2805782.
[27] C.P. J. H. Ryu, J. Artigas, "A passive bilateral control scheme for a teleoperator with time-varying communication delay," Mechatronics, **20**, 812–823, 2010, doi:doi:10.1016/j.mechatronics.2010.07.006.
[28] L. Sheng, U. Ahmad, Y. Ye, Y.-J. Pan, "A Time Domain Passivity Control Scheme for Bilateral Teleoperation," Electronics Mdpi.Com, **8**(3), 325, 2019, doi:10.3390/electronics8030325.
[29] X. Xu, B. Cizmeci, C. Schuwerk, E. Steinbach, "Model-Mediated Teleoperation: Toward Stable and Transparent Teleoperation Systems," IEEE Access, **4**, 425–449, 2016, doi:10.1109/ACCESS.2016.2517926.
[30] E.J. RodrIguez-Seda, Dongjun Lee, M.W. Spong, "Experimental Comparison Study of Control Architectures for Bilateral Teleoperators," IEEE Transactions on Robotics, **25**(6), 1304–1318, 2009, doi:10.1109/TRO.2009.2032964.
[31] S. Hirche, M. Buss, "Human-Oriented Control for Haptic Teleoperation," Proceedings of the IEEE, **100**(3), 623–647, 2012, doi:10.1109/JPROC.2011.2175150.
[32] L. Zhao, L. Liu, Y. Wang, H. Yang, "Active Disturbance Rejection Control for Teleoperation Systems with Actuator Saturation," Asian Journal of Control, **21**(2), 702–713, 2019, doi:10.1002/asjc.1767.
[33] Q. Wang, M. Ran, C. Dong, "On Finite-Time Stabilization of Active Disturbance Rejection Control for Uncertain Nonlinear Systems," Asian Journal of Control, **20**(1), 415–424, 2018, doi:10.1002/asjc.1558.
[34] C. Zhang, J. Yang, S. Li, N. Yang, C. Zhang, N. Yang, J. Yang, S. Li, "A generalized active disturbance rejection control method for nonlinear uncertain systems subject to additive disturbance," Nonlinear Dynamics, **83**(4), 2361–2372, 2016, doi:10.1007/s11071-015-2487-1.
[35] S. Frey, A. Dadalau, A. Verl, "Expedient modeling of ball screw feed drives," Producition Engineering, **6**(2), 205–211, 2012, doi:10.1007/s11740-012-0371-0.
[36] L. Luo, W. Zhang, Electromechanical Co-Simulation for Ball Screw Feed Drive System, IntechOpen, 2019, doi:10.5772/intechopen.80716.
[37] Y.D. YV Wen-bin, "Modeling and Simulation of an Active Disturbance Rejection Controller Based on Matlab/Simulink," International Journal of Research in Engineering and Science (IJRES), **3**(7), 62–69, 2015, doi:www.ijres.org.
[38] Z. Gao, Q. Zheng, L.Q. Gao, "On Validation of Extended State Observer Through Analysis and Experimentation ADRC: a new paradigm of the science of automatic control View project ADRC: Problem Solving View project On Validation of Extended State Observer Through Analysis and Experimentation," Article in Journal of Dynamic Systems Measurement and Control, **134**(2), 2012, doi:10.1115/1.4005364.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 5, 920-926 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Using Big Data Analytics to Predict Learner Attrition based on First Year Marks at a South African University

Gcobisile Matafeni[*,1], Ritesh Ajoodha[1]

[1]*School of Computer Science and Applied Mathematics. The University of the Witwatersrand, Johannesburg. 2050. South Africa.*

ARTICLE INFO

Article history:
Received: 17 August, 2020
Accepted: 21 September, 2020
Online: 12 October, 2020

Keywords:
Expert system
predicting student success
statistical models
learner attrition

ABSTRACT

Due to high failure rates many students end up spending unnecessary years struggling to qualify and subsequently accumulate unnecessary debt. In this paper, our principal contribution is to provide an expert system that statistically predicts the success of a first year student in an undergraduate Science programme given only academic merit in their subject matter. Over the past decades, much work has been done in the field of predicting student success in first year computer science and in other first year courses. Historically, other authors focused on using linear statistical models to predict student success. These models had limitations as the prediction was designed for inference as compared to machine learning techniques. This paper presents an approach of using the naïve Bayes classifier, support vector machines and decision trees as models that can be used to predict the completion of an undergraduate science degree. This was done by firstly training the classifiers and then testing them. The support vector machine achieved the best accuracy (87%) in predicting the completion of a science degree based only on first year marks, this was followed by the naïve Bayes model (86.36%) and the decision tree (65.62%) came last.

1 Introduction

The prevalence of high failure rates in South African universities results in students spending unnecessary years struggling to qualify and subsequently accumulate unnecessary debt. A view that is generally agreed upon is that acceptance into a university programme often proves to be life changing. The prospect of obtaining a degree is sometimes a promise of higher income and improves ones standard of living [1]. The idea of students having the ability to know after their first year of study which courses to take in order to maximize their chances of succeeding is helpful. Historically, institutions of higher learning have been struggling to improve their throughput rates over the years. Since the dawn of the democratic dispensation, enrolment rates in higher education institutions have sky rocketed while the dropout rates have increased significantly [2]. To mitigate drop-out rates and improve success, this paper will attempt to build a model for calculating the probability of a student completing an undergraduate degree at Wits University in order to promote early interventions focused on student success.

Some useful features to recognize the success of a student in first year computer science include using the Scholastic Aptitude Test (SAT) and the American College Testing (ACT) scores across different subjects [3, 4]. The approach in this paper will have the ability to calculate the alternative streamlines that might align a student with better options, so that they can reconsider their academic standing. In [3], the authors used statistical analysis in order to build a linear model that described the correlations using all linear combinations of the dependent variables. [5] made use of a step-wise discriminant analysis to identify all significant factors in a study group of 87 freshman computer science majors. The above papers demonstrate the methodologies that rely on using statistical models to perform the classification task. The purpose of this study is to explore the correlation between first year academic results and the probability of completing a Science degree at a South African university.

In this paper we trained the following three machine learning classification models: naïve Bayes classifiers, support vector machines, and decision trees to predict student attrition given their first year marks. The support vector machine achieved the best accuracy (87%) in predicting the completion of a science degree based only on first year marks, this was followed by the naïve Bayes model (86.36%) and the decision tree (65.62%) came last.

Our principal contribution from this paper, to stand along other contributions of this kind [6]–[8], is to provide a predictive model that statistically predicts the success of a first year student to complete their undergraduate Science degree given their first year marks. This will enable and promote early interventions to encourage stu-

[*]Corresponding Author: Gcobisile Matafeni, The University of the Witwatersrand, Johannesburg, +27 65 9940 231 & matafenigcobisile@gmail.com

dent success.

This paper is structured as follows: section II reviews the related work; section III discusses the research methodology that was used to conduct the relevant experiments; section IV presents the qualitative results of the experiments and a discussion. Lastly, section v provides a summary of this paper, it also suggests some future work that can be explored.

2 Related Work

In [9], the authors attempt to deduce student attrition at a South African higher-education institution with the aim of identifying students who are likely to be in need of academic support so that a focus could be provided on improving their academic performance. The techniques used in the above paper are; Decision Trees, K-Star, naïve Bayes, Support Vector Machines, Random Forests, and Linear Logistic Regression. In [10], the authors attempt to provide a data-driven solution to the data-congested environment of attributes related to student success and contribute towards preventing the increased dropout rates at South African higher education institutions. The techniques used in the above paper are; Decision Trees, K-Star, naïve Bayes, Support Vector Machines, feed-forward neural networks, and linear regression models. In this research we are concerned with predicting the completion of an undergraduate science degree based only on first year marks. This research will be adapted from the work in the literature survey which focused on predicting student success in first year computer science by using marks from high school. This section provides the necessary background and related work that was used during the research project.

2.1 Predicting success at first year computer science

In this section, we look at different approaches that were used to predict the performance of first year students who majored in computer science. We first evaluate how multiple factors were used for this purpose, we then move into how mathematics as a single subject was used for prediction purposes. Lastly, we look at how English as another single subject was used for prediction purposes.

2.1.1 Predicting success using multiple factors

There are many approaches that can be used to look at what affects student performance in first year courses. One approach is to use multiple variables to compute the probability of succeeding in a freshmen major. In [3], 13 variables were used and these were independent to each other, these included the American College Testing (ACT) scores for mathematics, the American College Testing (ACT) scores for English, the American College Testing (ACT) scores for natural science and the student's class rank to name a few. The American College Testing (ACT) are assessments that are used to measure the college readiness of high school students in America. The authors focused their study on a sample size of 269 students. For every variable; the mean, standard deviation, minimum and the maximum were calculated. Utilizing the correlations that were developed, a linear model was used to fit the data and the results indicated that it was possible to calculate the probability of a pass or a fail in an introductory computer science course. In [4], the authors used multiple factors to predict the performance in a computer science major and some of their work which focused on ACT English and other scores was extended in [3]. In [4], the authors consider the SAT scores, the sex of the student and their high school grades. By combining all of these, they developed a linear discriminant as a function to perform classification. In a data sample of 256 students, they were able to successfully classify 175 students which is 68.4% of the data into the correct group.

In [5], the author studied 10 factors and used these to find their relationship to student success in a first year computer science major. The ACT scores were used similarly to what [3] used in their variables. According to [5], the placement factors that reduced the failure rate in first year were the ACT English scores and the UTM mathematics placement scores, with the former scores being the best predictor. To identify the factors that contribute to the forecasting of success, a discriminate function was used. In [5], the author found that the other factors were not significant to be used as predictors. [4] found SAT mathematics and verbal scores as best predictors but did not indicate which one was the best predictor. In [11], the author builds a model that forecasts the success in a programming course, with the hope of counseling students to make informed decisions. On top of using past academic achievement as used in [3]–[5], the author includes certain cognitive skills and personality traits. The study focused on a sample size of 120 students that was randomly selected from a population of 600 students. The variables that were used in this paper are both independent and dependent, and they are 21 in total. This is significantly more variables in comparison to the number of variables that were used in other papers; [3]–[5]. The multiple regression equation that was developed was able to classify 61 students out of 79 (77.2%). The approach of using multiple factors to build a model for predicting the success of students worked in the papers mentioned above but there are other approaches.

2.1.2 Predicting success using Mathematics

In [12], the author uses the IBM Aptitude Test for Programmer Personnel (ATTP) scores to classify the students who are doing different computing courses, one based on COBOL and the other based on FORTRAN. The study focused on 46 students who had written the ATTPs. A positive correlation between the FORTRAN course and the arithmetical component of the ATTP was found, another correlation was found between the letter series and the COBOL course. The only factor that was used for forecasting the performance in an introductory computer course were the ATTP scores which is a different approach from using a combination of variables. In a paper that focuses on a commerce degree, the author [13] argues for the use of mathematics as a single variable rather than the matric aggregate to predict the performance in first year. The study focused on results that spanned a period of 4 years. The approach used was to take the individual mathematics scores and compare them to pass rates in first year. According to [13], a good mark in matric mathematics is the best predictor of first year success compared to the matric aggregate scores.

2.1.3 Predicting success using English

Similar to the approaches used in [12] and [13], in [14] the authors use a single factor to predict the success in first year computer sci-

ence. The paper studies the correlation between success in English and the actual performance in first year. Analysis took place by use of qualitative data from a survey and an in-depth quantitative look at matric results. Results show that language scores from matric results are better predictors of success. In contrast, a recent publication by [13] argues about an important relationship between mathematics results in matric and the performance in a commerce degree. The argument is of great importance as it bases its findings on the same matric results as in [14]. Also, in [13] the author did not consider language in their study and according to [14] there is a possibility that it might have been a factor. As alluded to earlier, [4] states that it is inappropriate to only use a single matric subject as a predictor of success for a first year computer science major. A holistic approach is to use the matric aggregate as a predictor. It is worthy revisiting the findings of [5], the research found that amongst the several variables that were used in the study, the best single predictor for success were the ACT English scores. The overall best predictor for success in computer science was the combination of UTM mathematics and ACT English placement scores. The predictor that was developed in [5] was able to to successfully reduce the failure rate in computer science the following year from 28% to 18% as stated in [14].

3 Research Methodology

Predicting the success of a first year computer science student is affected by many variables as demonstrated in [3]. The background to the problem has been presented in the previous section and it also lays out the different approaches that have been used by other authors who focused on first year performance by using previous student grades. It must be noted that the previous authors who are surveyed in the literature had their focus on predicting success in first year computer science and other first year courses. This research will be taking this work further by trying to predict the completion of a science degree based only on first year marks. The naïve Bayes classifier, support vector machines and decision trees can be used to build classification models for this particular problem area as presented in section II. These classification models will use the first year marks as training data.

3.1 Data and Preprocessing

The study participants were learners who studied at a South African higher-education institution. The study ethics application has been approved by the Human Research Ethics Committee of the University (Non-Medical). The ethics application addresses key ethical issues of protecting the identity of the learners involved in the study and ensuring the security of data. The clearance certificate protocol number is H19/03/02. The dataset had a total of 216 features with 8557 observations. The features were selected according to their relevance in order to serve our aims and objectives. The data was received in a spreadsheet format and this had to be converted into a format that Weka can work with. Weka can work with comma separated values (csv) or attribute-relation format files (arff), the spreadsheet was converted into a csv file. The spreadsheet was read into an iPython notebook which uses Python and a package called pandas to work with stored data, this platform makes it easy to perform exploratory data analysis. The student numbers were anonymized in order to protect the students identity. The dataset contained students who did their year of study 1,year of study 2 and year of study 3 and in our case since we were working with year of study 1 so we dropped the other records (YOS2 and YOS3). After dropping the unnecessary rows, the rows that were left contained the subject and its mark for each student, say student x, the first 6 rows would be student x with the 6 subjects they did in first year with the corresponding marks and Progression Outcome Type. Since each student had more than one row identifying them we performed a transpose on the given table so that each record is one unique student with the columns having the subject matters and marks. The last column was the class label which was either a Yes or a No for the Progression Outcome Type. There was 216 features in total after the data was prepared correctly. The data was then saved onto a Weka readable format in order to run the supervised learning algorithms.

3.2 Classification Models

After the implementation of the classifiers the focus shifted to evolving the models in order to improve their accuracy. This was done by performing feature extraction and selection techniques. A 10-fold cross validation scheme was used.

3.2.1 Naïve Bayes

The naïve Bayes pre-defines a finite set of mutually exclusive classes and assumes that all of the features are conditionally independent given the class label of each instance [15], [16]

3.2.2 Support Vector Machine

SVMs are a type of supervised learning algorithm that can be applied to both regression and classification problems. In SVMs, the required hyperplanes and weights are learned during training, these have the required optimal plane and correctly chosen support vectors. The default kernel used by Weka is the RBF with gamma equal 1/k. [17], [18]

3.2.3 Decision Tree

The decision tree is based on the J48 algorithm that is also known as Iterative Dichotomiser 3, in this research this will construct a tree from top down using the ID3 algorithm. [19]

After performing the feature extraction techniques we chose one supervised learning algorithm in order to see if the accuracies were comparable to the generalized models. Training was done on the Correlation, Information Gain and Wrapper Evaluation attributes. The tailored degrees have their specified subjects so for each category the training was performed as well. After learning the parameters of the classifiers, these will be used to either confirm or reject the hypothesis through testing.

4 Experimental Results

In the previous section we discussed how the hypothesis of the research project would be tested and in this section we look at the performance of the machine learning techniques that were mentioned in section III. Firstly, we do simple plots of the data distribution in order to have a better picture of our data and then we evaluate the performance of the 3 main techniques. This evaluation occurs on the full dataset that has 216 features and the results will be subsequently recorded. We then perform feature extraction using the techniques that were also mentioned in section III, we use one of them (of the three algorithms) in order to observe if there's any improvement on the model after feature extraction. The last batch of experiments involve training models for tailored degrees in the Faculty of Science, this is done due to the fact that in the initial models we generalized students in the Faculty. The tailored degrees fall under four categories which are; Physical Sciences, Mathematical Sciences, Biological and Life Sciences and the last one is Earth Sciences. The algorithms were ran on Weka (Waikato Environment for Knowledge Analysis) which is tool developed at the University of Waikato, New Zealand [20].

4.1 Graphs

The stacked plot in Figure 1 represents the distribution of the Progression Outcome Type, and this is referred to as the class label. The color blue represents students that have successfully completed their degrees in record time (3 years) and red represents the class of students that have not completed their degrees. The total number of unique students is 8557 for a period of 7 years.

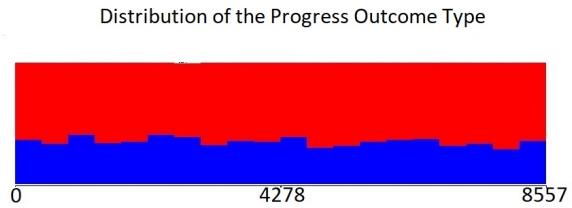


Figure 1: Stacked plot of the distribution of Progress Outcome Type

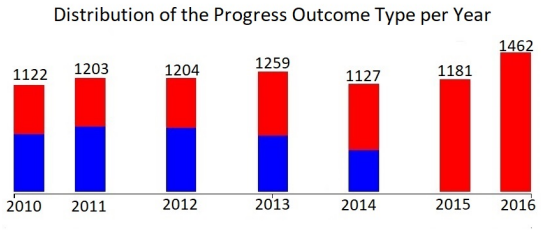


Figure 2: Stacked bar plot of the Progress Outcome Type per year

Figure 2 represents a stacked bar plot that compares the Progress Outcome Type for each year from 2010 until 2016. Blue represents students that have completed their degrees in record time and the color red represents students that have not completed their degrees. The last two years, 2015 and 2016 have students that have not completed their degrees which is a strange phenomenon because we would expect that in every 3 year cycle there would be students that complete their degrees. Students that enrolled in 2014 should have completed their degrees in 2016, and the same applies to students that enrolled in 2013, these students should have completed their degrees in 2015. This strange observation in Figure 2 can be accounted to the second phase of the research methodology, we might have dropped students who completed their degrees in 2015 and 2016 by mistake. This might have been caused by the format in which we received the dataset.

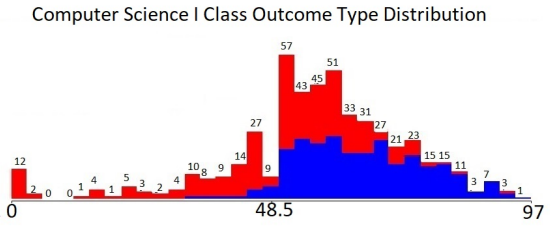


Figure 3: A plot that represents the distribution of Computer Science I marks

Figure 3, we have a stacked plot of the distribution of marks in Computer Science I. The plot resembles a normal distribution (Gaussian) which is what we expected for a set of marks for any given course. There are outliers in the Computer Science I marks and these are located at the far left in the plot.

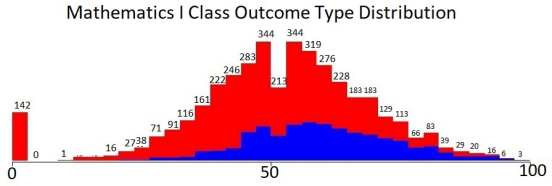


Figure 4: A plot that represents the distribution of Mathematics I marks

Figure 4 illustrates a stacked plot of the distribution of marks in Mathematics I major. The plot resembles a normal distribution (Gaussian) more than the Figure 3, this is because we have more data points than before. There are also outliers to the far left of the stacked plot. Given that there is 216 possible courses that are features of the data, the same phenomenon of normally distributed marks can be generalized to these but we plot only 2 diagrams for the purpose of illustration.

Figure 5a represents the results of the 3 different machine learning algorithms that were used to test the hypothesis. The support vector machine outperformed the naïve Bayes and the decision tree. The results of the decision tree are the worst from the first two algorithms as an accuracy of 65.62% cannot be deemed as reasonable given that the support vector machine achieved an accuracy of 87%.

Figure 5b, Figure 5c and Figure 5d represent the Confusion Matrices for the respective algorithms that are in Figure 5a. These tables describe the performance of each algorithm and this is another way of representing the accuracy of each algorithm.

Algorithm	Correctly Classified	Incorrectly Class.
naïve Bayes	7391 [**86.3636**%]	1167 [**13.6364**%]
Support Vector Machine	7446 [**87.0063**%]	1112 [**12.9937**%]
Decision Tree	5616 [**65.9663**%]	2942 [**34.3772**%]

(a) A table that compares the accuracies of the algorithms

P	F	
2622	363	P = Pass
804	4769	F = Fail

(b) Confusion Matrix of the naïve Bayes

P	F	
2380	605	P = Pass
507	5066	F = Fail

(c) Confusion Matrix of the support vector machine

P	F	
2979	6	P = Pass
2936	2637	F = Fail

(d) Confusion Matrix of the decision trees

Figure 5: Tabulated results of the different algorithms

Extraction Technique	Correctly Classified	Incorrectly Class.
Correlation	7393 [**86.387**%]	1166 [**13.613**%]
Information Gain	7387 [**86.3169**%]	1171 [**13.6831**%]
Wrapper Evaluation	6968 [**81.4209**%]	1590 [**18.5791**%]

(a) A table that compares the accuracies of the algorithms

P	F	
2696	289	P = Pass
876	4697	F = Fail

(b) Confusion Matrix of the naïvee Bayes (Correlation)

P	F	
2640	345	P = Pass
826	4747	F = Fail

(c) Confusion Matrix of the naïve Bayes (Information Gain)

P	F	
2437	548	P = Pass
1042	4531	F = Fail

(d) Confusion Matrix of the naïve Bayes (Wrapper Evaluator)

Figure 6: Tabulated results of the feature extraction techniques

Degree	Correctly Classified	Incorrectly Class.
Physical Sciences	7392 [**86.3753**%]	1166 [**13.6247**%]
Mathematical Sciences	7198 [**86.3169**%]	1360 [**15.8916**%]
Biological and Life Sciences	7357 [**81.4209**%]	1201 [**14.0337**%]
Earth Sciences	7396 [**81.4209**%]	1162 [**13.5779**%]

(a) A table that compares the accuracies for the different degree fields

P	F	
2753	232	P = Pass
934	4639	F = Fail

(b) Confusion Matrix of the naïve Bayes (Physical Sciences)

P	F	
2771	214	P = Pass
1146	4427	F = Fail

(c) Confusion Matrix of the naïve Bayes (Mathematical Sciences)

P	F	
2784	201	P = Pass
1000	4573	F = Fail

(d) Confusion Matrix of the naïve Bayes (Earth Sciences)

P	F	
2746	239	P = Pass
923	4650	F = Fail

(e) Confusion Matrix of the naïve Bayes (Wrapper Evaluator)

Figure 7: Tabulated results of the different degree fields

In Figure 8 we have a bar plot that compares the accuracy of the three algorithms. The legend of the plot indicates that the red color represents the category of the correctly classified instances and the cyan represents the incorrectly classified instances.

4.2 Results with Feature Extraction using Naïve Bayes

Figure 6a represents the results of the naïve Bayes classifier that was used on the 3 feature extraction techniques namely, correlation, information gain, and wrapper evaluation.

The Figure6b, 6c and 6d represent the Confusion Matrices for the respective algorithms that are in Table 6a.

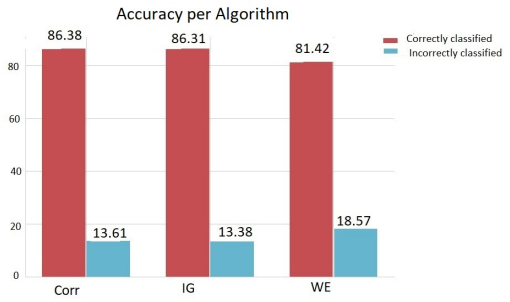


Figure 8: Bar plot

In Figure 8 we have a bar plot that compares the accuracy of the three feature extraction techniques.

4.3 Predicting the completion of a field specific Bachelor of Science degree based on first year marks

Figure 7a represents the results of the field specific degrees in the Faculty of Science. The naïve Bayes classifier was used for each of the 4 fields. The correctly classified instances for each of the fields have comparable accuracies and these are also comparable to the generalized case which uses the 216 features. Given that some students when they start their first year's of study they pursue field specific degrees, the model was also tested for field specific degrees in order to see whether we could test our hypothesis for these cases.

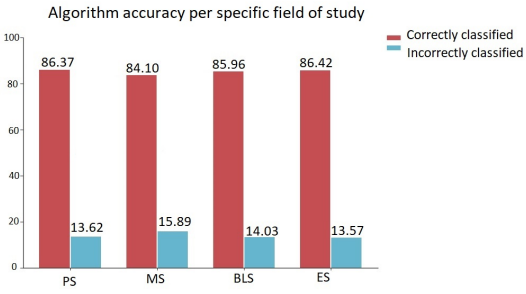


Figure 9: Bar Plot

In Figure 9 we have a bar plot that compares the accuracy of the the naïve Bayes when used on each of the four field specific degrees in the Faculty of Science. The graphical representation gives a better picture than Figure 7a and the Confusion Matrix tables. Here we are able to visually analyze the accuracy of each field specific degree when the naïve Bayes classifier is applied.

5 Conclusion

Our primary objective of this research was to build a model that can successfully predict the completion of a student's Science degree based only on their first year marks. To achieve this, we first conducted research on different machine learning techniques, analyzed their advantages and limitations of each. According to the literature survey, we selected the naïve Bayes classifier, support vector machine and the decision tree. These have reasonable performances in theory that is why they were selected to be used in this research. Data preprocessing was used in order to package the data into the correct format and the feature extraction techniques were used in order to minimize the effect of over-fitting. Following a certain methodology, we conducted experiments on the selected models and on their variations. The methodology we performed can be easily extended to build other models that can be used for prediction purposes.

To evaluate the predictive performance of the models, this research applied a combination of features to the given models and their accuracies were determined. The computation time of each model is comparable except to the support vector machine that took some time do build. The experiment results show that our hypothesis is true and the support vector machine is considered as the most efficient model that produces results of reasonable accuracy. This was followed by the naïve Bayes and the decision tree had the worst performance. The naïve Bayes classifier was chosen purely on discretion when it was used for the feature extraction techniques and the field specific degrees. The support vector machine could have been chosen as well for this purpose as it had a comparable accuracy. The results of the feature extraction techniques and the field specific degrees were comparable to the general case.

A lot of factors were not taken into account when building the necessary models, i.e. the social circumstances of a particular student, whether a student is on financial aid, personal preferences and this affects the future work of the research. It can also be argued that the gender of the student might also affect the performance of a particular student, so this can improve future models.

It would be interesting to apply these models on a dataset with more variables that describe the learners academic trajectory. It would also be interesting to compare the predictive across models trained from various South African institutions in order to discover if the same features drive success and failure at difference institutions. Further work could explore optimising the hyper-parameters used in this study to optimise on accuracy. Deep neural networks or other probabilistic graphical models, such as Bayesian networks, could be used to provide better explanatory power of the influence that these features play in determining student success. These models do not make a lot of assumptions about our data and they can also be used to achieve the research objectives.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgement This work is based on the research supported in part by the National Research Foundation of South Africa (Grant number: 121835).

References

[1] L. Thurow, "Education and economic equality," in In The Public Interest, 66, 1972.

[2] I. Scott, "A proposal for undergraduate curriculum reform in South Africa: The case for a flexible curriculum structure," in In Council on Higher Education, 2013.

[3] D. Butcher, W. Muth, "Predicting success in an introductory computer science course," in Communications of the ACM, 263–268, ACM, 1985.

[4] P. Campbell, G. G. McCabe, "Predicting the success of freshmen in a computer science major," in Oxford: Clarendon, 1102–1113, 1984.

[5] E. Gathers, "Screening freshmen computer science majors," in ACM SIGCSE Bulletin, 44–48, ACM, 1986.

[6] N. Ndou, R. Ajoodha, A. Jadhav, "A Case Study: Educational Data-mining to Determine Student Success at Higher Education Institutions," in The International Multidisciplinary Information Technology and Engineering Conference, 2020.

[7] N. Mngadi, R. Ajoodha, A. Jadhav, "A Conceptual Model to Identify Vulnerable Undergraduate Learners at Higher-Education Institutions," in The International Multidisciplinary Information Technology and Engineering Conference, 2020.

[8] N. Philippou, R. Ajoodha, A. Jadhav, "Using Machine Learning Techniques and Matric Grades to Predict the Success of First Year University Students," in The International Multidisciplinary Information Technology and Engineering Conference, 2020.

[9] R. Ajoodha, A. Jadhav, S. Dukhan, "Forecasting Learner Attrition for Student Success at a South African University," in In Conference of the South African Institute of Computer Scientists and Information Technologists 2020 (SAICSIT '20), September 14-16, 2020, Cape Town, South Africa. ACM, New York, NY, USA, 10 pages., ACM, 2020, doi:https://doi.org/10.1145/3410886.3410973.

[10] R. Ajoodha, S. Dukhan, A. Jadhav, "Data-driven Student Support for Academic Success by Developing Student Skill Profiles," in International Multidisciplinary Information Technology and Engineering Conference. ISBN: 978-1-7281-9519-9., IEEE, 2020.

[11] T. Hostetler, "Predicting student success in an introductory programming course," in ACM SIGCSE Bulletin, 40–43, 1983.

[12] C. Capstick, J. Gordon, A. Salvadori, "Predicting performance by university students in introductory computing courses," in ACM SIGCSE Bulletin, 21–29, 1975.

[13] D. Tewari, "Is matric math a good predictor of student's performance in the first year of university degree? A case study of faculty of management studies, University of Kwazulu-Natal, South Africa," in In International Journal of Science Education, 233–237, 2014.

[14] S. Rauchus, B. Rosman, G. Konidaris, I. Sanders, "Language performance at high school and success in first year computer science," in ACM SIGCSE Bulletin, 2927–2948, 2006.

[15] Z. Ghahrami, "An Introduction to Hidden Markov Models and Bayesian Networks," in Journal of Pattern Recognition and Artificial Intelligence, 9–15, 2001.

[16] D. Pham, G. Ruz, "Unsupervised training of bayesian networks for data clustering," in In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 2927–2948, 2009.

[17] D. Isa, L. Lee, V. Kallimani, R. Rajkumar, "Text document preprocessing with the bayes formula for classification using the support vector machine," in IEEE Transactions on Knowledge and Data engineering, 1265–1272, 2008.

[18] D. Boswell, "Introduction to Support Vector Machines," Carlifonia Institute of Technology, 2002, doi:http://www.work.caltech.edu/~boswell/IntroToSVM.pdf.

[19] D. Isa, L. Lee, V. Kallimani, R. R. Rajkumar, "Decision Tree Learning," in In Machine Learning, 52–57, 1997.

[20] R. Remco, E. Bouckaert, M. Frank, "Weka Manual," in In Weka Manual for version 3.6.9, University of Waikato, New Zealand, 2008.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 951-958 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Investigating the Optical Behavior of Single/Multi-Dimensional Photonic Crystal Structures for Photovoltaic Applications

Gehad Ali Alsayed, Zahraa Ismail, Sameh O. Abdellatif *

FabLab, Centre for Emerging Learning Technologies (CELT) and Electrical Engineering Department, Faculty of Engineering, the British University in Egypt (BUE), Cairo, 11387, Egypt

ARTICLE INFO

Article history:
Received: 28 September, 2020
Accepted: 25 November, 2020
Online: 08 December, 2020

Keywords:
Third generation solar cells
Light trapping structures
Optical modeling
Distributed Bragg reflectors
Optimized Photonic crystals
Experimental fabrication
Optical Characterization

ABSTRACT

The results investigated in this work are toward the optimization of the photonic crystal structures in 1D and 2D scale. One-dimensional distributed Bragg reflectors (DBRs) have demonstrated substantial potential in various optoelectronic applications, due to the observed tunable optical band-gap. Herein, the use of DBRs in light trapping solar cells was simulated and validated, representing its effect as a back reflector structure. In terms of the layer thickness, material selection and number of layer, the optimized DBR structure was modeled and evaluated with respect to previously published numerical and experimental data. The proposed model is capable of designing photonic crystal structures with tunable band-gap varies from 400 nm to 700 nm while controlling the pass-band in both Visible and Near Infra-red regions. On the other hand, 2D grating structure has been simulated where the transmission spectra under various design dimensions have been investigated. Finally, thin film deposition is utilized for experimental validation to our proposed optical model.

1. Introduction

This paper is an extension of the work originally presented at the IEEE Conference on Power Electronics and Renewable Energy (CPERE) in 2019 [1].

Nowadays, new techniques of solar cells have shown a perfect solution to provide low cost alternative compared with conventional methods that used in solar cells [2-10]. These methods depend on using low cost active layer with less than 1 μm thickness and also minimizing the fabrication cost [2, 11-14]. By observing the literature, the reported structures showed low absorption in the active layer [15]. Based on that, the researchers began to find method to avoid the limitation in the absorption layer by using light trapping structures [8, 16, 17] (cf. Figure1). The light trapping structures can be divided into two main classifications, either by function or composition of the structure. The composition of the light trapping structure can be in the form of photonic crystal (PCs) [18] or plasmonics structures [1, 16, 17, 19-23]. The fabrication of PCs structures depends on the contrast between dielectric materials and semiconductor materials (refractive index contrast) which influences the optical properties of the structure [24]. This paper will discuss 1D PCs [25] as well as 2D grating structure for solar cells utilization

Additionally, other new generations of solar cells, namely dye sensitized solar cells and perovskites solar cells, showed enhanced performance with embedding two-dimensional colloidal crystal structures [26-29]. Such colloidal crystal structures are not only acting to enhance the photon capturing of the active layer of solar cells, but can be integrated for decorative purposes [29]. Over and above, light trapping techniques can have a positive role while harvesting diffused light in underwater applications [30, 31]

Light trapping structures can be applied in new technologies of solar cells, by trapping the light in the active layer through embedding nano-structures either above or below the active layer [8]. The technique of applying light trapping in the active layer depends on minimizing the reflected light in the upper layer by using anti-reflecting coating (ARC) while maximizing the reflected light in the back layer. The angle of reflection in the back reflector must exceed the critical angle, targeting total internal reflection in the active layer, which attributes to the absorption [8] (see Figure 1).

Additionally, other new generations of solar cells, namely dye sensitized solar cells and perovskites solar cells, showed enhanced

*Corresponding Author: Sameh O. Abdellatif, British University in Egypt (BUE), Cairo, 11387, Egypt, +201222194077, sameh.osama@bue.edu.eg

performance with embedding two-dimensional colloidal crystal structures [26-29]. Such colloidal crystal structures are not only acting to enhance the photon capturing of the active layer of solar cells, but can be integrated for decorative purposes [29]. Over and above, light trapping techniques can have a positive role while harvesting diffused light in underwater applications [30, 31].

The demonstrated optical model introduces an advanced optimization tool for 1D/2D PCs, seeking for maximum absorption in the active layer utilized in new generation of solar cells. Using Matlab scripting, the developed tool was essentially deployed and proved using a free optical simulator platform called "Open-Filter" [32], while absorption profiles were simulated numerically using Comsol multi-physics. The optimization procedure was applied on particular parameters such as the number of layers constricting the PC structure, the difference in refractive index and the layer thickness. The output results showed complete consistency with the literature data previously measured [20] and simulated [8, 17].

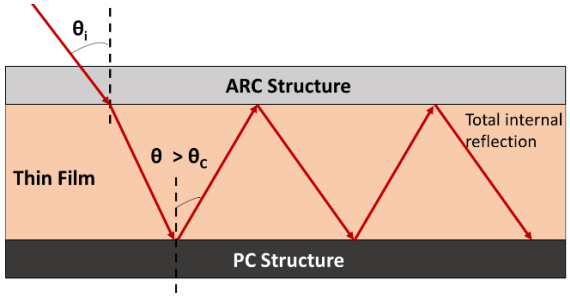


Figure 1: Schematic diagram of sandwiched light trapping structure for solar cells.

2. Optimization Model

2.1. Analytical modeling

The thin film optical modeling represented by characteristics matrix analytical method [33] is utilized in this paper. This method depends on Maxwell's equations which describe by electric field that cross continuously from medium to another. When the incidence of layer was given by electric and magnetic field, the field of both electric and magnetic can derived by matrix that given above [33]:

$$\begin{bmatrix} E_1 \\ H_1 \end{bmatrix} = M \begin{bmatrix} E_2 \\ H_2 \end{bmatrix} \quad (1)$$

where the intensity of electric and magnetic fields in the input medium are E_1 and H_1, while the electric and magnetic fields in the output medium are E_2 and H_2. Herein, M is given by:

$$M = \begin{bmatrix} \cos k_o h & i \sin k_o h/\gamma_1 \\ \gamma_1 i \sin k_o h & \cos k_o h \end{bmatrix} \quad (2)$$

and γ_1 is given by:

$$\gamma_1 = \sqrt{\frac{\varepsilon_o}{\mu_o}} \frac{n_1}{\cos \theta_i} \quad (3)$$

where h is thickness of thin film, k_o represented as propagation coefficient ($k_o = 2\pi/\lambda$), ε_o and μ_o are represented as air permittivity and permeability, refractive index of material called n_1, angle of incidence called θ_i.

A concatenation of layers can be referred to as a system matrix, represented by the multiplication of individual layer matrices. This matrix used to calculate the overall coefficients of reflection and transmission. Open-filters [32] is used as open source of analytic thin film optical modeling where we can validate our characterization matrix results.

Distributed Bragg Reflector (DBR) is a one-dimensional structure of simple photonic crystal. Optical band-gap is formed for Distributed Bragg Reflector as observed by solving Maxwell's equation (more detailed analytical solution for 1D and 2D photonic crystals can be accessed in [16]). The angle of incident and polarization should directly affect the width of band-gap. By incrementing the incidence angle to oblique, the TE modes band-gap becomes wider, while the band-gap of TM mode becomes narrower. This finding leads us to the independence of the band-gap from the polarizations and angle of incidence, as presented later in the discussion section. The photonic band-gap would exist, when contrast of dielectric occurred. In this manner, DBR could be shaped by layers of silicon with thickness 550 nm (n = 3.48) and SiO_2 of thickness 550 nm (n =1.46). The Lorentzian-Drude (LD) model introduced in [9] is used, which explain the absorption and dispersion of silicon as a wavelength dependent function.

2.2. Material modeling

The coefficient of dispersion should be calculated to modal dispersive and absorptive materials. Herein, Lorentzian-Drude (LD) model through FDTD simulations are enrolled [34]. The LD modal divided into two models. The Drude model is defined as the first model that, due to intra-band transitions, directly affects background absorption and dispersion. The second modal that represents dispersive media by multiplying resonant frequencies is the Lorentz model.

The model of semi-quantum reflects the electronic inter-band transitions. A material using the LD model's complex permittivity could be expressed by:

$$\varepsilon(\omega) = \varepsilon_{\bowtie} + \sum_n \frac{\sigma_n \omega_n^2}{\omega_n^2 - \omega^2 - i\omega\Gamma_n} \quad (4)$$

where $\varepsilon_{\bowtie}$ is the instantaneous dielectric response (DC), Γ_n and ω_n are the LD coefficients and σ_n is a position function that defines the power of the n^{th} resonance. Figure 2 shows the optical model for the DBR structure to be optimized. The optimization process is considering the number of DBR layers, reflective index (n_1– n_2), and the thickness of the layer (h_1, h_2).

2.3. Structure optical modeling using FDTD

The finite difference time-domain (FDTD) algorithm is used to make simulation of light interaction with various structure of layers [35]. The (FDTD) algorithm is considered as a standard numerical technique in classical electromagnetism. To simulate (FDTD) algorithm, open source algorithm is used called MEEP [36].The simulated volume is reflected on the required simulation time and RAM that required to make simulation on MEEP. A runtime of nearly 10 ns per pixel per time step of time is needed with RAM of almost 100 bytes per pixel [37].

MEEP was used to implement simulation of reflection and transmission of light by using different structure of light trapping. 10 μm × 10 μm is volume that used in simulation with unit step and resolution of 10 nm and 10 respectively. The wavelength is set as 550 nm to simulate sun irradiance for TM and TE modes. Figure 3, shows the modal with two detectors integrated for the calculation of transmitted and reflected field.

The maximum intensity of AM1.5G is determined to be 1000 W/m^2. Two identical Gaussian light sources with a width from 200 nm to 1500 nm and a mean wavelength of 550 nm, were utilized to represent TE and TM modes. It is important to model and integrate in MEEP the wavelength dependent variation in the real and imaginary parts of the refractive index for different materials used. Here, we integrate the published LD model to illustrate the complex refractive indices using the Lorentzian-Drude (LD) coefficients for the chosen materials with specific focus on Si as a semiconductor material. The compatibility of the LD model with MEEP is considered as the main reason for selecting it [38].

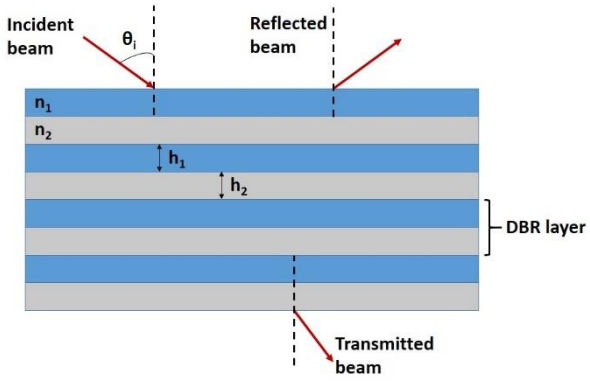


Figure 2: Schematic diagram of the structure of the DBR under test.

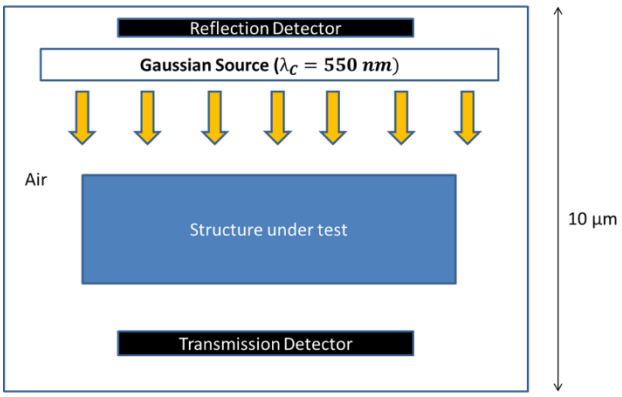


Figure 3: Schematic diagram used for the calculation of transmission and reflection spectra for the MEEP model.

2.4. 2D optical modeling using Comsol Multi-physics

Using the commercial simulator Comsol Multiphysics 5.0 [37], optoelectronic modeling is carried out. Comsol utilize finite element method (FEM) solver such that a tetrahedral mesh is selected based on the minimum feature length in the geometry under investigation. Various 2D optical structures devices could be simulated by optical and electromagnetic libraries and are modeled by the continuity, Poisson, drift-diffusion and Maxwell's equations assuming 2D solution. The targeted solar cells with light trapping structure are described as a p-i-n junction [39]. Consideration is also provided to the bulk, surface recombination, and contact effects. The primary governing equations are given by [8]:

$$-\nabla.\left(\frac{\varepsilon_o \varepsilon_r}{q}\nabla V\right) = -n + p + N_D(1 - f_D) - N_A(f_A) + \sum_j N_{tj}(\delta_j - f_{tj}) \quad (5)$$

$$\nabla.J_n - \sum_j R_n^{tj} - R_{sp} - R_{st} - R_{av} + G_{opt}(t) = \frac{\partial n}{\partial t} + N_D\frac{\partial f_D}{\partial t} \quad (6)$$

$$\nabla.J_p - \sum_j R_p^{tj} - R_{sp} - R_{st} - R_{av} + G_{opt}(t) = \frac{\partial p}{\partial t} + N_A\frac{\partial f_A}{\partial t} \quad (7)$$

where p and n are the excess hole concentration and minority-electron concentration with respect to equilibrium, N_A and N_D are the doping levels of acceptors and donors, J_p and J_n are the current density for holes and electrons, R_p and R_n represent the bulk recombination of holes and electrons, G_{opt} is the term for optical generation, R_{sp} represents the recombination of surfaces, and R_{av} and R_{st} are trapping recombination. The required optoelectronic parameters per each layer are captured from the simulation tool library and the optical generation term is calculated using an Matlab based optical model, taking into confederation a doping level of 10^{18} cm^{-3} in both the n and p sides [10].

Herein, the optical generation parameter G_{opt} plays a critical role in investigating the optical impact of the proposed photons management structure in enhancing the performance of the solar cell. The expected enhancement can be attributed to the elongation of the optical path length inside the active materials under simulation. This elongation is directly reflected on the generation term.

3. Simulation Results

3.1. Optimizing number of layer in DBR structure

Si-SiO$_2$ DBR structures have been simulated following the characteristics matrix analytical approach defined in the section two. Based on our previous work, these DBR structures were taken into account as an initial model to be optimized [8]. Firstly, the influence of the DBR layers on the presence of the optical band-gap was originally discussed in Figure 4. The spectra showing the reflectance are plotted and the structures of two, four and six DBR layers have been simulated. Here, the ratio between the the reflected light and the incident light is termed as the optical reflectivity (dimensionless unit), cf. Figures 2 and 3.

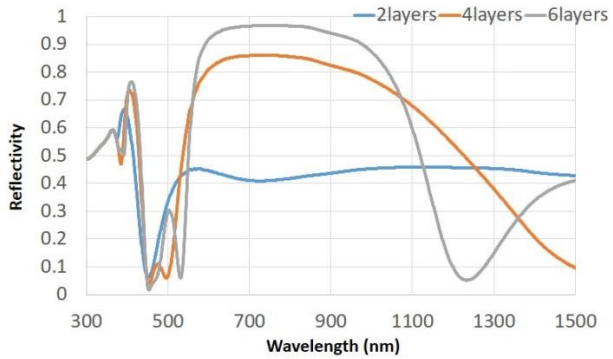


Figure 4: Si-SiO2 Reflectivity of DBR for different layer numbers.

It is evident from Figure 4 that the occurrence of the optical band-gap is associated to the number of layers constructing the one dimensional photonic crystal structure. 550 nm optical band-gap

(from 550 nm to 1100 nm) was formed with reflectively of around 0.95, while utilizing six DBR layers structures. In contrast, attention should be given to the trade in parameters concerned to the overall DBR dimensions and to the fabrication procedure sophistications. For instance, a sharper optical band-gap is anticipated by increasing the number of layers, more than six, however the complicity and fabrication cost will increase as well. Thus, the optimal structure in terms of layer count for such Si-SiO_2 DBR structure is the six layer DBR structure. It is worth noting that this layer number is not a global optimum, as the observation of the optical band-gap depends not only on the DBR layers count, but also plays the refractive index contrast plays an critical role.

3.2. DBR thickness optimization

Following the first optimization stage demonstrated in section 3.1, Figure 5 shows the influence of the thickness variance on reflectivity. The DBR structure of six layers optimized in section 3.1 is used. To reduce the variables included in the optimization problem under investigation, the thickness of both layers, SiO_2 and Si, are considered equal. Figure 5 examined 550 nm, 250 nm, and 125 nm reflectivity spectra. The nominated wavelengths are selected to fit the peak wavelength for the input source at 550 nm. Consequently, λ_o , $\lambda_o/2$ and $\lambda_o/4$ thicknesses have been simulated. As shown in Figure 4, the presence of the photonic-gap is associated at thin film thickness of $\lambda_o/4$. This is attributed to the maximum local point at which the maximum reflection occurs. The thin film thickness is nominated to be a 1/4 of the optical wavelength where maximum reflection is observed [14]. In view of the projection of a quarter wavelength layer of Si with a refractive index of n_1 and light on a wavelength layer of λ_o, the thickness h_1, which induces maximum reflection, is determined by [14]:

$$h_1 = \frac{\lambda_o}{4\, n_1} \quad (8)$$

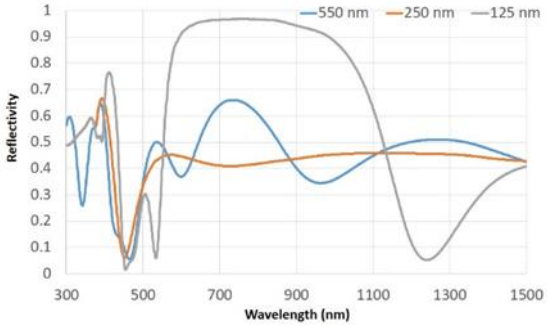


Figure 5: Six Si-SiO2 DBR reflectivity layers with various layer thicknesses.

3.3. Refractive index contrast optimization

Finally, Titanium dioxide is integrated as a direct substitution for Si, in order to examine the impact of refractive index contrast variation on the photonic-gap. TiO_2 is chosen as an easily prepared, low-cost and efficient n-type semiconductor material, knowing that we have already investigated its refractive index experimentally in [28]. Manual as well as automated Dr Blade screen printing can be utilized, as a sol-gel method, to deposit thin film of TiO_2 layer in micrometer range as previously discussed in [28]. In such context, diluted paste with isopropanol can be managed to tune the thin film thickness [27]. The large variation between the refractive index of TiO_2, nearly 1.78, and that of Si, around 3.9 [26], negatively contribute to the detected width of the band-gap. This narrowing can be observed in Figure 6, as a reduced, less sharp optical band-gap with maintaining the number of layers in both DBRs under test.

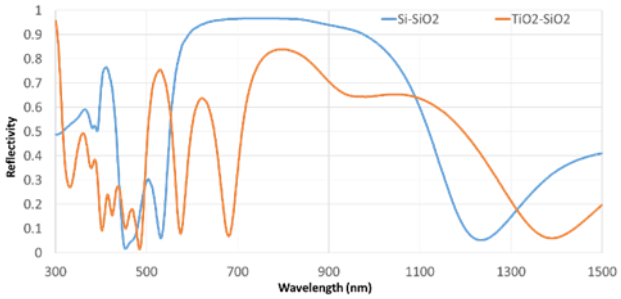


Figure 6: 125 nm, six layers Si-SiO2 and TiO2-SiO2 DBR reflectivity.

3.4. Absorption profile simulation

Towards more optical investigation in the DBR structure optimized in the previous versions, Comsol Multi-physics was utilized. Herein, six layers Si-SiO_2 DBR modeled in the previous section were simulated. The electric field components are used to describe the absorption profile following the Poyting vector method illustrated in [8]. The demonstrated optical absorption profile shown in figure 7, at 550 nm, illustrates the trapping functionality of the DBR structure to trap light inside the active material seeking for optical path length elongation. Accordingly an enhancement of nearly 38% is recorded in the optical generation term. The simulation has been conducted over an optical wavelength range from 200 nm (UV-range) to 2000 nm (IR-range) with a peak at 550 nm.

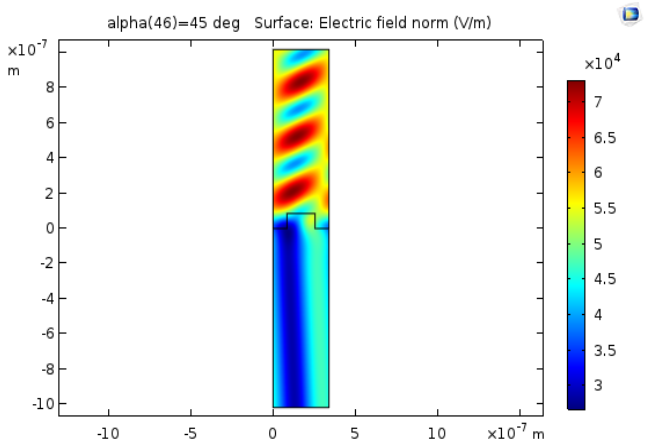


Figure 7: 125 nm, six layers Si-SiO_2 absorption profile.

3.5. Two-dimensional grating light trapping structure

As shown in figure 8a, a 2D grating is characterized as structures that are periodic with a periodicity L in two-dimensions, (e.g. x-direction and y-direction), and are uniform in only one dimension (z-direction). We consider the propagation of incident light in the xz-plane. Since the structure considered is two-dimensional, the reaction of the structures can be separately considered for the s-polarization, which has its electric field polarized along the y-direction, and for the p-polarization, which, as shown in figure 8b, has its magnetic field polarized along y-direction.

With a periodicity L of 1 µm, SiO_2 thickness d_2 of 100 nm, and an aspect ratio (a/b) of 1 (figure 8b), a 2D Si/SiO_2 grating structure is simulated. The minimum x-direction dimension b = 500 nm is used to allow such a device to be fabricated in an area of 200 µm x 200 µm. Furthermore, another experiment was carried out in order to simulate the same structure but with aspect ratio of 2.4. The reflectance of this structure is shown in Figure 9 in the near infrared region with an optical band-gap. With respect to the DBR structure, this shifting in the optical band gap structure is due to the scaling up of the minimum dimension length from 50 nm to 500 nm.

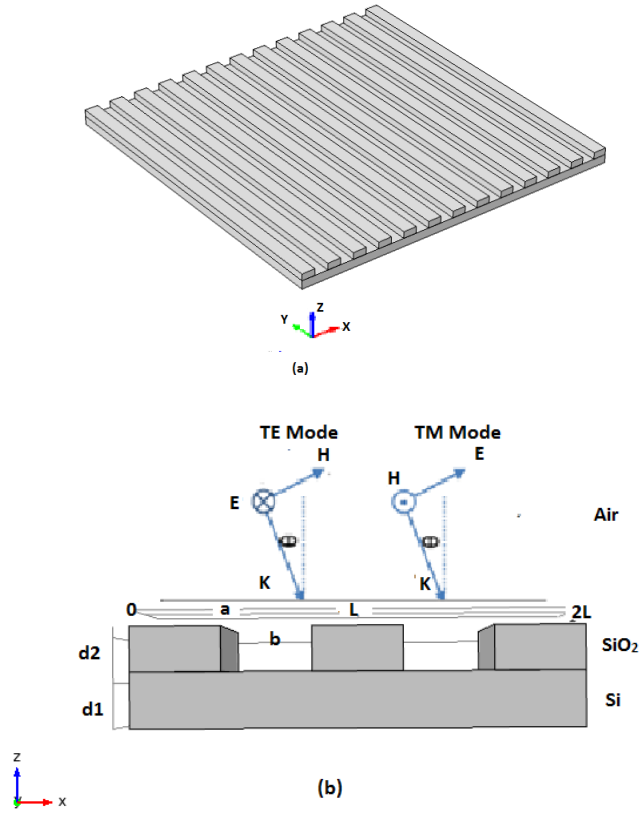


Figure 8: (a) 3D schematic for the structures of 1D grating. (b) X-Z cross section for 1D grating structure.

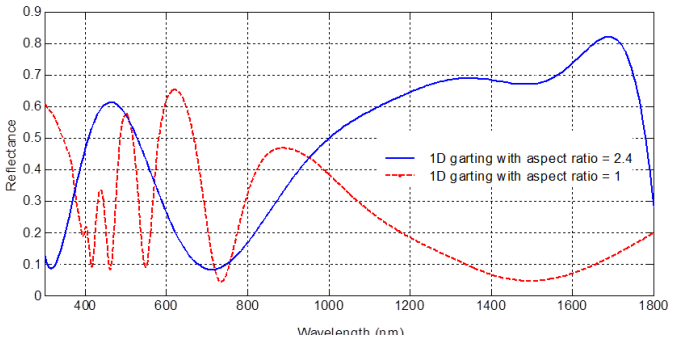


Figure 9: 1D grating structure with aspect ratio of 1 (red curve) and 2.4 (blue curve).

4. Fabrication Process

As mentioned in the introduction, one of our targets here is to fabricate thin film to validate our proposed optical model. Herein, thermal evaporator and spin coater are suggested to be used for this process, while UV-Vis spectrometer is used for the optical characterization purposes.

The thermal evaporator consists of different subsystems. One of the subsystems is the vacuum chamber, the control valves, the vacuum pumps are integrated to satisfy the vacuum itself, and the gauges is utilized to evaluate the pressure level in the chamber. Another obvious subsystem is the filament and the current source required to heat the filament and to control its temperature. Another subsystem measures the thickness of the deposited film so that we can know when the targets are evaporating and when the film deposition is complete. A more trivial subsystem is the framework used physically to hold the substrates and to control whether the evaporated material reaches the substrates.

The Spin coater is equipment enrolled in deposing thin films to flat surfaces substrates. A coating material, usually called suspension, is drooped in the center of the sample before starting the spinning process. The sample is then rotated either in one step or multi-steps with a certain rotating speed and acceleration. This rotation produces the thin film with the aid of centrifugal force.

The speed and the acceleration chosen play an important role in adjusting the thickness of the produced film. The used suspension is usually chosen to be easily evaporated and volatile. It is also worth to mention that not only the speed and acceleration of rotation tune the thickness but also the viscosity and concentration of the suspension. However, in principle, the faster the rotation speeds, the thinner the produced film. Typically, spin coater are used in deposing thin films of few nanometers thick to micrometers range such as the photoresist materials.

For optical characterization, the UV-Vis spectrometer used to scan the ultraviolet to near infrared regions including the visible portion of spectrum. The basic setup of the UV-VIS spectrometer is consisting of a light source which includes two lamps. The first is a tungsten-halogen lamp for the visible and the near infrared (NIR) region. For the UV region, the deuterium lamp is used A NIR and a visible light detector are used in detecting light. Finally, fibers optics cables are used to connect the source detectors and sample holder (cf. Figure 10).

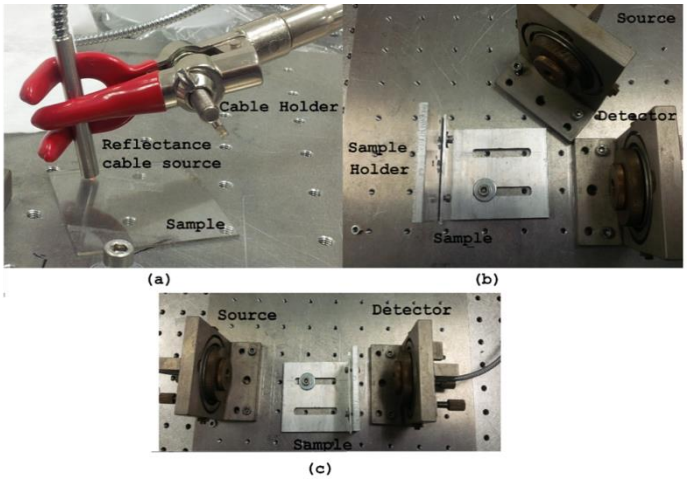


Figure 10: (a) Measuring reflectance using three terminals fibres cable. (b) Measuring reflectance using source/detector setup. (c) Measuring the transmission

5. Experiential validation

In order to validate order optical proposed model, experimental validation take place using thin film based samples. A thermal

evaporation process is used to evaporate a 70 nm of Ag on a highly P-doped Si wafer. The highly P-doped Si wafer is chosen to maintain its optical characteristics during the fabrication process. The thermal evaporation process is done under a current rate of 12% where the tooling factor was previously determined using a dektak profilometer and it was found to be 0.551. The output evaporation rate is 0.1 nm/sec. Then a characterization process takes place using UV-VIS spectrometer to measure the reflectance of the sample in the desired portion of the spectrum as shown in figure 10. The effect of adding the collimating lens is also considered here showing a more accordance with the simulation results than the without collimating lens results. The simulation results shown here are based on our optical numerical model described in the previous section.

The results shown in figure 11 is not only considered as a new validation step to our optical model but it can also be considered as a validation to the new suggested measurement setup of the UV-VIS spectrometer with adding collimating lens. The results show that the adjusted lens doesn't cause any undesirable effect in the wavelength portion of interest making it possible to use such a setup in measuring the reflectance of small area samples.

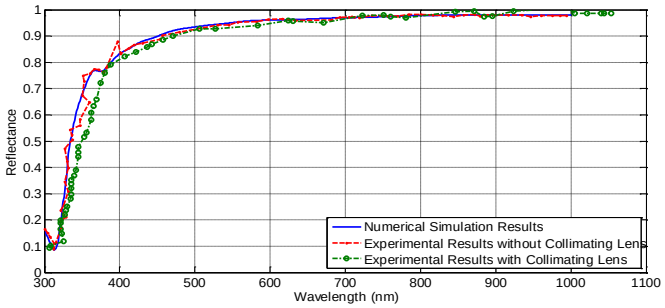


Figure 11: Comparison between the simulated and the measured reflectance of a 70 nm Ag thin film.

A 40 nm thin film of SiO_2 is deposited on a Si <100> wafer using thermal evaporator. The tooling factor is calculated and it was 0.1812 while the current rate is set to be 35%. It is observed that the current needed to evaporate SiO_2 is much higher than that used for Ag, this is due to the high evaporating temperature of SiO_2 with respect to that of Ag. The evaporation rate is varying around 0.01 nm/sec. The Reflectance of the thin film is measured using the UV-VIS spectrometer as observed in figure 12. The figure shows a high accordance between the numerical simulated and the measured reflectance. It was found that due to the high evaporating temperature of SiO_2 (about 1800 ^oC); it is not efficient to evaporate thin film of SiO_2 using thermal evaporation technique. As deposing thin film of SiO_2 must be done by more than one iteration, due to the very low evaporating rate, which causes a poor uniformity of the surface.

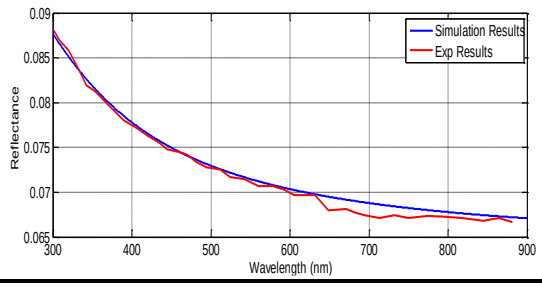


Figure 12: Comparison between the simulated and the measured reflectance of a 40 nm SiO_2 thin film.

It's worth to mention here that the optical validation process implemented on two sample of thin films behaving as optical filter showed perfect agreement with the data captured from our proposed optical model. However, the validation process demonstrated in this paper is limited to thin films samples due to constrain associated with photonic crystals fabrication. Accordingly, we consider the validation step with fabricated photonic crystal structure as a future extension of the current work.

6. Discussions

After the iterations used to optimize DBR structure the output for three steps of optimization that shown in section 3 and also take in consideration the previous measurement in [8], shown in figure 13. By our requested of optimization modal, it is shown the suitable match. Our represented structure is suggested structure to back reflector of thin film of solar cells. This model of structure lead to optical band-gap in near infrared and far visible region, (from 800 nm to 1100 nm) the low absorption occurred at the layer of Si. We attribute such extended band-gap to the refractive index contrast with respect to the DBR shown in figure 5. This structure showed below describe the performance of non-absorbed photons which cannot escape from active layer, after that the light trapping occurred inside material. It is worth to mention that the height (y-axis reflectance value) is inversely proportional to the optical band gap-width [16]. The higher the reflectance the lower the band-gap.

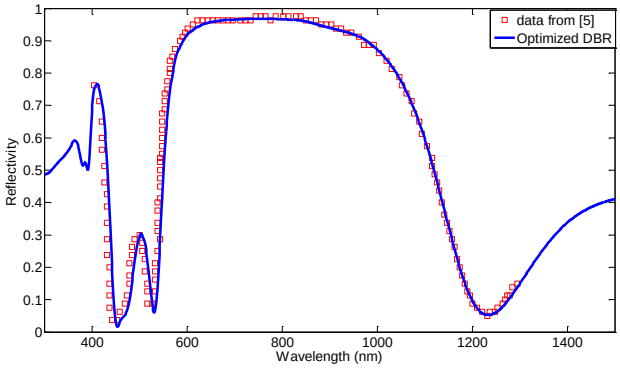


Figure 13: 125 nm, six layers Si-SiO_2 DBR verification with previously published data in [5].

A primary downside within the proposed silicon-slica DBR structure is that the nonappearance of metal which may act as a counter electrode. Consequently, another altered DBR structure with Indium tin-oxide (ITO) rather than Si can be suggested. Though, the proposed planned structure will lack of a condensed optical band-gap because of the less index of refraction distinction with the ITO-SiO_2 DBR with respect to the optimized Si-SiO_2 DBR structure. Additionally, the proposed ITO-SiO_2 DBR structure will show more refined fabrication process, particularly within the deposition of SiO_2 layer the thin ITO layer. Regularly the SiO_2 deposition is achieved by chemical vapor deposition (CVD) at moderately high temperature. Such testimony conditions might influence the ITO electrical properties yet as morphology.

Fundamentally, Nano-lithography technology can be essentially used to fabricate higher dimensions (2D and 3D) photonic crystal structure. In this manner, more refined manufacturing processes and costs would be expected. This specifically influences the cost of delivered energy (Wh/$) by TFSCs. An inclusive analysis showing the effect of the 1D, 2D, and 3D PCs structure in terms of cost versus increase in conversion

efficiency can be carried out as a future task. In principle, alternative low-absorption solar cells can be upgraded using the same light trapping process.

7. Conclusion

The optimization in DBR structure that represented in the paper depends on thickness of layers, refractive index contrast and number of layers. The optimization of DBR structure implement in two step, firstly analytical implantation by MATLAB and secondly verification by using Open-filter. In order to simulate the solar spectrum (AM1.5G), Gaussian sources and FDTD were used.

By using 6 layers of Si-SiO_2 of DBR as result of optimization, the band-gap occurred in 550 nm. This band-gap represents a massive potential in Si as reflected near-infrared of photonics and this reflection avoid light from escaping from material. After comparing between previous data that published and our simulation results, the suitable match was chosen. On another hand, TiO_2 was examined as alternative of Si to observe the effect of changing refractive index in the optical band-gap. At the end the effect of adding metal layer of ITO on Si as substitution was discussed in this paper and released the effect of it on the optical band-gap. Additionally, 2D grating structure was also examined based on the developed optical mode, where transmission spectra were simulated. Fabricated thin films were utilized to validate the optical model while a perfect matching was recorded.

The proposed optical model has showed a great potential in designing optimal 1D and 2D photonic crystal structures with a tunable optical band-gap varying from 400 nm to 700 nm capable of integrating in various solar cells applications. Moreover, the study investigates the optical absorption profile for various photonic crystals seeking for minimum parasitic absorption.

Acknowledgment

In this work, the authors would like to acknowledge the contribution and support of the STDF. As part of the STDF Project entitled, “Mesostructured Based Solar Cells for Smart Building Applications”, Project ID#33502. Additionally, the authors would like to thank the Centre for Emerging Learning Technology (CELT), directed by Prof. Hani Ghali, in BUE for providing all the fabrication facilities needed.

References

[1] S. O. Abdellatif and G. A. Alsayed, "Optimizing 1D Photonic Crystal Structures for Thin Film Solar Cells," in 2019 IEEE Conference on Power Electronics and Renewable Energy (CPERE), 333-337, 2019, doi:10.1109/CPERE45374.2019.8980046.

[2] M. Mujahid, C. Chen, W. Hu, Z.-K. Wang, and Y. Duan, "Progress of High-Throughput and Low-Cost Flexible Perovskite Solar Cells," Solar RRL, 1900556, 2020, doi:10.1002/solr.201900556.

[3] M. Sajadnia, S. Dehghani, Z. Noraeepoor, and M. H. Sheikhi, "Highly improvement in efficiency of Cu (In, Ga) Se2 thin film solar cells," World Journal of Engineering, **17**(4), 527–533 2020, doi:10.1108/WJE-02-2020-0068.

[4] S. Amiri and S. Dehghani, "Design and Simulation of Single-Junction and Multi-junction Thin-Film Solar Cells Based on Copper Tin Sulfide," Journal of Electronic Materials, **49**, 5895–5902, 2020, doi.org/10.1007/s11664-020-08382-6.

[5] F. Sadegh, S. Akin, M. Moghadam, V. Mirkhani, M. A. Ruiz–Preciado, Z. Wang, et al., "Highly efficient, stable and Hysteresis–less planar perovskite solar cell based on chemical bath treated Zn_2SnO_4 electron transport layer," Nano Energy, 105038, 2020, doi:10.1016/j.nanoen.2020.105038.

[6] S. Liu, J. Yuan, W. Deng, M. Luo, Y. Xie, Q. Liang, et al., "High-efficiency organic solar cells with low non-radiative recombination loss and low energetic disorder," Nature Photonics, **14**, 300–305, 2020, doi.org/10.1038/s41566-019-0573-5.

[7] F.-I. Lai, J.-F. Yang, Y.-C. Hsu, and S.-Y. Kuo, "Enhanced omnidirectional light harvesting in dye-sensitized solar cells with periodic ZnO nanoflower photoelectrodes," Journal of Colloid and Interface Science, **562**, 63-70, 2020, doi:10.1016/j.jcis.2019.12.003.

[8] S. Abdellatif, K. Kirah, R. Ghannam, A. Khalil, and W. Anis, "Enhancing the absorption capabilities of thin-film solar cells using sandwiched light trapping structures," Applied optics, **54**, 5534-5541, 2015, doi:10.1364/AO.54.005534.

[9] S. Abdellatif, R. Ghannam, and A. Khalil, "Simulating the dispersive behavior of semiconductors using the Lorentzian–Drude model for photovoltaic devices," Applied optics, **53**, 3294-3300, 2014.

[10] S. Abdellatif and K. Kirah, "Numerical modeling and simulation for a radial pin nanowire photovoltaic device," Energy Procedia, **36**, 488-491, 2013, doi:10.1364/AO.53.003294.

[11] M. M. Hassan, A. Sahbel, S. O. Abdellatif, K. A. Kirah, and H. A. Ghali, "Toward low-cost, stable, and uniform high-power LED array for solar cells characterization," in New Concepts in Solar and Thermal Radiation Conversion III, 114960Q, 2020, doi:10.1117/12.2572874.

[12] G. R. Li and X. P. Gao, "Low-cost counter-electrode materials for dye-sensitized and perovskite solar cells," Advanced Materials, **32**, 1806478, 2020, doi:10.1002/adma.201806478.

[13] B. Pashaei, S. Bellani, H. Shahroosvand, and F. Bonaccorso, "Molecularly engineered hole-transport material for low-cost perovskite solar cells," Chemical Science, **11**, 2429-2439, 2020, doi:10.1039/c9sc05694g.

[14] E. G. Durmusoglu, G. S. Selopal, M. Mohammadnezhad, H. Zhang, P. Dagtepe, D. Barba, et al., "Low-Cost, Air-Processed Quantum Dot Solar Cells via Diffusion-Controlled Synthesis," ACS Applied Materials & Interfaces, **12**, 32, 36301–36310, 2020 doi:10.1021/acsami.0c06694.

[15] J. H. Werner, J. Mattheis, and U. Rau, "Efficiency limitations of polycrystalline thin film solar cells: case of Cu (In, Ga) Se2," Thin Solid Films, **480**, 399-409, 2005, doi:10.1016/j.tsf.2004.11.052.

[16] R. Biswasa and D. Zhoua, "Enhancing Light-trapping and Efficiency of Solar Cells with Photonic Crystals," in Materials Research Sociaty Spring Meeting - Symposium A – Amorphous and Polycrystalline Thin-Film Silicon Science and Technology, Germany, 0989-A03-02, 2007, doi:10.1557/PROC-0989-A03-02.

[17] S. Abdellatif, K. Kirah, R. Ghannam, A. Khalil, and W. Anis, "Comprehensive study of various light trapping techniques used for sandwiched thin film solar cell structures," in Physics, Simulation, and Photonic Engineering of Photovoltaic Devices VII, 10527152018, doi:10.1117/12.2291613.

[18] J. Hou, M. Li, and Y. Song, "Patterned colloidal photonic crystals," Angewandte Chemie International Edition, **57**, 2544-2553, 2017, doi:10.1002/anie.201704752.

[19] W. Zhu, R. Esteban, A. G. Borisov, J. J. Baumberg, P. Nordlander, H. J. Lezec, et al., "Quantum mechanical effects in plasmonic structures with subnanometre gaps," Nature communications, **7**, 1-14, 2016, doi:10.1038/ncomms11495 .

[20] M. Mariano, P. Mantilla-Pérez, P. Romero-Gómez, A. Martínez-Otero, X. Elias, R. Betancur, et al., "One-Dimensional Photonic Crystals for Light Management in Organic Solar Cells," in Organic and Hybrid Photonic Crystals, ed: Springer, 303-320, 2015, doi:10.1007/978-3-319-16580-6_13.

[21] F. Enrichi, A. Quandt, and G. C. Righini, "Plasmonic enhanced solar cells: Summary of possible strategies and recent results," Renewable and Sustainable Energy Reviews, **82**, 2433-2439, 2018, doi: 10.1016/j.rser.2017.08.094.

[22] C. Zhao, P. Tong, L. Shi, T. Ji, F. Zhu, Y. Hao, et al., "Performance improvement in inverted organic solar cells by incorporating core-shell SiO_2@ Au plasmonic structures," in 9th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Micro-and Nano-Optics, Catenary Optics, and Subwavelength Electromagnetics, 108400N, 2019, doi:10.1117/12.2510899.

[23] R. S. Moakhar, S. Gholipour, S. Masudy-Panah, A. Seza, A. Mehdikhani, N. Riahi-Noori, et al., "Recent Advances in Plasmonic Perovskite Solar Cells," Advanced Science, **7**, 1902448, 2020, doi:10.1002/advs.201902448.

[24] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, "Photonic crystals: putting a new twist on light," Nature, **386**, 143-149, 1997, doi:10.1038/386143a0.

[25] A. Fu, H. Gao, P. Petrov, and P. Yang, "Widely tunable distributed Bragg reflectors integrated into nanowire waveguides," Nano letters, **15**, 6909-6913, 2015, doi:10.1021/acs.nanolett.5b02839.

[26] S.-P. Yu, J. A. Muniz, C.-L. Hung, and H. Kimble, "Two-dimensional photonic crystals for engineering atom–light interactions," Proceedings of

the National Academy of Sciences, **116**, 12743-12751, 2019, doi: 0.1073/pnas.1822110116.
[27] S. Abdellatif, S. Josten, P. Sharifi, K. Kirah, R. Ghannam, A. S. Khalil, et al., "Optical investigation of porous TiO_2 in mesostructured solar cells" 10526: SPIE, 2018, doi:10.1117/12.2288376.
[28] S. Abdellatif, P. Sharifi, K. Kirah, R. Ghannam, A. Khalil, D. Erni, et al., "Refractive index and scattering of porous TiO_2 films," Microporous and Mesoporous Materials, **264**, 84-91, 2018, doi: 10.1016/j.micromeso.2018.01.011.
[29] S. O. Abdellatif, S. Josten, A. S. Khalil, D. Erni, and F. Marlow, "Transparency and Diffused Light Efficiency of Dye-Sensitized Solar Cells: Tuning and a New Figure of Merit," IEEE Journal of Photovoltaics, **10**, 522-530, 2020, doi:0.1109/JPHOTOV.2020.2965399.
[30] S. Kamal, G. M. Al-sayyad, R. Abdelmoteleb, M. Abdellatif, and S. O. Abdellatif, "Submerged Solar Energy Harvesters Performance for Underwater Applications," in 2019 International Conference on Innovative Trends in Computer Engineering (ITCE), 444-449, 2019, doi:10.1109/ITCE.2019.8646459..
[31] M. M. Abdellatif, S. M. Maher, and M. Ghazal, "Implementation of a low cost, solar charged RF modem for underwater wireless sensor networks," International Journal on Smart Sensing and Intelligent Systems, **13**, 1, 1-11, doi:10.21307/ijssis-2020-015.
[32] S Larouche and L Martinu, "OpenFilters: open-source software for the design, optimization, and synthesis of optical filters," Appl. Opt. **47**, C219-C230, 2008, doi:10.1364/AO.47.00C219.
[33] E. Hecht, OPTICS, Fourth eddition. Adelphi university, San Francisco, Boston, New York, 2002.
[34] A. D. Rakić, A. B. Djurišić, J. M. Elazar, and M. L. Majewski, "Optical properties of metallic films for vertical-cavity optoelectronic devices," Applied optics, 37, 5271-5283, 1998, doi:10.1364/AO.37.005271.
[35] K. Kunz and R. Luebbers, Finite differnce time domain method for electromagnatics. United State of America: CRC Press, 1993.
[36] J. D. Joannopoulos, "MEEP," ed: Massachusetts Institute of Technology, https://meep.readthedocs.io/en/latest/, accessed October 2020.
[37] A. F. Oskooi, D. Roundy, M. Ibanescu, P. ermel, J. D. Joannopoulos, and S. G. Johnson, "Meep: A flexible free-software package for electromagnetic simulations by the FDTD method," Computer Physics Communications, ELSEVER, **181**, 687–702, 2010, doi:10.1016/j.cpc.2009.11.008.
[38] "Reference solar spectral irradiance: Air mass 1.5," ed: National Renewable Energy Lab.
[39] C. Multiphysics, "version 4.2," Heat Transfer Module, https://www.comsol.com/products/4.2/, accessed October 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 960-965 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Total Family Risk in Families who go to Popular Dining Rooms in a Vulnerable Area of Collique, Comas

Hernan Matta-Solis[1,*], Rosa Perez-Siguas[1], Eduardo Matta-Solis[1], Melissa Yauri-Machaca[2]

[1]*Research and Intellectual Creativity Direction, Universidad María Auxiliadora, 15408, Lima-Perú*

[2]*Research and Technology Direction, Business on Making Technologies, 15307, Lima-Perú*

ARTICLE INFO

Article history:
Received: 04 August, 2020
Accepted: 28 September, 2020
Online: 20 October, 2020

Keywords:
Family
Family relationships and risk
Popular dining rooms
Threatened family

ABSTRACT

The family can be referred to as a basic nucleus of society, where it must be fully formed as a group and guarantee the safety and development of its members. Increasing social inequality affects society and also the families that comprise it. It is a study with a quantitative approach, with a non-experimental, descriptive and cross-sectional design. The population consisted of 240 heads of families who go to 12 Popular Dining Rooms of a Vulnerable Area of Collique. The data collection technique was the survey-interview and the instrument used was the RFT 5-33 questionnaire of 5 dimensions and 33 items. The total family risk in families who go to Popular Dining Rooms of a Vulnerable Area of Collique in Comas, it is presented as follows, 172 participants representing 72% are threatened families; 41 participants representing 17% are families with low risk and 27 participants representing 11% are families with high risk. Regarding dimensions, threatened families predominated in all, in psycho-affective conditions with 94%, in health services and practices with 91%, in housing and neighborhood conditions with 62%, in socioeconomic situation with 85% and in child management with 85%. The total family risk that predominated is threatened families, followed by families with high risk and families with low risk. Regarding the dimensions of the main variable, threatened families predominate in all of them. The dimension with the highest high-risk value is housing and neighborhood conditions.

1. Introduction

The family, as the basic nucleus of society, must develop plenty as a group and guarantee the safety and development of its members. Growing social inequality affects society and the families that make it up, this problem affects and destabilizes family dynamics, exposing and making it vulnerable to existing social problems [1].

Similarly, it not only makes families vulnerable to social problems but also to psycho-affective emotional and motivational problems where these aspects compromise the development of self-esteem, self-confidence and the fulfillment as a person of each family member, altering the development of values and optimal functionality that benefits the family [2].

In the family, there is a diversity of problems that are accentuated more frequently, and there are no ways to solve them, couples who are at the head of the family nucleus choose divorce or separation as a solution, a situation that affects the Family structure and functioning [3], also, family ruptures bring relevant psychosocial consequences for their members, many times these affect the personality of their children [4].

According to the National Program Against Family and Sexual Violence of the Ministerio de la Mujer y Poblaciones Vulnerables, it indicates the frequency of cases of different types of violence, where the main type of violence is psychological, followed by physical, sexual and economic. In relation to the age groups with the types of violence given in women, cases of violence were attended in the first semester of 2018, where it was found that 72% of young and adult women said they had suffered some type of violence, followed 22% to girls and adolescents and 5% to older women [2].

In the same way, the Instituto Nacional de Estadística e Informática mentions the complaints registered by the National Police on family violence due to marital and family problems, according to department, during 2016, they show that Lima has

*Corresponding Author: Hernan Matta-Solis, Email: hernan.matta@uma.edu.pe

the highest number of complaints with 28 699 reports, followed by Arequipa with 9,748 reports [5].

In the research study carried out in Colombia, it shows the results obtained in families of schoolchildren where the level of total family risk shows that 55% of families are at a threat risk level, 37% at a low risk level and 8% at high risk level; when evaluating the individual risk categories of the five factors studied according to RFT 5-33, all the factors are classified as threatened and the factor related to housing and neighborhood is located in its highest frequency at 21% [6].

In the research study carried out in Colombia, the results of the characterization of the total family risk show that in the group of adolescents, 57% of the total are threatened families and 8% are families with high risk. From the measurement of adolescents, 57% are threatened families and 8% are families with high risk, from it is measured by caregivers, 66% are threatened families and 15% are families with low risk [7].

In the research study carried out in Peru, they mention the results of the total family risk, they found 35 families that represent 31% of the total are family with high risk, 58 families that represent 51% of the total are threatened and 20 families that represent 18% of the total have low risk. Regarding the dimensions of the total family risk, those that present a high risk are the housing and neighborhood conditions, they have 46 families that represent 41%, and in health services and practices they have 41 families that represent 36% [8].

The objective of the research work is to determine the total family risk in families who go to Popular Dining Rooms of a Vulnerable Area of Collique in Comas, which it will allow to observe what the family risk is in families. This study is important since it will provide with relevant and real data about the vulnerability of families exposed to risky conditions that can compromise all family members.

In the present research work, the Total Family Risk questionnaire also known as RFT Questionnaire 5-33 will be used as a collection instrument.

The data collection was processed through the survey of families who go to Popular Dining Rooms of a Vulnerable Area of Collique in Comas, the data to be entered was carried out in a data matrix that will be designed in the statistical program IBM SPSS Statistics Base 25.0, it proceeded to its corresponding analysis, which will allow to make statistical tables so that they can be described and interpreted in results and discussions, respectively.

The following research work is structured as follows: In section II, the development of the data collection process of the families who go to a popular dining room of Collique in Comas will be presented and also the guidelines to consider. In section III, the results will show the total family risk in families who go to Popular Dining Rooms of a Vulnerable Area of Collique in Comas according to the measures based on the variable. In section IV, it presents the discussions of the research work and in section V, the conclusions.

2. Methods

In this part, the type and design of the research will be developed, also the population and the sample that will be carried out in the research work, the inclusion and exclusion criteria in detail and finally the technique and the instrument for collecting data.

2.1. Research type and design

The present study has the characteristics of the quantitative approach. Regarding the research design, this study is non-experimental, descriptive and cross-sectional [9].

2.2. Population

This study was carried out in the total population. This is made up of 240 family managers who go to a popular dining room of Collique in Comas. The family heads identified in their entirety is women.

Inclusion criteria

- Family managers who live in the area at least 1 year old
- Family managers who consume food in a popular dining room of Collique in Comas.
- Family managers who are physically and/or mentally able to answer an interview.
- Family managers who wish to participate and sign the informed consent for participation in the study.

Exclusion criteria

- Family managers who do not live in the study jurisdiction.
- Family managers who do not consume food in the Popular Dining Room linked in the study.
- Family managers who are not physically and/or mentally capable of answering an interview.
- Family managers who do not wish to participate and do not sign the informed consent to participate in the study.

2.3. Technique and instrument

The data collection technique to be used in this study will be the survey, since a structured instrument is used to capture a good amount of information from the participants in the data collection process [10].

In the present research work, the Total Family Risk questionnaire also known as RFT Questionnaire 5-33 will be used as a data collection instrument, which is made up of 5 factors or dimensions, in which 33 main items are distributed.

The 5 factors or dimensions are: psycho-affective conditions, health practices and services, housing and neighborhood, socioeconomic situation and child management. The answers are all dichotomous for each of the items, if it is at risk, it is given a value of 1 and if there is an absence of it, it is valued at 0. Between 5 to 12 risks of 33 possible, it has families threatened, between 13 and 33 risks of 33 possible, it is families with high risk, and 0 to 4 risks of 33 possible, it has families with low risk. The final value

of the variable Total family risk has three values: families with high risk, threatened families and families with low risk [11].

The values as families with high risk, is based as a situation where the family is exposed to dangers in which it can compromise all the members; threatened families is the situation where the family is vulnerable to dangerous situations where one or all the members of the family would be compromised and families with low risk, refers to the situation that the family is at some type of risk but that can be controlled by the family members.

2.4. Place and Application of the Instrument

The survey to measure total family risk was carried out in families who go to a popular dining room of Collique in Comas.

In order to start data collection, administrative procedures were previously carried out at the university in order to obtain the cover letter stating the title of the study. With this document, steps were taken regarding authorizations to access the popular dining room of Collique in Comas. In a first contact, the people in charge of the Popular Dining Room were happy to support the development of the study.

By identifying the number of families who go to Popular Dining Rooms, it coordinated with the representatives of the popular dining room to date the most appropriate time to collect the data.

The data collection was carried out in the second quarter of the year (April to June of the year), this was carried out mainly during the hours in which family managers went to the popular dining rooms to coordinate the purchase of their food. In some cases, the family heads suggested that the interview be conducted at their home. An average time of approximately 15 to 25 minutes was taken with each family head.

It is important to mention the nursing staff present at the research site since they carry out their controls in popular dining room, weight and height screening is done in families, especially in children and adolescents since they are in a growth stage and development which is very important in their diet, so that they are not prone to get diseases.

3. Results

Below, there are graphs of the surveys carried out following the guidelines corresponding to the research work:

In Figure 1, it has the data of the study participants, where it can observe regarding to the family risk in families who go to popular dining room, 172 participants that represent 71.7% of the total are threatened families, 41 participants that represent the 17.1% of the total are families with low risk and 27 representing 11.3% are families with high risk.

In the Figure 2, it can see regarding to the Total Family Risk in its psycho-affective conditions dimension, in families who go to Popular Dining Rooms, that 226, which represents 94%, are threatened families and 14, which represents 6%, represents Families with high risk. In its dimension, health services and practices, in families who go to Popular Dining Rooms, 219, representing 91% are threatened families and 21, representing 9%, are families with high risk, in its dimension, housing and neighborhood conditions, in families that go to Popular Dining Rooms, that 149 participants representing 62% are threatened families and 91 participants that represent 38% are families with high risk, in their socioeconomic situation dimension, in families that go to Popular Dining Rooms, that 203 participants that represent 85% are threatened families and 37 participants that represent 15% are families with high risk and in their dimension child management, in families who go to Popular Dining Rooms, that 204 participants representing 85% are threatened families and 36 participants representing 15% are families with high risk.

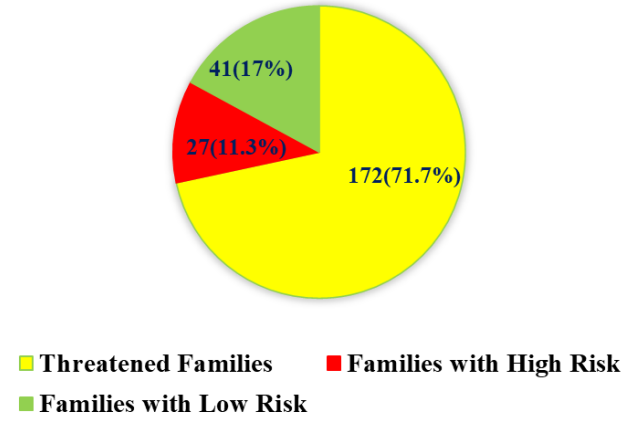


Figure 1: Total Family Risk in Families who go to Popular Dining Rooms in a Vulnerable Area of Collique, Comas – 2018

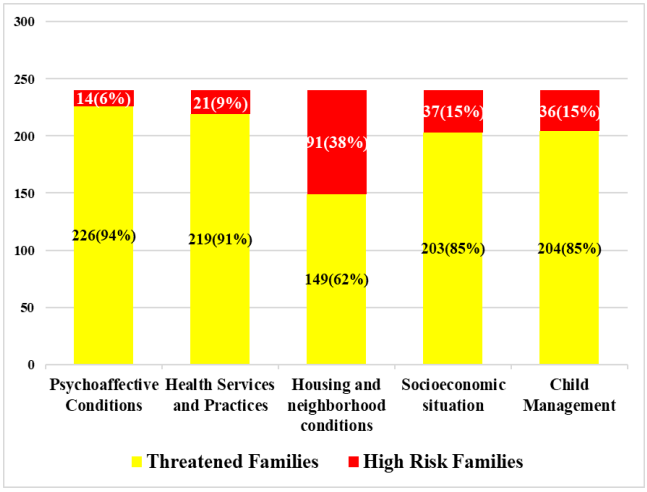


Figure 2: Total family risk according to its dimensions in families who go to Popular Dining Rooms in a Vulnerable Area of II of Collique, Comas - 2018

In Table 1, it has the study participants from popular dining rooms in a vulnerable area of Collique, Comas, where it can see that, in the "Los Diamantes" popular dining room 100% represent the threatened families, in the " Nuestra Señora de Luz" popular dining room 94.1% represents threatened families, 5.9% represents families with high risk, in the "Cruz de Motupe" popular dining room 92% represents threatened families, 8% represents families with low risk, in the "El Buen Pastór" popular

Table 1: Total Family Risk in Families who go to Popular Dining Rooms in a Vulnerable Area of Collique, Comas.

			TOTAL FAMILY RISK - Risk level			
			Threatened Families	Families with High Risk	Families with Low Risk	Total
Popular Dining Rooms	Virgen de Fatima	Count	13	0	3	16
		% in the dining room	81,2%	0,0%	18,8%	100,0%
	Micaela Bastidas	Count	13	2	1	16
		% in the dining room	81,2%	12,5%	6,3%	100,0%
	Corazón de Jesus	Count	18	9	4	31
		% in the dining room	58,1%	29,0%	12,9%	100,0%
	Iglesia Evangelica	Count	13	4	0	17
		% in the dining room	76,5%	23,5%	0,0%	100,0%
	Nuestra Sra. de la Luz	Count	16	1	0	17
		% in the dining room	94,1%	5,9%	0,0%	100,0%
	Victor Raul Haya de la Torre	Count	10	1	13	24
		% in the dining room	41,6%	4,2%	54,2%	100,0%
	Corazon de Jesus ll	Count	17	4	1	22
		% in the dining room	77,3%	18,2%	4,5%	100,0%
	Cruz de Motupe	Count	23	0	2	25
		% in the dining room	92,0%	0,0%	8,0%	100,0%
	Los Diamantes	Count	11	0	0	11
		% in the dining room	100,0%	0,0%	0,0%	100,0%
	Milagros de Jesus	Count	14	1	13	28
		% in the dining room	50,0%	3,6%	46,4%	100,0%
	El buen Pastor	Count	19	2	0	21
		% in the dining room	90,5%	9,5%	0,0%	100,0%
	Los Diamantes ll	Count	5	3	4	12
		% in the dining room	41,7%	25,0%	33,3%	100,0%
Total		Count	172	27	41	240
		% in the dining room	71,7%	11,3%	17%	100,0%

dining room 90.5% represent threatened families, 9.5% represent families with high risk, in “Micaela Bastidas” popular dining room 81.2% represent to threatened families, 12.5% represent families with high risk and 6.3% represent families with low risk, in the "Virgen de Fátima" popular dining room 81.2% represent threatened families, 18.8% represent families with low risk in the “Corazón de Jesús II” popular dining room 77.3% represent threatened families, 18.2% represent families with high risk and 4.5% represent families with low risk, in the "Iglesia Evangélica" popular dining room 76.5% represents threatened families, 23.5% represents families with high risk, in the “Corazón de Jesús” popular dining room 58.1% represents threatened families, 29% represents families with high risk and 12.9% represent families with low risk, in the “Milagros de Jesús” popular dining room 50% represent threatened families, 3.6% represent families with high risk and 46.4 % represents families with low risk, in the "Los Diamantes II" popular dining room 41.7% represents threatened families, 25% represents families with high risk and 33.3% represents families with low risk and in the “Victor Raúl Haya de la Torre” popular dining room 41.6% represents threatened families, 4.2% represents families with high risk and 54.2% represent families with low risk.

The results of this study can be interpreted that the total family risk of families who go to popular dining room is linked to the promotion of family health, since it is very important to know the consequences that exist in families, whether due to disorganization within the family, family dysfunction, or diseases, that allow us to act on it and have the ability to make changes in families that present some type of difficulties.

4. Discussion

The total family risk assessment allows guiding family health care, ensuring the health of each member of the family group and the context in which they take place [12]. For this reason, the importance of putting into practice the proposal of Pilar Amaya, who involves a series of factors and dimensions that allow to see the family's problems in an integral way [11].

The findings proposed by Pilar Amaya, who emphasizes in her instructions that threatened families present between 5 and 12 risks in relation to the 33 items that are part of the risk assessment; these families are characterized by the presence of emotional and affective alterations that result in an adverse environment for the optimal development of the family members. These family groups have difficulties in accessing public health services, in addition to their precarious situation being this situation a threat to the members of the family group. Total family risk is a dynamic concept, that is, it is changing and can be altered over time [11,13]. We agree with what was pointed out by Barboza E. and collaborators, who in their respective studies indicate that this group of threatened families must be given due attention, since they can become families with high risk, which would worsen their situation [14].

In relation to its psycho-affective conditions dimension, threatened families predominated, followed by families with high risk. A threatened family must be monitored to see its evolution over time, in order to prevent the family from becoming involved in more risks, this dimension is closely linked to the mental and emotional health of families, in which it is altered in the interpersonal relationships between family members. Becerra et al. point out in their results in relation to psycho-affective conditions that affective needs are important to satisfy and preserve the mental health of families, therefore, at an emotional level, it will depend on how the family can face situations where the role and the dynamics in the family are not altered [8].

In terms of health services and practices, threatened families predominated, followed by families with high risk. In the study, the basic needs of health benefits are increasingly difficult because most families do not predispose with time because they are mainly engaged in activities that generate income to cover their basic needs, where the health is prioritized in the background, because many of them, due to their educational level, do not value health care, also, they do not know about the care provided by health facilities in their jurisdiction. We agree with what was stated by Giraldo M., who in his study that the distance and lack of involvement of the health establishment with its population has led to the underutilization of the health services available there, where families do not perceive the related risks with their health, also their members, this often causes health problems that evolve over time and can generate unfavorable sequelae in the family [7].

In terms of housing and neighborhood conditions, threatened families predominated, followed by families with high risk; Compared to the other dimensions, this presents a higher risk in families. Health or illness is the result of a dynamic interrelation between families and the environment that surrounds them, although it is true that the environment has a direct link with being healthy in different ways, which is why it improves living conditions and improves the health levels of family members. However, the study showed environmental sanitation problems, where the lack of support from the municipality, the lack of organization and participation of families in caring for their environment has been very high, but that there is no culture in themselves that allow them to prioritize that aspect. In the study by Soto A. and collaborators, findings were evidenced in relation to deficiencies in this aspect and that attributes the socioeconomic conditions and the lack of support from the authorities to improve the environment and adapt the place to live [12]. In the same way, Millan M. and collaborators also point it out in their study associate the socioeconomic situation with the development and adequacy of homes and their sanitation, where unsanitary conditions constitute a risk for all family members [15].

In its socioeconomic situation dimension, threatened families predominated, followed by families with high risk, on the analysis of this dimension conditions the development of a family, because it explains the shortcomings and constraints that families go through in their daily lives, where the Precarious work and education are points that compromise the family and its members. In the same way, Pérez R. and collaborators highlight that education and job insecurity make it difficult for family members to know and access their social rights as citizens, therefore social assistance is important since it allows them to fulfill a palliative function given the needs of families, and that it is important that government institutions and the state promote decent and sustained employment for the population [16].

Finally, in its dimension child management, threatened families predominated followed by families with high risk, in the study the results were alarming, because the health of the child in this community is prioritized as second plane. The risk that minors are exposed to where the conditions of housing, neighborhood and

the economic situation of the family are determinants that sustain the well-being of the child, coincides with the study by Cabello E. et al. [17].

5. Conclusions

It is concluded that the total family risk in families who go to popular dining room of Collique in Comas, where threatened families predominated, followed by families with low risk and families with high risk, is suggested to the health establishment of the jurisdiction, implement as part of its extramural activities, a program of visits to families with high risk, so that it can specifically identify their problems and be able to address them comprehensively.

It is concluded that more variables should be introduced in future studies, in order to establish new possibilities of relating the main variable total family risk with other variables that allow a better understanding of family health and the risks of families in all their dimensions.

It is concluded that studies with qualitative and mixed approaches should be developed that allow a deeper understanding of the problem that these families are going through, this will give a better picture when analyzing family health and the care that is oriented to improve it.

It is suggested that the nursing professional continuously monitor the area through a program of home visits directed from the results of this study, focusing on the growth and development of children and adolescents, their diet, care, and the family relationship.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] M.E. Benítez-Pérez, "La familia: Desde lo tradicional a lo discutible," Revista Novedades En Población, **13**(26), 58–68, 2017.

[2] Ministerio de la Mujer y poblaciones Vulnerables, Observatorio Nacional de las Familias, MIMP, 1–12, 2018.

[3] Superintendencia Nacional de Registros Públicos, Crece inscripción de divorcios a nivel nacional, Lima - Perú, 2017.

[4] M. Jiménez Arrieta, M. Amarís Macías, M. Valle Amarís, "Afrontamiento en crisis familiares: El caso del divorcio cuando se tienen hijos adolescentes," **28**(1), 99–112, 2012.

[5] Instituto Nacional de Estadística e Informática, Denuncias registradas por la Policía Nacional sobre violencia familiar, según departamento, Lima - Perú, 2017.

[6] S. Diaz Cardenas, K. Ramos Martinez, K.M. Arrieta Vergara, "Asociacion del nivel de riesgo familiar total y caries dental en escolares de La Boquilla, Cartagena," Avances En Enfermería, **31**(2), 43–52, 2013.

[7] D.M. Giraldo Ocampo, Caracterización de riesgo familiar total en familias con adolescentes escolarizados, Universidad nacional de Colombia, Bogota - Colombia, 2014.

[8] W.I. Becerra Cubas, J.S. Diaz Pareja, J.P. Rocha Durand, Riesgo familiar total en familias del sector 1 del Asentamiento Humano Juan Pablo II Los Olivos - 2013, Universidad de Ciencias y Humanidades, Lima-Perú, 2015.

[9] R. Hernández Sampieri, C. Fernández Collado, P. Baptista Lucio, Metodología de la Investigación, 6a Edición, McGraw-Hill Education, México, 2014.

[10] F. Alvira-Martín, La encuesta: una perspectiva general metodológica, 2a edición, Centro de Investigaciones Sociológicas, Madrid-España, 2011.

[11] P. Amaya, Instrumento de riesgo familiar total RFT : 5-33 : manual aspectos teoricos, psicometricos de estandarización y de aplicación del instrumento., 1st ed., Universidad Nacional de Colombia, Bogotá, 2004.

[12] A.M. Soto González, V.F. Velásquez G., "Riesgo Familiar Total de Familias desplazadas residentesen el Municipio de Bugalagrandade (Valle de Cauca, Colombia)," Hacia La Promoción de La Salud, **20**(1), 126–139, 2015, doi:10.17151/hpsal.2015.20.1.9.

[13] N.C. Charrys Bravo, "Riesgo familiar total en familias con mujeres diagnosticadas con neoplasia de mama - Dialnet," Revista Ciencia y Cuidado, **14**(2), 8–21, 2017.

[14] E.L. Barboza Delgado, L.I. Prado Pérez, P. Vega Tarazona, Riesgo familiar total en familias de niños menores de 5 años del Asentamiento Humano 19 de Mayo, Los Olivos - 2018, Universidad de Ciencias y Humanidades, Lima-Perú, 2018, doi:10.22258/uch.thesis/217.

[15] M.E. Millan Ambrocio, Riesgo Familiar Total y tipo de familia en el Asentamiento Humano Santa Rosa Alta sector VII del distrito del Rímac, Lima 2017, Universidad Norbert Wiener, 2018.

[16] R.E. Pérez Siguas, H.H. Matta Solís, T.M. Espinoza Moreno, C.R. Paredes Tafur, "Caracterización del riesgo familiar total y tipo de familia en pobladores de un asentamiento humano de un distrito de Lima Norte | Pérez Siguas | Peruvian Journal of Health Care and Global Health," Peruvian Journal of Care and Global Health, **1**(1), 2017.

[17] E.Y. Cabello Huerta, J.B. Francisco Yauri, S.J. Mayta Guanilo, Riesgo familiar total de las familias con niños preescolares de la institución educativa inicial Semillitas de Amor I Puente Piedra – 2015, Universidad de Ciencias y Humanidades, Lima - Perú, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 966-972 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Health-Related Quality of life in Students of an Education Institution of Ventanilla

Lucia Silva-Bueno[1], Brian Meneses-Claudio[2], Hernan Matta-Solis[1,*], Lourdes Matta-Zamudio[1]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 28 September, 2020
Online: 20 October, 2020

Keywords:
Adolescents
Quality of life
Students
Health
Adolescent psychology

ABSTRACT

The Health-related quality of life corresponds to the perception that people have of their level of well-being, considering aspects of their lives and their impact on their state of health. The studies were initially carried out in a population that presented both acute and chronic pathologies and covered practically all areas of medical specialties; subsequently, the interest focused on studying health-related quality of life in the general population and then in the population of children and adolescents. That is why, in this research work, the evaluation of the adolescent to prevent future problems was raised; to carry out the evaluation, the questionnaire called KIDSCREEN 52 was used, which consists of 52 closed questions divided into 10 dimensions, providing an idea of how the adolescent feels. The following results were obtained: The Health-related quality of life in adolescents of an Educational Institution of Ventanilla, where 430 students representing 59.1% have high quality of life and 298 students representing 40.9% have medium quality of life; according to its dimensions, adolescents from the Educational Institution of Ventanilla, its most affected dimension was economic resources, with 64 students representing 8.8% having a low level. In the relationship with parents' dimension, it appreciates that 3.4% of students do not have a good family atmosphere since they do not live with their parents, but with uncles and/or grandparents.

1. Introduction

Health-related quality of life (HRQL) defined as "individual and group perceived physical and mental health". It has been used to evaluate medical care from the patient's point of view and as a subjective indicator of social well-being [1]. It corresponds to the perception that people have of their level of subjective well-being, considering various aspects of their lives and their impact on their state of health. Studies in HRQL were initially carried out in a population that presented both acute and chronic pathologies and covered practically all areas of medical specialties; subsequently, the interest was concentrated in studying HRQL in the general population and then in the population of children and adolescents. Currently, studies in school children and adolescents have increased [2]. HRQL in childhood and adolescence is useful for describing the health of a population and establishing comparisons between subgroups with certain characteristics (for example, sociodemographic, regional, etc.). In pediatric clinical practice, the study of HRQL can facilitate knowledge of health status, study the effect of different treatments, identify levels of morbidity and facilitate clinical decisions [3]. Adolescents are considered as individuals in constant development, that is why they need to grow in a balanced and healthy way within a safe and structured environment, in order to develop positively on a physical, educational, emotional and social level [4].

In the research work [5] in 2017, it was indicated that a total of 550 adolescents aged 16 ± 1 years completed the questionnaires. 39% consumed alcohol in the last month, 31% smoked, 33% used marijuana, and 33% admitted using multiple drugs. High-risk alcohol consumption was associated with a lower perception of psychological well-being, self-perception, and the school environment. This last dimension was affected in those who admitted the use of marijuana during the last month. Multiple drug use (three substances) was associated with a lower perception of physical and psychological well-being, self-perception, relationship with parents, family life and the school environment; in conclusion, high-risk alcohol consumption and multiple drug use (three substances) have a negative impact on the HRQL of school-age adolescents.

*Corresponding Author: Hernan Matta-Solis, Mr., +51 1 999751065 & hmatta@uch.edu.pe

In [6] in 2019, the authors explain that of the 723 adolescents, 13% reported having been bullied; there were no gender differences within this finding. However, more boys than girls reported that they had bullied others. Both bullying and intimidation of others were associated with a lower HRQL; however, being bullied was associated with lower scores. Greater self-efficacy was associated with better HRQL. Self-efficacy contributed significantly to predicting variation in HRQL. In conclusion, being involved in bullying, as a victim or bully, is associated with a lower HRQL. The association between general self-efficacy and HRQL indicates that self-efficacy could be a resource to increase HRQL in adolescents. Their findings highlight the importance of focusing on self-efficacy beliefs as an intervention strategy to improve self-efficacy and HRQL in adolescents involved in bullying.

In [7] in 2018, the authors presented that girls and adolescents aged 17-18 years presented significantly higher psychosomatic symptoms, both groups also scored worse in all dimensions of HRQL, although only the dimensions related to physical well-being, mental well-being and mood and stress reached significance. All psychosomatic symptoms were inversely associated with the 10 dimensions of the KIDSCREEN-52. The regression models performed showed that sadness, difficulty concentrating and difficulty sleeping were the predictors of the worst HRQL in both sexes and age groups, and these variables explained between 30% and 41% of the variance in the HRQL of adolescents. In conclusion, psychosomatic symptoms are more frequent in girls and older adolescents, and predictors of worse HRQL. It is important to differentiate them from medical conditions to avoid unnecessary interventions. As expressions of emotional distress, they must be evaluated and treated comprehensively because they interfere with daily life and increase the vulnerability of adolescence.

The general objective of the study is to determine the health-related quality of life in students of an educational institution. To do this, reviewing various studies verifies that there is no research on this area, neither from this population that currently has problems most of them in the psychological field. In the review of national and international scientific literature, it is evident that there are few scientific researches that use the KIDSCREEN 52 in their work on health-related quality of life, and it is evident that most of the researchers are international. In Callao, and specifically in the place where the research work is carried out, there is no evidence of other research on this topic, which is why this research wants to fill the knowledge gap.

The technique used is the survey. The data collection instrument was the KIDSCREEN 52, which consists of 52 items distributed in its ten dimensions (physical well-being, psychological well-being, mood, self-perception, autonomy, relationship with parents, financial resources, friends and social support, social environment and social acceptance). After completing the field work, the data was entered into a matrix in Microsoft Excel 2016, and then exported to the IBM SPSS Statistics Base 23.0 statistical package, in which the mainly descriptive analysis was performed.

2. Methodology

2.1. Population and sample

In the present study, it worked with the total population (N=728), which is made up of adolescents between 12 and 18 years old of the 2019 school year from a public school in the Ventanilla district, the 21 classrooms have around 700 students, each participant was selected according to the inclusion and exclusion criteria of the study; in addition, each participant was informed about the objectives and scope of the study, and then invited to bring the corresponding informed consent and assent signed.

Table 1: Distribution of classrooms at the secondary level of the Jorge Portocarrero Rebaza educational institution

Grade	Shifts	Classrooms	Total students
First	Afternoon	5	164
Second	Afternoon	5	177
Third	Morning	4	162
Fourth	Morning	4	119
Fifth	Morning	3	106
Total		21	728

2.2. Inclusion and Exclusion Criteria

Inclusion criteria:

- All students who present the consent of their parents and informed assent by them.
- Students between 8 and 18 years old.
- Students who have attended the day of the survey.

Exclusion criteria:

- Students over 18 years old.
- Students who do not wish to take the survey.

2.3. Analysis of the variable

The present study presents Health-related quality of life as the main variable, according to its nature, it is a qualitative variable and its measurement scale is ordinal.

The health-related quality of life (HRQL) corresponds to the perception that people have of their level of subjective well-being, considering various aspects of their lives and their impact on their state of health [2]. It is the perception that high school adolescents of II.EE Jorge Portocarrero Rebaza has about their position in life, in the cultural context and value systems in which they operate and in relation to their objectives, expectations, standards and concerns also, it is expressed by physical well-being, psychological well-being, mood, self-perception, autonomy, relationship with parents, financial resources, friends and social support, social environment and social acceptance, which will be measured with the KIDCREEN 52 questionnaire.

The technique used in the study for the data collection process was the survey, this technique contemplates a set of standardized procedures that wants to obtain a significant amount of data quickly and efficiently.

Regarding the data collection instrument that allowed to measure the main variable Health-related quality of life, the KIDSCREEN 52 questionnaire was used. It is financed by the European Commission within the "Quality of life and management of living resources" Fifth Framework Programmed (QLC-CT-2000-00751) and was held for 3 years; they are a group of instruments developed and standardized for the topography of Health-related quality of life (HRQL - Health related to quality of life), in children and adolescents aged 8 to 18 years [8].

The KIDSCREEN-52 tool was developed jointly by 13 European countries, including Spain, following a simultaneous intercultural protocol. This ensures that future comparisons can be made between different countries. Later it has been adapted to different countries in Latin America, such as Venezuela and Argentina. The instrument was constructed using psychometric studies based on classical test theory and more advanced models, such as Item Response Theory (IRT). Based on the KIDSCREEN implementations, a computer-adaptive trial version was also developed using an IRT-based methodology. The original Spanish version of the KIDSCREEN-52 (KIDSCREEN-52-ES) was culturally adapted according to the International Test Commission (ITC) guidelines. We carry out the adaptation following the instructions of the KIDSCREEN Group. There were two stages: First, an adjusted linguistic validation of the instrument was performed. This consisted in the forward and backward translation of the instrument and harmonization until obtaining a semantic and conceptually equivalent version to the original. To maintain the validity of the content of the instrument, the items had to be formulated and expressed in an acceptable and culturally relevant way for Colombian children and adolescents. Subsequently, a pilot study was conducted to examine the validity of the instrument and correct any content or format errors [9].

This instrument is made up of 52 main items which are distributed in 10 dimensions, being these:

Physical well-being: This dimension measures the physical activity of the person, the sensation of energy and everything about the physical part of the person.

- High: Indicates that the participant performs well in physical activities and maintains an adequate physical level.
- Medium: Indicates that the participant alternately performs physical activities but also maintains a good level of physical well-being.
- Low: Indicates that the participant performs minimal physical activities.

Psychological well-being: In this dimension, the person's emotions and satisfaction with life are measured.

- High: Indicates that the participant has their emotions in constant control.
- Medium: Indicates that the person may at times be unable to control their emotions.
- Low: Indicates that the person cannot control their emotions when faced with any situation.

Mood: The level of stress that the person presents and if it presents depressive states are measured.

- High: The person can control and maintain a balance regarding stress.
- Medium: The person may at times not be able to control stress adequately.
- Low: The person may present depressive states due to not being able to adequately control stress.

Self-perception: The perception of its physical appearance and body image and if it feels satisfied with it.

- High: The person appreciates and values their body appearance and feels satisfied.
- Medium: The participant may sometimes not appreciate their body appearance.
- Low: The participant is not satisfied with their image and physical appearance.

Autonomy: Opportunities for activities in their free time are perceived.

- High: The participant has opportunities to do activities in their free time.
- Medium: The participant has some opportunities for activities in their free time.
- Low: The participant has few opportunities to do activities in their free time.

Relationship with parents: The family atmosphere is perceived.

- High: The participant has a very good family relationship.
- Medium: Presents a medium family atmosphere, there may be some problems between family members.
- Low: It presents a bad family atmosphere.

Economic resources: The perception of the family's financial capacity is measured.

- High: Presents adequate family financial capacity to meet the family's needs.
- Medium: Presents an adjusted family financial capacity but still satisfies the basic expenses of the family.
- Low: Presents an adjusted financial capacity.

Friends and social support: The relationship with classmates and/or friends is measured.

- High: Good relationship with classmates.
- Medium: Good relationship with some classmates.
- Low: Bad relationship with classmates.

Social environment: The perception of the ability to concentrate and cognition is measured, in addition to feelings about school.

- High: The participant presents a good concentration and cognition within the school.
- Medium: Presents a regular cognition within the school.

- Low: The participant does not show adequate concentration and cognition in school.

Social acceptance: Measures the feeling of rejection from friend or people around.

- High: The participant does not present any rejection from its colleagues.
- Medium: Presents some rejection of its colleagues.
- Low: Presents rejection by their peers, the participant suffers from bulling by their peers.

Reliability of the KIDSCREEN 52 data collection instrument

The reliability of the instrument was determined based on the Cronbach's alpha statistic, which presents an internal consistency index of 0.885 (α> 0.6). Therefore, a high level of general reliability of the instrument is identified.

Reliability	Cronbach's alpha	0,885
	No. of elements	52

Validity of the data collection instrument Health-related quality of life KIDSCREEN 52

To determine the validity of the instrument, the Kaiser-Meyer-Olkin (KMO) sample adequacy index and the Barttlet sphericity test were used. After the analysis, a sample adequacy coefficient of 0.928 and a significance value of 0.000 (x2 = 20481.828; g.l. = 1326; p <0.05) were obtained in the Barttlet sphericity test. Given the above data, an acceptable validity of the instrument can be identified.

	Measure of sampling adequacy of Kaiser-Meyer-Olkin		0,928
Validity	Bartlett's test of sphericity	X^2 approximate	2048
		gl	1,828
		Sig.	1326
			0,000

2.4. Data Collection

Authorization and prior coordination for data collection

In order to start the data collection process, steps were taken to access the institution where the field work will be carried out.

At first during the month of September of the year 2019, a request letter of introduction to the Professional School of Nursing of the Universidad de Ciencias y Humanidades to formally manage the permit and be able to carry out the surveys at the Jorge Portocarrero Rebaza educational institution.

In a second moment, it coordinated with the director of the Jorge Portocarrero Rebaza educational institution, Mr. Rafael Salvador Mimbela, and the importance of the research was explained to him, authorization was requested for the development of the work, to which he agreed.

Application of data collection instrument

On October 4th, 2019, the survey was carried out in the students who attended the educational institution, they were 11 classrooms in the morning shift and 10 classrooms in the afternoon shift with 40 students for each classroom. The time used in each classroom was approximately 25 minutes. In each classroom, an authorization document was requested from the parent or guardian of the adolescent, since the informed consent had to be signed. At the time of the delivery of the questionnaire, each classroom was informed how it should be filled out correctly since they were closed-response questions based on marking with an "X" in each question, the students were clarified in their doubts having more doubts the students from first years of secondary study. At the conclusion of the data collection procedure, the correct filling and coding of each of the sheets with the instrument was verified. It is important that a professional carry out this type of research, in this case I as a nursing professional went to explain the correct filling of the questionnaire since nursing is prepared in all areas of health such as the psychological area among others and therefore prepared for any doubt of the group of questions presented by each dimension of the instrument. As the reader can see in the Figure 1, it shows the application of the survey in the Education Institution.



Figure 1: Application of the surveys in the Educational Institution

2.5. Statistical analysis methods

At the end of the data collection process, a matrix prepared in the Microsoft Excel spreadsheet will be entered, then these will be exported to the statistical program IBM SPSS Statistics Base 23.0, it will proceed to its corresponding analysis. As it is a descriptive design research, the methods of descriptive statistics will be used for data analysis; where the tables of absolute and relative frequencies will be used, the measures of central tendency, among other tests for the analysis of the sociodemographic information, also corresponding to the variables and dimensions involved. Finally, tables and graphs will be obtained according to the sociodemographic data, general and specific objectives of the study; for later description and discussion.

3. Results

Below are three summary tables; the first will show the demographic data of the students who participated in the study, the second table shows the result of the variable Quality of life related to health, and the third table shows the results of each dimension.

Table 2: Sociodemographic data in students of an Educational Institution of Ventanilla (N = 728)

Age	**Minimum – Maximum**	**Media (D.T.)**
	12 – 18	14.06 (± 1.505)
Sex	**Frequency (f)**	**Percentage (%)**
Female	357	49.0
Male	371	51.0
Degree of the interviewee	**Frequency (f)**	**Percentage (%)**
First high school	164	22.5
Second high school	177	24.3
Third high school	162	22.3
Fourth high school	119	16.3
Fifth high school	106	14.6
Study shift	**Frequency (f)**	**Percentage (%)**
Morning	389	53.4
Afternoon	339	46.6
Type of Family	**Frequency (f)**	**Percentage (%)**
Nuclear	424	58.2
Single parent	176	24.2
Extended	33	4.5
Expanded	58	8.0
Reconstituted	26	3.6
Family equivalent	6	.8
Single person	5	.7
Disability, illness or medical problem	**Frequency (f)**	**Percentage (%)**
No	622	85.4
Yes	106	14.6

In Table 2, it has the sociodemographic data of the study participants, a total of 728 students. Regarding the age of the study participants, it was determined that they fluctuate in a range from 12 to 18 years old, with an average of 14.06 (+ 1.505) years old. Regarding sex, it was found that 51.0% are male. Fourth and fifth year of secondary school have the lowest percentages of students with 16.3% and 14.6% respectively. The majority of respondents, equivalent to 53.4%, study in the morning shift. 58.2% of the sample claims to be part of a nuclear family. Only 14.6% claim to have a disability, difficulty or some medical problem.

Table 3: Health-related Quality of life in adolescents of an Educational Institution of Ventanilla (N = 728)

		N	**%**
Health-related quality of life	**Low**	0	0.0
	Medium	298	40.9
	High	430	59.1
Total		728	100

In Table 3, it can see the health-related quality of life in adolescents of an Educational Institution of Ventanilla, where 430 students representing 59.1% have high quality of life and 298 students representing 40.9% have medium quality of life.

Table 4: Health-related Quality of life according to its dimensions, in adolescents of an Educational Institution of Ventanilla (N = 728)

		N	**%**
Physical well-being	**Low**	23	3.2
	Medium	360	49.5
	High	345	47.4
Psychological well-being	**Low**	6	0.8
	Medium	239	32.8
	High	483	66.3
Mood	**Low**	25	3.4
	Medium	219	30.1
	High	484	66.5
Self-perception	**Low**	7	1.0
	Medium	260	35.7
	High	461	63.3
Autonomy	**Low**	27	3.7
	Medium	333	45.7
	High	368	50.5
Relationship with parents	**Low**	25	3.4
	Medium	267	36.7
	High	436	59.9
Economic resources	**Low**	64	8.8
	Medium	400	54.9
	High	264	36.3
Friends and social support	**Low**	14	1.9
	Medium	254	34.9
	High	460	63.2
Social environment	**Low**	18	2.5
	Medium	321	44.1
	High	389	53.4
Social acceptance	**Low**	18	2.5
	Medium	166	22.8
	High	544	74.7

In Table 4, it can see the Health-related quality of life according to its dimensions in adolescents of an Educational Institution of Ventanilla, where the most affected dimension was that of economic resources with 64 students representing 8.8% having a level low. In the physical well-being dimension, it is noted that it presents a medium quality of life with a percentage

of 49.5%, giving an idea that adolescents from that educational center practice physical activity both inside and outside the educational institution. In the psychological well-being dimension, it is noted that they present a high quality of life of 66.3%, revealing that adolescents from the educational institution present emotions where they prioritize the positive ones and satisfaction with life. In the mood dimension, it is noted that it presents a high level with a total of 66.5%, thus expressing that the students present a balanced level of stress. In the self-perception dimension, it can observe that it presents a result of 63.3%; which corresponds to a high level, demonstrating that students have a good perception of their body image and physical appearance. In the autonomy dimension, it is noted that it presents a high level with a percentage of 50.5% showing that the educational institution does allow them to have time for their activities in their free time. In the relationship with parents' dimension, it appreciates that 3.4% of students do not have a good family atmosphere since they do not live with their parents, but with uncles and/or grandparents. In the economic resources dimension, it is noted that they present a medium level with a percentage of 54.9%, reflecting that the financial capacity of the parents is not low but that it does not cover the maximum with respect to the expenses of the students, since it was observed that most students have one or more siblings studying in lower grades. Regarding the friends and social support dimension, it presents a high level with a total of 63.2%, reflecting that the students have good communication with their classmates or friends who do not study with them. In the social environment dimension, it reflects a high level with a total of 53.4%, which shows a good feeling about school, also a good ability to concentrate and a positive feeling about school. Finally, in the social acceptance dimension, it is observed that a total of 2.5% presents a low level, reflecting that in the educational institution there are some students who express rejection of their peers.

4. Discussion

In the following study, the issue of health-related quality of life is raised from the point of view of promoting adolescent health, which wants to contribute to educational institutions which are part of adolescent educational training, which must include programs that allow adolescents to enhance their ability to interrelate and adequately shape their personality.

The participants in this study were high school students from a public school in the province of Callao, who had an average age of 14 years, from a lower-middle socioeconomic level in the district where the school is located, who mostly belonged to a nuclear family. Most of the students in this study did not have any significant physical disabilities, illnesses, or medical problems.

At a general level, it was found that the students who participated in this study had high levels of HRQL, this because they were very clear about their personal goals and objectives, optimistically seeing their future, as in other studies on the same subject, such as the study of Hidalgo et al [1], in this study it is evident that the present one focuses on school adolescents, which results in a low level HRQL since the students affirmed that they presented bullying among their classmates, thus demonstrating a Health-related quality of life from medium to low, therefore it is asked to delve into the dimensions that present difficulties.

Compared to other studies, this research only worked with the variable Health-related quality of life since the instrument used focuses directly on that variable and the type of population, therefore the KIDSCREEN 52 compared to its two existing types, holistically assesses the age group of people who carried out the research work.

The research work wants to demonstrate the related quality of life in adolescents since in these times, that group of people present problems most of them in the psychological part. In the following research work, it wants to demonstrate that the adolescent is a holistic person and therefore it must be studied in all its dimensions; In this study group, it verified that the educational institution presents few problems due to the result found that was between a medium and high level with respect to the variable Health-related quality of life.

5. Conclusions

It is concluded that the Health-related quality of life is the set of manifestations that a person presents most of them in the subjective part (emotions, sensations, among others); this is influenced by the state of health that is presented at the time; this already gives an idea of how the adolescent feels at the school stage. In this case, what we are looking for with this study is to give us an idea of how the adolescent feels at that moment.

It is concluded that the educational institution, the director provided the support for the realization of the survey and later the processing of the information given by the study population. The study was also successfully carried out thanks to the teachers of the institution since they offered class time and especially to the students who were the study population, since they were predisposed to collaborate with the study.

As a future research work, other health professionals are expected to carry out research work in localities where populations are vulnerable, such as the case of Pachacutec-Ventanilla; since other research could be carried out with other groups of people and thus obtain relevant results to complement the improvement of the Peruvian population.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] C. Hidalgo, T. Molina, R. Molina, R. Sepúlveda, V. Martínez, Montaño, Bullying y calidad de vida relacionada con la salud en adolescentes escolares Chilenos. Rev Med Chil. 2015; **143**(6):716–23. http://dx.doi.org/10.4067/S0034-98872015000600004

[2] E. González et al., "Diferencias de género en la calidad de vida relacionada con la salud en adolescentes escolarizados chilenos Gender differences in health-related quality of life of Chilean adolescent students. Rev Med Chil. 2016; **144**(3):298–306. http://dx.doi.org/10.4067/S0034-98872016000300004

[3] L. Rajmil, "Advances and challenges in the measurement of health related quality of life in the child and adolescent population. An Pediatr [Internet]. 2019; **90**(5) :261–2.: https://doi.org/10.1016/j.anpedi.2019.01.011

[4] I. Bica , J. Costa, "Original Article Sociodemographic infl uence in health-related quality of life in adolescents. Acta Paul Enferm. 2020;**33**:1–7. https://doi.org/10.37689/acta-ape/2020ao0054

[5] F. Vilugrón, C.A. Hidalgo et al., "Uso de sustancias psicoactivas y calidad de vida relacionada con la salud en adolescentes escolarizados. Rev Med Chil. 2017; **145**(12):1525–34. http://dx.doi.org/10.4067/s0034-98872017001201525

[6] K. Haraldstad et al., "Associations between self-efficacy, bullying and

health-related quality of life in a school sample of adolescents: a cross-sectional study. BMC Public Health. 2019; **19**(1):1–9. https://doi.org/10.1186/s12889-019-7115-4

[7] R.M. Fuentes-Chacón et al., "Psychosomatic symptoms as an expression of the deterioration of the health-related quality of life in adolescents. Aten Primaria. 2018; **50**(8):493–9. doi: 10.1016/j.aprim.2017.06.009

[8] KIDSCREEN-52 - kidscreen.org [Internet]. Available from: https://www.kidscreen.org/español/cuestionario-kidscreen/kidscreen-52/

[9] J. Valencia, P. Galvañ, C. Martínez, C. García, R. Martínez. Adjusted linguistic validation and psychometric properties of the Colombian version of KIDSCREEN-52. J Child Heal Care. 2019; **23**(1):20–34. https://doi.org/10.1177/1367493518777291

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1340-1344 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Quality of Life in Patients with Type 2 Diabetes of the Central Hospital of the Peruvian Air Force, 2019

Jared Zavala-Izaguirre[1], Fanny Mego-Llanos[1], Sarita Cornejo-Quispitongo[1], Brian Meneses-Claudio[2,*], Hernan Solis-Matta[1], Lourdes Matta-Zamudio[1]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 05 December, 2020
Online: 16 December, 2020

Keywords:
Quality of life
Type 2 diabetes
Diabetes control
Social burden
Sexual functioning

ABSTRACT

This research shows the study carried out on the Health Related to Quality of Life (HRQL) focuses on aspects related to the perception of health experienced and declared by the person, in different dimensions such as physical, mental, social, general perception of health and satisfaction achieved measured at different levels; the objective of the study is to determine the quality of life in patients with type 2 diabetes at the Central Hospital of the Peruvian Air Force, 2019. To determine the validity of the instruments, the Kaiser-Meyer-Olkin (KMO) sample adequacy index was used and the Bartlett sphericity test, which result in a significance value of 0.005, being an acceptable validity of both instruments. Among the most relevant results is that the following dimensions: Energy and Mobility, 147 patients representing 73.5% have low level. Likewise, regarding to Social Burden, 156 patients representing 78.0%, have a low level. Lastly, regarding the Sexual Functioning dimension, 178 patients representing 89.0%, have a low level; patients significantly affect quality of life. "Non-adherence to drug treatment" represents 159 patients, of which 115 who represent 72.3% have a low level with respect to their quality of life, of 44 patients who represent 27.7% have a high level regarding their quality of life; "Adherence to pharmacological treatment" 41 patients representing 92.7% have a low level regarding their quality of life and 3 patients who represent 7.3% have a high level regarding their quality of life.

1. Introduction

According to the International Diabetes Federation, 382 million adults worldwide had diabetes and 5.1 million died from the disease. About half (48%) of those who died were under the 60 years old. Three of four people living with diabetes (352 million) are in an active age (that is, between 20 and 64 years old) this number is projected to increase to 417 million by 2030 and 486 million by 2045 [1].

They also estimate that by 2035, if current trends continue, the global prevalence of diabetes will rise to 592 million, and the World Health Organization predicts that by 2030 it will become the seventh leading cause of death in the world [2].

This disease is considered as one of the main reasons for deaths worldwide, due to the poor quality of life without measuring the complications that it will bring them in the long term [3].

This non-communicable pathology is a worldwide concern, since due to the multiple organic complications it causes deaths, mostly related and complex with the organs of the cardiovascular system [4].

Diabetes is a chronic disease that is released when the pancreas does not produce insulin necessary for the body (a hormone that regulates the level of sugar, or glucose in the blood) or when the body cannot vigorously use the insulin it produces; they have a substantially increased risk of cardiovascular disease. It has also been related to hypertension and hyperlipidemia, it has vascular complications and is considered one of the four non-communicable diseases [5], [6].

Its own symptoms such as thirst, polyuria, blurred vision, weight loss, and sometimes polyphagia is recurrent. Perennially, the symptoms are not severe or may be distant; consequently,

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & Email: bmeneses@uch.edu.pe

hyperglycemia can provoke functional and pathological changes for a long time before diagnosis [7].

Diabetes is a chronic disease, and the associated stress of long-term symptoms, medical appointments, research, and daily treatments can affect sexual function through a person's mood, anxiety levels, fatigue, and well-being. In general, people with diabetes may have comorbidity of hypertension, hyperlipidemia, cardiovascular disease, or other endocrine dysfunction, and their associated treatments, may affect sexual function [8].

In [9], these problems cannot be addressed without a solid medical and scientific base that allows local solutions to be identified and proposed; unfortunately, the scientific production on diabetes in Peru is limited with only 81 scientific articles published in Web of Science and Scopus between 1996 and 2015.

In [10], diabetes was positioned as the fifth cause of death with 10 528 (7.11%) of the cases, in addition, the numbers of patients with this disease who died from causes of kidney complications were from 2 266 cases (1,191 men and 1,075 women) in 2011 to 3,190 cases (1,638 men and 1,552 women) in 2016.

In [11], this study shows that the mean time of evolution of the disease was 11.37 years old. Quality of life was self-perceived as deficient in 67.57% of older adults at the expense of satisfaction with their health (75.68%).

In [12], it is believed that, in Peru, more than 2 million people suffer from diabetes. The national average is 8% and in Lima, up to 10%. However, it is good to clarify that, in many regions, up to at least 50% of the population does not know if they have diabetes.

According to the [13], 4.0% of the population residing on the Coast reported having diabetes mellitus, being the urban area with the highest incidence of this disease (4.1%) and in the rural area only 2.7%; followed by the jungle regions with 1.9%, the urban area registers 2.7% and rural 1.1%; while the highlands area register 1.6 %, urban area 2.5% and rural with 1.0%.

The objective in this study is to determine the quality of life of patients with diabetes mellitus who come to the consultation of the Central Hospital of the Peruvian Air Force, using the Diabetes questionnaire 39.

2. Methodology

2.1. Research Design Approach

The present study has a quantitative, descriptive, and cross-sectional approach. A survey was carried out as a data collection technique, by applying a questionnaire as a documentary instrument. The study population was made up of patients with diabetes mellitus from the Central Hospital of the Peruvian Air Force [14].

2.2. Inclusion Criteria

All continuous patients over 18 years old with type 2 diabetes mellitus, regardless of the sex treated at the Central Hospital of the Peruvian Air Force.

2.3. Exclusion Criteria

- Patients with type 1 diabetes mellitus or gestational diabetes.
- Patients seen for the first time.
- Resigned with intellectual inability to recognize the question, and those who refuse to report the study.

2.4. Measurement Techniques and Instruments

Data collection technique: The technique used for this study was the survey, it is not the same as the questionnaire, however, the questionnaires are usually used in a survey, which requires a systematic way of collecting the data for the later to be able to analyze them statistically and obtain conclusions about the study population [15].

To measure the quality of life of patients diagnosed with type 2 diabetes mellitus, it will be applied with the diabetes 39 instrument, which includes 39 items that evaluate five domains of the patient's life: energy and mobility (15 items), control of diabetes (12 items), anxiety and worry (4 items), social burden (5 items) and sexual functioning (3 items). Each item is calculated from the evaluation made by the patient itself, in relation to its quality of life, how much was affected during the last month by action or activity that each item expresses- placing an "X" on a scale that is presented as a continuous line, with vertical marks that delimit spaces in which the numbers from 1 to 7 are identified, in which, on the value scales, the number 1 represents the quality of life that was not affected at all, and number 7, extremely affected. Two final items are incorporated into the five domains that rate the patient's perception about their quality of life in general, with a range from 1 (minimum) to 7 (maximum), and the severity of diabetes mellitus with a 1 (no severity) to 7 (extremely serious) [16].

The data collected was entered into a Microsoft Excel 2013 data matrix, to be later exported to the statistical program IBM SPSS Statistics Base 24.0.



Figure 1: Endocrinology area for conducting surveys at the Central Hospital of the Peruvian Air Force.

3. Results

To determine the validity of the instrument, the Kaiser-Meyer-Elkin sample adequacy index (KMO) and the Bartlett's test of sphericity were used. For the Quality-of-Life Test for patients with type 2 diabetes, a sample adequacy coefficient of 0.834 and a significance value of 0.005 (x^2 = 2633,620; gl = 741; p <0.05) were obtained in the test of Bartlett's sphericity.

The reliability of the instrument was determined based on the Cronbach's Alpha statistic. For the instrument Quality of life test

for patients with type 2 diabetes, it was determined that it presents an internal consistency index of 0.741 (α> 0.6); while the Morisky-Gree test instrument of adherence to pharmacological treatment has a coefficient of 0.591 (α <0.6). Therefore, a high level of general reliability of the Quality-of-Life Test is identified for patients with type 2 diabetes.

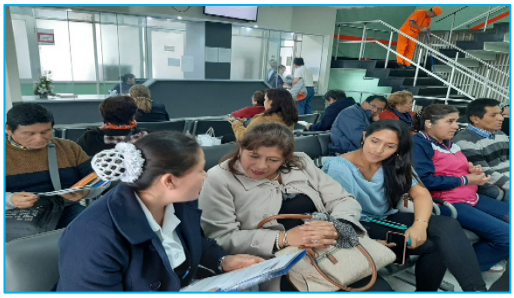




Figure 2: Application of the Survey in the Central Hospital of the Peruvian Air Force

3.1. Sociodemographic characteristics

Regarding the age of the study participants, it was determined that they fluctuate in an interval of 20 to 90 years, with an average of 56.45 (± 13,847) years old. Regarding sex, 60% of the sample is male. Likewise, 72.1% of the sample lives with their parents.

Table 1: Quality of life in patients with type 2 diabetes of the Central Hospital of the Peruvian Air Force, 2019 (N = 200)

Quality of life	N	%
Low Level	153	76,5
High Level	47	23,5
Total	200	100,0

In Table 1, it can observe the quality of life in patients with type 2 diabetes of the Central Hospital of the Peruvian Air Force, where 153 patients representing 76.5% have a low level followed by 47 patients representing 23, 5% have a high level.

In Table 2, it can see the quality of life according to its dimensions, in patients with type 2 diabetes of the Central Hospital of the Peruvian Air Force, where the most affected dimension was sexual functioning with 178 patients representing 89% have low level followed by 22 patients representing 11% have high level.

Table 2: Quality of life in patients with type 2 diabetes according to its dimensions, from the Central Hospital of the Peruvian Air Force, 2019 (N = 200)

Dimensions		N	%
Energy-mobility	**Low Level**	147	73,5
	High Level	53	26,5
Diabetes control	**Low Level**	132	66,0
	High Level	68	34,0
Control of anxiety - concern	**Low Level**	85	42,5
	High Level	115	57,5
Social burden	**Low Level**	156	78,0
	High Level	44	22,0
Sexual Functioning	**Low Level**	178	89,0
	High Level	22	11,0

The importance of obtaining these results in continuing patients with diabetes mellitus is to be able to identify in which of the dimensions are most affected by this disease and to be able to act in a timely manner in the dimension and above all to teach how to be responsible and aware of the disease. So, they could learn that diabetes is a chronic disease that involves the variation of lifestyles. For its treatment, it is essential that people learn to manage it correctly to have a good quality of life and avoid possible complications.

Table 3: Contingency table between the variables Quality of life and Adherence to pharmacological treatment and chi-square tests

			Quality of life		Total
			Low Level	High Level	
Adherence to pharmacological treatment	**Non-adherent**	Count	115	44	159
		% within APT	72,3%	27,7%	100,0 %
	Adherent	Count	38	3	41
		% within APT	92,7%	7,3%	100,0 %
Total		Count	153	47	200
		% within APT	76,5%	23,5%	100,0 %
Pearson's chi-square			Valor X^2	g.l.	Sig. (p)
			7,513	1	0,006

APT = Adherence to Pharmacological Treatment

In Table 3; on adherence, it can observe in patients with type 2 diabetes of the Central Hospital of the Peruvian Air Force, where the "Non-adherence" represents 159 patients, 115 of which is equivalent to 72.3% have low level and 44 patients 27.7% with

high level; "Adherence" 41 patients representing 92.7% have low level and 3 patients representing 7.3% have high level.

It is important that people with diabetes and their families should receive education about adherence and non-adherence; ignorance is one of the reasons for not achieving the necessary benefits that medications can provide in their disease; diabetes is a chronic disease that affects different organs of the human body.

4. Discussions

Quality of Life is a representative variable that allows us to measure the impact it has on a certain pathology (in this case Diabetes Mellitus type II) in the person's life, which has an impact on various biological and psychosocial levels in the individual health.

The results obtained according to the Diabetes 39 questionnaire demonstrate that of a population made up of 200 diagnosed continuing patients who are treated at the Central Hospital of the Peruvian Air Force, the dimensions most affected are determined that, regarding the Energy and mobility dimension, 147 is equivalent to 73.5% have low level. Likewise, regarding to social burden, 156 is equivalent to 78.0%, have low level. Lastly, regarding the Sexual Functioning dimension, 178 is equivalent to 89.0% have low level.

These results coincide and are reinforced by the study carried out in [17]. They show that patients with Diabetes Mellitus regarding their Quality of Life, who were surveyed with 101 diabetics; the most affected domains in order of score were: energy and mobility (60) equal 59.4%, anxiety-concern (62) equals 61.3% and sexual functioning (66) equals 65.3%.

Regarding adherence, it can observe in patients with type 2 diabetes of the Central Hospital of the Peruvian Air Force, where non-adherence represents 159 patients, 115 representing 72.3% have low level and 44 patients representing 27.7% have high level; adherence 41 patients representing 92.7% have low level and 3 patients representing 7.3% have high level. No relationship was found between the percentage of adherence to treatment with sex, age, educational level, and employment status; it is the lack of awareness about its treatment.

According to [18], it was found, the low level was 21 (32.8%) and with high level adherence 24 (37.5%). No relationship was found between the percentage of adherence to treatment with sex, age, educational level, and employment status.

In the analysis of each dimension of the instrument, it was observed that there were significant differences. In the dimension between quality of life and adherence to treatment in patients with treatment, as a strength of the study, it can be highlighted that there was no rejection by the patients, all the questionnaires were correctly filled out, with no missing data or loss in the sample. It is also important to note that a widely used instrument with a good internal consistency index was used.

Among the limitations of the study, the selection of patients is found, since as it is not a probabilistic sampling, there may be the possibility of selection bias. For future studies on the subject in primary care, it would be interesting to carry out a random sampling that provides greater reliability in the results obtained. Another important aspect to control would be trying to match the sample by sex.

5. Conclusions

It is concluded that the dimensions that highly affect patients are Energy and Mobility, Social Burden, Sexual Functioning, having a low level of more than 70% of the sampled cases.

In addition, the results of therapeutic adherence show that 72.3% of patients present a low level with respect to their quality of life, therefore, compliance with treatment, physical activities and a balanced diet help diabetic patients a better recovery and a balanced quality of life.

6. Recommendation

A restructuring of promotional preventive intervention programs targeting such patients with type II diabetes is recommended. This intervention program must involve a multidisciplinary team of professionals, who see aspects related to the disease and the environment of the person and are accelerators for their recovery from physical, biological, and psychosocial health.

It is suggested an awareness program, educational talks for both the patient and family members for their support and control, home health personnel visit programs.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] R. Williams, Guía de incidencia política de la novena edición del atlas de la diabetes de la FID 2019, International Diabetes Federation, **9**, 1–28, 2019.

[2] G. Chen, A. Iezzi, J. McKie, M. Khan, J. Richardson, "Diabetes and quality of life: Comparing results from utility instruments and Diabetes-39.," Diabetes Research and Clinical Practice, **109**(2), 326–333, 2015, doi:10.1016/j.diabres.2015.05.011.

[3] C.F.D.C. and Prevention, Informe nacional de estadísticas de la diabetes, National Center for Chronic Disease, 1–33, 2017.

[4] Organización Panamericana de la Salud, El número de personas con diabetes en las Américas se triplicó desde 1980, 2016.

[5] P. Miguel, Y. Sarmiento, A. Mariño, Y. Llorente, T. Rodriguez, M. Peña, "Prevalencia de enfermedades crónicas no transmisibles y factores de riesgo en adultos mayores de Holguín," Revista Finlay, **7**(3), 155–167, 2017.

[6] R. Wang, P. Zhang, Z. Li, X. Lv, H. Cai, C. Gao, Y. Song, Y. Yu, B. Li, Y. Cui, "The prevalence of pre-diabetes and diabetes and their associated factors in Northeast China: a cross-sectional study," Scientific Reports, **9**(1), 1–8, 2019, doi:10.1038/s41598-019-39221-2.

[7] J. French, A. Godoy, ¿Cuál es la definición de diabetes? Criterios diagnósticos (prueba/s a realizar [glucemia plasmática en ayunas, sobrecarga, etc.]) y puntos de corte, Fundación RedGDPS, 1–3, 2015.

[8] E. Holloway, "Sexual problems in diabetes," Medicine (United Kingdom), **47**(2), 106–109, 2019, doi:10.1016/j.mpmed.2018.11.004.

[9] R. Carrillo, A. Bernabé, "Diabetes mellitus tipo II en Perú: una revisión sistemática sobre la prevalencia e incidencia en población general.," Rev Peru Med Exp Salud Publica, **36**(1), 36–36, 2019, doi:10.17843/rpmesp.2019.361.4027.26.

[10] J. Guerrero, L. Parra, J. Mendoza, "Autoeficacia y calidad de vida en pacientes con diabetes mellitus tipo 2 sometidos a hemodiálisis," Revista Cubana de Salud Publica, **42**(2), 193–203, 2016.

[11] D. Jiménez, P. Casado, R. Santos, D. Jiménez, G. Hernández, "Percepción de la calidad de vida en pacientes adultos mayores con diabetes mellitus tipo II," Revista Electrónica Medimay, **26**(1), 54–62, 2019.

[12] V. Arrunátegui, "Despistaje de diabetes mellitus tipo 2 en una población adulta urbana del distrito de Coishco, Ancash, Perú.," Revista Medica Herediana, **26**(3), 173, 2015, doi:10.20453/rnp.2015.2585.

[13] Instituto Nacional de Estadística e Informática, En el Perú 3 de cada 100

personas de 15 y más años reportan tener diabetes., INEI, 2015.
[14] C. Fernández, P. Baptista, Metodología de la Investigación. 6ta ed. México: Mc Graw-Hill/Interamericana., 2015.
[15] R. Ortega, R. Veloso, O. Hansen, "Percepción y actitudes hacia la investigación científica," Revista de Investigación En Ciencias Sociales y Humanidades, **5**(2), 101–109, 2018, doi:10.30545/academo.2018.jul-dic.2.
[16] J. López, R. Rodríguez, "Adaptación y validación del instrumento de calidad de vida Diabetes 39 en pacientes Mexicanos con diabetes mellitus tipo 2.," Salud Publica de Mexico, **48**(3), 200–211, 2006, doi:10.1590/S0036-36342006000300004.
[17] M. Brítez, E. Torres, "Calidad de vida en pacientes con diabetes mellitus tipo 2.," Rev. Nac (Itauguá), **9**(1), 78–91, 2017, doi:10.18004/rdn2017.0009.01.078-091.
[18] Y. Ramos, R. Morejón, Y. Cabrera, D. Herranz, W. Rodríguez, "Adherencia terapéutica, nivel de conocimientos de la enfermedad y autoestima en pacientes diabéticos tipo 2," Gaceta Médica Espirituana, **20**(3), 13–23, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 967-971 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Hand Gesture Classification using Inaudible Sound with Ensemble Method

Jinwon Cheon[1], Sunwoong Choi[*,2]

[1]*Department of Security Enhanced Smart Air Mobility, Kookmin University, Seoul, 02707, South Korea*

[2]*School of Electrical Engineering, Kookmin University, Seoul, 02707, South Korea*

ARTICLE INFO

Article history:
Received: 07 September, 2020
Accepted: 02 November, 2020
Online: 08 December, 2020

Keywords:
Hand gesture classification
Short-Time Fourier Transform (STFT)
Ensemble

ABSTRACT

Recognizing the human behavior and gesture has become important due to the increasing use of wearable devices. This study classifies hand gestures by creating sound in the inaudible frequency range from a smartphone and analyzing the reflected signals. We convert the sound using Short-Time Fourier Transform to magnitude and phase. We trained two types of data on Convolutional Neural Network model. And then we propose a method applying soft voting, an ensemble technique, to improve classification accuracy taking the average of two models' result. In this paper, the classification accuracy of the Mag model is 96.0% and the classification accuracy of the Phase model is 90.0% for 8 hand gestures. While the ensemble model showed 96.88%, which is better than Mag and Phase models.

1 Introduction

With the advancement of IT technology, the use of wearable devices such as smartwatches and IoT-based devices is becoming more popular. However, as devices become smaller for portability, there are some limitations, including the difficulty of controlling these devices using buttons or touches. To date, lots of studies are being conducted to overcome this limitation. Most studies collect and recognize data from sensors like cameras and controllers [1]–[4]. However, it is a fatal flaw that requires a sensor or product to recognize movement. Therefore, in this study, we propose a hand gesture classification using a smartphone without a separate sensor.

When the smartphone's speaker makes an inaudible sound, the microphone records this sound. A subject performs a hand gesture while recording is in progress, the signal hit by the subject's hand and reflected changes in frequency domain because of the Doppler effect. There are several methods to analyze a signal into time and frequency domain, among them, we analyzed the signal using the Short-Time Fourier Transform (STFT) [5]. This is because STFT converts the data into a 2D grid-like topology with time and frequency domains. It can be easy to input in 2D Convolutional Neural Network (CNN) [6]. In this way, the sound signal is transformed into magnitude and phase. We propose a method to increase the classification accuracy of the independently trained magnitude and phase CNN models by applying soft voting, an ensemble technique. As a result, the proposed method showed a classification accuracy of 96.88% for 8 hand gestures.

The rest of this paper is organized as follows. It introduces the hand recognition and classification research using ultrasound and ensemble technique in Chapter 2, describes the method proposed in Chapter 3, evaluates its performance in Chapter 4, and finally concludes this paper in Chapter 5.

2 Related Work

In this chapter, we briefly introduce how hand gesture recognition and classification are being used, and studies using ultrasonic signals and studies to improve the performance of a model using multiple CNN models are described.

First, the current status of research on hand gesture recognition and classification is introduced [7]. Data for hand gesture recognition and classification are collected from various sensors, for example, mount-based sensors, multi-touch screen sensors, vision-based sensors, and ultrasonic-based sensors. The collected data are trained on machine learning models used in various fields, such as human-robot interaction, virtual manipulation, sign language, and gaming.

Second, studies on recognizing or classifying hand motions using ultrasonic signals will be described. There is a study called

[*]Corresponding Author: Sunwoong Choi, Kookmin University, School of Electrical Engineering, 82-2-910-4416 & schoi@kookmin.ac.kr

FingerIO that uses smartphone and smartwatch to find fingers using the microphone and speaker [8]. In this study, they use sound in a inaudible frequency range and track the fingers' position by Orthogonal Frequency Division Multiplexing (OFDM). Other study proposes a deviceless gesture tracking scheme. This is called LLAP (Low-Latency Acoustic Phase) [9]. In this study, they use microphones and speakers that built-in on mobile devices to track hands and fingers. This study measures the phase changes that is formed by movement and converts the changes into the distance of the movement. There is a paper that classifies hand gestures using two smartphones [10]. One used as a microphone, while the other used as a speaker. In this paper, they used STFT. These window size is 500, and overlapped size is 475. The frequency resolution is 2,048. As a result, their proposed CNN model classified 8 hand gestures with a classification accuracy of 87.75%.

Finally, studies using multiple CNN models will be described. These two papers [11, 12] explained ensemble method. In paper [12], the performance of the model for classifying hand gestures was improved using the ensemble technique. They ensembled three types of classifiers: a polynomial classifier, a multi-layer perceptron, and a support vector regression. So, they obtained median recognition rates per moment in time of more than 86%. Another paper [13] is a paper that improves performance by fusion of GoogLeNet, VGGNet, and ResNet. The performance was improved by extracting feature layers (100-dimensional and 40-dimensional fully connected and softmax) to each model and attaching them. This paper implemented a model that recognizes the motions of a person in an image, and performance improved in order when only one of the three models mentioned above was used, when two were used, and when all three were used.

3 Proposed Method

Our system can be divided into three steps (data acquisition, data pre-processing and model training) like Figure 1. First, we collect sound data using our own application. Then, the data are pre-processed by STFT on MATLAB. Lastly, We train the pre-processed data on the proposed CNN model.



Figure 1: Overall system architecture.

3.1 Data Acquisition

We collected data with a smartphone application developed in Android Studio. This application produces a single frequency of 20 kHz on the smartphone's speaker. At the same time, the application starts recording on the smartphone's microphone. The speaker and microphone are activated for 3 seconds, and a subject makes a hand gesture within that time at a distance of 1–5 cm from the smartphone. The application stores the sound reflected by gestures. When collecting data, only one subject in the office performing hand gestures to reduce the noise.

3.2 Data Pre-processing

We have used STFT to convert sound data because STFT informs the change of frequency over time. A window size of 5,000, an overlap size of 4,750, and a frequency resolution of 2,048 are detailed information of STFT we used. The microphone recorded for 3 seconds, but the speaker only produced up to 2 seconds. So, we cut off the last second of the data. We also trimmed the data around 20 kHz to suppress the background noise and reduce the number of data. As shown in the Figure 2, the raw data graph represents amplitude value only, however the STFT spectrogram represents the signal intensity using color. STFT data contain magnitude and phase parts. Figure 3 shows magnitude of STFT data, and Figure 4 shows phase of STFT data. So, we pre-processed the sound data to magnitude and phase data.

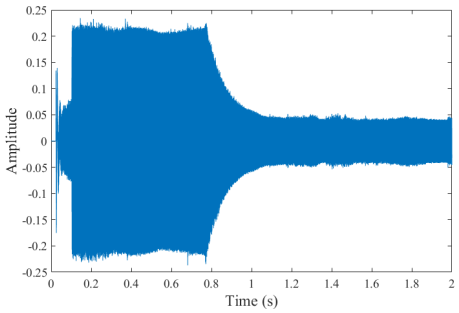


Figure 2: Recorded raw data amplitude graph.

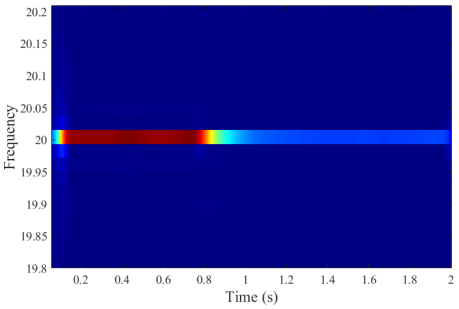


Figure 3: STFT spectrogram (magnitude).

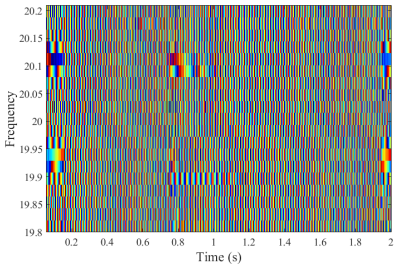


Figure 4: STFT spectrogram (phase).

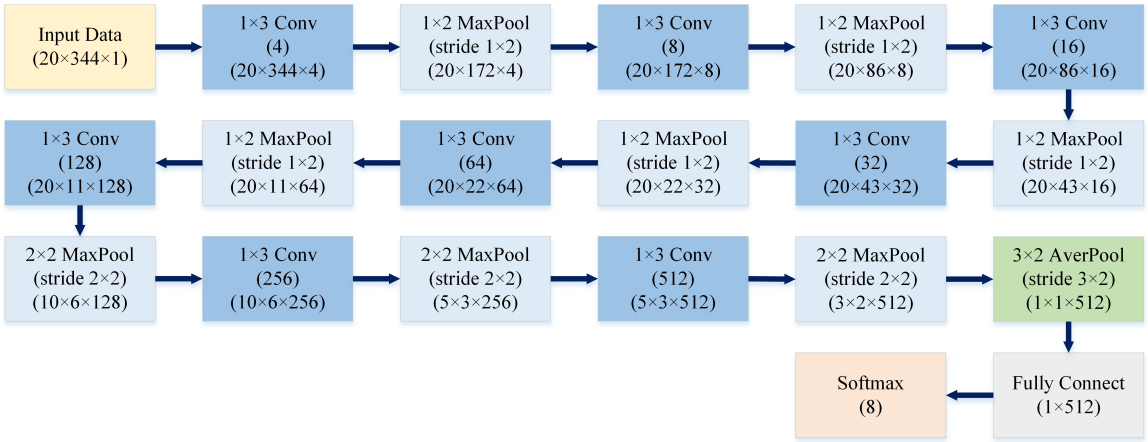


Figure 5: The structure of Mag and Phase models.

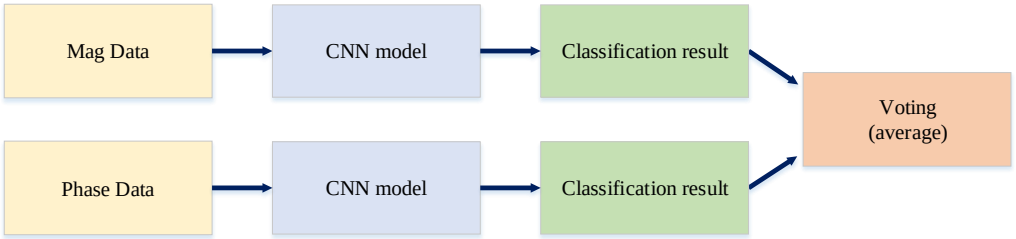


Figure 6: The structure of ensemble model.

3.3 Model Training

We propose two CNN models which called Magnitude (Mag) model and Phase model. Mag and Phase models have the same architecture. The models trained with one part of pre-processed data (magnitude or phase). The data were trimmed to the specific frequency range of 19.8–20.2 kHz through STFT. The models are largely composed of 5 parts: convolution layers, max pool layers, an average pool layer, a fully connected layer and a softmax function. There are 8 pairs of a convolutional layer and a max pool layer in the models to extract data feature. The filter size of the convolutional layer is 1×3, and the filter size of the max pool layer is 1×2 and 2×2. The output of pairs of the convolutional layer and the max pool layer passes through an average pool layer of 3×2 and then through a fully connected layer. Figure 5 shows the structure of our CNN models.

After that, the models were ensembled by taking the average of the softmax function outputs (soft voting). When Mag and Phase models train, the ensemble model not affected training but is only used when testing. Two results of softmax function are calculated by inputting magnitude data and phase data for one hand gesture into each model. The size of the softmax result of each model is the same as the number of classes, and a new softmax result is created by averaging two softmax results corresponding to each hand gesture. The operation is as (1). In this equation, S means the result of softmax funcion.

$$S^{ensemble} = \frac{S^{mag} + S^{phase}}{2} \quad (1)$$

Figure 6 shows how to progress the ensemble model. The magnitude part of the STFT is input to the Mag model and the phase part is input to the Phase model to calculate the average of the softmax matrix outputs from each model. The class that has the largest value among the values of the average matrix is the predicted class of the ensemble model.

4 Evaluation

4.1 Experiment Setup

In the experiment, we have collected the data to TG&Co.'s LUNA smartphone in an office where only one subject performed hand gestures. The application was developed in the Android Studio environment. The UI of developed application is as Figure 7. The smartphone was placed on a flat desk and recording started when the REC-START button was pressed. The application activated the speaker to produce a single frequency at 20 kHz and the microphone to start recording. After 3 seconds, recording stopped automatically.

The saved sound data were transferred from the smartphone to the computer. The data were converted to STFT on MATLAB R2019b. The models were implemented using the Tensorflow 1.6.0, and Python 3.4.8 for running the Tensorflow. Training and test were performed on a GPU server which operates CentOS 7.5 and uses Intel Xeon Silver 4114 CPU. The learning progressed about 4,000 epochs. Besides, a 5-fold cross validation technique was used to increase the reliability of the models.

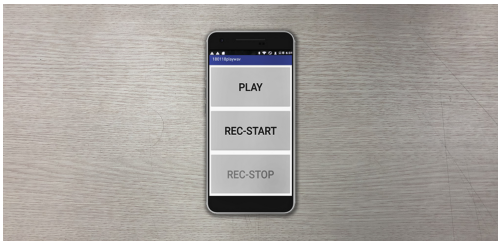


Figure 7: UI of developed application.

4.2 Gesture Dataset

In this study, 8 hand gestures were classified. Figure 8 shows 8 hand gestures that were used as a dataset. The subject gestured at a distance of 1–5 cm from the phone without touching the screen. This action was performed within 3 seconds and repeated 100 times per gesture. Through this process, we have collected a total of 800 data. The data were divided into training data and test data at a ratio of 8:2.

- Do nothing : Do not move the hand while recording.
- Block : Block the microphone with palm while recording.
- Move : With the palm open, record perpendicular to the screen. Place the hand at the beginning of the arrow and move it in the direction of the arrow while recording (to right, to bottom, to left, to top).
- Draw : Stretch only index finger and draw on the screen once in the direction of the arrow (circle and triangle).

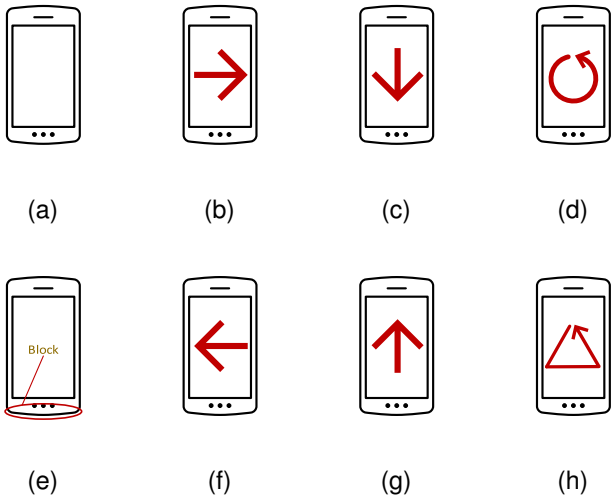


Figure 8: Used hand gestures.

4.3 Experiment Result

When performing STFT, the STFT result value also changes depending on the window size, so that we evaluated this. We tested the classification accuracy by increasing the window size from 2,000 to 8,000 in steps of 1000 using only magnitude data. In all window sizes, the frequency resolution and overlap size were unified to 2,048 and 95% during the evaluation. The classification accuracy by window size is shown in Figure 9. Therefore, the evaluations of the Phase and ensemble model were proceeded when window size is 5,000.

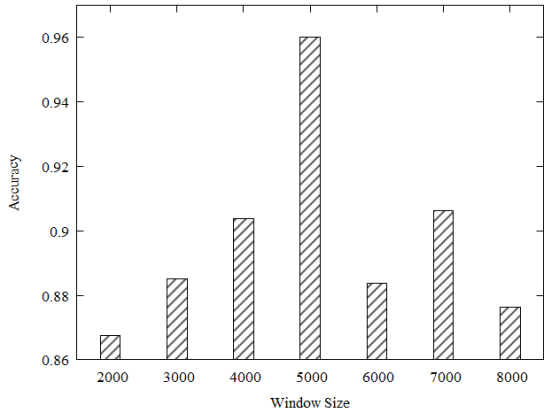


Figure 9: Classification accuracy by window size.

In conclusion, we presented the ensemble model that classified 8 hand gestures as 96.88% classification accuracy. We estimated Mag, Phase, and the ensemble model when window size is 5,000. Mag model showed 96.0% classification accuracy, and Phase model showed 90.0%, slightly lower than Mag model. However, when these two models were ensembled with our method, the performance improved, showing a classification accuracy of 96.88%. Not only classification accuracy but also various evaluation indicators were tested. Figure 10 shows precision, recall, F1-score, and accuracy of each model. It can be seen that the ensemble model better than the two models in all evaluation indicators.

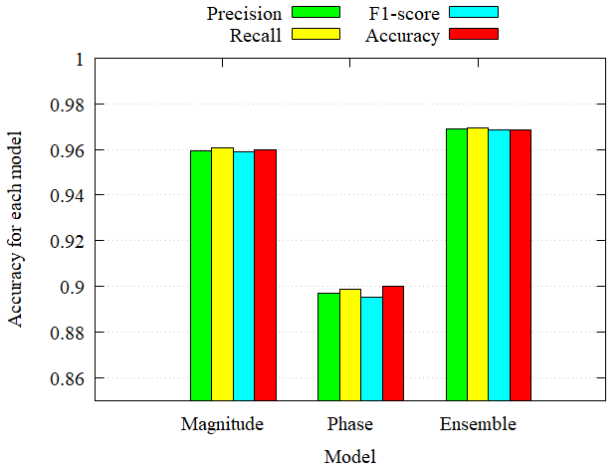


Figure 10: The performance comparison of each model.

We plotted the confusion matrices to also check the accuracy corresponding to each class. The confusion matrices are shown in figures 11– 13. Comparing the confusion matrices, it can be seen that the Mag and Phase models differ in the class that predicts well. For example, in Mag model, 'do nothing' gesture shows the highest accuracy, but not in Phase model. Since the accuracy of Mag model

is higher than that of Phase model, the ensemble model shows a similar pattern to Mag model. Also, almost classes of the ensemble model show better accuracy than Mag and Phase models.

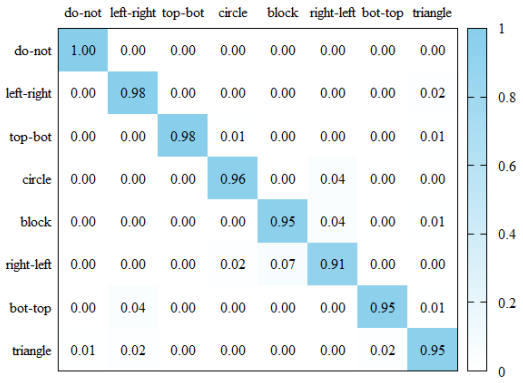


Figure 11: The confusion matrix of Mag model.

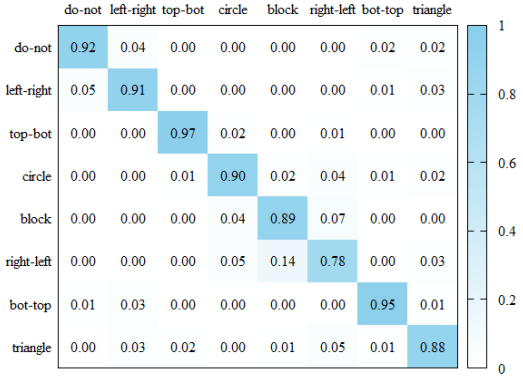


Figure 12: The confusion matrix of Phase model.

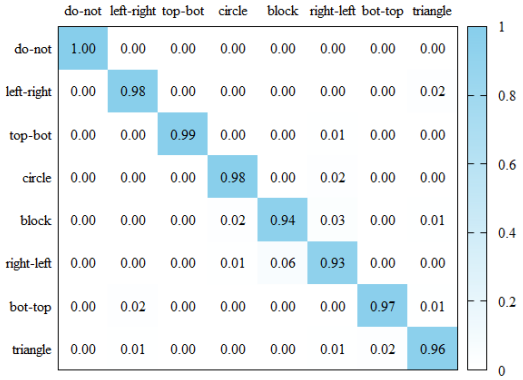


Figure 13: The confusion matrix of the ensemble model.

5 Conclusion

This study classified hand gestures using inaudible sound using a smartphone's speaker and microphone. We used our own application to produce and record 20 kHz sound while gesturing. The sound data were converted to STFT as magnitude and phase data. These data used for training and test in our CNN models. We proposed Mag and Phase models, and also suggested ensemble method using soft voting to improve classification accuracy. As a result, our model showed 96.88% classification accuracy for 8 hand gestures.

Acknowledgment This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIP) (No.2016R1A5A1012966).

References

[1] G. Zhu, L. Zhang, P. Shen, J. Song, "Multimodal gesture recognition using 3-D convolution and convolutional LSTM," IEEE Access, **5**, 4517–4524, 2017, doi:10.1109/ACCESS.2017.2684186.

[2] W. Lu, Z. Tong, J. Chu, "Dynamic hand gesture recognition with Leap Motion controller," IEEE Signal Processing Letters, **23**(9), 1188–1192, 2016, doi:10.1109/LSP.2016.2590470.

[3] T. Schlömer, B. Poppinga, N. Henze, S. Boll, "Gesture recognition with a Wii controller," in Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, 11–14, 2008, doi:10.1145/1347390.1347395.

[4] M. Sathiyanarayanan, S. Rajan, "MYO armband for physiotherapy healthcare: A case study using gesture recognition application," in 2016 8th International Conference on Communication Systems and Networks (COMSNETS), 1–6, 2016, doi:10.1109/COMSNETS.2016.7439933.

[5] R. Schafer, L. Rabiner, "Design and simulation of a speech analysis-synthesis system based on short-time Fourier analysis," IEEE Transactions on Audio and Electroacoustics, **21**(3), 165–174, 1973, doi:10.1109/TAU.1973.1162474.

[6] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning," Nature, **521**(7553), 436–444, 2015, doi:10.1038/nature14539.

[7] H. Cheng, L. Yang, Z. Liu, "Survey on 3D hand gesture recognition," IEEE Transactions on Circuits and Systems for Video Technology, **26**(9), 1659–1673, 2016, doi:10.1109/TCSVT.2015.2469551.

[8] R. Nandakumar, V. Iyer, D. Tan, S. Gollakota, "FingerIO: Using active sonar for fine-grained finger tracking," in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 1515–1525, 2016, doi:10.1145/2858036.2858580.

[9] W. Wang, A. X. Liu, K. Sun, "Device-free gesture tracking using acoustic signals," in Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, 82–94, 2016, doi:10.1145/2973750.2973764.

[10] J. Kim, S. Choi, "Hand gesture classification based on nonaudible sound using convolutional neural network," Journal of Sensors, **2019**, 1084841, 2019, doi:10.1155/2019/1084841.

[11] T. G. Dietterich, "Ensemble methods in machine learning," in Multiple Classifier Systems, 1–15, 2000, doi:10.1007/3-540-45014-9_1.

[12] J. Schumacher, D. Sakič, A. Grumpe, G. A. Fink, C. Wöhler, "Active learning of ensemble classifiers for gesture recognition," in Pattern Recognition, 498–507, 2012, doi:10.1007/978-3-642-32717-9_50.

[13] Y. Lavinia, H. H. Vo, A. Verma, "Fusion based deep cnn for improved large-scale image action recognition," in 2016 IEEE International Symposium on Multimedia (ISM), 609–614, 2016, doi:10.1109/ISM.2016.0131.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 972-980 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Docker-C2A : Cost-Aware Autoscaler of Docker Containers for Microservices-based Applications

Mohamed Hedi Fourati[*,1], Soumaya Marzouk[1], Mohamed Jmaiel[1,2], Tom Guérout[3]

[1] *ReDCAD Laboratory, National Engineering School of Sfax, University of Sfax, Sfax, 3038, Tunisia*

[2] *Digital Research Center of Sfax, University of Sfax, Sfax, 3021, Tunisia*

[3] *Laboratory for Analysis and Architecture of Systems, University of Toulouse, Toulouse, 31031, France*

ARTICLE INFO

Article history:
Received: 14 September, 2020
Accepted: 26 November, 2020
Online: 08 December, 2020

Keywords:
Cloud Computing
Microservices
Docker
Kubernetes
Autoscaling
PSO algorithm

ABSTRACT

*This article proposes a cost-aware autoscaler for microservices-based applications deployed with docker containers. This autoscaler decreases the cost of the application deployment as it reduces computing resources. In elastic treatment, microservice resources are scaled when the used metric as the central processing unit (CPU) exceeds the threshold. In case of threshold exceeding, an autoscaler adds many instances of docker containers in order to satisfy the need of the application. In many studies, the autoscaler adds many containers without selecting the appropriate microservices for scaling and without in advance calculation of the adequate number of containers. This may lead to allocating additional resources to inappropriate microservices and a non optimal number of containers. For this reason, we propose our autoscaler "*Docker-C2A*" which identifies the adequate microservices to add resources. It also calculates the optimal number of needed containers. "*Docker-C2A*" analyses the state of the application, uses the execution history and uses a Particle Swarm Optimization (PSO) algorithm to identify the adequate microservices for scaling resources and to determine the optimal number of containers. As a result, "*Docker-C2A*" helps to reduce computing resources and to save extra costs. Experimental measurements were conducted on a microservices-based application as a concrete use-case demonstrating the effectiveness of our proposed solution.*

1 Introduction

Microservices is an architectural style of development consisting of a collection of small independent and loosely coupled components [2]. Nowadays, several companies like Amazon, Linkedin, Spotify and Netflix have migrated towards microservice architecture in order to increase the efficiency and the scalability of computing resources.

Currently, the most of microservices-based applications are deployed in docker containers [3] and orchestrated with kubernetes tool [4]. Docker is a container technology designed for creating, deploying and running applications using containers. The latter allow developers to package up an application with all needed parts such as libraries and deploy it as one package. Each instance of a microservice is a docker container. They are orchestrated and deployed in a cluster of VMs (Virtual Machines) configured by kubernetes, an open-source container orchestration system created by Google for automation of application deployment, scaling and management.

When there is a rise on the workload, microservices resources are overloaded and thus additional resources should be allocated. One of the most important stakes in Cloud environment is to minimize computing resources so as to reduce the deployment cost of the application. We can resolve this issue and optimize computing resources using autoscaling techniques. Over-provisioning of computing resources leads to allocating unused resources. Whereas, under-provisioning causes performance degradation of the application. Thus, an autoscaling technique should identify the needed resources for the application in order to allocate quasi-exact needed computing resources.

In literature, few autoscalers are proposed for containers and microservices-based applications, and most existing solutions have the same treatment as kubernetes autoscaler. The autoscaler of kubernetes is called Horizontal Pod Autoscaler HPA [5]. HPA scales the number of containers based on CPU usage. CPU threshold is

[*]Corresponding Author: Mohamed Hedi Fourati, ReDCAD Laboratory Sfax, Tunisia. Email : mohamed-hedi.fourati@redcad.org
This paper is an extension of work originally presented in [1].

specified by the user. When the application receives high workload due to the increase of requests number, many microservices are overloaded and the CPU threshold is exceeded. Consequently, kubernetes HPA adds one container instance to each overloaded microservice. If the threshold is still exceeded, the autoscaler continues adding other containers until reaching the normal state of the CPU or the maximum number of containers. This autoscaling process has several issues that allocate many resources and thus charge a high deployment cost.

The first issue is the usage of CPU metric. In fact, using CPU metric in autoscaling is not effective and may lead to allocating unnecessary resources. Indeed, it was proved in [6] that a high CPU usage means that the container instance is fully utilized, but it can still provide acceptable response time without adding more resources. So, it is possible to execute further requests and to have acceptable response time even if we have high CPU and high workload. This is also why many elastic solutions tend to use response time metric instead of CPU usage to benchmark their autoscaler. Moreover, it is highly recommended to use response time metric instead of resource metrics as CPU when launching autoscaling actions.

The second issue is that existing autoscalers do not select the adequate microservices for adding resources. In fact, the microservices of an application are in relationship, and an issue in a given microservice may be propagated to other ones. Similarly, if this issue is resolved for a microservice, it is automatically resolved for related microservices. Thus, an autoscaler should select and find the eligible microservices for scaling. To do that, the autoscaler should analyze the whole application with an overview of all microservices. Also, it should use the execution history of the application to find eligible microservices.

The third issue is that existing autoscalers do not calculate the needed number of containers for each microservice and do not react fast. For example, the autoscaler of kubernetes adds containers one by one until reaching the normal state. In this case, the normal state is recovered after a period of time, while in Cloud applications, the time constraint is critical and fast scaling actions are requisite. Then, the autoscaler should calculate and estimate the needed number of containers for each microservice from the beginning to launch fast actions. To do this, the autoscaler should analyze the application and its microservices to estimate the adequate number of containers for each microservice. Moreover, it should consider the execution history and the architecture of the application as the relationship between microservices to estimate the adequate number of containers. As a result, the autoscaler calculates precisely the needed number of container and launches fast actions. This leads to optimize computing resources and deployment cost.

This paper proposes "Docker-C2A" a **C**ost **A**ware **A**utoscaler for microservices-based applications deployed with docker containers. This autoscaler focuses on existing issues in order to optimize computing resources and reduce deployment costs. Firstly, "Docker-C2A" is based on response time metric instead of CPU metric. Also, it selects eligible microservices for scaling and calculates the optimal number of containers to be added to each microservice. To do this, during autoscaling treatment, "Docker-C2A" considers the architecture of the application and the relationship between microservices. In addition, it uses the execution history of the application and executes a PSO algorithm. As a result, "Docker-C2A" selects appropriate microservices for scaling and calculates the needed number of containers for each microservice. Experimental results conducted on a concrete use case demonstrates that we can satisfy high workload with minimum computing resources by allocating a small number of containers. Thus, "Docker-C2A" optimizes computing resources and reduces the deployment cost. And that is the main purpose of this autoscaler.

This paper is organized as follows: Section 2 presents the related work. Section 3 presents the motivating example of our autoscaler. Section 4 presents in details Docker-C2A autoscaler. Section 5 details the PSO algorithm in Docker-C2A. Section 6 discusses the experimentations. We conclude the paper in Section 7.

2 Related Work

Autoscaling is dedicated for optimizing resource allocation in order to avoid allocating unused resources and to reduce the cost of deployment. Few studies were proposed for the autoscaling of microservices-based applications focusing on containers. As it was explained in section 1, these solutions have many issues. In fact, most of these solutions focus on CPU metric. Also, the resource allocation is not accurate as many resources are allocated without selecting the adequate microservices for scaling and without previous calculating the needed amount of resources. This may lead to allocating unused and unnecessary resources and thus charging the user extra costs. In this section we review these studies and explain the shortcomings of each one.

In [7], the author designed a container-based autoscaling policy with a mathematical model dedicated for Iot applications based on containers and microservices. This autoscaler uses two main metrics, number of requests and resource related metrics as CPU or memory usage. Also, it uses the history of the application to take adequate actions for similar cases recorded in the history. After each scaling action, the autoscaler waits for a cooling period defined by the user to avoid launching other scaling actions until the application becomes stable. The main objective of this autoscaler is to minimize computing resources and the deployment costs. Using the history, this autoscaler calculates and estimates the optimal number of containers used to scale each microservice. Although this solution estimates the needed resources for allocation, it does not select the eligible components for scaling.

In [8], the author presented a message queue as a specific use case in IoT context. They considered metrics related to message queue mi- croservices. Two classes of microservices were considered, compute-intensive and I/O-intensive. In [9], the author aimed to adjust computing resources to the incoming load. In scaling process, a new container is added when the response time exceeds 1000 ms in a 15 s window. A container can be removed if it uses less than 10% of cpu. These two studies neither select containers and components for scaling nor calculate the adequate quantity of resources to be allocated.

In [10], the author proposed a genetic algorithm for resources allocation and elasticity management dedicated for microservices-based applications deployed on a containerized environment. The main objectives of the proposed solution is the container allocation

in VMs and elastic treatment of containers. This study models the whole system such as applications, microservices, the relationship between microservices, the required resources of each microservice, containers, the needed resources, physical machines and the network. Then, it optimizes this model using a meta heuristic genetic algorithm to get optimal results for resource allocation of containers and the number of containers for each microservice. This autoscaler considers the architecture of the application and the relationship between microservices in scaling process. However, it does not calculate the needed amount of resources for scaling.

In [11], the researchers developed an autoscaler called Microscaler for microservices-based applications which identifies the scaling-needed services using a customized metric called Service Power. When detecting a response time violation, Microscaler launches the process of detecting services that should be scaled using the customized metric Service Power. This metric is based on response time and used to identify the microservices that need to be scaled with additional resources. Besides, this autoscaler calculates the adequate number of containers to be added for each selected microservice. This autoscaler uses a Bayesian Optimization (BO) and an heuristic approach to determine the optimal number of instances for selected services. Although this autoscaler selects the appropriate microservices to be scaled and calculates the number of additional containers, it is strongly related to applications based on service mesh that make this approach not applicable in other contexts.

In this paper, we propose "Docker-C2A" autoscaler for microservices-based applications using docker containers. We focus on docker container technology because it is the most used in the context of containerization and microservices-based applications. This autoscaler is based on response time metric. Also, "Docker-C2A" selects eligible microservices for scaling and calculates the optimal amount of resources to be allocated. This autoscaler uses the execution history of the application and a PSO heuristic algorithm. The execution history helps to determine which of application microservices need to be scaled. Then, the PSO algorithm calculates the optimal number of containers to be added to each selected microservice.

"Docker-C2A" can be distinguished over existing solutions as it allocates the optimal amount of resources to the adequate microservices. "Docker-C2A" is a fast and accurate autoscaler that gives more precise scaling actions. This leads to optimizing application resources and deployment costs.

3 Motivating Example

To illustrate existing issues, we detail the behavior of kubernetes autoscaler on a concrete microservices-based application. We take the example of kubernetes autoscaler as most existing autoscalers have the same treatment. In fact, we consider Bookinfo [12] a microservices-based application. Figure 1 illustrates Bookinfo application composed of four microservices.

This application displays the data related to each book. It is similar to a unique catalog entry of an online book store. It displays book details as ISBN, number of pages, and book reviews. Bookinfo application is composed of four separate microservices: **Productpage** microservice calls details and reviews microservices to get the book data. **Details** microservice gives the book information. **Reviews** microservice contains the book reviews and calls ratings microservice. **Ratings** microservice contains ranking formation deducted from the book review. Reviews microservice has three versions as three types of instance : **Reviews v1** does not call Rating microservice. **Reviews v2** invokes Rating microservice and displays rating stars as 1 to 5 black stars. **Reviews v3** invokes Rating microservice and displays rating stars as 1 to 5 red stars. This application is deployed with docker containers and deployed on a kubernetes cluster.

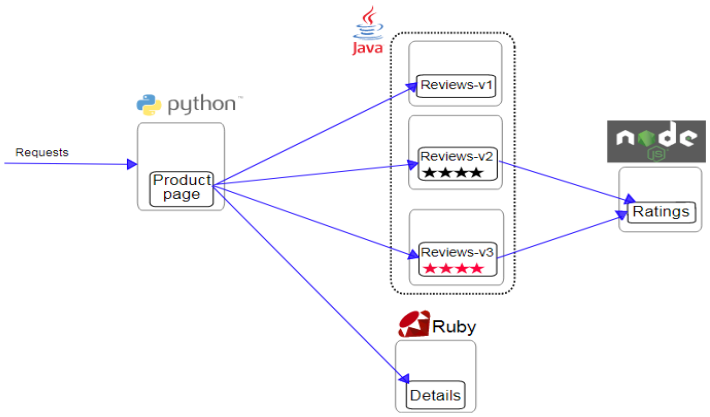


Figure 1: Bookinfo Application

Kubernetes is the most known and efficient orchestrator for docker containers. In the following, we use the pod concept of kubernetes. A pod in kubernetes is considered as a single docker container of a microservice. A microservice is composed by a set of pods (a set of containers). The elasticity controller of kubernetes is called Horizontal Pod Autoscaler HPA [5] and is based on CPU metric. It adds pods when the CPU threshold is exceeded. CPU threshold is specified by the user. If the maximum number of pods (containers) is reached, HPA autoscaler cannot add any more pods.

Many types of workload can be launched in this application. For example we can manage a workload that routes all incoming requests to Reviews v3 of microservice Review. Another example of workload routes incoming requests in 50% to Reviews v2 and 50% to Reviews v3. Technically, it is difficult to route requests flow to a specific container of a microservice. However, it is easy to manage the request flow with Istio [13] service. Istio is based on service mesh which facilitates the creation and management of the network of deployed services. Istio allows to manage network communication between microservices as well as the traffic of requests. In fact, it helps to create configuration rules for traffic routing in order to control the traffic flow between microservices.

In this paper we consider a workload that routes all incoming requests of microservice review to reviews-v3 instances. This workload launches a flow of 70 Req/s. It is not a very high workload but it saturates resources and generates overloaded microservices. This section presents the behavior of kubernetes HPA autoscaler using this workload.

Figure 2 shows the behavior of kubernetes HPA autoscaler using the presented workload with a request flow of 70 Req/s. Many pods were saturated and Kubernetes HPA launches autoscaling ac-

tions. Kuberenetes HPA adds 4 pods for productpage microservice resulting in a total of 5 pods. It adds also 1 pod to reviews-v3 microservice. From t=70, we have a total of 12 pods of reviews-v3. Ratings and details microservices were not saturated and we have no additional pods. The curves of ratings and details are superimposed as they have 1 pod all the time. The total number of pods required for this workload is 19 pods as 5 pods for productpage, 12 pods for reviews-v3, 1 pod for ratings and 1 pod for details. This autoscaler is based on CPU. If the CPU usage of a microservice exceeds threshold, the autoscaler adds many pods until recovering of the normal state.

However, as explained in section 1, using CPU usage in autoscaling treatment is not effective and may lead to allocating unnecessary resources. Therefore, we should also analyze the behavior of kubernetes autoscaler towards the response time. The red curve in Figure 2 illustrates the evolution of the response time in kubernetes autoscaler.

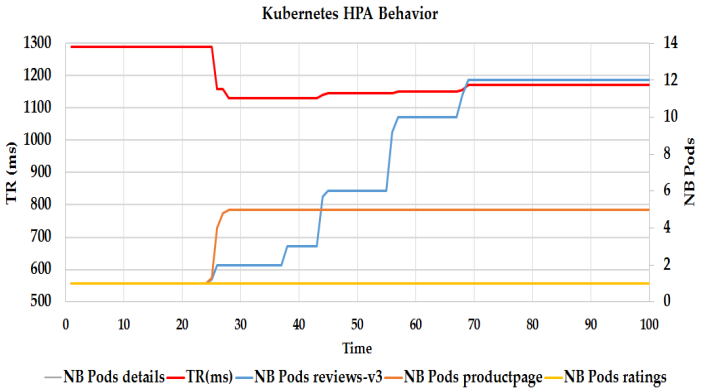


Figure 2: Kubernetes HPA Autoscaler behavior

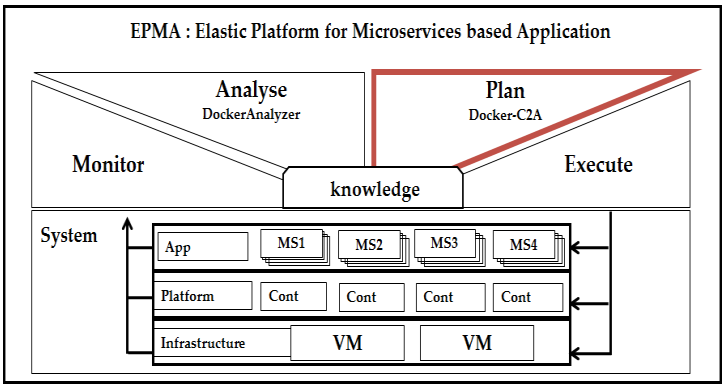


Figure 3: EPMA : Elastic Platform for Microservices based Application : Autoscaling Engine

Approximately, the response time starts with 1300 milliseconds (ms). At time t=25, we have 5 pods for productpage microservice and 2 pods for reviews-v3 microservice. We can notice a considerable evolution of the response time with approximately 1100 milliseconds. From t=37 to t=100, kubernetes autoscaler adds pods for the microservice reviews-v3 until reaching 12 pods. Following this addition, we notice no improvement in response time, but we can see a slight increase starting from t=43. It is concluded that during this interval of time, kubernetes autoscaler added 10 pods for reviews-v3 microservice, while the response time is not improved. So the added pods for reviews-v3 are not useful and kubernetes allocates unnecessary resources without any benefit.

With our autoscaler Docker-C2A, based on response time metric, we aim at optimizing the addition of resources and avoid allocating unnecessary resources. Docker-C2A uses PSO heuristic algorithm and the execution history of the application to optimize resource allocation. Using Docker-C2A, we can satisfy the same workload used in kubernetes example with a very small number of pods and with a better application performance.

4 Docker-C2A Autoscaler

4.1 EPMA platform

Docker-C2A is part of our EPMA platform (Elastic Platform for Microservices based Applications) illustrated in Figure 3. This platform is based on the autonomic MAPE-K loop proposed by IBM [14]. MAPE-K loop is composed of four components that share the Knowledge: Monitor, Analyze, Plan and Execute. EPMA manages the entire system with all levels: VM level, container level and microservice level. The analyze component DockerAnalyzer was previously presented [1] while this paper details the plan component Docker-C2A.

The monitor component collects data from different levels: VM, container and Microservice. This component collects metrics related to resources as CPU usage and memory usage and metrics related to the application performance as the response time. These parameters are filtered and correlated in order to determine symptoms that should be analyzed. As example of a symptom, an overload state caused by the CPU violation. If the monitor detects a symptom, DockerAnalyzer, the analyze component, is invoked to check whether this symptom is caused by a normal behavior as resource saturation or by an abnormal behavior as VM problem. If an abnormal behavior is detected, the analyzer had to check the root cause of the problem. Then, the plan component Docker-C2A generates the appropriate change plan describing the desired set of changes, and deliver the change plan to the execute component. The execute component applies the change plan and executes adequate actions for each layer. These four components share the knowledge containing particular data as policies and symptoms. This paper explains deeply the plan component Docker-C2A of EPMA platform.

Docker-C2A is based on the execution history of the application and a PSO heuristic algorithm. When the application state needs additional resources, Docker-C2A selects the best microservices for scaling and calculates the optimal amount of needed resources for each microservice.

4.2 Execution History

The main purpose of the execution history is to identify the microservices that should be scaled, and those that should not be scaled as there is no need to add resources. To do that, we attribute a score value for each microservice. The score value is a number varying from 0 to 5 depending on the utility of adding resources to the

microservice. For example, if we have a microservice that the addition of pods (containers) improves greatly the response time, the attributed score will be 5. If we have a microservice where adding pods will not improve the response time, the attributed score will be 0. As a result of the execution history, we generate the score value for each microservice. This helps to scale the appropriate microservices.

We account for the score attribution using bookinfo application with the same workload explained in section 3 which routes the incoming traffic of reviews microservice to reviews-v3 instances and uses a flow of 70 Req/s. Bookinfo application is composed of 4 microservices. We present in the following an extract from the history related to two microservices in order to demonstrate the score attribution.

NB Pods Productpage	NB Pods Reviews-v3	NB Pods Ratings	NB Pods Details	Response Time (ms)
1	1	1	1	1289
5	1	1	1	1134

Figure 4: An extract of Productpage history

In Figure 4, we change the number of pods in microservice productpage and we keep the same number of pods for other microservices. In the second row in this table of Figure 4, we have 5 pods of productpage and the response time has been improved considerably compared to the first line with 1 pod. The response time decreased effectively with 155 milliseconds. So requests are much faster with 5 pods than if we had 1 pod of productpage. In this case, we notice that adding pods to productpage microservice improves greatly the response time. We attribute the high score with value 5 for this microservice.

NB Pods Productpage	NB Pods Reviews-v3	NB Pods Ratings	NB Pods Details	Response Time (ms)
1	1	1	1	1289
1	5	1	1	1292

Figure 5: An extract of Reviews-v3 history

In Figure 5, we change the number of pods in microservice reviews-v3 and we keep the same number of pods for other microservices. There is no remarkable difference in response time between 1 and 5 pods. There is a very small increase in response time with 3 milliseconds, which is not profitable when we add 4 pods. In this case, we notice that adding pods to reviews-v3 microservice has no effect on the response time. Then, there is no change in the result. So we attribute the lowest score to reviews-v3 with value 0.

With the same process, we analyze the evolution of the response time when we add pods to ratings and details microservices. In fact, there is a weak improvement of the response time when we add pods to these two microservices. Then, we attribute the score value of 1 for ratings and details microservices while adding pods improves the results slightly.

As a conclusion for the execution history, we can admit that adding more resources to productpage microservice can effectively improve the response time of the application. Adding pods to reviews-v3 microservice has no effect on the response time. So adding pods to this microservice is unnecessary. Adding pods to ratings and details microservices improves the results slightly. They are not privileged in autoscaling treatment.

We illustrate the score values for each microservice in Figure 6.

	Microservice Index	Score value of scaling
productpage microservice →	0	5
reviews-v3 microservice →	1	0
ratings microservice →	2	1
details microservice →	3	1

Figure 6: Score value for each microservice

The execution history is very helpful when adding resources in scaling actions. In fact, it helps to select the appropriate microservices for scaling which leads to optimize resource allocation. After selecting eligible microservices for scaling, Docker-C2A calculates the optimal number of pods to be added to each microservice. To do this, Docker-C2A uses PSO algorithm to estimate the needed amount of resources for each microservice.

4.3 PSO Algorithm

Particle Swarm Optimization (PSO) is a computational technique based on the behavior of herds of animals that optimizes a given problem. It tries iteratively to improve candidate solutions with regard to a given quality measure. It was developed by Kennedy and Eberhart [15] in 1995 and widely used and researched ever since.

In PSO algorithm we define a particle which is analogous to a bird in a flock of birds. The flock of birds forms the search space. The movement of each particle is guided by the velocity vector. Each particle moves according to its best position pbest and the best position of the swarm gbest.

In this algorithm, we use a fitness function to measure the performance of each particle. In each iteration, the algorithm changes the velocity of each particle towards pbest and gbest positions. The algorithm stops when a stopping criterion is met or the number of iterations is reached. The velocity and positions equations are represented with Eqs. (1) and (2), respectively. The pseudo-code of the algorithm is presented in Algorithm 1.

$$V_i^{k+1} = wV_i^k + c_1 r_1 (pbest_i - X_i^k) + c_2 r_2 (gbest - X_i^k) \quad (1)$$

$$X_{i+1}^k = X_i^k + V_i^{k+1} \quad (2)$$

where

- w : inertia weight

- c_1 and c_2 : acceleration coefficients
- X_i^k : position of particle i at iteration k
- X_i^{k+1} : position of particle i at iteration k+1
- V_i^k : velocity of particle i at iteration k
- V_i^{k+1} : velocity of particle i at iteration k+1
- $pbest_i$: best position of particle i
- $gbest$: best position of the swarm

Algorithm 1 PSO Algorithm

```
Initialize a population of particles with random values positions and velocities from D dimensions in the search space
while Termination condition or Number of iterations not reached do
    for Each particle i do
        Update the velocity of the particle using Equation 1
        Update the position X_i of the particle using Equation 2
        Calculate the fitness value f(X_i)
        if f(X_i) < f(pbest_i) then
            pbest_i ← X_i
        end if
        if f(X_i) < f(gbest) then
            gbest ← X_i
        end if
    end for
end while
```

5 Docker-C2A : PSO Algorithm

Docker-C2A uses PSO algorithm to calculate the optimal number of pods (containers) to be added to each microservice. Two main objects should be defined in PSO algorithm: particle and fitness function.

5.1 Particle definition in Docker-C2A

In PSO algorithm, a particle represents a solution candidate of the problem. In our context, a solution is a particle that contains the adequate number of pods for each microservice. The dimension of the particle is defined by the number of microservices of the application. We consider the same example and workload explained in section 3. In this example, we use bookinfo application composed by 4 microservices: productpage, reviews-v3, ratings and details. In this case we define particles of 4 dimensions. Each microservice is referenced by an index. For example, we consider the productpage microservice with index 0, reviews-v3 with index 1, ratings with index 2 and details with index 3. We illustrate the particle model for bookinfo application in Figure 7.

	Particle_i	
	Microservice Index	Number of Pods
productpage microservice →	0	Nb_pods_ms_0
reviews-v3 microservice →	1	Nb_pods_ms_1
ratings microservice →	2	Nb_pods_ms_2
details microservice →	3	Nb_pods_ms_3

Figure 7: Particle model of bookinfo application

In Figure 8. below we present an example of two particles. Particle_0 contains 3 pods for productpage microservice, 1 pod for reviews-v3 and ratings microservices, 3 pods for details microservice. Particle_1 contains 1 pod for productpage microservice, 5 pod for reviews-v3 microservice, 1 pod for ratings details microservices.

Particle_0	
Microservice Index	Number of Pods
0	3
1	1
2	1
3	3

Particle_1	
Microservice Index	Number of Pods
0	1
1	5
2	1
3	1

Figure 8: Example of Particles

The PSO algorithm 1 evaluates each particle using the fitness function. The fitness value helps to decide about the best and which one is better than another by searching pbest and gbest positions as described in the algorithm.

5.2 Fitness function in Docker-C2A

The fitness function helps to evaluate particles and update pbest and gbest positions. We aim to minimize the fitness value to get the best solution. In Docker-C2A, the fitness function depends on the score based on the execution history as depicted in section 4.2 and the cost of the particle. We aim to maximize the score of particles in order to improve the response time and to minimize the cost of the particles to save deployments charges.

The fitness value of each particle P_i is calculated following equation 3 which is the cost of the particle divided by the score of the particle.

$$fitness(P_i) = \frac{Cost(P_i)}{Score(P_i)} \tag{3}$$

If we reduce the cost and increase the score then we minimize the fitness value and obtain a better solution. The particle with the lowest fitness value is the best particle, while it has a low cost with high score. The high score reflects a great improvement of the response time.

The score value of the particle is calculated following equation 4 which is the sum of the score of each microservice MS_k multiplied by the number of pods of the microservice.

$$Score(P_i) = \sum Score(MS_k) * NB_Pods(MS_k) \tag{4}$$

The cost of the particle is calculated according to equation 5 which is the sum of the cost unit of each microservice MS_k multiplied by the number of Pods of the microservice.

$$Cost(P_i) = \sum Cost(MS_k) * NB_Pods(MS_k) \quad (5)$$

5.3 PSO Example

In this section, we explain a concrete example of executing PSO algorithm. We consider the bookinfo application with the same workload presented in section 3 which routes requests of reviews microservice to reviews-v3 pods. In this example, we use 3 particles and 3 iterations. We follow the algorithm 1.

	Best position of the swarm gbest	Fitness value of gbest position
After Particle Initialization	Particle_1 Microservice Index / Number of Pods 0 / 2 1 / 2 2 / 2 3 / 1	Fitness = 0,538
After Iteration 0	Particle_0 Microservice Index / Number of Pods 0 / 3 1 / 2 2 / 2 3 / 1	Fitness = 0,444
After Iteration 1	Particle_0 Microservice Index / Number of Pods 0 / 4 1 / 2 2 / 2 3 / 1	Fitness = 0,391
After Iteration 2	Particle_0 Microservice Index / Number of Pods 0 / 5 1 / 2 2 / 2 3 / 1	Fitness = 0,357

Figure 9: Summarize of the execution treatment

In the first step, the algorithm initializes the number of pods needed to the 3 particles randomly. Initially, the best position of each particle pbest is the particle itself. The fitness value of each particle is calculated following the equation 3. When we calculate the cost in fitness function, we consider that all pods have the same cost as 1 unit of cost. In the first step, the fitness values of particles 0, 1 and 2 are ”0.75”, ”0.538” and ”0.714” respectively. The particle with the lowest fitness value is the best particle which is Particle_1 with a fitness value of ”0.538”. PSO algorithm selects Particle_1 as the best particle having the best position gbest in the swarm.

In each iteration, the algorithm calculates the fitness value for each particle, selects the gbest position and moves particles to new positions in order to ameliorate particle positions with better fitness values.

Figure 9. summarizes the execution treatment in this example. This figure illustrates the best position gbest after each iteration and its fitness value. We remember that we use 3 iterations in this example. It is clear that the fitness value of the best solution gbest is improved in each iteration, as the purpose of the algorithm is to minimize this value. In fact, if we minimize the fitness value, we guarantee that the particle position has the minimum cost with significant gain in response time.

At the beginning, the fitness value of gbest was ”0.538”, this value was decreased after each iteration, so our algorithm is improving the solution correctly with the decrease of fitness value. The lowest value of fitness ”0.357” is presented in the final iteration and it is the output of the algorithm. This particle has 5 pods for microservice 0 which is productpage, 2 pods for microservices 1 and 2 which are reviews and ratings microservices respectively, and 1 pod for microservice 3 which is details microservice.

In the next section, we discuss the results of our autoscaler Docker-C2A based on PSO algorithm and execution history using this example and we discuss the difference and the contribution compared to Kubernetes HPA autoscaler.

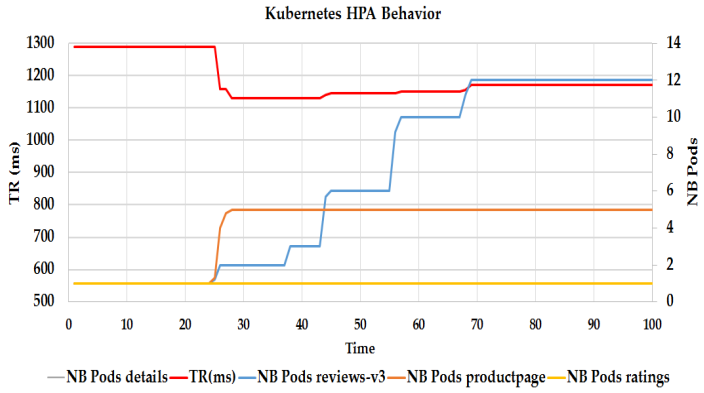
6 Experimentations

In this section, we discuss the impact of our autoscaler Docker-C2A compared to Kubernetes HPA autoscaler.

In our experimentations, we use a cluster Kubernetes composed by 4 virtual machines (VMs). Each VM runs on Ubuntu OS and contains 4 cores CPU, 8GB memory and 100 GB for disk storage. We use the same microservices-based application bookinfo detailed in section 3 composed by 4 microservices: productpage, reviews, ratings and details. Each microservice is deployed in Kubernetes cluster as a deployment containing a set of similar pods. We use the same workload used previously which routes the incoming traffic of reviews microservice to reviews-v3 instances with a flow of 70 Req/s.

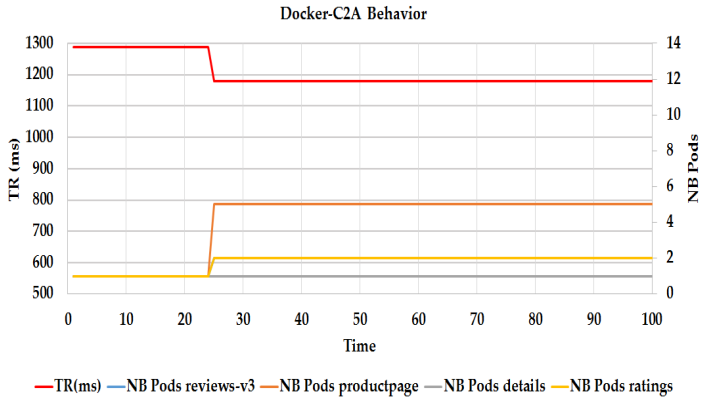
Figure 10 illustrates the difference between Kubernetes HPA and Docker-C2A behaviors. At the time t=24, we launch these two autoscalers. Kubernetes behavior is illustrated in Figure 10a. We detailed the behavior of Kubernetes HPA in section 3 using the workload of bookinfo application. When Kubernetes HPA detects a high load it waits 15 seconds to launch autoscaling actions. In each period of 15 seconds, it calculates and launches the needed number of pods. Every period, it updates the number of needed pods and monitor the situation. Kubernetes HPA stops adding pods when the state of the application becomes stable or it reaches the maximum number of pods. As a result, HPA autoscaler launches 5 pods for productpage microservice, 12 pods for reviews-v3 microservice and 1 pod for both details and ratings microservices. So for this workload Kubernetes HPA launches 19 pods.

Whereas, with Docker-C2A autoscaler, based on the execution history and the PSO algorithm, we can manage the same workload with reduced number of pods. Docker-C2A behavior is illustrated in Figure 10b. When it detects a high workload, it waits 15 seconds and launches the action plan issued by PSO algorithm. In this example, the output of PSO algorithm is presented in Figure 9. The output is the gbest particle of the last iteration. This particle contains 5 pods for productpage microservice, 2 pods for both reviews-v3 and ratings microservices and 1 pod for details microservice. Docker-C2A



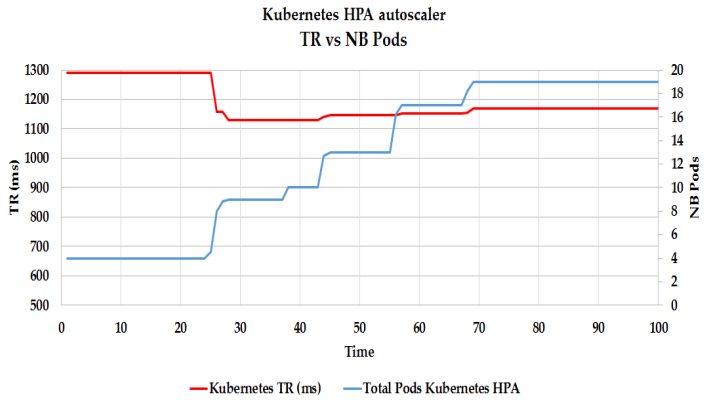
a

(a) Kubernetes HPA Autoscaler behavior



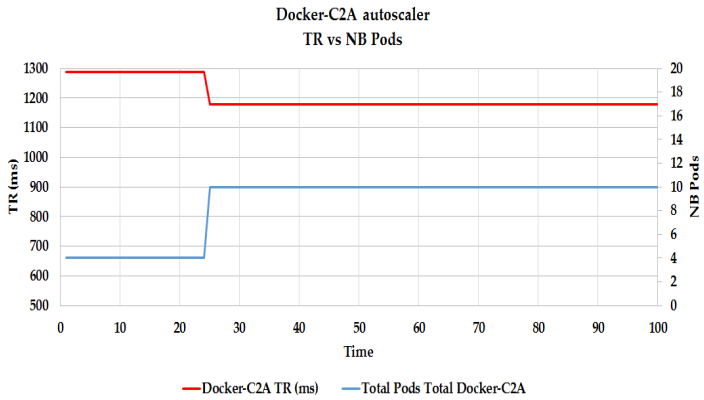
b

(b) Docker-C2A autoscaler behavior



c

(c) Kubernetes HPA Response Time vs NB of Pods



d

(d) Docker-C2A Response Time vs NB of Pods

Figure 10: Kubernetes HPA autoscaler vs Docker-C2A autoscaler

launches in total 10 pods.

If we consider the response time of requests which is presented with red curves in Figures 10a and 10b, we notice that the response time is almost the same in Kubernetes or Docker-C2A autoscalers. While Docker-C2A launches only 10 pods which is 9 pods less than Kubernetes HPA that launches 19 pods. We summarize the number of pods and the response time of these two autoscalers in Figure 11. The response time of requests with Docker-C2A is 9 milliseconds higher than the response time with Kubernetes. This difference is negligible and is not considerable as the difference is appreciated if it is in the order of tens or even hundreds of milliseconds. So, we have certainly reached our goal with Docker-C2A using the execution history and the PSO algorithm which gives an optimized amount of resources with good performance.

Autoscaler	NB Pods Productpage	NB Pods Reviews-v3	NB Pods Ratings	NB Pods Details	Total NB of Pods	Response Time (ms)
Kubernetes HPA	5	12	1	1	19	1170
Docker-C2A	5	2	2	1	10	1179

Figure 11: Kubernetes HPA autoscaler vs Docker-C2A autoscaler

Also, Docker-C2A takes in average 60 milliseconds to generate the decision of scaling, which is a short period of time that does not disturb the elastic treatment. However, Docker-C2A uses a high amount of resources of about 1.2 cores of CPU to process the PSO algorithm. The high resource usage is the cost of problem optimization and among the weak points of our autoscaler. Moreover, to run our algorithm, it is recommended to execute up to 4 iterations as a maximum. Indeed, in our context, PSO algorithm generates erroneous results from 5 iterations.

7 Conclusion and Future Works

In this paper, we presented our autoscaler Docker-C2A a cost aware autoscaler for microservices-based applications deployed with docker containers. Docker-C2A is based on response time metric. It uses the execution history of the application and PSO algorithm. Docker-C2A selects appropriate microservices for scaling and calculates the optimal number of containers for each microservice. We clearly presented the effectiveness of our autoscaler compared to Kubernetes HPA autoscaler using bookinfo application as a real case study. We compared with Kubernetes autoscaler as most of existing autoscalers have the same issues. They neither select appropriate microservices for scaling nor calculate the optimal number of additional containers. As a result, Docker-C2A optimizes computing resources and reduces deployment cost. And that is the main purpose of our autoscaler. However, there are some

limitations of Docker-C2A. In fact, it consumes a high amount of resources to generate the solution. This fact is caused by the PSO algorithm which requires high resources and this is the case of several optimization algorithms.

In future work, we improved Docker-C2A to consume less resources, either by proposing another particle definition in PSO algorithm, or by integrating another optimization algorithm instead of PSO that consumes less resources and generates good results. Also, in future work, we improved the execution history using other performance metrics and more execution data in order to analyze the application behavior in each workload. This greatly helps to estimate the needed resources with more precision and allows better optimization of computing resources.

References

[1] M. H. Fourati, S. Marzouk, K. Drira, M. Jmaiel, "DOCKERANALYZER : Towards Fine Grained Resource Elasticity for Microservices-Based Applications Deployed with Docker," in 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), IEEE, 2019, doi:10.1109/pdcat46702.2019.00049.

[2] J. Lewis, M. Fowler, "Microservices," https://martinfowler.com/articles/microservices.html, march 2014.

[3] "Docker," https://www.docker.com/.

[4] "Kubernetes," https://kubernetes.io/, june 2019.

[5] "Kubernetes Autoscaler," https://kubernetes.io/docs/tasks/run-application/horizontal-pod-autoscale/, june 2019.

[6] G. Yu, P. Chen, Z. Zheng, "Microscaler: Automatic Scaling for Microservices with an Online Learning Approach," in 2019 IEEE International Conference on Web Services (ICWS), IEEE, 2019, doi:10.1109/icws.2019.00023.

[7] F. Zhang, X. Tang, X. Li, S. U. Khan, Z. Li, "Quantifying cloud elasticity with container-based autoscaling," Future Generation Computer Systems, **98**, 672–681, 2019, doi:10.1016/j.future.2018.09.009.

[8] M. Gotin, F. Lösch, R. Heinrich, R. Reussner, "Investigating Performance Metrics for Scaling Microservices in CloudIoT-Environments," in Proceedings of the 2018 ACM/SPEC International Conference on Performance Engineering, ACM, 2018, doi:10.1145/3184407.3184430.

[9] G. Toffetti, S. Brunner, M. Blöchlinger, J. Spillner, T. M. Bohnert, "Self-managing cloud-native applications: Design, implementation, and experience," Future Generation Computer Systems, **72**, 165–179, 2017, doi:10.1016/j.future.2016.09.002.

[10] C. Guerrero, I. Lera, C. Juiz, "Genetic Algorithm for Multi-Objective Optimization of Container Allocation in Cloud Architecture," Journal of Grid Computing, **16**(1), 113–135, 2017, doi:10.1007/s10723-017-9419-x.

[11] G. Yu, P. Chen, Z. Zheng, "Microscaler: Automatic Scaling for Microservices with an Online Learning Approach," in 2019 IEEE International Conference on Web Services (ICWS), IEEE, 2019, doi:10.1109/icws.2019.00023.

[12] "Kubernetes Autoscaler," https://istio.io/latest/docs/examples/bookinfo/, august 2020.

[13] "Istio," https://istio.io/.

[14] IBM, "An architectural blueprint for autonomic computing," 2005.

[15] J. Kennedy, R. Eberhart, "Particle swarm optimization," in Proceedings of ICNN'95 - International Conference on Neural Networks, IEEE, doi:10.1109/icnn.1995.488968.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 984-989 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Rectification Circuit with Co-Planar Waveguide Antenna for 2.45 GHz Energy Harvesting System

Nuraiza Ismail*, Ermeey Abd Kadir

Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, 40450, Malaysia

ARTICLE INFO

Article history:
Received: 24 January, 2021
Accepted: 07 April, 2021
Online: 22 April, 2021

Keywords:
Co-Planar Waveguide
RF Energy Harvesting
AC-DC Conversion

ABSTRACT

A new approach for designing RF energy harvester with a single stage converter circuit is presented in this paper. The proposed converter configuration is integrated with an antenna that is based on the coplanar waveguide (CPW) transmission line with improved gain resonated at 2.45 frequency ISM band. The CPW patch antenna as a harvester antenna is designed in a rectangular shape that uses FR-4 substrate with a loss tangent and relative permittivity of 0.025 and 4.3 respectively. The output from the harvester antenna is connected to the converter circuit using only two Schottky diodes. The rectifier design achieves between 0.1% to 37% of RF-DC power conversion efficiency over the ambient RF input signal range from -20 dBm to 0 dBm and the antenna exhibits a directivity of 3.896 dBi as well as a return loss of -48.85 dB. For an input power of 0 dBm, the proposed circuit can rectify an AC signal up to 6.09 V. Moreover, the proposed CPW antenna that is integrated with a converter circuit agrees for the harvesting of ambient electromagnetic energy to power low power electronic devices.

1. Introduction

Advancements in energy harvesting from the environment or ambient have grown a lot of attention in recent years as ambient energy harvesting delivers a green self-sustainable operation for powering low power electronics devices [1–3]. This technology exploits energy from several sources from the environment, for example solar light, heat, vibration, thermal energy and electromagnetic waves [4, 5]. It has become an alternative for powering low power electronic devices instead of using a conventional method which is battery plug in mainly used in wireless sensor networks of the Internet of Things (IoT), wearable electronic devices and implantable biomedical devices[6–9].

Today, batteries remain the vital source to power up these system devices. Nevertheless, the method has various inherent disadvantages. For instance, the battery size and weight are huge, limitation of battery lifespan and also requires more cost of battery replacement for these devices [8, 10–12]. To over come these drawbacks, engineers have exploited a new approach called energy harvesting , to extend the battery life and empower an autonomous manoeuvre of the IoT end nodes [13, 14]. The energy harvesting has rendered an attractive approach for energizing low power electronic devices [2, 15]. It empowers wireless charging, battery-less and a reasonable alternative for supplying the wireless nodes even in cold, dark, and static environments without requiring power cables or battery replacement [14, 16–18].

The past decade has seen an increase in the use of an ambient environment as the source of radio frequency (RF) energy harvesting. This harvesting system is the most prominent of its accessibility and easy scavenging system compared to other sources. Other sources are far from human control as thermal energy requires heat, vibration needs motion while solar requires light present [19, 20]. This alternative approach used the concept of capturing the ambient power, including the wasted power that hovering in the surrounding environment without any distraction to the environment [21–23]. In contrast, the naturally low ambient RF energy can be significantly challenging. Also, it is easily carried or wear, especially when the RF energy harvester is combined with completely autonomous systems due to its low input power and voltage necessities [24, 25]. To overcome this issue, the Schottky diode has been commonly used for RF energy harvesting as it has a minimum threshold voltage and fast switching diode speed [10, 13].

The RF energy harvesting system comprises an antenna that captures an RF incoming signal, a network that represents impedance matching for exploiting power transfer from the antenna and the rectifier or converter circuit. The rectifier circuit is a key component in the system block which determines the system efficiency. It converts the RF signal to DC voltage and stores the energy in holding load storage devices such as capacitor

*Corresponding Author: Nuraiza Ismail, Faculty of Electrical Engineering, UiTM, Email: nuraizaismail@gmail.com

and battery. Figure 1 illustrates the basic block diagram of energy harvesting structures. The following sections describe four main parts of this system.

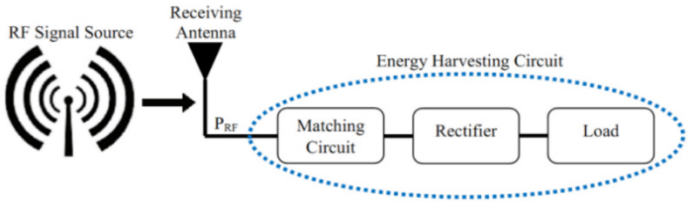


Figure 1: Block diagram of RF energy harvesting system [13]

The major challenge in designing an ambient RF energy harvesting system is the requirement to operate with low power RF signals. The works published in [13,26,27] have achieved the maximum power conversion efficiency (PCE) of the rectifier at relatively high input power levels, which is above 5 dBm while the available ambient RF signals are much lower levels in the environment which are lower than -15 dBm. In order to reach high maximum PCE, multistage rectifier is the most commonly used technique among previous reseachers [28,29]. Nevertheless, the high power efficiency results from high input RF signals that are applied to the multistage rectifier. The circuit designed in [30] needs four stages in rectifier configuration for their circuit optimization when -15 dBm input power signal is applied. They observed that the number of stages does not help to improve efficiency for low power RF input signal. As the number of stages increases, the peak of the efficiency curve decreases toward higher powers.

This paper is an expansion of work originally reported in Proceedings of the 2019 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE) [31]. In the previous work, a square patch antenna was designed and analysed with a matching converter circuit. In this proposed work, the antenna and the rectifier of the converter circuit have been enhanced for both parts. This work also highlights a single RF energy harvesting circuit that aims to convert the power within the 2.45 frequency band RF signal into usable DC supply. As compared with previous work [31], the gain of the enhancement antenna has been improved. Meanwhile, the ripple voltage of the output signal that occurred in previous work has been resolved with a new arrangement of components in converter circuit design. In the end, a novel RF energy harvesting system using only one stage converter circuit is presented which is able to harvest low input power signal.

2. Harvesting Antenna Design

An antenna is required in the receiving part for harvesting RF energy as it captures energy from ambient. The RF energy is then rectified and converted into DC voltage. The proposed harvesting antenna design as illustrated in Figure 2 is a rectangular patch antenna with co-planar waveguide feed. This arrangement is preferred due to its simple impedance matching and high circuit density on one layer. Besides, its active and passive elements are easy to integrate and it has low dispersion and low radiation loss. All these make the co-planar waveguide preferable for rectifying circuit design.

The low-cost FR-4 substrate with a dielectric constant $\varepsilon_r = 4.3$, tan δ = 0.025 and thickness of 1.6 mm are used for designing the proposed CPW antenna. The area of 30 × 38 mm^2 where the width and length area calculated using equation (1) and (2) respectively. This antenna is excited by a feeding line of 50 Ω characteristic impedance with a width of 4.235 mm. Both gaps between the feeding line and the ground plane are g = 0.5 mm. The detailed geometry structure of the proposed antenna design after optimization is shown in Table 1.

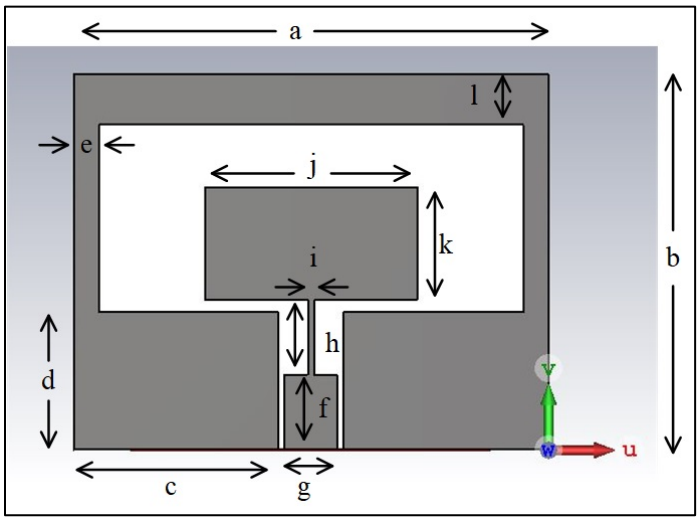


Figure 2: Geometry of the proposed antenna based on co-planar waveguide feed

$$W = \frac{c}{2f_r}\sqrt{\frac{2}{\varepsilon_r+1}} \quad (1)$$

$$L = \frac{c}{2f_r\sqrt{\varepsilon_r}} \quad (2)$$

Table 1: The Proposed Antenna Parameters

Parameter	(mm)
a	38.0
b	30.0
c	16.4
d	11.0
e	2.0
f	6.0
g	4.4
h	6.0
i	0.5
j	17.0
k	9.0
l	4.0

3. Converter Circuit Design

In the RF energy harvesting system, a rectifier circuit is required for converting RF input energy into funtional DC energy. The rectifier that acts as a converter circuit converts the AC input signal from the RF source. AC to DC conversion is required for this system as the RF source is being received in the form of a sine wave signal. Hence, the RF source is represented by a sinusoidal voltage source Vs as depicted in Figure 3. The arrangement of

rectifier circuit contains three parts which are a matching network, a conversion circuit and a load. The converter circuit and the circuit components with parameters after optimization are shown in Figure 3 and Table 2 respectively. Two Schottky diodes of MBR0520L are used for the rectification process. The Schottky diode is preferred due to its attractive feature of low substrate losses and quick switching. Moreover, the Schottky diode has a high sensitivity to the very small ambient signal.

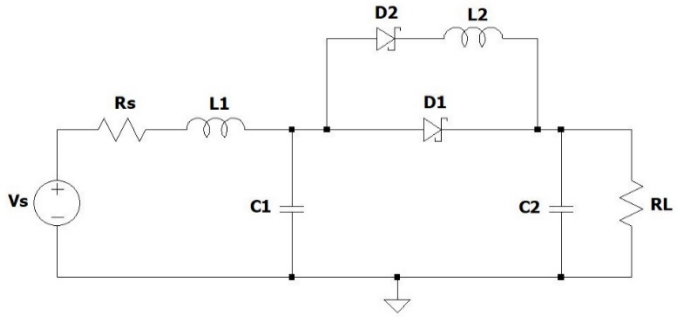


Figure 3: The proposed RF harvesting converter circuit using Schottky connected diodes

For simulation study, Vs is connected to the converter circuit in a sinusoidal source that represents an RF energy harvester from the receiving antenna. The antenna port impedance represented by internal resistance, Rs is set to 50 Ω. In this circuit topology, the matching circuit is simply constructed using capacitor C1 and inductor L1. This matching circuit is required to match their impedances between the receiving antenna and converter circuit where a good matching network will allow a maximum power transfer between them. The vital part in this harvesting circuit is the conversion where it rectifies RF input signal to DC signal. Both Schottky diodes and inductor L2 are needed for this circuit to convert AC to DC signal. In this circuit design, the capacitor C2 is important in order to eliminate ripple DC voltage that occurs at the output terminal. After the DC signal is passed through the DC bypass filter, then it appears as a useable voltage across the load, RL.

Table 2: Circuit Components and Parameters

Components	Value	Unit
R_s	50	Ω
L1	33	nH
L2	3	nH
C1	1	pF
C2	33	uF
RL	100	kΩ

The circuit analysis examines the constant input ac voltage with low amplitude sinusoidal voltage and it is separated into four operating modes for each cycle.

Mode 1: During the positive half input cycle, the inductor, L and capacitor, C_1 are charged by the sinusoidal voltage source. Both D_1 and D_2 are in OFF state. The load, R_L is feed by the energy stored in the C_2 that acts as a filter capacitor.

Mode 2: The mode starts when $V_{C1} = V_{IN} - V_L > V_{OUT}$, where Diode D_2 is switched to ON state. The capacitor C_2 is energized. All the energy stored in L_1 and C_1 during *Mode 1* is fed to the load. The saturated current of inductor, L_1 and capacitor, C_1 drops linearly. The changing state of D_2 influence the loss during this mode.

Mode 3: This mode happens as $V_{IN} = V_{C1} + V_{L2} > V_{OUT}$, D_2 is switched to ON state. As capacitor C_2 is energizing, the energy stored in L_2 and C_1 during *Mode 3* is transferred to the load. The changing state of D_1 influences the loss during this mode.

Mode 4: When V_{C1} and $VL_1 + VC_1 > V_{OUT}$, the diodes, D_1 and D_2 are switched to ON state. The load is driven by the energy stored in L_1 and C_1 as they transferred to the load through L_2 and C_2.

4. Results and Discussion

The simulation results are discussed in two parts; the receiving antenna and the converter circuit design.

4.1. Receiving Antenna

The receiving antenna design has been simulated using CST Microwave Studio Software. Before connecting to the converter circuit, the proposed CPW antenna is analysed separately by evaluating the S-parameters. This is to ensure the performance of the proposed antenna achieve a good return loss with perfect matching impedance (≈ 50 Ω) and better gain compared to previous work. Figure 4 indicates the simulation result for return loss, S_{11} is -48.85 dB at 2.4976 GHz resonating frequency. The operating bandwidth of 309 MHz in the range between 2.3618 to 2.707 GHz, covers the IEEE 802.11 standard in the 2.45 GHz wireless ISM band. Matching impedance is approximate to 50 Ω while VSWR is 1.007 as shown in Figure 5 and Figure 6 respectively. Since the value of VSWR is under 2, the proposed antenna is worth mentioning that it has a good match and is applicable for most antenna applications.

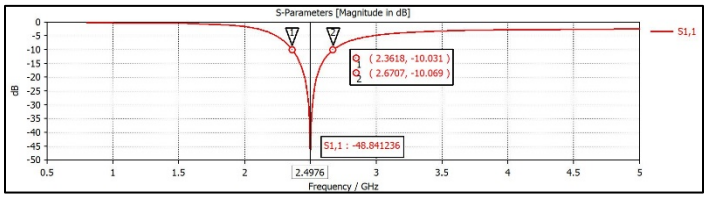


Figure 4: Return loss simulation result of proposed CPW antenna

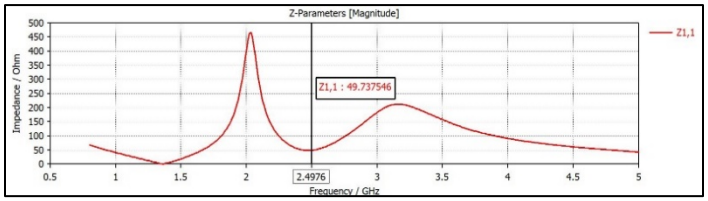


Figure 5: Line impedance at 2.4976 resonant frequency

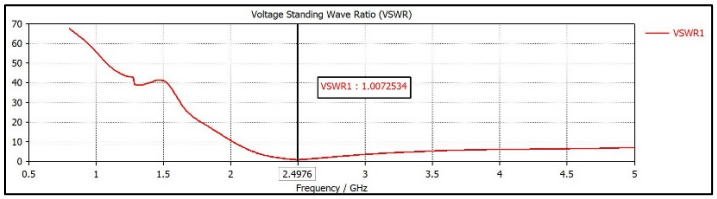


Figure 6: Voltage standing wave ratio (VSWR) for the CPW antenna

The good values of S_{11} and VSWR are insufficient to prove an antenna has good radiation. Antenna gain is one of the key elements that needs to be considered in antenna design as it is an essential metric. The gain of an antenna highlights how much transmitted power in the direction of peak directivity to that of radiation energy. The simulated far-field radiation pattern can be observed in Figure 7 where it shows the gain and the 3-D forward directional pattern of the proposed CPW antenna. The gain of this antenna is 3.896 dBi and it is better compared to previous work [31] which is less than 1 dBi.

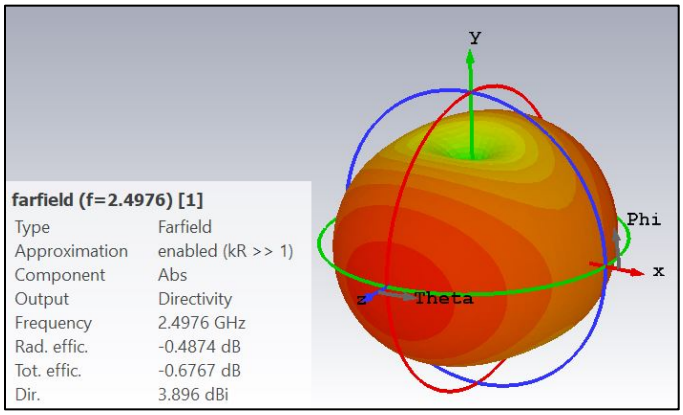


Figure 7: Simulated 3-D radiation pattern of CPW antenna

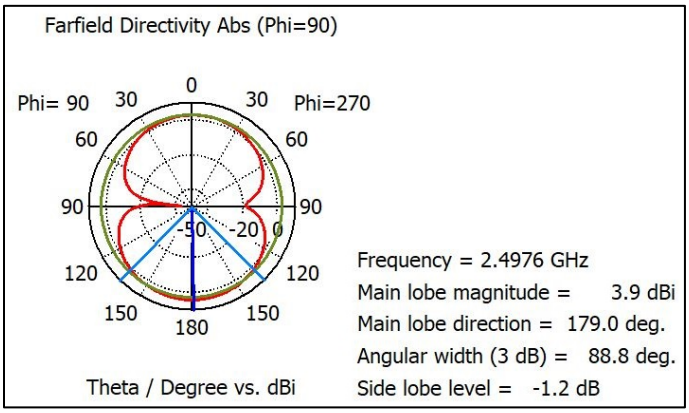


Figure 8: Simulated 2-D radiation pattern

Figure 8 indicates the main and side lobes of the proposed antenna. From this simulated 2-D radiation pattern figure, it indicates that the dipole antenna has an elevation plane beamwidth of 88.8 degrees. Since all metal patches of the substrate for this proposed CPW antenna are designed on one side, the antenna radiates in both front and back directions.

4.2. Converter Circuit

To determine the harvested voltage and current, LTSpice software is used to simulate the converter circuit. The output voltage and current for 0 dBm input power are depicted in Figure 9. From this output, it shows that both output current and voltage are simultaneously harvested to 6.09 V and 0.06 mA respectively. In this work, the ripple output voltage is successfully eliminated as compared to previous work that had large ripple voltage [31].

Equation (3) is the RF to DC power conversion efficiency (PCE) which is defined as the ratio of the DC power P_{DC} and the RF power P_{RF}, where P_{DC} is the dc power produced at the load resistance while P_{RF} is the power received by the antenna.

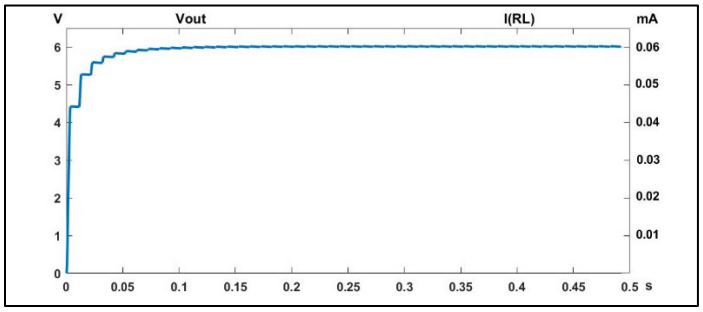


Figure 9: The output voltage and current obtained from the converter circuit for 0 dBm RF input power

$$\eta\% = \frac{P_{DC}}{P_{RF}} \times 100\% \quad (3)$$

Table 3 shows the output voltage and the PCE for every single input power level. Both the output voltage and efficiency of the circuit are increased as the input power varies from -20 dBm to 0 dBm. As seen in Figure 10, the DC output voltage exceeds 6.09 V with 37.0% efficiency when the input power is 0 dBm. Hence, the energy stored in capacitor C2 is 100.5 µJ.

Table 3: The output voltage and PCE with respect to five different input power levels

Input Power (dBm)	Output Voltage (V)	PCE (%)
-20	0.02	0.1
-15	0.12	0.5
-10	0.52	2.7
-5	1.85	10.9
0	6.09	37.0

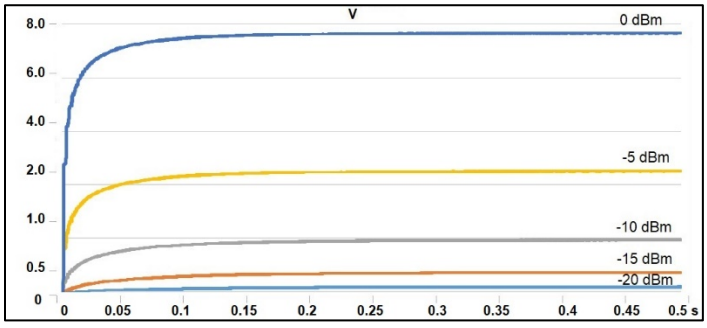


Figure 10: The output voltages with respect to different input power level

4.3. S-Parameter Analysis

The antenna and the converter circuit are matched to maximize the stored energy. Results from S-parameter are used to analyze the signal performance of the converter circuit regarding the difference in circuit component value. Figure 11 shows the connection between the CPW antenna and the converter circuit by using CST Microwave Studio integrated with CST Design Studio.

Table 4: Performance summary and distinction with related works

Reference	This work	[7]	[16]	[24]	[32]	[33]
Frequency [GHz]	2.45	0.90	2.40	0.90	0.90	1.81
Rectifier	Schottky	Schottky	CMOS (10 Cascoded)	CMOS (10 Stages)	CMOS	Schottky
Antenna Substrate	FR-4	Paper	NA	NA	FR-4	Rogers
Input Power (dBm)	-15	-20	-18	-16	-15	-20
Output Voltage (V)	0.12	1.02	1.50	2.32	3.20	0.42
Load Resistance (Ω)	100 k	3 k	88 k	0.5 M	1 M	8.2 k
Power Conversion Efficiency (PCE) %	0.5	57.0	23.9	42.8	32.0	40.8

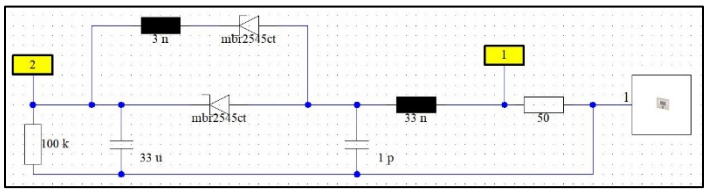


Figure 11: The integrated of the proposed CPW antenna and converter circuit

The result of the S-parameter for the proposed antenna integrated with the converter circuit is simulated by using CST Design Studio as illustrated in Figure 12. The results for both return loss, $|S_{11}|$ = 0.197 and transfer coefficient $|S_{21}|$ = 0.4 at 2.4659 GHz frequency are good since the magnitude value of $|S_{11}|$ and $|S_{21}|$ are ideal to be close to 0.1 and 1 respectively.

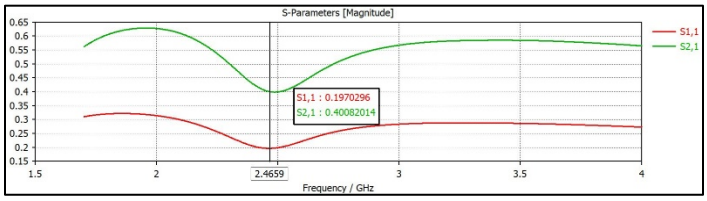


Figure 12: The circuit return loss $|S_{11}|$ and transfer coefficient $|S_{21}|$

4.4. Comparison with Previous Work

Table 4 summarizes a distinction between the proposed study and previous researchers' work based on rectifier used. Since the incident power can never be 0 dBm for 900 MHz and 2.45 GHz ISM band, the comparison is made up of input power between -15 dBm to -20 dBm. As observed at the input power level of -15 dBm, this work provides a low cost antenna substract with lower PCE compared to [32]. However, the load resistance in [32] is 10 times greater than this work, which contributes large power efficiency. In terms of the PCE, this proposed work indicates the lowest performance as compared to the same type of rectifier used in [7] and [33]. Nevertheless, the output voltage in this work is not much different compared to both of them. Meanwhile, reference [24] shows the second highest power conversion effiency among other work, but the load resistance used is also high, which is 5 times greater than this work. Besides, the proposed circuits in [16] and [24] are much more costly since they use 10 cascoded and 10 stages of CMOS respectively. The power efficiency in this proposed work needs to be optimized to give higher efficiency. This can be done by optimizing the impedance matching between the antenna and the rectifier circuit. Moreover, diode characteristics need to be considered since this work uses only non-zero bias diode which degrades the power efficiency. Besides, the influence of saturation current of schottky diode to output voltage needs to be investigated.

5. Conclusion

In this work, a coplanar waveguide antenna with a gain of 3.896 dBi at the resonated frequency of 2.4976 GHz is analyzed for RF energy harvesting applications. A setup has been employed here to study the behaviour of the co-planar waveguide antenna integrated with two Schottky diodes in converter circuit design. Our proposed design has returned very satisfactory results that are applicable for the 2.45 GHz ISM band. The proposed converter circuit specifies the input power can be rectified up to 6.09 V DC signal with power conversion efficiency of 37% and energy of 100.5 µJ for input power of 0 dBm. The findings indicated that a sinusoidal RF incoming signal can be converted to a smooth DC voltage using a proposed RF energy harvesting circuit for powering wireless sensor nodes. For future improvements, matching impedance, diode characteristics and saturation current are essential for improving power efficiency.

References

[1] B.S. Kim, R. Vyas, J. Bito, K. Niotaki, A. Collado, A. Georgiadis, M.M. Tentzeris, "Ambient RF Energy-Harvesting Technologies for Self-Sustainable Standalone Wireless Sensor Platforms," Proceedings of the IEEE, **102**(11), 2014, doi:10.1109/JPROC.2014.2357031.

[2] T.A. Elwi, H.S. Ahmed, "A UWB Monopole Antenna Design based RF

Energy Harvesting Technology," 2018 Third Scientific Conference of Electrical Engineering (SCEE), 111–115, 2018, doi:10.1109/SCEE.2018.8684112.

[3] R. Ren, J. Huang, H. Sun, "Investigation of Rectenna's Bandwidth for RF Energy Harvesting," 2020 IEEE MTT-S International Microwave Workshop Series on Advanced Materials and Processes for RF and THz Applications, IMWS-AMP 2020 - Proceedings, 2020, doi:10.1109/IMWS-AMP49156.2020.9199653.

[4] S. Adami, V. Marian, N. Degrenne, C. Vollaire, B. Allard, F. Costa, "Self-Powered Ultra-low Power DC-DC Converter for RF Energy Harvesting," 2012 IEEE Faible Tension Faible Consommation, 1–4, 2012, doi:10.1109/FTFC.2012.6231746.

[5] H. Ulusan, K. Gharehbaghi, O. Zorlu, A. Muhtaroglu, H. Kulah, "An efficient integrated interface electronics for electromagnetic energy harvesting from low voltage sources," 2013 Transducers and Eurosensors XXVII: The 17th International Conference on Solid-State Sensors, Actuators and Microsystems, TRANSDUCERS and EUROSENSORS 2013, (June), 450–453, 2013, doi:10.1109/Transducers.2013.6626800.

[6] S.M. Noghabaei, R.L. Radin, M. Sawan, "Efficient dual-band ultra-low-power RF energy harvesting front-end for wearable devices," Midwest Symposium on Circuits and Systems, 2018-Augus, 444–447, 2019, doi:10.1109/MWSCAS.2018.8623832.

[7] V. Palazzi, J. Hester, J. Bito, F. Alimenti, C. Kalialakis, A. Collado, P. Mezzanotte, A. Georgiadis, L. Roselli, M.M. Tentzeris, "A Novel Ultra-Lightweight Multiband Rectenna on Paper for RF Energy Harvesting in the Next Generation LTE Bands," IEEE Transactions on Microwave Theory and Techniques, **66**(1), 366–379, 2018, doi:10.1109/TMTT.2017.2721399.

[8] U. Muncuk, K. Alemdar, J.D. Sarode, K.R. Chowdhury, "Multiband ambient RF energy harvesting circuit design for enabling batteryless sensors and IoT," IEEE Internet of Things Journal, **5**(4), 2700–2714, 2018, doi:10.1109/JIOT.2018.2813162.

[9] C.H. Lin, C.W. Chiu, J.Y. Gong, "A Wearable Rectenna to Harvest Low-Power RF Energy for Wireless Healthcare Applications," Proceedings - 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2018, 1–5, 2019, doi:10.1109/CISP-BMEI.2018.8633222.

[10] J. Bito, R. Bahr, J.G. Hester, S.A. Nauroze, A. Georgiadis, M.M. Tentzeris, "A Novel Solar and Electromagnetic Energy Harvesting System with a 3-D Printed Package for Energy Efficient Internet-of-Things Wireless Sensors," IEEE Transactions on Microwave Theory and Techniques, **65**(5), 1831–1842, 2017, doi:10.1109/TMTT.2017.2660487.

[11] L. Xia, J. Cheng, N.E. Glover, P. Chiang, "0.56 V, -20 dBm RF-powered, multi-node wireless body area network system-on-a-chip with harvesting-efficiency tracking loop," IEEE Journal of Solid-State Circuits, **49**(6), 1345–1355, 2014, doi:10.1109/JSSC.2014.2305074.

[12] F. Yuan, Q.T. Zhang, S. Jin, H. Zhu, "Optimal harvest-use-store strategy for energy harvesting wireless systems," IEEE Transactions on Wireless Communications, **14**(2), 698–710, 2015, doi:10.1109/TWC.2014.2358215.

[13] M.M. Mansour, H. Kanaya, "High-Efficient Broadband CPW RF Rectifier for Wireless Energy Harvesting," IEEE Microwave and Wireless Components Letters, **29**(4), 288–290, 2019, doi:10.1109/LMWC.2019.2902461.

[14] A. Ercan, M.O. Sunay, I.F. Akyildiz, "RF Energy Harvesting and Transfer for Spectrum Sharing Cellular IoT Communications in 5G Systems," IEEE Transactions on Mobile Computing, **17**(7), 1680–1694, 2018, doi:10.1109/TMC.2017.2740378.

[15] D. Mishra, S. De, D. Krishnaswamy, "Dilemma at RF Energy Harvesting Relay: Downlink Energy Relaying or Uplink Information Transfer?," IEEE Transactions on Wireless Communications, **16**(8), 4939–4955, 2017, doi:10.1109/TWC.2017.2704084.

[16] M. Stoopman, K. Philips, W.A. Serdijn, "An RF-Powered DLL-Based 2.4-GHz Transmitter for Autonomous Wireless Sensor Nodes for Autonomous Wireless Sensor Nodes," IEEE Transactions on Microwave Theory and Techniques, **65**(7), 1–10, 2017, doi:10.1109/TMTT.2017.2651817.

[17] R.K. Sidhu, J. Singh Ubhi, A. Aggarwal, "A Survey Study of Different RF Energy Sources for RF Energy Harvesting," 2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019, 530–533, 2019, doi:10.1109/ICACTM.2019.8776726.

[18] W. Ni, S. Member, X. Dong, S. Member, "Energy Harvesting Wireless Communications With Energy Cooperation Between," IEEE Transactions on Communications, **63**(4), 1457–1469, 2015, doi:10.1109/TCOMM.2015.2404777.

[19] I. Adam, M.N.M. Yasin, M.E.A. Aziz, M.I. Sulaiman, "Rectifier for RF energy harvesting using stub matching," Indonesian Journal of Electrical Engineering and Computer Science, **13**(3), 1007–1013, 2019, doi:10.11591/ijeecs.v13.i3.pp1007-1013.

[20] S. Keyrouz, H.J. Visser, A.G. Tijhuis, "Rectifier analysis for Radio Frequency energy harvesting and Power Transport," 428–431, 2012, doi:10.23919/eumc.2012.6459081.

[21] N.A. Eltresy, D.N. Elsheakh, E.A. Abdallah, H.M. Elhennawy, "Tri-Band Antenna for Energizing IoT Low Power Devices," 2018 IEEE Global Conference on Internet of Things, GCIoT 2018, 1–5, 2019, doi:10.1109/GCIoT.2018.8620145.

[22] J.C. Wang, M. Leach, Z. Wang, K.L. Man, E.G. Lim, "Rectanna Design for Energy Harvesting," 2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), 456–457, 2016, doi:10.1109/APCCAS.2016.7804001.

[23] J.C.S. Kadupitiya, T.N. Abeythunga, P.D.M.T. Ranathunga, D.S. De Silva, "Optimizing RF energy harvester design for low power applications by integrating multi stage voltage doubler on patch antenna," 2015 8th International Conference on Ubi-Media Computing, UMEDIA 2015 - Conference Proceeedings, 335–338, 2015, doi:10.1109/UMEDIA.2015.7297481.

[24] S.M. Noghabaei, R.L. Radin, Y. Savaria, M. Sawan, "A High-Efficiency Ultra-Low-Power CMOS Rectifier for RF Energy Harvesting Applications," 2018 IEEE International Symposium on Circuits and Systems (ISCAS), 18–21, 2018, doi:10.1109/ISCAS.2018.8351149.

[25] D. Misra, G. Das, D. Das, "An IoT based Wireless Energy Harvesting using Efficient Voltage Doubler Stages in a RF to DC Converter," 2018 4th International Conference on Computing Communication and Automation (ICCCA), 1–5, 2019, doi:10.1109/ccaa.2018.8777712.

[26] S. Kim, H. Abbasizadeh, B.S. Rikan, S.J. Oh, Y. Park, D. Khan, T. Thi, K. Nga, K. Kang, S. Member, Y. Yang, K. Lee, S. Member, "A -20 to 30 dBm Input Power Range Wireless Power System with a MPPT-based Reconfigurable 48 % Efficient RF Energy Harvester and 82 % Efficient A4WP Wireless Power Receiver with Open Loop Delay Compensation," Transactions on Power Electronics, **8993**(c), 1–31, 2018, doi:10.1109/TPEL.2018.2872563.

[27] W. Liu, K. Huang, S. Member, T. Wang, "A Broadband High-Efficiency RF Rectifier for Ambient RF Energy Harvesting," IEEE Microwave and Wireless Components Letters, **30**(12), 50–53, 2020, doi:10.1109/LMWC.2020.3028607.

[28] N. Pournoori, M.W.A. Khan, L. Ukkonen, T. Björninen, "RF Energy Harvesting System Integrating a Passive UHF RFID Tag as a Charge Storage Indicator," 2018 IEEE Antennas and Propagation Society International Symposium and USNC/URSI National Radio Science Meeting, APSURSI 2018 - Proceedings, 685–686, 2018, doi:10.1109/APUSNCURSINRSM.2018.8608530.

[29] M.M. Mansour, S. Yamamoto, H. Kanaya, S. Member, "Reconfigurable Multistage RF Rectifier Topology for 900 MHz ISM Energy-Harvesting Applications," IEEE Microwave and Wireless Components Letters, **30**(12), 1–4, 2020, doi:10.1109/LMWC.2020.3029252.

[30] U. Muncuk, K. Alemdar, J.D. Sarode, K.R. Chowdhury, "Multiband ambient RF energy harvesting circuit design for enabling batteryless sensors and IoT," IEEE Internet of Things Journal, **5**(4), 2700–2714, 2018, doi:10.1109/JIOT.2018.2813162.

[31] N. Ismail, M.T. Ali, E.A. Kadir, M. Firdaus, M. Zin, F.N. Mohd, "A Power Converter Circuit of Co-Planar Waveguide RF Harvester Antenna for Low Power Wireless Applications," 2019 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE), (November), 25–27, 2019, doi:10.1109/APACE47377.2019.9020999.

[32] Z. Hameed, K. Moez, "Design of impedance matching circuits for RF energy harvesting systems," Microelectronics Journal, **62**(April 2016), 49–56, 2017, doi:10.1016/j.mejo.2017.02.004.

[33] D.K. Ho, V. Ngo, I. Kharrat, T.P. Vuong, Q.C. Nguyen, M.T. Le, "A Novel Dual-Band Rectenna for Ambient RF Energy Harvesting at GSM 900 MHz and 1800 MHz," 2016 IEEE International Conference on Sustainable Energy Technologies (ICSET), **2**(3), 612–616, 2017, doi:10.1109/ICSET.2016.7811800.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 988-994 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Mobile Money Wallet Attack Resistance using ID-based Signcryption Cryptosystem with Equality Test

Seth Alornyo[*,1], Mustapha Adamu Mohammed[1,2], Francis Botchey [1,3], Collinson Colin M. Agbesi [1]

[1]*Koforidua Technical University, Computer Science Department, KTU, Koforidua, Ghana*

[2]*Kwame Nkrumah University of Science and Technology, Computer Science Department, KNUST, Kumasi, Ghana*

[3]*University of Electronic Science and Technology of China, School of Information and Software Engineering, UESTC, Chengdu, China*

ARTICLE INFO

Article history:
Received: 18 September, 2020
Accepted: 17 November, 2020
Online: 08 December, 2020

Keywords:
ID-based encryption
Signcryption
Equality test

ABSTRACT

This paper is an extension of a research work presented at ICSIoT 2019. An attack continuum against the insider attack in mobile money security in Ghana using a witness based cryptographic method proposed by Alornyo et al. resisted the service provider from peddling with users data for economic gains. Our improved scheme achieves a simultaneous benefit of digital signature in public key encryption (PKE). The adoption of signcryption cryptosystem in our scheme achieved a desired security property of EUF-CMA using the random oracle model.

1 Introduction

Ghana and other African countries have seen a tremendous growth in mobile money service patronage in recent times. The use of the mobile money services platform requires the use of a less-resource constraint mobile devices that do not require access to internet or mobile app for mobile payment. Thus, any mobile device that does not have access to internet but can access the mobile money service provider cellualr network can perform financial transactions and other utility payments. This and many other reasons has lead to the high patronage of mobile money services in Ghana. However, the mobile money system is not immune to data forgery and re-play attacks. The system model of our scheme is depicted in 1

Our research work is an extension of a publication presented at ICSIoT 2019. The construction presented at ICSIoT 2019 [1] considered the service provider as the adversary who could launch the insider attack. The service provider had access to users token information and transaction details and was possible for the service provider to peddle with users data via the following:

- The mechant receives valid transaction details such as transaction ID, td_{ID} and finds out the content M_1 as well as the token information td_{ID} from the transaction details TSD.
- The adversary (insider) attains the ciphertext CT_1 of a guessed message $M_1^{'}$ and token tdk_{ID}.
- The adversary checks if the test $Test(CT_1, TSD, td_{ID}^{'}) = 1$

In the above scenario, the insider wins the game if it succeeds in deriving the token information from the transaction details. However, the scheme utilized the witness based cryptographic primitive with an added pairing operation to resist the insider attack continuum.

In spite of their construction, there were a possible attack during data transmission such as forgery and re-play attacks in mobile money payment systems. To solve the problem, we propose a scheme using a signcryption cryptographic primitive to achieve the simultanoeus benefit of digital signature and PKE in mobile money systems deployed in Ghana (West Africa). This approach resist data forgery and re-play attacks.

The first generic construction of identity-based signcryption was unveiled [2]. Their work fulfilled a dual function of digital signature and public key cryptosystem (PKE). Other traditional approaches that employed signature-then-encrypt had a high cost computations as compared to [2]. However, the scheme [2] adopted data encapsulation method instead of key encapsulation method [3]–[6] to achieve confidentiality and unforgeability instantiation in the standard model.

Signcryption scheme unveiled in [7] did not propose a secu-

[*]Corresponding Author: Seth Alornyo, Koforidua Technical University, Computer Science Department, +233 554936729 & sabigseth@ktu.edu.gh

rity proof even though their construction were based on discrete logarithm assumptions. In view of this, several research in signcryption and signature schemes have been unveiled [8]–[15] with other functional extensions [16]–[19]. In 2011, a survey of identity-based signcryption cryptosystem was examined by Li et al.[20] to examine the security models as well as a comparative study of their security properties and efficiency. Analysis of other flavours of signcryption constructions [4], [11]–[13], [21]–[24], threshold signcryption [25]–[30], proxy signcryptions [31]–[37] and ring signcryption [38]–[42] have been proposed and constructed. Subsequently, Xiong et al. [43] did a cryptanalysis on the scheme [21]. The security notion of CPA in their work was diffused by Xiong et al. [43]. Hence, [43] showed that their scheme did not achieve chosen plaintext attack (CPA) security as they claimed.

Due to the need to resist the cloud server from peddling users outsourced data, efficient signcryption cryptosystem was unveiled in [44]. According to him, some existing signcryption schemes in [45]–[55] had certain lapses such as lack of data integrity, authentication and non-repudiation. Hence the need to construct a scheme to deny the insider adversary in the cloud from data peddling and modification for economic gains became paramount. Furthermore, the emergence of distributed computing and interconnected systems propelled [56] to unveiled a signcryption scheme in heterogeneous systems. Although in [57], the author earlier discussed the above problem, they could not construct a scheme to solve the problem of signcryption in heterogeneous systems. However, the constructions in [57, 58] and [59] achieved insider and outsider attack resistant respectively.

Recently, secure identity-based cryptosystem has been unveiled in [60]. Their security improvement was based on certain proposed signcryption algorithms [11]–[14], [61] constructed using the random oracle, and as well as schemes designed using the standard model [21, 23, 24, 62] with certain deficiencies such as IND-CCA2 and existential unforgeable chosen message attack (EUF-CMA). However, an attack was launched in [63] to unveil a new functional secure identity-based signcryption cryptosystem in [60]. A scheme to curtail an attack continuum in mobile money wallet system in Ghana to prevent message forgeability and re-play attack is still an open problem.

1.1 Our Contribution

By considering that the integrity, authentication and non-repudiation of the data in mobile money wallet system in Ghana, we proposed the mobile money wallet attack resistant scheme using the ID-based signcryption cryptographic primitive with equality test (known as MWAR-ID-SET) to achieve a simultaneous benefit of digital signature with public key encryption, and with equality test. Concretely, the formal definition, security model and concrete construction of MWAR-ID-SET are proposed in this paper. Further, our proposed scheme was shown to achieve the security property of existential unforgeable chosen message attack (EUF- CMA) by using the formal security proof.

1.2 Paper Organization

The rest of this work is organized as follows; In Section 2, our scheme outlines preliminaries for the construction and formulate the definitions of MWAR-ID-SET. In Section 3, the definitions of our scheme are outlined, section 4 outlines the security model of MWAR-ID-SET. Section 5 details the construction of our scheme, and proof the security in Section 6. Section 7 compares our work with existing schemes. Section 8 concludes our work.

2 Preliminaries

Definition 1: Bilinear Map. Let G and G_T be two multiplicative cyclic groups of prime order p. Suppose that q is agenerator of G. A bilinear map $e : G \times G \rightarrow G_T$ satisfies the following properties:

1. Bilinearity: For any $g \in G$, and $b \in Z_p$, $e(g^a, g^b) = e(g, g)^{ab}$.
2. Non-degenerate: $e : (g, g) \neq 1$.
3. Computable: There is an efficient algorithm to compute $e(g, g)$ for any $g \in G$.

Definition 2: Bilinear Diffie-Hellman (BDH) problem. Let G and G_T be two groups of prime order q. Let $e : G \times G \rightarrow G_T$ be an admissible bilinear map and let q be a generator of G. The BDH problem in (q, G, G_T, e) is as follows: Given (q, q^a, q^b, q^c), for random $c, d, f \in Z_p^*$, for any randomized algorithm. A computes the value $e(q, q)^{cdf} \in G_T$ with advantage:
$ADV_A^{BDH} Pr[A(q, q^c, q^d, q^f) = e(q, q)^{cdf}]$.

We say that the BDH assumption holds if for any polynomial-time algorithm A, its advantage ADV_A^{BDH} is negligible.

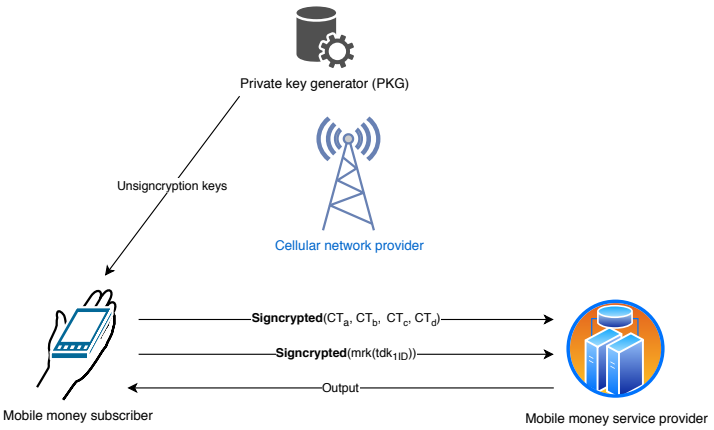


Figure 1: System model of our scheme

3 Definitions

In this section, the formal definition of our proposed scheme is outlined. MWAR-ID-SET achieves a formal security property of EUF-CMA. The construction specifies six(6) steps: $Setup$, $MW - Extract$, $TokenGen$, $MW - Signcrypt$, $MW - Unsigncrypt$, $MW - Test$. M_T and C_T are considered as the plaintext space and ciphertext space respectively.

1. $Setup$: The construction on input a security parameter k, output public parameter K with MSK as master secret key.

2. $MW - Extract$: The scheme on input MSK, $ID \in \{0,1\}$ arbitrary and gives out a message recovery key mrk corresponding to an identity.

3. $TokenGen$: With the input message recovery key mrk, arbitrary $ID \in \{0,1\}^*$ and it return token tdk_1 corresponding to an identity (ID_1).

4. $MW - Signcrypt$: The scheme on input $ID \in \{0,1\}^*$, random witness $w \in W$, a chosen plaintext $m_1 \in M_1$, and output ciphertext $CT_1 = (x_1, c_1)$ where $x \in X$ from generated witness $WInsGen(w) = x$ with the relation R satisfied [1].

5. $MW - Unsigncrypt$: On input the ciphertext CT_1, message recovery key mrk with a random chosen witness $w \in W$, the plaintext $m_1 \in M_1$ is uncovered provided CT_1 is deemed as a valid ciphertext derived from a witness relation R.

6. $MW - Test$: The scheme on input the ciphertext $CT_A \in CT_1$ with its corresponding receiver ID_A derived from the token tdk_{1_A} with its corresponding ID_A, the ciphertext $CT_A \in CT_1$ with its corresponding receiver ID_B derived from the token tdk_1 with its corresponding ID_B. The scheme respond 1, provided CT_A and CT_B have same message. Otherwise it respond as $\perp$.

4 Security Model

We assume $\sqcap = \{Setup, MW - Extract, TokenGen, MW - Signcrypt, MW-Unsigncrypt, MW-Test\}$ as the scheme and polynomial time algorithm adversary A. $MWAR - ID - SET$ achieves two main security notion of $IND - CCA2$ and $EUF - CMA$. However, our scheme adds the security notion of ID-based indistinquishability to IND-CCA2 and is coined as IND-ID-CCA2, similarly presented in [55] via the standard model..

1. $Setup$: The challenger A executes the security parameter k and outdoors K with a randomly chosen witness $w \in W$ and generates $x \in X$ of the relation R. It gives out the relation R to the adversary A.

2. Phase 1: The mobile merchant adversary A then issues $(\wedge_1, \wedge_2, ..., \wedge_{n-1})$. It is assumed that such query is of:

   - $MW-Query(ID_i)$: The merchant executes $H(.)$ and generates mrk_i which corresponding to ID_i as the identity. The recovered mrk_i is forwarded to A.
   - $TokenGen(ID_i)$: The merchant executes the $MW - Unsigncryption$ to generates tdk_i via the witness relation R. The algorithm then forwards tdk_i to A.

3. $MW - Challenge$: When the phase comes to an end, two messages (m_1, m_2) with equal-length and ID^* is submitted by A to the challenger and wishes to be challenged. But (m_1, m_2) were both not issued during signcryption and ID^* was never extracted during phase 1 section. Given this, the challenger randomly chooses $t \in \{0,1\}$ and returns $CT_1 = MW - Signcrypt(m_t, ID^*, w^*)$. Again, a challenge ciphertext token $mrk^* = (ID^*, x^*)$ is issued by the execution of $TokenGen$ phase $mrk^* \leftarrow mrk(tdk, m_t, x^*)$ and it output mrk^* to A.

4. $MW - Unsigncrypt$: The merchant executes an unsigncryption algorithm to unsigncrypt CT_1. The merchant obtains mrk_1 which corresponds to a public key of ID_i. The plaintext m_i is forwarded to A.

5. Phase 2: The merchant adversary issue the query $(\wedge_1, \wedge_2, ..., \wedge_n)$. The query is of the form:

   - $MW - Query$. A similar response as in phase 1 is given because $ID_i \neq ID^*$.
   - $TokenGen(ID_1)$. Given $x \neq x^*$, the merchant respond same as in phase 1.

6. Result. The adversary A does a guess v' of v. Since $v' = v$, then the adversary will win the game.

   The adversary advantage in breaking the scheme is noted as: $ADV_{MWAR-ID-SET}(k) = Pr[v' = v] - \frac{1}{2}$ as a negligible probability.

5 Construction

A detailed construction of our proposed scheme is outlined as follows:

1. On input a secured parameter k, the algorithm output public parameter K, with MSK as master secret key.

   - Two multiplicative group G and G_T are generated by the system with the same order of lemgth θ bits with a bilinear map $e : G \times G \rightarrow G_T$. The group generater $P \in G$ is selected.
   - The algorithm exploits keyed permutation $F : \{0,1\}^s \times \{0,1\}^n \rightarrow Z_p^*$ with a positive interger $D = k(i)$ and $L = b(i)$, it then activate a random value r_1 from $\{0,1\}^L$. Message authentication code (MAC), $MAC = GSV$. Thus, generate, sign, verify (GSV). After executing $G(i)$, it obtains r_2. It then set master token key as $MTK_1 = (r_1, r_2)$.
   - A hash functions $H_a : \{0,1\}^\tau \rightarrow Z_p^*$, $H_b : \{0,1\}^* \rightarrow G$, $H_c : A \times G \times G_T \rightarrow \{0,1\}^{\tau + r_1}$, where r_1 represent random numbers and τ as length of message. $(\tau_1, \tau_2) \in Z_p^2$ is randomly chosen and $R_v = P^{\tau_1}$, $R_m = P^{\tau_2}$. The paramrter $K = (A, \tau, G, G_T, P, R_v, R_m, MAC, H_a, H_b, H_c)$ is made public (published)

2. $MW-Extract$: With an $ID \in \{0,1\}$ as string, the system computes $Q_{ID} = H_b(ID) \in G$ and private key $mrk_{ID} = (Q_{ID}^{\tau_1}, Q_{ID}^{\tau_2})$. It should therefore be noted that (τ_1, τ_2) are secret value randomly chosen by the algorithm.

3. $TokenGen$: The algorithm with an input string $ID \in \{0,1\}^*$, the computation $Q_{ID} = H_b(ID) \in G$ is executed and the token $tdk_{1_{ID}} = (Q_{ID}^{\tau_2})$ is generated.

4. $MW - Signcrypt$: The algorithm with an input public parameter K, string ID, it executes $Q_{ID} = H_b(ID) \in G$ and a **signcryption** $m_1 \in G$ is triggered by choosing two random values $(q_1, q_2) \in Z_p^*$. The ciphertext is set as $CT_1 = (CT_a, CT_b, CT_c, CT_d)$ as :

$$CT_a = H_b(e(Q_{ID}, R_m)^{q_1}) \cdot m_1^{q_1}, \qquad CT_b = P^{q_1}$$

$$CT_c = P^{q_2} \quad CT_d = (m_1 \| w) \oplus H_b(CT_a \| CT_b \| Z \| e(Q_{ID}, R_v)^{q_2}).$$

The signature $Z \leftarrow S(r_2, CT_c)$ is used to signcrypt the ciphertext of the employed MAC. The signcrypted tag Z is used to verify signcrypted ciphertext CT_c

5. $MW - Unsigncrypt$: To **unsigncrypt**, the algorithm with an input the ciphertext CT_1, private **unsigncryption** key $mrk = (Q_{ID}^{\tau_1}, Q_{ID}^{\tau_2})$ with $CT_1 = (CT_a, CT_b, CT_c, CT_d)$ corresponding to an identity ID. The algorithm executes $(m_1 \| w^{'}) = CT_b \oplus H_b(CT_a \| CT_b \| CT_c \| e(CT_c, Q_{ID}^{\tau_1}))$. The algorithm do a check of $CT_a = (m_1^{'}) \| x^{'}$ and $CT_b = P^{q_1^{'}}$ to determine whether they are equal. If they are equal, the algorithm returns m_1, contrarily, it returns $\perp$.

6. $MW - Test$: With a given **signcrypted** ciphertext CT_{1_A} with trapdoor tdk_{1_A} and a different **signcrypted** ciphertext CT_{1_B} with a trapdoor tdk_{1_B}. The algorithm determines whether $m_{1_A} = m_{1_B}$ is equal or not. The algorithm does this by executing:

$$TD_A = \frac{CT_{a_A}}{H_b(e(CT_{a_A}, R_{m_A}))}, \quad TD_B = \frac{CT_{a_B}}{H_b(e(CT_{a_B}, R_{m_B}))}.$$

If the above equation holds, the algorithm then output 1 on success and 0 on failure.

Construction Correctness.
We assume that $CT_1 = (CT_a, CT_b, CT_c, CT_d)$.). Test algorithm executes:

$$CT_{1_A} = \frac{CT_{a_A}}{H_b(e(CT_{a_A}, R_{m_A}))}, \quad CT_{1_B} = \frac{CT_{a_B}}{H_b(e(CT_{a_B}, R_{m_B}))}$$

$$CT_{1_A} = \frac{H_b(e(Q_{ID}, P^{\tau_2})^{q_1}) \cdot m_{1_A}^{q_1}}{H_b(e(Q_{ID}, P^{\tau_2})^{q_1})}, CT_{1_B} = \frac{H_b(e(Q_{ID}, P^{\tau_2})^{q_1}) \cdot m_{1_B}^{q_1}}{H_b(e(Q_{ID}, P^{\tau_2})^{q_1})}$$

$$CT_{1_A} = m_{1_A}^{q_1}, \quad CT_{1_B} = m_{1_B}^{q_1}$$

Hence, $m_{1_A}^{q_1} = m_{1_B}^{q_1}$

From the above, the algorithm output 1 on success and 0 on failure.

Therefore:

$e(CT_{b_A}, R_{m_B}) = e(CT_{b_A}, R_{m_B})$.

$$e(CT_{b_A}, R_{m_B}) = (P^{q_1}, P^{\tau_2}) = e(P, P)^{q_1 \tau_2},$$

$$e(CT_{b_A}, R_{m_B}) = (P^{q_1}, P^{\tau_2}) = e(P, P)^{q_1 \tau_2}.$$

This implies that if, $m_{1_A} = m_{1_B}$, then

$e(CT_{b_A} = R_{m_B}) = e(CT_{b_A} = R_{m_B})$.

6 Security Analysis

In this section, we consider a formal security property of IND-CCA2 and EUF-CMA [60]. Our scheme adds the notion of ID-based indistinquishability to IND-CCA2 and referred to as IND-ID-CCA2.

6.1 IND-CCA2 Security

Our MWAR-ID-SET is $(SET_\varepsilon, t_s, q_{ks}, q_{ns}, q_{us})$-IND-CCA2 secure if $(\varepsilon_{mdbdh}, t_s) - mDBHDH$ assumption holds. Thus, H_1 and H_2 serves as (ε_{H_1}) and (ε_{h_2}) are both collision resistant hash functions, such that:

$$\varepsilon_{SET} \leq \varepsilon_{mdbdh} + \varepsilon_{H_1} + \varepsilon_{H_2} + \frac{q_{ks}+q_{us}+3}{p} + \frac{q_{ns}}{p^2}.$$

Where t_s refers to index period time, q_{ks} refers to number of extraction key queries, q_{ns} represent number of signcryption queries and q_{us} represents number of unsigncryption queries. Therefore, the security analysis with a collission resistant hash function proves our scheme secured.

6.2 EUF-CMA Unforgeability

Proof Theorem: This section outlines the security proof of unforgeability against adaptive CMA derived from the security constructions in Chow[41] ID-based **signcryption** cryptosystem. Thus, it is expected that the adversary can forge a ciphertextof a message m_1, if the assumption $CT_1 = (CT_a, CT_b, CT_c, CT_d)$ corresponding to a user identity ID holds. Thus, $CT_d = (m_1 \| w) \oplus H_3(CT_a \| CT_b \| CT_c \| e(Q_{ID}, R_v))^{q_2}$ is regarded as the signature of the message $m_1 \| w$, where $e(Q_{ID}, R_v)^{q_2}$ is regarded as the pairing of a corresponding user with secret **signcryption** key R_v. It is however noted that the difficulty of CDH problem makes the scheme unforgeable via the random oracle model.

6.3 Security Analysis of Token Key

Further details on the token security analysis can be accessed in [1]. The token security analysis experiment to our scheme is defined as: $EXP_{MWAR-ID-SET_A}^{IND-ID-CCA2}(k)$.

With a security parameter k, a master token key $MTK_1 = (r_1, r_2)$ and A adversary against token security. According to [1], the adversary A win the game if $b^{'} = b$, which shows that the output of the experiment is 1 on success and 0 on failure. Adversary A advantage in the experiment is defined as:

$Adv_{MWAR-ID-SET}^{IND-ID-CCA2}(w) = |Pr[Exp_{MWAR-ID-SET}^{IND-ID-CCA2}] - \frac{1}{2}|$

However, the probability for the adversary to win the game is negligible, hence proves our construction secured.

7 Comparison

A security strength comparison computations with existing **signcryption** schemes are outlined in Table 1. Constructions in ID-based **Signcryption** cryptosystem in [21, 23, 60, 64] are compared to with respect to security strength. The parameters for our comparison includes group multiplication, group exponentiation, inverse computations, pairing operation, test for equality, and support for EUF-CMA.

Table 1: The performance computational cost and Communication overheads

Scheme	G_{Mult}	G_{Exp_a}	$G_{T_{Mult}}$	$G_{T_{Exp_a}}$	$G_{T_{Inv}}$	Pr	IND-CCA2	EUF-CMA	ET
[21]	$2n_{u_a} + 2n_{m_a} + 1$	3	5	1	1	7(+2)	⊠	√	No
[23]	$2n_{u_a} + 2n_{m_a} + 1$	3	5	1	1	7(+2)	⊠	⊠	No
[60]	$2n_{u_a} + 2n_{m_a} + 1$	7	5	1	1	7(+2)	√	√	No
[64]	$2n_{u_a} + 2n_{m_a} + 3$	7	5	1	1	7	√	√	No
Ours	$2n_{u_a} + 2n_{m_a} + 1$	7	3	2	2	8	√	√	Yes

Legend: $''G_{Mult}$": multiplication in group G, $''G''_{Exp_a}$: exponentiation in group G, $''G''_{T_{Exp_a}}$: exponentiation in group G_T, $G_{T_{Inv}}$: inverse computation in group G_T, $''n_m, n''_u$: length of identity in bits string, $''Pr$": pairing operations in the form $x(+y)$ as in [23], $''ET''$: equality test, $''\boxtimes''$: not supportive, $''\surd''$: supportive .

Table 2: Running Times with Symbols

Symbols	Description	Times
G_{Exp_a}	G exponentiation operation	6.3937
$G_{T_{Exp_a}}$	G_T exponentiation operation	1.9518
T_{p_a}	Operation with pairing	11.4173
T_{h_a}	Hash functions	000853
G_{Mult}	Multiplication operation in G	0.047
$G_{T_{Mult_a}}$	Multiplication operation in G_T	0.0119

The Pairing-Based Cryptography (PBC) Library [65] is used to quantify the time consumption of signcryption, unsigncryption and test operations. We use the code of a program in VC++ 6.0 and executed on a computer (Windows 10 Pro, operating system), Capacity of Intel(R) Core (TM) i5-4460 CPU with 3.20GHZ and 4Gb RAM. The code was executed several times and average time of execution extracted (see Table 2). With respect to the scheme in [66], and other pairing-based constructions with a security level of 1024-bit RSA, a supersingular curve $z^2 = x^3 + x$ with an embedded degree of 2 is adopted. Also, $q = 2^{159} + 2^{17} + 1$ noted as a 160 bit Solinas prime noted as a $512 - bit$ prime. With regards to ECC-based schemes, an equivalent security level of Koblitz elliptic curve of $y = x^3 + ax^2 + b$ defined on a $F_{2^{163}}$ is used to provide the same security level in the ECC group. The computational units are in millisecond (ms) and bytes respectively. The execution times of each respective algorithm were calculated and Matlab program was used to generate Table 1 and computational results in Table 2.

Therefore, it is clear that our scheme support equality test as compared with other schemes without equality test. The computation of trapdoor and trapdoor delegation to the tester is equally achieved in our scheme. In terms of computational cost, our scheme has a lower computation cost as compared to existing schemes. Also, the support for IND-ID-CCA2 is featured in our sccheme as compared to other existing scheme with IND-CCA support. However, the inverse compuatation to recover the message achieves a ligh computational cost as compared to others in Table 1. This is pardonable due to the additional trapdoor computations featured in our scheme.

8 Conclusion

The construction $MWAR - ID - SET$ achieves a security improvement in mobile money service security in Ghana via the adoption signcryption cryptosystem. Data forgery and re-play attack is curtailed in our construction and the simultaneous benefit of PKE and digital signature is achieved in our scheme using the random oracle model. A desirable security property of EUF-CMA is achieved in our construction. In spite of other applications and extentions of identity based cryptosystem [67]–[70], our scheme achieves a remarkable achievement in indentity based cryptosystem.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We will like to use this opportunity to thank the anonymous reviewers for their kind support and consideration.

References

[1] S. Alornyo, K. K. Mireku, A. Tonny-Hagan, X. Hu, "Mobile Money Wallet Security against Insider Attack Using ID-Based Cryptographic Primitive with Equality Test," in 2019 International Conference on Cyber Security and Internet of Things (ICSIoT), 82–87, IEEE, 2019, doi:10.1109/ICSIoT47925.2019.00021.

[2] F. Li, H. Xiong, Y. Liao, "A generic construction of identity-based signcryption," in 2009 International Conference on Communications, Circuits and Systems, 291–295, IEEE, 2009, doi:10.1109/ICCCAS.2009.5250509.

[3] K. Bentahar, P. Farshim, J. Malone-Lee, N. P. Smart, "Generic constructions of identity-based and certificateless KEMs," Journal of Cryptology, **21**(2), 178–199, 2008, doi:10.1007/s00145-007-9000-z.

[4] L. Chen, Z. Cheng, J. Malone-Lee, N. P. Smart, "An Efficient ID-KEM Based On The Sakai-Kasahara Key Construction." IACR Cryptol. ePrint Arch., **2005**, 224, 2005.

[5] E. Kiltz, D. Galindo, "Direct chosen-ciphertext secure identity-based key encapsulation without random oracles," in Australasian Conference on Information Security and Privacy, 336–347, Springer, 2006, doi:https://doi.org/10.1007/11780656_28.

[6] A. Shamir, "Identity-based cryptosystems and signature schemes," in Workshop on the theory and application of cryptographic techniques, 47–53, Springer, 1984, doi:https://doi.org/10.1007/3-540-39568-7_5.

[7] Y. Zheng, "Digital signcryption or how to achieve cost (signature or encryption) cost (signature)+ cost (encryption)," in Annual international cryptology conference, 165–179, Springer, 1997, doi:https://doi.org/10.1007/BFb0052234.

[8] A. Fiat, A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems," in Conference on the theory and application of cryptographic techniques, 186–194, Springer, 1986, doi:https://doi.org/10.1007/3-540-47721-7_12.

[9] L. C. Guillou, J.-J. Quisquater, "A "paradoxical" indentity-based signature scheme resulting from zero-knowledge," in Conference on the Theory and Application of Cryptography, 216–231, Springer, 1988, doi:https://doi.org/10.1007/0-387-34799-2_16.

[10] J. Malone-Lee, "Identity-Based Signcryption." IACR Cryptol. ePrint Arch., **2002**, 98, 2002.

[11] P. S. Barreto, B. Libert, N. McCullagh, J.-J. Quisquater, "Efficient and provably-secure identity-based signatures and signcryption from bilinear maps," in International conference on the theory and application of cryptology and information security, 515–532, Springer, 2005, doi:https://doi.org/10.1007/11593447_28.

[12] X. Boyen, "Multipurpose identity-based signcryption," in Annual International Cryptology Conference, 383–399, Springer, 2003, doi:https://doi.org/10.1007/978-3-540-45146-4_23.

[13] S. S. Chow, S.-M. Yiu, L. C. Hui, K. Chow, "Efficient forward and provably secure ID-based signcryption scheme with public verifiability and public ciphertext authenticity," in International Conference on Information Security and Cryptology, 352–369, Springer, 2003, doi:https://doi.org/10.1007/978-3-540-24691-6_26.

[14] B. Libert, J.-J. Quisquater, "A new identity based signcryption scheme from pairings," in Proceedings 2003 IEEE Information Theory Workshop (Cat. No. 03EX674), 155–158, IEEE, 2003, doi:10.1109/ITW.2003.1216718.

[15] T. H. Yuen, V. K. Wei, "Constant-Size Hierarchical Identity-Based Signature/Signcryption without Random Oracles." IACR Cryptol. ePrint Arch., **2005**, 412, 2005.

[16] Y. Zheng, H. Imai, "How to construct efficient signcryption schemes on elliptic curves," Information processing letters, **68**(5), 227–233, 1998, doi:https://doi.org/10.1016/S0020-0190(98)00167-7.

[17] F. Bao, R. H. Deng, "A signcryption scheme with signature directly verifiable by public key," in International Workshop on Public Key Cryptography, 55–59, Springer, 1998, doi:https://doi.org/10.1007/BFb0054014.

[18] J.-B. Shin, K. Lee, K. Shim, "New DSA-verifiable signcryption schemes," in International Conference on Information Security and Cryptology, 35–47, Springer, 2002, doi:https://doi.org/10.1007/3-540-36552-4_3.

[19] D. H. Yum, P. J. Lee, "New signcryption schemes based on KCDSA," in International Conference on Information Security and Cryptology, 305–317, Springer, 2001, doi:https://doi.org/10.1007/3-540-45861-1_23.

[20] F. Li, M. K. Khan, "A survey of identity-based signcryption," IETE Technical Review, **28**(3), 265–272, 2011, doi:https://doi.org/10.4103/0256-4602.81236.

[21] Y. Yu, B. Yang, Y. Sun, S.-l. Zhu, "Identity based signcryption scheme without random oracles," Computer Standards & Interfaces, **31**(1), 56–62, 2009, doi:https://doi.org/10.1016/j.csi.2007.10.014.

[22] K. G. Paterson, J. C. Schuldt, "Efficient identity-based signatures secure in the standard model," in Australasian Conference on Information Security and Privacy, 207–222, Springer, 2006, doi:https://doi.org/10.1007/11780656_18.

[23] Z. Jin, Q. Wen, H. Du, "An improved semantically-secure identity-based signcryption scheme in the standard model," Computers & Electrical Engineering, **36**(3), 545–552, 2010, doi:https://doi.org/10.1016/j.compeleceng.2009.12.009.

[24] F. Li, Y. Liao, Z. Qin, "Analysis of an identity-based signcryption scheme in the standard model," IEICE transactions on fundamentals of electronics, communications and computer sciences, **94**(1), 268–269, 2011, doi:https://doi.org/10.1587/transfun.E94.A.268.

[25] F. Li, J. Gao, Y. Hu, "ID-based threshold unsigncryption scheme from pairings," in International Conference on Information Security and Cryptology, 242–253, Springer, 2005, doi:https://doi.org/10.1007/11599548_21.

[26] S. Duan, Z. Cao, R. Lu, "Robust ID-based threshold signcryption scheme from pairings," in Proceedings of the 3rd international conference on Information security, 33–37, 2004, doi:https://doi.org/10.1145/1046290.1046298.

[27] C. Peng, X. Li, "An identity-based threshold signcryption scheme with semantic security," in International Conference on Computational and Information Science, 173–179, Springer, 2005, doi:https://doi.org/10.1007/11596981_26.

[28] F. Li, Y. Yu, "An efficient and provably secure ID-based threshold signcryption scheme," in 2008 International Conference on Communications, Circuits and Systems, 488–492, IEEE, 2008, doi:10.1109/ICCCAS.2008.4657820.

[29] S. S. D. Selvi, S. S. Vivek, C. P. Rangan, N. Jain, "Cryptanalysis of Li et al.'s identity-based threshold signcryption scheme," in 2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, **2**, 127–132, IEEE, 2008, doi:10.1109/EUC.2008.187.

[30] F. Li, X. Xin, Y. Hu, "ID-based threshold proxy signcryption scheme from bilinear pairings," International Journal of Security and Networks, **3**(3), 206–215, 2008, doi:https://doi.org/10.1504/IJSN.2008.020095.

[31] X. Li, K. Chen, "Identity based proxy-signcryption scheme from pairings," in IEEE International Conference onServices Computing, 2004.(SCC 2004). Proceedings. 2004, 494–497, IEEE, 2004, doi:10.1109/SCC.2004.1358050.

[32] M. Wang, H. Li, Z. Liu, "Efficient identity based proxy-signcryption schemes with forward security and public verifiability," in International Conference on Networking and Mobile Computing, 982–991, Springer, 2005, doi:https: //doi.org/10.1007/11534310 103.

[33] Q. Wang, Z. Cao, "Two proxy signcryption schemes from bilinear pairings," in International Conference on Cryptology and Network Security, 161–171, Springer, 2005, doi:https://doi.org/10.1007/11599371 14.

[34] S. Duan, Z. Cao, Y. Zhou, "Secure delegation-by-warrant ID-based proxy signcryption scheme," in International Conference on Computational and Information Science, 445–450, Springer, 2005, doi:https://doi.org/10.1007/11596981-65.

[35] M. Wang, Z. Liu, "Identity based threshold proxy signcryption scheme," in The Fifth International Conference on Computer and Information Technology (CIT'05), 695–699, IEEE, 2005, doi:10.1109/CIT.2005.129.

[36] J. Zhang, J. Mao, "A novel identity-based multi-signcryption scheme," Computer Communications, **32**(1), 14–18, 2009, doi:https://doi.org/10.1016/j.comcom.2008.07.004.

[37] S. S. D. Selvi, S. S. Vivek, C. P. Rangan, "Breaking and fixing of an identity based multi-signcryption scheme," in International Conference on Provable Security, 61–75, Springer, 2009, doi:https://doi.org/10.1007/978-3-642-04642-1 7.

[38] M. Zhang, B. Yang, S. Zhu, W. Zhang, "Efficient secret authenticatable anonymous signcryption scheme with identity privacy," in International Conference on Intelligence and Security Informatics, 126–137, Springer, 2008, doi:https://doi.org/10.1007/978-3-540-69304-8 14.

[39] F.-g. Li, S. Masaaki, T. Tsuyoshi, "Analysis and improvement of authenticatable ring signcryption scheme," Journal of Shanghai Jiaotong University (Science), **13**(6), 679–683, 2008, doi:https://doi.org/10.1007/s12204-008-0679-2.

[40] Z. Zhu, Y. Zhang, F. Wang, "An efficient and provable secure identity-based ring signcryption scheme," Computer standards & interfaces, **31**(6), 1092–1097, 2009, doi:https://doi.org/10.1016/j.csi.2008.09.023.

[41] Y. Yu, F. Li, C. Xu, Y. Sun, "An efficient identity-based anonymous signcryption scheme," Wuhan University Journal of Natural Sciences, **13**(6), 670–674, 2008, doi:https://doi.org/10.1007/s11859-008-0607-1.

[42] L. Zhun, F. Zhang, "Efficient ID–based ring signature and ring signcryption schemes," in 2008 International Conference on Computational Intelligence and Security, volume 2, 303–307, IEEE, 2008, doi:10.1109/CIS.2008.51.

[43] H. Xiong, Z. Qin, F. Li, "Cryptanalysis of an Identity Based Signcryption without Random Oracles," Fundamenta Informaticae, **107**(1), 105–109, 2011, doi:10.3233/FI-2011-395.

[44] F. Li, B. Liu, J. Hong, "An efficient signcryption for data access control in cloud computing," Computing, **99**(5), 465–479, 2017, doi:https://doi.org/10.1007/s00607-017-0548-7.

[45] V. Kirtane, C. P. Rangan, "RSA-TBOS signcryption with proxy re-encryption," in Proceedings of the 8th ACM workshop on Digital rights management, 59–66, 2008, doi:https://doi.org/10.1145/1456520.1456531.

[46] J. Malone-Lee, W. Mao, "Two birds one stone: signcryption using RSA," in Cryptographers' Track at the RSA Conference, 211–226, Springer, 2003, doi:https://doi.org/10.1007/3-540-36563-X_14.

[47] C. Wang, X. Cao, "An improved signcryption with proxy re-encryption and its application," in 2011 Seventh International Conference on Computational Intelligence and Security, 886–890, IEEE, 2011, doi:10.1109/CIS.2011.200.

[48] W. Huige, W. Caifen, C. Hao, "ID-based proxy re-signcryption scheme," in 2011 IEEE International Conference on Computer Science and Automation Engineering, volume 2, 317–321, IEEE, 2011, doi:10.1109/CSAE.2011.5952478.

[49] X. Tian, X. Wang, A. Zhou, "DSP RE-Encryption: A flexible mechanism for access control enforcement management in DaaS," in 2009 IEEE International Conference on Cloud Computing, 25–32, IEEE, 2009, doi:10.1109/CLOUD.2009.65.

[50] Q. Liu, C. C. Tan, J. Wu, G. Wang, "Reliable re-encryption in unreliable clouds," in 2011 IEEE Global Telecommunications Conference-GLOBECOM 2011, 1–5, IEEE, 2011, doi:10.1109/GLOCOM.2011.6133609.

[51] S. Yu, C. Wang, K. Ren, W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in 2010 Proceedings IEEE INFOCOM, 1–9, Ieee, 2010, doi:10.1109/INFCOM.2010.5462174.

[52] M. Li, S. Yu, Y. Zheng, K. Ren, W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption," IEEE transactions on parallel and distributed systems, **24**(1), 131–143, 2012, doi:10.1109/TPDS.2012.97.

[53] M. Nabeel, N. Shang, E. Bertino, "Privacy preserving policy-based content sharing in public clouds," IEEE Transactions on Knowledge and Data Engineering, **25**(11), 2602–2614, 2012, doi:10.1109/TKDE.2012.180.

[54] K. Yang, X. Jia, "Expressive, efficient, and revocable data access control for multi-authority cloud storage," IEEE transactions on parallel and distributed systems, **25**(7), 1735–1744, 2013, doi:10.1109/TPDS.2013.253.

[55] J. Hur, "Improving security and efficiency in attribute-based data sharing," IEEE transactions on knowledge and data engineering, **25**(10), 2271–2282, 2011, doi:10.1109/TKDE.2011.78.

[56] F. Li, H. Zhang, T. Takagi, "Efficient signcryption for heterogeneous systems," IEEE Systems Journal, **7**(3), 420–429, 2013, doi:10.1109/JSYST.2012.2221897.

[57] Y. Sun, H. Li, "Efficient signcryption between TPKC and IDPKC and its multi-receiver construction," Science China Information Sciences, **53**(3), 557–566, 2010, doi:https://doi.org/10.1007/s11432-010-0061-5.

[58] Q. Huang, D. S. Wong, G. Yang, "Heterogeneous signcryption with key privacy," The Computer Journal, **54**(4), 525–536, 2011, doi:10.1093/comjnl/bxq095.

[59] J. H. An, Y. Dodis, T. Rabin, "On the security of joint signature and encryption," in International Conference on the Theory and Applications of Cryptographic Techniques, 83–107, Springer, 2002, doi:https://doi.org/10.1007/3-540-46035-7_6.

[60] F. Li, T. Takagi, "Secure identity-based signcryption in the standard model," Mathematical and Computer Modelling, **57**(11-12), 2685–2694, 2013, doi:https://doi.org/10.1016/j.mcm.2011.06.043.

[61] L. Chen, J. Malone-Lee, "Improved identity-based signcryption," in International Workshop on Public Key Cryptography, 362–379, Springer, 2005, doi:https://doi.org/10.1007/978-3-540-30580-4_25.

[62] B. Waters, "Efficient identity-based encryption without random oracles," in Annual International Conference on the Theory and Applications of Cryptographic Techniques, 114–127, Springer, 2005, doi:https://doi.org/10.1007/11426639_7.

[63] Y. Ming, Y. Wang, "Cryptanalysis of an Identity Based Signcryption Scheme in the Standard Model." IJ Network Security, **18**(1), 165–171, 2016, doi:10.6633/IJNS.

[64] G. Zhu, H. Xiong, Z. Qin, "Fully secure identity based key-insulated signcryption in the standard model," Wireless personal communications, **79**(2), 1401–1416, 2014, doi:https://doi.org/10.1007/s11277-014-1936-3.

[65] B. Lynn, et al., "The stanford pairing based crypto library," Privacy preservation scheme for multicast communications in smart buildings of the smart grid, **324**, 2013.

[66] H. Xiong, Q. Mei, Y. Zhao, "Efficient and provably secure certificateless parallel key-insulated signature without pairing for IIoT environments," IEEE Systems Journal, **14**(1), 310–320, 2019, doi:10.1109/JSYST.2018.2890126.

[67] S. Alornyo, M. Asante, X. Hu, K. K. Mireku, "Encrypted Traffic Analytic using Identity Based Encryption with Equality Test for Cloud Computing," in 2018 IEEE 7th International Conference on Adaptive Science & Technology (ICAST), 1–4, IEEE, 2018, doi:10.1109/ICASTECH.2018.8507063.

[68] S. Alornyo, E. Aidoo, K. K. Mireku, B. Kwofie, X. Hu, M. Asante, "ID-Based Outsourced Plaintext Checkable Encryption in Healthcare Database," in 2019 International Conference on Cyber Security and Internet of Things (ICSIoT), 48–53, IEEE, 2019, doi:10.1109/ICSIoT47925.2019.00016.

[69] S. Alornyo, Y. Zhao, G. Zhu, H. Xiong, "Identity Based Key-Insulated Encryption with Outsourced Equality Test." IJ Network Security, **22**(2), 257–264, 2020, doi:10.6633/IJNS.860.

[70] S. Alornyo, A. E. Mensah, A. O. Abbam, "Identity-based Public Key Cryptographic Primitive with Delegated Equality Test Against Insider Attack in Cloud Computing," International Journal of Network Security, **22**(5), 743–751, 2020, doi:10.6633/IJNS.202009_22(5).04.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 1002-1012 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Assessment of the Municipal Solid Waste Transfer Stations Suitability in Harare, Zimbabwe

Trust Nhubu[1,*], Edison Muzenda[1,2], Mohamed Belaid[1], Charles Mbohwa[3]

[1]*Department of Chemical Engineering Technology, University of Johannesburg, Johannesburg, 0027, South Africa*

[2]*Department of Chemical, Materials and Metallurgical Engineering, Botswana International University of Science and Technology, Palapye, 00267, Botswana*

[3]*Department of Quality and Operations Management, University of Johannesburg, Johannesburg, 0027, South Africa*

ARTICLE INFO

Article history:
Received: 22 December, 2020
Accepted: 19 February, 2021
Online: 22 April, 2021

Keywords:
Municipal solid waste
Life cycle impacts
Waste transfer station
Integrated solid waste management
Citizens drop off points
Buy back centres
Recycling
Harare
Zimbabwe

ABSTRACT

The suitability of incorporating waste transfer stations (WTS) in likely future Municipal Solid Waste Management (MSWM) systems for Harare city and neighbouring urban centres was assessed under this study. WTS will bring about location of landfills and other MSWM facilities further away from population centres, increasing recycling, reducing waste collection costs and burden on the overall MSWM budget, an increase in waste collection effectiveness and efficiency, reduction in waste collection derived greenhouse gases (GHG) emissions and other associated impacts. Life cycle Assessment (LCA) on contribution of waste collection to human health impact potential of 34 DALYs as well as acidification, global warming and eutrophication impact potentials of 0.012, 0.065 and 0.0002 species.year respectively under all the six MSWM options were observed. Highest impacts in the species extinction impact categories was realised in the global warming impact category resultant of GHG emissions from fuel combustion during waste collection. The unavailability of land and the above factors support the incorporation of WTS in future MSWM options for Harare City and surrounding urban centres under a separation at source waste collection system to derive maximum benefits. Citizens drop off centres (CDOPs) and buy-back centres (BBCs) could also compliment the WTS leading to increased recycling. Though there is a relative sound supportive legislative, regulatory and policy framework that supports the need for waste recycling consequently supporting WTS, CDOPS and BBCs due to their recycling promotion capabilities, there is need for specific legislation, regulation and policies that supports the development and operation of such facilities that will bring interest amongst would be operators, effectiveness and efficiency resultantly reducing associated human health and environmental impacts. Further studies that determine the breakeven distance and LCA studies that specifically assess the associated environmental loads of incorporating WTS within the likely future MSWM systems are recommended.

1. Introduction

This paper is a revised and extended version of a paper entitled, "Suitability of municipal solid waste transfer stations in Harare, Zimbabwe", that was presented at the 7th International Renewable and Sustainable Energy Conference (IRSEC' 19), Sofitel Agadir Royal Bay, Agadir, Morocco, November 27-30, 2019.

The continued population increases together with the prevailing rapid urbanization has resulted in astronomical increase in MSW generation rendering the management of municipal solid waste (MSW) a challenging and difficult task for urban local authorities (ULAs) especially in low income countries. In [1], the authors projected annual global MSW generation to increase from 1.3 x 10^9 tons (1.42kg/capita/day) in 2012 to 2.2 x 10^9 tons in 2025. This results in an evident mismatch between increased population and urbanisation rate with the necessary corresponding municipal solid waste management (MSWM) investment, infrastructure,

*Corresponding Author: Trust Nhubu, Department of Chemical Engineering Technology, University of Johannesburg, P O Box 17011, 2028, South Africa, 2094, +263777342295, nhubutrust@gmail.com

equipment and expertise in the face of the MSW generation spike [2]. Significant increase in urbanization from general population increases coupled with rural to urban migration in pursuit of better life and jobs [3] has been a global phenomenon characterizing the recent past decades. In [4], the author reported global urbanization levels of 29.6% and 55.3% in 1950 and 2018 respectively, simultaneously projecting global urbanisation to increase to 68.4% by the middle of the 21st century. Increasing urbanisation rates were reported in the 1982, 1992, 2002 and 2012 Zimbabwe national census reports [5]. The increase in the urbanisation is also confirmed by [4] which reported urbanisation levels of 32.2% in 2018 and 45.9% in 2050 for Zimbabwe as well as 40.4% in 2018 and 58.1% in 2050 for Sub Saharan Africa (SSA) implying that by mid-21st century, urban population in SSA would have outstripped rural population constituting over 58% of the total population. Increased population densities together with the corresponding increases in MSW generation characterise these reported increasing urbanization rates [6, 7]. The increase in MSW generation is significantly not matched with the required corresponding expertise and infrastructural (equipment and technology) development with regards to MSW storage, collection and transportation, management, treatment and disposal.

Increased MSW generation associated challenges are realised or manifest first in the first stages of a MSWM system namely storage, collection and transportation. MSW collection figures of slightly above 40% were reported by [1] for developing nations as shown in Figure 1. MSW collection figures of below half of the MSW generated in the Zimbabwean ULAs have been reported with the collected waste disposed of at official open dumpsites whilst the over 50% non-collected MSW generated is subjected to any of the following illegal methods namely; burning, burying, dumping in undesignated areas like road sides, vacant spaces, alleys, storm water drains. All these MSW disposal methods including the dumping of the officially collected MSW at official dumpsites pose potential and imminent human health and environmental disasters already manifesting or in the making. MSW disposal thus becomes as equally vital stage as the first stages of a MSWM system. Figure 2 provides a comparative overview of the dominant MSW disposal methods in Harare (Zimbabwe) as reported by [8] and those reported by [1] for low income countries.

In Zimbabwe, ULAs including municipalities, city and town councils, rural district councils as well as local boards are authorised entities responsible for MSWM [9] largely under centralised and conventional models inherited from the colonial error which has no clear mechanisms for cooperation between the ULAs and other stakeholders especially residents to be served by such a MSWM system who have developed a ***'we dumb they collect attitude'*** [10]. ULAs are authorised also to engage other players for the purposes of managing waste on their behalf. MSWM was singled out by [11] amongst the enormous challenges facing ULAs within the vicinity of Harare the Capital City of Zimbabwe including Chitungwiza municipality, Epworth, Ruwa and Norton local boards largely due to the astronomical increase in MSW generation presently estimated at above 400,000 tons [12] which has not been matched with corresponding increases in capacity development with regards to financial, infrastructure (equipment and technology) and technical expertise. The population increases together with the associated increased population density and the ensuing MSWM challenges from the mismatch between the population dynamics and developments in MSWM with regards to infrastructure (technology and equipment), finance and investment and technical expertise to manage the ever-increasing amount of MSW being generated evidently manifest in the form of water pollution (both surface and groundwater) together with the perennial water borne diseases outbreaks namely Cholera, Typhoid and Dysentery [13-15].

MSW collection and transportation (MSW-CT) from its sources to final disposal dumpsites in Harare city, Chitungwiza, Ruwa, Epworth and Norton consume more than the expected reported proportion on MSWM budget of between 20 to 50% [16] and 50 to 70% [17]. Therefore, there is need to redesign future MSW-CT to possibly allow for significant cost reductions on MSW-CT cost which will result in savings on the MSWM budget with the savings possibly channeled towards other developments in other MSWM downstream processes, infrastructure and technical expertise. The incorporation of appropriately designed, sited and constructed WTS in MSWM systems is such a strategy towards bringing cost effective and savings in MSW-CT [18-21]. WTS apart from their costs abatement capabilities they could possibly bring about reduced environmental impacts of future MSWM systems that could be implemented in Harare city, Chitungwiza, Ruwa, Epworth and Norton. In [22], the authors observed that the recycling of atleast 20% of the recoverable materials contributes to significant reductions of environmental impacts (acidification, eutrophication, human health and global warming) of MSWM options that could be earmarked for future implementation in Harare city, Chitungwiza, Ruwa, Epworth and Norton. WTS have largely been reported that they can act as platforms for recycling and materials recovery hence their incorporation in future MSWM systems as platforms for recycling and materials recovery will result in reduced environmental impacts. This study therefore complements extensive literature review with LCA of the suitability of WTS for Harare city, Chitungwiza, Ruwa, Epworth and Norton considering the need to redesign and implement future sustainable MSWM systems along the integrated waste management model, the likely capacity exhaustion by 2020 of the only dumpsite for Harare (Pomona dumpsite), key national priorities and objectives under the Zimbabwe National Integrated Solid Waste Management Plan (ZNISWMP) that came into effect in July of 2014 and the Zimbabwean Government waste sector Low Emission Development Strategies (LEDS) for the Nationally Determined Contributions (NDCs) to its Paris Agreement differential obligations or responsibilities under the United Nations Framework Convention on Climate Change (UNFCCC).

2. Waste Transfer Stations

WTS provide linkages between MSW generation, storage, collection and transportation and final disposal (landfill or MSW processing facility) for a specific spatial scale (locality or community over a given temporal scale [23]. WTS consolidate MSW from different primary small or low volume MSW collection vehicles from MSW sources into secondary large or high volume collection vehicles to effect low MSW transportation cost by using the secondary large or high volume collection vehicles to haul MSW to distantly located waste treatment (anaerobic digestion and or a composting, incineration, pyrolysis,

gasification, waste to energy facilities etc) and disposal sites mainly landfills. Material Recovery Facilities (MRFs) are specialised WTS where recyclables are segregated and consolidated prior to waste transportation to landfills or other MSW treatment and disposal facilities [19, 21]. Therefore WTS operating as MRFs diverts waste sent to landfills thereby bringing along with significant environmental benefits such as reduced land sizes for landfills and demand for land considering its competing interests amongst its uses.

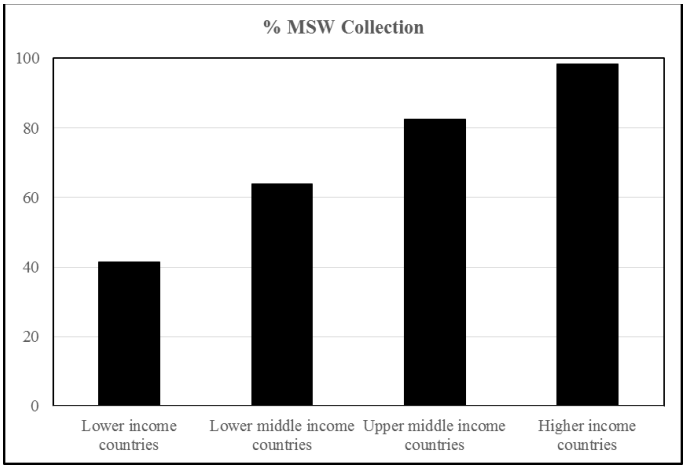


Figure 1: MSW collection according to country income levels [1]

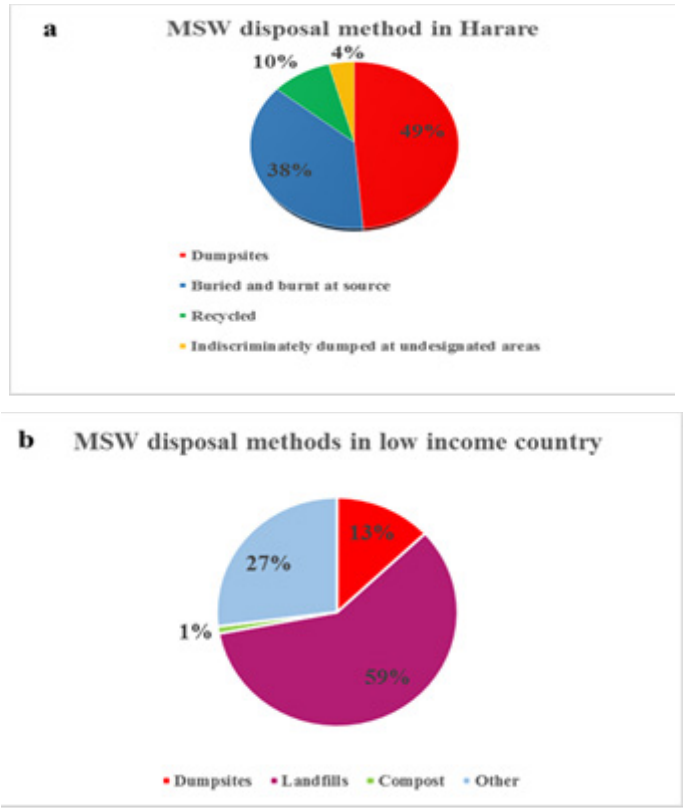


Figure 2: Dominant disposal methods in Harare (a) [8, 24] and low income countries (b) [1]

2.1. Reasons for Waste Transfer Stations

WTS bring about a number of benefits to a MSWM system including reduced transport costs, reduced pollution, easing of selecting new sites for MSW disposal or processing facilities, enhanced timeous collection and improved MSW collection efficiency. The construction of MSW-WTS results from the need for remote and regional landfills which are large and require tracks of land and have to be located further away from urban population centres where there is no land for constructing such large landfills coupled with the human health concerns of having landfills near population centres emanating from flies, rodents, mosquitoes as well as stray animals [2, 18]. Low tipping fees can be charged at large regional landfills due to their economies of scale which enable them to handle high MSW volumes from different sources under wider spatial scales thereby attracting more clients. Where MSW disposal and processing plants are over 15km from MSW sources from which MSW is collected, there is a possibility of establishing WTS at strategic sites within an urban area to cut on MSW transportation costs. The general rule with regards to the siting of MSW disposal sites and processing facilities is that they have to be established distant away from the urban population centres due to their associated potential human health and environmental impacts. This makes the primary MSW collection vehicles to travel long distances from MSW sources to the disposal sites and processing facilities leading to increased duration of travelling to and from disposal and processing facilities, increased number of primary MSW collection vehicles and cost in the MSW collection system. The increased number of primary MSW collection vehicles emanates from increased MSW generation coupled with the increased duration the vehicles take to and from MSW disposal sites and processing facilities since primary waste collection vehicles are light weight carrying vehicles capable to navigate the streets collecting waste at sources a difficult undertaking using heavy vehicles. WTS therefore provide the necessary infrastructure or platform for the transfer of MSW from light weight carrying vehicles to heavy weight carrying vehicles commonly referred to as secondary waste carrying vehicles. This combination of low weight and heavy weight carrying vehicles in a MSW collection system brings about costs reduction compared to the use of primary waste collection vehicles alone.

Apart from waste transportation cost reduction, the consolidation of light waste weight or smaller waste volumes from light weight carrying vehicles into heavy weight carrying vehicles increases the effectiveness and efficiency of a waste collection system through the reduction of the time waste collection teams take transporting MSW between MSW sources and disposal and or processing facilities giving more time for waste collection [25]. Ultimately this reduces the number of light weight carrying vehicles subsequently reducing pollution, noise and traffic congestion within an urban centre. Considering the atleast 40% organic waste composition in MSW generated in Harare [12] and high temperatures for Harare which speeds up the decomposition of biodegradable waste, WTS provide mechanisms for frequent and timeous biodegradable MSW collection prior to its decomposition resultantly avoiding potential human health and environmental risks, odours etc as argued by [26]. The timeous or just in time MSW collection thus is enhanced by the increased collection frequency brought about by the WTS emanating from the reduced durations taken and distances travelled by primary waste collection vehicles collecting waste from sources to WTS.

MSW transportation is associated with the consumption of fuel by waste collection vehicles subsequently causing air pollution and other accompanying environmental and human health impacts. The Global Warming Potential (GWP) emanates from the

emission of greenhouse gas (GHG) from waste collection vehicles travelling to and from waste disposal and processing facilities as well as during waste collection at sources. Therefore, the associated reductions in times spent and distances travelled by primary waste collection vehicles or at source collection vehicles brought about by incorporating WTS in MSWM systems reduce the associated potential global warming impacts and other environmental impacts. This is so because WTS reduces the total distance MSW-CT vehicles travel resulting in reduced fuel consumption, GHG emissions and costs of maintaining MSW-CT vehicles.

The United Nations Environmental Program (UNEP) reported over ten thousand climate change induced disasters to have occurred and been witnessed, destroyed and killed property and millions respectively, affecting many more globally. MSW was cited by [27] amongst the primary contributors to the emissions of GHG that cause global warming or climate change that has led to these disasters. MSW derived GHG emissions significantly emanate from MSW-CT. Two hundred and fifty million tons of MSW were reported by [28] to have been processed in the United States of America, producing GHG emissions that accounted for over 8% of non-energy derived GHG emissions and 2% of total net GHG emissions. The projected increase in global MSW generation from 1.3 billion tons to 2.2 billion tons in 2011 and 2025 respectively [1, 28], has brought and anticipated to result in increased MSW-CT derived GHG emissions. Efforts towards reducing MSW derived GHG emissions require the design, development and implementation of technically feasible and sound, economically viable, environmentally sustainable and socially acceptable MSW-CT systems [29] that makes the reduction of carbon dioxide emission feasible [30]. In [31], the authors therefore, noted the appropriateness with regards to financial savings and environmental benefits include land use planning of incorporating WTS in the MSW-CT systems within developing nations.

WTS also ease the selection of new sites for waste disposal and processing by eliminating the overreliance on at source collection vehicles accessibility to the sites giving preferences and considerations to human health and environmental concerns. Therefore, WTS allows for waste disposal and processing facilities location further away from human settlements thereby reducing human health risks from disease causing vectors like mosquitoes, flies, stray animals, rodents and leachates. WTS further provide platforms for waste screening before final disposal, the needed flexibility to choose MSW treatment and disposal options as well as serving as citizen drop off station.

2.2. Siting WTS

The selection or identification of a suitable location of a WTS is a tedious and difficult process. Environmental and human health potential hazards need to be taken into consideration when assessing all potential WTS locations. Such factors include but not limited to potential nuisance, odours, noise from primary and secondary waste collection vehicles as well as WTS equipment, proximity and accessibility to sources with high waste throughput. A number of potential locations are identified first after which a comparative evaluation is undertaken to identify the best site that has minimum impacts with regards to some or all of the factors highlighted above. A number of studies have been conducted during the past twenty years using different approaches and considering different factors to identify the best WTS locations in different countries. The need for waste management cost minimisation as a factor for best WTS location was observed in most of the studies namely; [32]-[40] etc. In [37], the authors further observed other economic, environmental and social benefits of WTS incorporation and its suitable location in a waste management system that results in the simultaneous reduction of GHG emissions and resultant visual pollution apart from the reduced overall waste management costs in Tehran, Iran. In [34] and [40], the authors also noted revenue generation from waste processing facilities associated with WTS. For Harare therefore, MSWM costs and pollution reduction as well as revenue generation from MSW processing should be considered when deciding for the establishment and location of WTS. Thus, sustainability indicators for WTS in Harare will be based on cost effectiveness, reduction in water, land and water pollution, ability to accommodate the already at source collection vehicles being used by the ULAs, local expertise capable of designing, constructing and operating the WTS and energy efficiency.

2.3. Economic viability of WTS

As previously highlighted that WTS reduces the duration at source waste collection vehicles take collecting MSW and hence the reduction of primary MSW collection costs as well as the number of involved vehicles, economic viability of WTS is based on the breakeven point. This involves the comparison of WTS costs, hauling costs of primary or at source waste collection vehicles and secondary waste collection vehicles. Hauling costs have variable and fixed costs. Variable costs include fuel consumption, tyre, maintenance, depreciation, engine oil, and grease and waste collection crew wages. WTS costs encompass construction, operation and maintenance costs, calculated based on capacity basis with the distances between MSW final disposal and processing facilities and MSW sources being the only key consideration in determining WTS need. The following derivation by [41] identified the dependency and is used for assessing economic viability.

By having the following;

i. The hauling costs with primary or at source collection vehicles be a$/ton/km
ii. The hauling costs with secondary collection vehicles be c$/ton/km
iii. The WTS costs be b$/ton
iv. The distance between MSW sources and WTS be x km
v. The distance between WTS and MSW disposal and or processing facility be ykm
vi. Total hauling costs with WTS using primary or at source collection vehicles and secondary collection vehicles be TC_1 and without a WTS using only the primary or at source collection vehicles be TC_2 respectively

Therefore;

$$TC_1 = ax + b + cy \quad (1)$$

$$TC_2 = ax + ay \quad (2)$$

For WTS to be economically viable;

$$TC_1 \leq TC_2 \quad (3)$$

Thus,

$$ax + b + cy \leq ax + ay$$

$$y \geq \frac{b}{a-c} \quad (4)$$

Therefore, the breakeven distance between WTS and MSW disposal or processing facilities is site specific and depends on a number of different input parameters.

2.4. LCA in MSWM

Developments in the management of MSW has brought about the advent of the Integrated Municipal Solid Waste Management (IMSWM) model that entails combining socially acceptable, technically sound, economically feasible and environmentally sustainable MSW storage, collection and transportation, treatment and final disposal systems to manage MSW [42]. Such MSWM systems are complex to design and implement due to the difficulties associated with the need to optimize more than one variable hence the need for tradeoffs. LCA has come in handy as a computer based tool that can be applied during the design, development and implementation of such complex MSWM systems within the framework of the IMSWM model. LCA was described by [43] as a computer aided tool or decision support system that holistically quantifies estimates of potential environmental loads and or impacts of the entire life cycles of products and or processes. In [44], the authors described LCA further as a science-based method of assessing the environmental impacts of a system and or product though not being an entire scientific tool.

Three LCA methodologies have been described by [45] namely Social LCA (S-LCA), Environmental LCA (E-LCA) and Life Cycle Costing (LCC) based on the 3Ps - three-pillar approach towards sustainable development namely People, Planet, and Profit respectively. Despite the reported development of more integrated Life Cycle Sustainability Assessment (LCSA) approaches by [46] and [47], in [45] the authors noted the difficulties of combining the three methodologies due to overlapping issues during results interpretation thus it is advisable to undertake E-LCA, LCC and S-LCA separately. In [45], the authors however, reported the possibility of applying S-LCA individually or combined with E-LCA and/or LCC.

LCA application in informing decision making and policy formulation in MSWM systems design, development and implementation has been used since 1995 [48]. Its application in MSWM strategy development including decision making and policy formulation has expanded significantly during the past few years as a sustainability assessment tool that assists in solving the complexities and interdependencies of modern IMSWM systems [49]. In [50], the authors reported its potential in helping the designing of MSWM systems that reduce local pressures and associated MSWM costs, simultaneously considering the wider trade-offs and impacts of a MSWM system felt in other areas across a specific spatial scale. Spatial scale differences together with those differences with regards to MSW generation rates, composition and characteristics, sources of energy, available MSW treatment and disposal methods with their associated products and services renders LCA appropriate or suitable to aid MSWM strategy development, decision making and policy formulation [51, 52]. It has emerged therefore as an appropriate and suitable holistic method with increasing wider application in MSWM strategy development, decision making and policy formulation [52, 53]. Several MSWM based LCA studies have therefore been carried out in the past with [54] having reviewed one hundred and fifty three studies carried out and published globally from 2013 to 2018, [55] analysed 91 studies carried out from 2006 to 2017 in Asian nations, [56] and [57] reviewed two hundred and twenty two studies they accessed including two hundred and seventeen from peer reviewed journal articles and fifteen public reports.

MSWM based LCA studies assesses different MSW management processes usually referred to as life cycle stages of MSW individually or a combination of processes within a MSWM system with regards to storage, collection, transportation or transfer, treatment and final disposal. These MSW life cycle stages includes the collection and transportation or transfer of MSW including WTS operations. For example, in [52], the authors incorporated WTS in their MSWM based LCA in Greece assuming that all the MSW generated is collected and transported to WTS prior to being sent to the different MSWM, treatment and disposal facilities. Other studies do not incorporate WTS and MSW-CT under the different scenarios arguing that the impacts will be the same under the different scenarios considering same functional unit. However, by considering the reducing effect that WTS have on the impacts from MSW-CT, it is wise to estimate the impacts associated with MSW-CT as it is likely to give an indication of its contribution to the overall MSWM scenario impacts. Higher contributions of MSW-CT will entail the need to redesign the collection and transportation system with the incorporation of WTS being a worthwhile consideration. Likewise MSWM based LCA studies were undertaken to inform possible least impactful MSWM systems that could be considered by decision makers and policy formulators for future implementation in and around Harare City, in Zimbabwe [15, 22, 58]. There is therefore need to assess the contribution of MSW-CT to the impacts of the overall MSWM options that were developed and assessed for their Life Cycle Impacts (LCI) to evaluate any need thereof of WTS considering their reducing effect on MSW-CT derived impacts.

2.5. MSW-CT contribution to LCI of MSWM in Harare

WTS incorporation in a waste management system brings about reductions in environmental or life cycle impacts. In [59], the authors observed a possible 16.8% average reduction in environmental or life cycle impacts when WTS are incorporated in a waste management system. MSWM based LCA studies were carried out to assess impacts of six developed MSWM scenarios that could be put for future implementation considerations in Harare. One of the MSW life cycle stages within the MSWM scenarios that was assessed for its impacts was the MSW-CT considering its contribution to Human Health Potential (HHP), Acidification Potential (AP), GWP or climate change and Eutrophication Potential (EP) of six developed MSWM scenarios for Harare City, Chitungwiza, Ruwa, Epworth and Norton.

The 467,303 tons of MSW generated in Harare City, Chitungwiza, Ruwa, Epworth and Norton [12, 22, 24, 60] are indiscriminately collected and transported for their subsequent landfilling and incineration under MSWM options 1 and 2 respectively with the recovery of energy and landfill leachate, incineration bottom ash and flue gas treatment. Zimbabwe's road infrastructural needs will likely provide a ready market for incineration bottom ash use in road construction. 196,167 tons of source separated biodegradable MSW fraction is subjected to anaerobic digestion (AD) with the recovery of energy from produced biogas whilst the remainder amounting to 271,036 tons mixed bag non-biodegradable MSW fraction undergo incineration under MSWM option 3 with incinerator bottom ash and flue gas treatment. The only distinction under MSWM option 3 and 4 is the landfilling of the remaining 271,036 tons mixed bag non-biodegradable MSW fraction associated with landfill leachate treatment and energy recovery. MSWM option 5 and 6 involve the same processes under MSWM options 3 and 4 respectively serve for the recovery of 20% of the 271,036 tons mixed bag non-biodegradable MSW fraction that are recovered for recycling and reuse purposes being the sole difference.

Several LCA softwares exist in the market. SimaPro software and its embedded Ecoinvent database has been widely used for MSWM based LCA studies having been used by a significant majority of studies reviewed by [53]-[57] and [61]. SimaPro has been and is widely applied because it is generic in nature with universal application capabilities. It also possess a broad database with embedded features from environmental product declarations generation, carbon and water footprints provision as well as sustainability and multi-language interface reporting [55]. SimaPro is standardized which enables it to generate universal and reliable results. It is capable of analysing complex MSWM systems and is based on the Ecoinvent database that has a wide parametric coverage [62]. SimaPro gives MSWM scenarios with description of MSW streams rather than product terms [63] and can be used under spatial scales characterised by general lack of information on MSWM practices [52]. Therefore SimaPro 8.5.2 software and update852 together with the embedded Ecoinvent 3 database (2018) were used [64]. The reference flow of 467,300 tons of MSW generated per annum was used as the functional unit as having been applied in other studies [65-67]. Results are shown in Figure 3 and Tables 1 to 4.

3. MSW Recycling Promotion

The MSW generated in Harare City, Chitungwiza, Ruwa, Epworth and Norton is reportedly suitable for recycling and reuse with [12] and [68] having observed respectively atleast 75% and 90% reusability and recyclability potentials. The potential reusability and recyclability of the MSW has remained an untapped low hanging fruit that could address some of the MSWM challenges considering the 10% recycling reported by [8] and [24] shown in Figure 2a. WTS have largely been reported to provide platforms towards materials recovery and recycling promotion hence their design, construction and operation could significantly increase the MSW recycling and reuse figures to levels that could reduce or result in net zero HHP, AP, EP, and GWP observed by [22] for the MSWM options that could be earmarked for future implementation in Harare City and its surrounding urban population centres of Chitungwiza, Epworth, Ruwa and Norton.

WTS could be operated under a mixed bag MSW collection system or a source separation MSW collection system. Incorporating source separation and WTS operation apart from easing and enhancing MSW recycling, was observed that it brings along with it environmental benefits due to the reduced environmental impacts. In [52], the authors identified that combining WTS with a separation at source MSW collection system and the recycling of plastics, metals, paper and glass MSW fractions as well as biological treatment of biodegradable MSW fraction yielded the overall optimum results with regards to GWP, Human Toxicity, AP, EP, Ozone Layer Depletion, and Photochemical Ozone Creation Potentials under the CML 2001 Method and HHP, Ecosystems and Resources Damage under Eco-indicator 99 method. Interestingly, the Zimbabwe National Integrated Solid Waste Management Plan provides for source separation of MSW thus providing a solid basis and foundation for having WTS under a separation at source MSW-CT systems whose opportunities and challenges are extensively discussed in [14].

Though they are not regarded WTS in the strictest terms, Citizens Drop Off Points (CDOPs) or Centres (CDOCs) sometimes referred to as Citizens Convenient Centres (CCCs) together with Buy Back Centres (BBCs) can also be considered to compliment WTS in aiding recycling and addressing some of the MSWM challenges associated with increased amounts of waste being sent to landfills. CDOPs are facilities were residents dispose of their recyclable MSW materials. CDOPs provides specific recyclable MSW receptacles where citizens or residents come and deposit or dispose of their recyclable MSW. CDOPs demands a great deal of citizens effort, willingness and time to transfer their recyclable MSW hence the need for them to be located within a radius easily accessible to residents. BBCs are facilities where individuals and or businesses sell recyclable MSW materials for BBCs to either process or sell to recycling companies. They provide the much needed linkage between MSW generators and recyclers. Unlike at WTS where those who come to drop waste are supposed to pay tipping fees, at BBCs only appropriate recycles are to be dropped off with those bringing in the recyclable materials paid. However waste pickers can also get paid for delivering recyclable MSW to the Materials Recovery Facilities based WTS. BBCs and MRFs therefore provide a source of livelihoods for both formal and informal waste pickers. BBCs however, must be appropriately located for ease access to waste pickers after ascertaining the availability of reasonable recyclable MSW quantities. The price of recyclable materials determines the sustainability of BBCs whereas tipping fees largely determines the sustainability of WTS. Therefore WTS, CDOPs and BBCs can be considered for incorporation within future MSWM strategies that seek to reduce the human health and environmental impacts currently being experienced in Harare city and its surrounding urban centres. Their consideration for development and operation will aid the diversion of MSW from landfills through materials recycling.

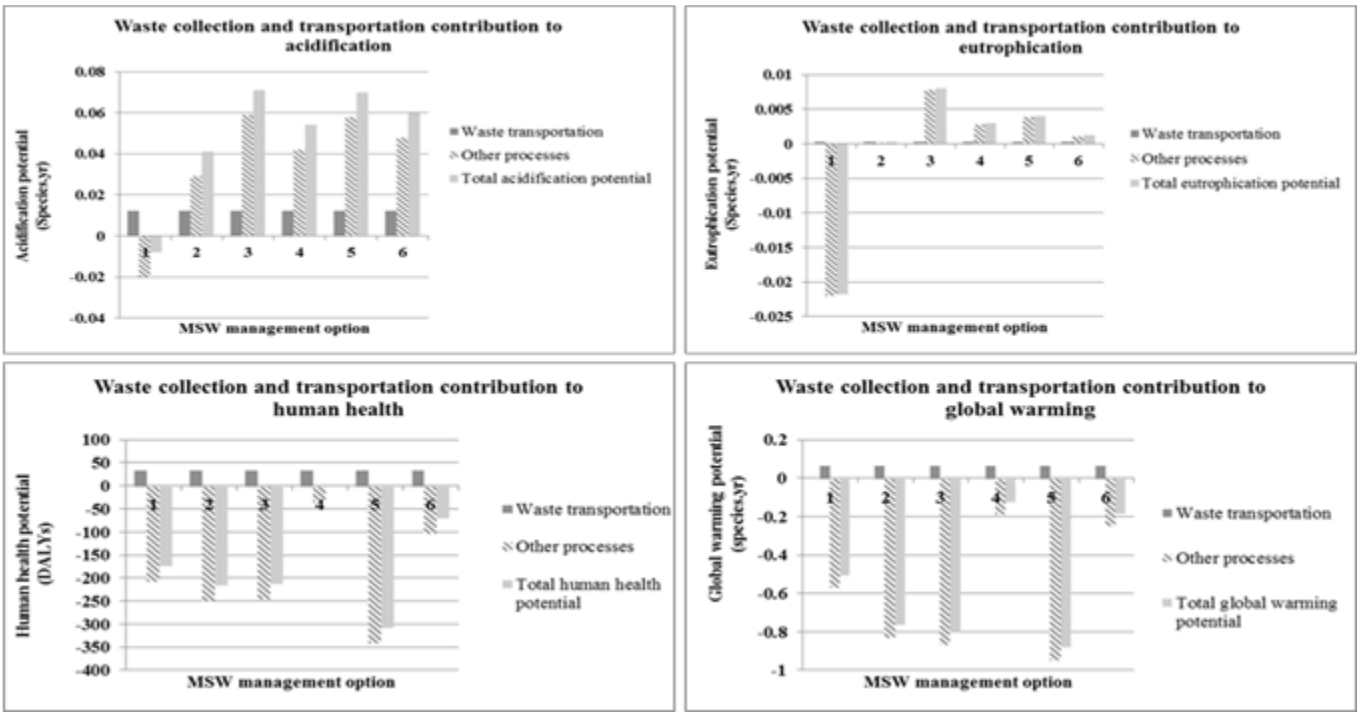


Figure 3: MSW-CT contribution to LCI

Table 1: MSW-CT contribution to AP (species.year)

Process	MSWM Option					
	1	2	3	4	5	6
MSW-CT	0.012	0.012	0.012	0.012	0.012	0.012
Other MSWM processes	-0.02	0.029	0.059	0.042	0.058	0.048
Total AP	**-0.008**	**0.041**	**0.071**	**0.054**	**0.070**	**0.060**

Table 2: MSW-CT contribution to EP (species.year)

Process	MSWM Option					
	1	2	3	4	5	6
MSW-CT	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
Other MSWM processes	-0.0220	0.00001	0.0078	0.0028	0.0038	0.0010
Total EP	**-0.0218**	**0.00021**	**0.0080**	**0.0030**	**0.0040**	**0.0012**

4. Supporting Legislation and institutional framework

The Environmental Management Act [Chapter 20:27] Subsection 1b(ii) under Section 10 of the Act mandates or authorizes the Environmental Management Agency in Zimbabwe to regulate and monitor the recycling of waste among other MSWM including MSW-CT and disposal. The Director General, inspectors and other officers of the Agency are authorised to ensure through monitoring that waste is either recycled and or re-used whenever possible and disposed only in a responsible manner under Subsection (d) of Section 36 of the Act. Every waste generator is obliged to take appropriate measures essential for waste minimisation through recycling amongst other actions under subsection 3 of Section 70 of the Act. Statutory Instrument (SI) 10 of 2007 which authorizes ULAs to manage MSW provides for MSW recycling amongst the MSWM functions and demand every waste generator including ULAS to prepare annual waste management plans that provides for waste recycling wherever practicable in an environmentally safe form and manner amongst the specific goals provided under Subsection 1b(iii) of Section 13. Environmental Impact Assessments (EIAs) are required prior to the establishment of MSW recovery/recycling plants under 11b of the First Schedule of the Act. This could bring challenges in

attracting private players in establishing WTS, CDOPs and BBCs considering the bottlenecks and associated bureaucratic challenges associated with EIA approvals.

The Zimbabwe Waste Sector Low Emission Development Strategy (LEDS) for the Nationally Determined Contributions to its Paris Agreement differential obligations or responsibilities under the UNFCCC does not include MSW recycling despite the need to increase the MSW recycling levels provided for in the ZNISWMP and the Zimbabwe National Climate Change Response Strategy (ZNCCRS). This is a clear policy inconsistence that brings about challenges towards policy implementation. Though the ZNISWMP did not specifically provides for the construction of WTS, CDOPs and BBCs, amongst its strategic objectives is the need to develop infrastructure for waste management including waste recycling centres. Therefore it could be argued that the plan do provides for the development and operation of WTS, CDOPs and BBCs under the recycling centres considering the role they play to promote recycling as facilities for recycling. Goal 4 of the plan seeks to optimize resource recovery through the creation of an enabling environment for recovery; expansion of markets for recyclable MSW and forging partnerships in the recyclable waste value addition chain. It is evidently clear that a relatively fair legislative and institutional framework is available for recycling promotion and the establishment of recycling facilities. However specific provisions with regards to the siting, construction and operation of WTS, CDOPs and BBCs are absent. The other ingredient which is stakeholder involvement is at an undesirably low to support the construction and operation of CDOPs and BBCs based on the prevailing "***We Dumb They Collect***" attitude and general lack of willingness to pay for MSWM services provided by the ULAs amongst the residents.

5. Results and Discussion

The factors that support the need for WTS incorporation in a MSWM system identified from literature justifies the need for WTS to contribute to addressing the significant MSWM challenges currently prevailing in Harare City and the surrounding urban centres of Chitungwiza, Ruwa, Epworth and Norton. The increased population densities and major infrastructural developments coupled with their demand for residential and commercial land bring about increased competition for land leading to the unavailability of land for MSWM facilities especially landfills which require large pieces of land. In [2] and [18], the authors reported that landfill facilities provides favorable environment and conditions for flies, mosquitoes, rodents and mosquitoes together with stray animals to thrive. Therefore, apart from their demand for land which has numerous competing interests, landfill facilities are not desirable in the vicinity of Harare city and the surrounding urban centres which are already experiencing perennial and annual outbreaks of water borne diseases (cholera, dysentery, typhoid etc). This is so because siting landfills near these population centres will further these current human health and environmental impacts, hence landfill facilities to manage MSW generated in Harare City and the surrounding urban centres of Chitungwiza, Ruwa, Epworth and Norton must be sited further away from these population centres due to the already prevailing human health and environmental impacts coupled with the unavailability of affordable land for landfill facilities and other IMSWM operations such as MSW incineration, pyrolysis, gasification and anaerobic digestion.

The location of MSW landfills further from the population centres entails an increased MSW-CT costs and by considering the high MSW-CT costs being reported for Harare and its dormitory urban centres, the incorporation of WTS in future MSWM systems pose to be an ideal strategy with the potential to bring significant MSW-CT costs reduction overally reducing the burden on the MSWM budget. However, quantitative assessments need to be carried out to determine the possible MSW-CT costs reductions that WTS will bring about. This will also entail the determination of the breakeven distance between MSW sources and the WTS that bring about effectiveness in MSW-CT cost. It has been reported that WTS lead to increased effectiveness and efficiency in MSW-CT through the reduction of time spent by waste collection teams moving to and fro MSW disposal facilities availing more time to be spent on MSW collection [25]. This will likely reduce the likelihood of organic MSW fraction decomposition prior to its collection and associated diseases outbreaks. Reductions in fuel consumptions and accompanying GHG emissions as well as waste collection vehicles maintenance costs are anticipated from incorporating WTS in a MSWM system. The reduction in GHG emissions will contribute to the global and national efforts on GHG emission reduction targets hence these reductions will contribute to the Zimbabwe LEDs towards fulfilling its NDCs within the Paris Agreement on differential responsibilities framework. Certain waste fractions such as stumps, fuel tanks despite being empty or filled, furniture, explosives, hazardous waste, electrical appliances, explosives, dead animals, sludge, oils and liquids are not ideal for deposition or disposal at WTS. Fortunately MSW generated in Harare and its surrounding urban centres does not include such waste as defined in the SI of 2007 under the Environmental Management Act [Chapter 20:27]. This further supports and justifies the need to incorporate WTS in future MSWM systems for Harare City, Chitungwiza, Ruwa, Epworth and Norton.

Results from LCA study on the contribution of MSW-CT to the AP, GWP, HHP and EP of MSWM options for possible future implementation considerations are given in Tables 1 to 4 and Figure 3. Results show that MSW-CT bring significant contributions to the AP, GWP, HHP and EP under all the six MSWM options. MSW-CT contributes to increased AP, EP, GWP and HHP of 0.012 species.year, 0.0002 species.year, 0.065 species.year and 34 DALYs respectively under all the six MSWM options. It is interesting to note that GWP brought the highest impact with regards to species extinction rate (species.year), largely emanating from the GHG emissions that cause climate change derived from fuel combustion in MSW collection vehicles during MSW-CT. The positive contributions of MSW-CT to all the four impacts categories entails the need to redesign future MSW-CT systems that could bring reductions to these impacts with the incorporation of distantly located WTS from population centres in such future systems likely to be an economically viable, environmentally sustainable and socially acceptable option. However, incorporating WTS with the source separation of MSW was found to bring optimal environmental performance of a MSWM system hence the need to design future MSW-CT under a source separation system to ensure the derivation of maximum environmental benefits. Consequently, the ZNISWMP provides for the need for separation at source thus providing an enabling framework or background to operate WTS under a separation at source MSW-CT system.

Table 3: MSW-CT contribution to GWP (species.year)

Processes	**MSWM option**					
	1	**2**	**3**	**4**	**5**	**6**
MSW-CT	0.065	0.065	0.065	0.065	0.065	0.065
Other processes	-0.57	-0.83	-0.867	-0.187	-0.947	-0.247
Total GWP	**-0.505**	**-0.765**	**-0.802**	**-0.122**	**-0.882**	**-0.182**

Table 4: MSW-CT Contribution to HHP (DALYs)

Process	**MSWM option**					
	1	**2**	**3**	**4**	**5**	**6**
MSW-CT	34	34	34	34	34	34
Other processes	-208	-250	-246	-34	-342	-104
Total HHP	**-174**	**-216**	**-212**	**0**	**-308**	**-70**

WTS, CDOPs and BBCs enhance recycling and divert a significant amount of waste from final disposal sites mainly landfills. The recycling of recyclable MSW fraction has been found to bring about significant environmental benefits from a review of LCA studies by [53]-[57] and [61]. In [22], the authors also noted significant environmental benefits that materials recovery and recycling bring about when it was considered under the study to determine least impactful MSWM option for possible future implementation in Harare City and surrounding urban centres. Incorporating WTS, CDOPs and BBCs in future MSWM systems in Harare City will likely to come with significant environmental benefits contributing to the addressing of the prevailing MSWM challenges and the low recycling figures of below 10% reported by [8] and [24]. Several legislations and policy documents provides for the need for MSW recycling namely the Environmental Management Act [Chapter 20:27], SI 10 of 2007, ZNCCRS and the ZNISWMP. Since WTS, CDOPs and BBCs facilitates the promotion and increase of recycling, it is therefore evident that they have a strong footing for their development and operation in Zimbabwean urban centres. However caution need to be noted in that there are no specific provisions for their construction within these pieces of legislative and policy documents serve for the ZNISWMP which generally provides for the need to develop infrastructure for waste management including waste recycling centres amongst its numerous objectives. Such lack of clarity on specific legislative, policy and institutional arrangements with regards to WTS, CDOPs and BBCs does not provide the needed confidence and certainty to would be operators of such facilities. The need for EIAs prior to the construction and operation of such facilities like WTS as provided for under the First Schedule of the Environmental Management Act [Chapter 20:27] will also bring its own challenges. Worryingly despite the reported provisions for the need for MSW recycling in a number of legislative and policy documents as well as WTS recycling promotion and GHG emission reductions, the National LEDs towards NDCs fulfillment under the Paris Agreement did not incorporate recycling which could justify WTS operationalization. This policy inconsistence brings challenges for institutional prioritisation of key action plans as decision makers and policy formulators at ULA level will justify none recycling to such inconsistencies.

6. Conclusion and Recommendations

Study findings have largely shown the suitability of incorporating WTS in future MSWM systems for Harare City and its surrounding urban centres as they will contribute to reducing the current human health and environmental impacts currently characterizing these urban environments. WTS will bring about location of landfills and other MSWM facilities further away from population centers, increasing recycling, reducing MSW-CT costs and burden on the overall MSWM budget, increase in MSW-CT effectiveness and efficiency, reduction in MSW-CT derived GHG emissions and other environmental and human health impacts etc. Combining WTS with a source separation MSW-CT system brings about maximum environmental benefits of a MSWM system hence future MSWM systems for Harare City and surrounding urban centres must incorporate WTS under a separation at source MSW-CT system. A sound background that could be used to justify WTS as well as CDOPs ad BBCs based on their potential to increase recycling is available based on the need to promote recycling as provided in the Environmental Management Act [Chapter 20:27], SI 10 of 2007, NCCRS and the ZNISWMP. However, there is need to strengthen the legislation with regards to the development and operationalization of WTS as well as CDOPs and BBCs because specific legislative, policy and institutional framework are absent. Such specific legislative framework complimented with a corresponding regulatory and institutional framework for the design, proper siting and location as well as operation of WTS together with CDOPs and BBCs is needed. This will increase their effectiveness and efficiency simultaneously reducing their associated negative human health and environmental impacts. Unfortunately, the absence of such an enabling legislative,

regulatory and institutional framework could have contributed to the evident silence on WTS incorporation in the proposed MSWM interventions in the ZNISWMP. Source reduction of MSW, materials recovery and recycling must be encouraged to lower the amount (volume and weight) of MSW that is earmarked for collection and transportation to distantly located facilities which will lead to fewer and smaller WTS and preservation of land. WTS have the potential of acting as CDOPs or CCCs resulting in reduced amount of MSW that is illegally dumped. Despite the observed suitability of incorporating WTS in the MSWM systems for possible future implementation in Harare City and its surrounding urban centres, there is however, need for further studies that determine the breakeven distance and LCA studies that specifically assess the associated environmental loads of incorporating WTS within the likely future MSWM systems.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Special thanks go to University of Johannesburg Postdoctoral fellowship that supported this study. The authors would like to thank the Life Cycle Initiative for the 2017 Life Cycle award in the form of SimaPro software that was used to carry out the LCIA.

References

[1] D. Hoornweg and P. Bhada-Tata, What a Waste : A Global Review of Solid Waste Management. Urban development series;knowledge papers no. 15. World Bank, 2012.

[2] V. Yadav, S. Karmakar, A. K. Dikshit, and S. Vanjari, "A feasibility study for the locations of waste transfer stations in urban centers: a case study on the city of Nashik, India," Journal of cleaner production, **126**, 191-205, 2016, doi:10.1016/j.jclepro.2016.03.017.

[3] C. Y. Cheng and J. Urpelainen, "Who should take the garbage out? Public opinion on waste management in Dar es Salaam, Tanzania," Habitat International, **46**, 111-118, 2015, doi:10.1016/j.habitatint.2014.11.001.

[4] UNDESA, World Urbanization Prospects: The 2018 Revision, United Nations Department of Economic and Social Affairs, 2018.

[5] Zimstat, 2012 Zimbabwe Census National Report, Zimbabwe National Statistics Agency, 2013.

[6] M. Sujauddin, S. M. S. Huda, and A. R. Hoque, "Household solid waste characteristics and management in Chittagong, Bangladesh," Waste management, **28(**9), 1688-1695, 2008, doi:10.1016/j.wasman.2007.06.013.

[7] K. Ravindra, K. Kaur, and S. Mor, "System analysis of municipal solid waste management in Chandigarh and minimization practices for cleaner emissions," Journal of cleaner production, **89**, 251-256, 2015, doi:10.1016/j.jclepro.2014.10.036.

[8] IES, A situational Analysis of Solid Waste Management in Zimbabwe's Urban Centres, Institute of Environmental Studies, University of Zimbabwe, 2013.

[9] EMA, Statutory Instrument (SI) 10 of 2007: Environmental Management (Hazardous Waste Management) Regulations, Environmental Management Agency, 2007.

[10] N. I. Sinthumule and S. H. Mkumbuzi, "Participation in community-based solid waste management in Nkulumane suburb, Bulawayo, Zimbabwe," Resources, **8**(1), 30-46, 2019, doi:10.3390/resources8010030.

[11] R. G. Tsiko and S. Togarepi, "A situational analysis of waste management in Harare, Zimbabwe," Journal of American Science, **8**(4), 692-706, 2012.

[12] L. Makarichi, R. Kan, W. Jutidamrongphan, and K. A. Techato, "Suitability of municipal solid waste in African cities for thermochemical waste-to-energy conversion: The case of Harare Metropolitan City, Zimbabwe," Waste Management & Research, **37**(1), 83-94, 2019, doi:10.1177/0734242X18804029.

[13] T. Nhubu, "A review of municipal solid waste data for Harare, Zimbabwe," in WASTES 2019, Lisbon, Portugal, 2019.

[14] T. Nhubu, E. Muzenda, C. Mbohwa, and B. Patel, "Opportunities and limitations for source separation of waste generated in Harare," in WASTES 2019, 2019.

[15] T. Nhubu, E. Muzenda, and C. Mbohwa, "Eutrophication impact potential of solid waste management options in Harare," in WASTES 2019, 2019.

[16] UN-HABITAT, Solid waste management in the world's cities, Nairobi, Kenya,United Nations Human Settlements Programme, 2010.

[17] U. Sonesson, "Modelling of waste collection: a general approach to calculate fuel consumption and time," Waste Management & Research, **18**(2), 115-123, 2000, doi:10.1177/0734242X0001800203.

[18] B. E. Washburn, "Avian use of solid waste transfer stations," Landscape and Urban Planning, **104**(3-4), 388-394, 2012, doi:10.1016/j.landurbplan.2011.11.014.

[19] M. D. Bovea, J. C. Powell, A. Gallardo, and S. F. Capuz-Rizo, "The role played by environmental factors in the integration of a transfer station in a municipal solid waste management system," Waste management, **27**(4), 543-553, 2007, doi:10.1016/j.wasman.2006.03.020.

[20] L. Zhen-shan, Y. Lie, Q. Xiao-Yan, and S. Yu-mei, "Municipal solid waste management in Beijing City," Waste management, **29**(9), 2596-2599, 2009, doi:10.1016/j.wasman.2009.03.018.

[21] EPA, Waste transfer stations: A manual for decision-making , Solid Waste and Emergency Response, Environmental Protection Agency, 2002.

[22] T. Nhubu and E. Muzenda, "Determination of the Least Impactful Municipal Solid Waste Management Option in Harare, Zimbabwe," Processes, **7**(11), 785, 2019, doi:10.3390/pr7110785.

[23] T. V. Ramachandra and S. Bachamanda, "Environmental audit of municipal solid waste management," International Journal of Environmental Technology and Management, **7**(3-4), 369-391, 2007, doi:10.1504/IJETM.2007.015152.

[24] EMA, Waste generation and management in Harare, Zimbabwe: Residential areas, commercial areas and schools, Environmental Management Agency, 2016.

[25] L. Cui, L. R. Chen, Y. P. Li, G. H. Huang, W. Li, and Y. L. Xie, "An interval-based regret-analysis method for identifying long-term municipal solid waste management policy under uncertainty," Journal of environmental management, **92**(6), 1484-1494, 2011, doi:10.1016/j.jenvman.2010.12.006.

[26] M. Sharholy, K. Ahmad, G. Mahmood, and R. Trivedi, "Municipal solid waste management in Indian cities: a review," Waste management, **28**(2), 459-467, 2008, doi:10.1016/j.wasman.2007.02.008.

[27] S. Consonni, M. Giugliano, and M. Grosso, "Alternative strategies for energy recovery from municipal solid waste: Part B: Emission and cost estimates," Waste management, **25**(2), 137-148, 2005, doi:10.1016/j.wasman.2004.09.006

[28] J. W. Levis, M. A. Barlaz, J. F. DeCarolis, and S. R. Ranjithan, "A generalized multistage optimization modeling framework for life cycle assessment-based integrated solid waste management," Environmental modelling & software, **50**, 51-65, 2013, doi:10.1016/j.envsoft.2013.08.007.

[29] C. A. Teixeira, C. Avelino, F. Ferreira, and I. Bentes, "Statistical analysis in MSW collection performance assessment," Waste management, **34**(9), 1584-1594, 2014, doi:10.1016/j.wasman.2014.04.007.

[30] L. I. Cioca, L. Ivascu, E. C. Rada, V. Torretta, and G. Ionescu, "Sustainable development and technological impact on CO2 reducing conditions in Romania," Sustainability, **7**(2), 1637-1650, 2015, doi:10.3390/su7021637.

[31] C. Mora, R. Manzini, M. Gamberi, and A. Cascini, "Environmental and economic assessment for the optimal configuration of a sustainable solid waste collection system: a 'kerbside'case study," Production planning & control, **25**(9), 737-761, 2014, doi:10.1080/09537287.2012.750386.

[32] V. Yadav, P. P. Kalbar, S. Karmakar, and A. K. Dikshit, "A two-stage multi-attribute decision-making model for selecting appropriate locations of waste transfer stations in urban centres," Waste Management, **114**(2), 80-88, 2020, doi:10.1016/j.wasman.2020.05.024.

[33] P. Rathore and S. P. Sarmah, "Modeling transfer station locations considering source separation of solid waste in urban centers: A case study of Bilaspur city, India," Journal of Cleaner Production, **211**, 44-60, 2019, doi:10.1016/j.jclepro.2018.11.100.

[34] V. Yadav, S. Karmakar, A. K. Dikshit, and S. Vanjari, "A feasibility study for thelocations of waste transfer stations in urban centers : a case study on the city of Nashik, India," Journal of Cleaner Production, **126**, 191- 205, 2016, doi:10.1016/j.jclepro.2016.03.017.

[35] H. A. Eiselt and V. Marianov, "A bi-objective model for the location of landfills for municipal solid waste," European Journal of Operational Research, **235**(1), 187-194, 2014, doi:10.1016/j.ejor.2013.10.005.

[36] C. Chatzouridis and D. Komilis, "A methodology to optimally site and design municipal solid waste transfer stations using binary programming,"

Resources Conservation and Recycling, **60**, 89-98, 2012, doi:10.1016/j.resconrec.2011.12.004.
[37] F. Habibi, E. Asadi, S. J. Sadjadi, and F. Barzinpour, "A multi-objective robustoptimization model for site-selection and capacity allocation of municipal solid waste facilities : a case study in Tehran," Journal of Cleaner Production, **166**, 816-834, 2017, doi:10.1016/j.jclepro.2017.08.063.
[38] H. Asefi and S. Lim, "A novel multi-dimensional modeling approach to integrated municipal solid waste management," Journal of Cleaner Production, **166,** 1131-1141, 2017, doi:10.1016/j.jclepro.2017.08.061.
[39] D. Komilis, "Conceptual modeling to optimize the haul and transfer of municipal solid waste," Waste Management, **28**(11), 2355-2365, 2008, doi:10.1016/j.wasman.2007.11.004.
[40] M. Badran and S. El-Haggar, "Optimization of municipal solid waste management in Port Said - Egypt," Waste Management, **26**(5), 534-545, 2006, doi:10.1016/j.wasman.2005.05.005.
[41] H. Phelps, G. Heinke, J. Jonker, E. Ouano, and C. Vandecasteele, Management of Solid Wastes. Technical Report. UNESCO, 1996.
[42] F. McDougall, P. R. White, M. Franke, and P. Hindle, Integrated solid waste management: a life cycle inventory. John Wiley & Sons, 2008.
[43] G. Rebitzer, T. Ekvall, R. Frischknecht, D. Hunkelerd, G. Norrise, T. Rydberg, W. P. Schmidt, S. Suh, B. P. Weidemai, D.W.Pennington, "Life cycle assessment: Part 1: Framework, goal and scope definition, inventory analysis, and applications," Environment International, vol. **30**(5), 701-720, 2004, doi:10.1016/j.envint.2003.11.005.
[44] J. R. Winkler and B. Bilitewski, "Comparative evaluation of life cycle assessment models for solid waste management," Waste Management, **27**(8), 1021-1031, 2007, doi:10.1016/j.wasman.2007.02.023.
[45] UNEP, Guidelines for Social Life Cycle Assessment of Products and Organizations. United Nations Environment Programme , 2020.
[46] J. B. Guinée, R. Heijungs, G. Huppes, A. Zamagni, P. Masoni, R. Buonamici, T. Ekvall, T. Rydberg, "Life Cycle Assessment: Past, Present, and Future," Environmental Science and Technology, **45**(1), 90-96, 2011, doi:10.1021/es101316v.
[47] T. Schaubroeck and B. Rugani, "A revision of what life cycle sustainability assessment should entail: Towards modeling the net impact on human well-being," Journal of Industrial Ecology, **21**(6), 1464-1477, 2017, doi:10.1111/jiec.12653.
[48] L. P. Güereca, S. Gassó, J. M. Baldasano, and P. Jiménez-Guerrero, "Life cycle assessment of two biowaste management systems for Barcelona, Spain," Resources, Conservation and Recycling, **49**(1), 32-48, 2006, doi:10.1016/j.resconrec.2006.03.009.
[49] G. A. Blengini, M. Fantoni, M. Busto, G. Genon, and C. Zanetti, "Participatory approach, acceptability and transparency of waste Management LCAs: Case studies of Torino and Cuneo" Waste Management, **32**(9), 112-1721, 2012, doi:10.1016/j.wasman.2012.04.010.
[50] K. Koneczny and D. W. Pennington, "Life cycle thinking in waste management: Summary of European Commission's Malta 2005 workshop and pilot studies," Waste Management, **27**(8), S92-S9, 2007, doi:10.1016/j.wasman.2007.02.020.
[51] M. R. Mendes, T. Aramaki, and K. Hanaki, "Comparison of the environmental impact of incineration and landfilling in Sao Paulo City as determined by LCA," Resources, Conservation and Recycling, **41**(1), 47-63, 2004, doi:10.1016/j.resconrec.2003.08.003.
[52] G. Banias, M. Batsioula, C. Achillas, S. L. Patsios, K. N. Kontogiannopoulos, D. Bochtis, N. Moussiopoulos, "A Life Cycle Analysis Approach for the Evaluation of Municipal Solid Waste Management Practices: The Case Study of the Region of Central Macedonia, Greece," Sustainability, **12**(19), 8221, 2020, doi:10.3390/su12198221.
[53] K. Abeliotis, "Life cycle assessment in municipal solid waste management, INTECH Open Access Publisher, 2011.
[54] H. Khandelwal, H. Dhar, A. K. Thalla, and S. Kumar, "Application of life cycle assessment in municipal solid waste management: A worldwide critical review," Journal of cleaner production, **209**, 630-654, 2019, doi:10.1016/j.jclepro.2018.10.233.
[55] P. Yadav and S. R. Samadder, "A critical review of the life cycle assessment studies on solid waste management in Asian countries," Journal of cleaner production, **185**, 492-515, 2018, doi:10.1016/j.jclepro.2018.02.298.
[56] A. Laurent, J. Clavreul, A. Bernstad, I. Bakas, M. Niero, E. Gentil, T. H. Christensen, M. Z. Hauschild, "Review of LCA studies of solid waste management systems–Part II: Methodological guidance for a better practice," Waste Management, **34**(3), 589-606, 2014, doi:10.1016/j.wasman.2013.12.004.
[57] A. Laurent, I. Bakas, J. Clavreul, A. Bernstad, M. Niero, E. Gentil, M. Z. Hauschild, T. H. Christensen, "Review of LCA studies of solid waste management systems–Part I: Lessons learned and perspectives," Waste Management, **34**(3), 573-588, 2014, doi:10.1016/j.wasman.2013.10.045.
[58] T. Nhubu, E. Muzenda, C. Mbohwa, and E. O. Agbenyeku, "Comparative assessment of compositing and anaerobic digestion of municipal biodegradable waste in Harare, Zimbabwe," Environmental Progress & Sustainable Energy, **39**(4), 2020, doi:10.1002/ep.13376.
[59] M. D. Bovea, J. C. Powell, A. Gallardo, and S. F. Capuz-Rizo, "The role played by environmental factors in the integration of a transfer station in a municipal solid waste management system," Waste Management, **27**(4), 545-553, 2007, doi:10.1016/j.wasman.2006.03.020.
[60] EMA, Zimbabwe's integrated solid waste management plan, Environmenntal Management Agency, 2014
[61] J. Cleary, "Life cycle assessments of municipal solid waste management systems: A comparative analysis of selected peer-reviewed literature," Environment international, **35**(8), 1256-1266, 2009, doi:10.1016/j.envint.2009.07.009.
[62] E. Moreno Ruiz, T. Lévová, J. Reinhard, L. Valsasina, G. Bourgault, and G. Wernet, Documentation of changes implemented in ecoinvent database v3.3, Ecoinvent, 2016.
[63] A. Vidalis, D. Malamis, K. Moustakas, K. Valta, D. Bolzonella, P. Grammelis, M. Loizidou., Development and implementation of a demonstration system on Integrated Solid Waste Management for Tinos in line with the Waste Framework Directive., Municipality of Tinos, 2011.
[64] E. Moreno Ruiz, L. Valsasina, F. Brunner, A. Symeonidis, D. FitzGerald, K. Treyer, G. Bourgault, G. Wernet, "Documentation of changes implemented in ecoinvent database v3.5," Ecoinvent, 2018.
[65] Y. Fernández-Nava, R. J.D, J. Rodríguez-Iglesias, L. Castrillón, and E. Marañón, "Life Cycle Assessment (LCA) of different municipal solid waste management options: A case study of Asturias (Spain)," Journal of Cleaner Production, **81**, 178-189, 2014, doi:10.1016/j.jclepro.2014.06.008.
[66] P. Beigl and S. Salhofer, "Comparison of ecological effects and costs of communal waste management systems," Resources, Conservation and Recycling, **41**(2), 83-102, 2004, doi:10.1016/j.resconrec.2003.08.007.
[67] F. Cherubini, S. Bargigli, and S. Ulgiati, "Life cycle assessment (LCA) of waste management strategies: Landfilling, sorting plant and incineration," Energy, **34**(12), 2116-2123, 2009, doi:10.1016/j.energy.2008.08.023.
[68] D. Tirivanhu and S. Feresu, A Situational Analysis of Solid Waste Management in Zimbabwe's Urban Centres,: Institute of Environment Studies-University of Zimbabwe, 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 1013-1028 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Review of Plastic Waste Management Practices: What Can South Africa Learn?

Zvanaka S. Mazhandu[1,*], Edison Muzenda[2,1], Mohamed Belaid[1], Tirivaviri A. Mamvura[2], Trust Nhubu[1]

[1]*University of Johannesburg, Department of Chemical Engineering Technology, Johannesburg, 2001, South Africa*

[2]*Botswana International University of Science and Technology, Department of Chemical, Materials and Metallurgical Engineering, Palapye, 00000, Botswana*

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 20 February, 2021
Online: 22 April, 2021

Keywords:
Plastic waste management practices
Separation of waste at source
Waste collection

ABSTRACT

Municipal Solid Waste (MSW) is composed of items that are discarded or disposed of daily including paper, plastics, glass, metals, used gadgets, paint and old furniture. The plastic waste stream has proven to be problematic to manage sustainably on a global scale. Various researchers are trying to come up with innovative ways of alleviating the detrimental effects of plastic on the environment. Examples include the production of liquid fuel and synthetic gas through pyrolysis and gasification of plastic waste, use of microbial strains that can break down polyethylene, manufacture of plastic-infused tar, use of plastic waste in cement and concrete as well as its use in the manufacture of bricks. Conducting public awareness and outreach programmes has also been found to be beneficial in reducing plastic littering. This paper reviews South Africa's strengths, weaknesses, and opportunities in plastic waste management as well as lessons from other jurisdictions that can be adopted in South Africa making it a role model for Africa with regards to plastic waste management. There exists an untapped opportunity for improvement of post-consumer plastic recycling rates to levels comparable to other recyclables in the country through compulsory separation of waste at source. Hence an enabling environment should be created to encourage this practice. Since this will require a fully functional waste management infrastructure, collection services should expand to cover rural areas and informal settlements while industries can assist municipalities to upgrade infrastructure through the extended producer responsibility (EPR) scheme. In addition, there is potential for more jobs to be created in the waste sector through recycling as compared to landfilling, thus urgent attention is needed to divert 100% waste from the landfill. Finally, the integration of informal waste pickers into the waste management chain should be prioritised.

1. Introduction

This paper is an extension of work originally reported in *Proceedings of the 7th International Renewable and Sustainable Energy Conference* and assesses plastic waste management implemented in Belgium, Australia, South Korea and Canada against that implemented in South Africa. Belgium and South Korea could potentially become the gold standards for South Africa to follow while Australia and Canada have a wealth of information and a clearly laid out vision of the direction the countries intend to take in fighting plastic pollution. Mini descriptive reviews of studies highlighting current plastic waste management practices as well as EPR studies from various countries have also been conducted. The aim of the paper is to determine the lessons that South Africa can learn as we navigate the path to zero waste in line with circular economy principles. MSW, loosely termed garbage is composed of items that are discarded or disposed of daily including paper, plastics, glass, metals, used gadgets, paint and old furniture [1-3]. Waste is defined as any item that has reached the end of its usefulness and needs to be discarded [4]. MSW is generated from households, institutions (e.g. schools), businesses and non-hazardous waste from industries [2, 3] evidencing a "resource-intensive" lifestyle by consumers [5]. The management of MSW is critical because unmanaged waste is an eyesore, can be a breeding ground for disease-causing organisms, can block water drains and sewer networks leading to flooding episodes as well as cause damage to the marine environment and animals [3, 4, 6, 7]. The characteristics or composition and quantity of MSW vary between communities

*Corresponding Author: Zvanaka S. Mazhandu, zvanaka@gmail.com

and countries [3,8]. Consequently, waste management practices employed within the different communities and countries will also differ. Figure 1 shows the various fundamental stages involved in waste management.

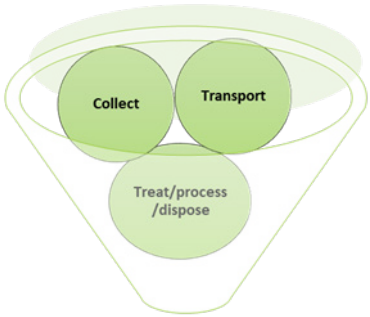


Figure 1: Stages in Waste Management [3, 9]

The plastic waste stream has proven to be problematic to manage sustainably on a global scale. In Sub Saharan Africa, plastic waste constitutes 13% of the total MSW [5, 7].

In an earlier study, [7] highlighted leakages of small plastic items unnoticeably into the environment post-consumer use in South Africa. This is rampant in areas where there are no systems in place to collect them [7]. Albeit small, their continuous accumulation in the environment poses danger to animals as well as humans [7] in the short to long term.

To put this into perspective using bread tags as an example, the annual bread consumption in South Africa with a population of 59.7 million [10] is reportedly 25.8 kgs per capita [11] implying that over a billion tags weighing 300 tonnes are used yearly. Dedicated collection points for bread tags will go a long way in preventing such leakages. The Bread Tags for Wheelchairs an initiative started in 2006 by Mary Honeybun in South Africa, collects bread tags and sells to recyclers in order to buy wheelchairs from the proceeds. Approximately 600,000 tags weighing 200 kgs are used to purchase one wheelchair [12]. Therefore, 300 tonnes worth of tags would result in approximately 1 500 wheelchairs annually, Figure 2.



Figure 2: Number of Wheelchairs Expected Annually from Sale of Bread Tags

Various researchers are trying to come up with innovative ways of alleviating the detrimental effects of plastic waste on the environment. Examples include the production of liquid fuel and synthetic gas through pyrolysis and gasification, respectively of plastic waste [13-16], use of microbial strains that can break down polyethylene [17], manufacture of plastic-infused tar [18], use of plastic waste in cement [19] and concrete [20] as well as its use in the manufacture of bricks [21].

Conducting public awareness and outreach programmes has also been found to be beneficial in reducing plastic waste littering [22]. In 2013, the authors in [23] advocated for plastic waste to be considered hazardous unlike the current status quo where plastic waste falls under solid waste. The authors argue that if this is undertaken, countries would be compelled to put more effort in finding ways to mitigate plastic waste leakage as well as finding replacements for single use plastics (SUPs).

2. Data Sources

Data for this study was gathered from peer reviewed journals and grey literature. Authors' observations are also included. Some of the academic databases accessed include Google Scholar, Web of Science, Scopus, Science Direct and Springer Link with studies undertaken between 2010 and 2021 considered. Phrases and key words used to acquire relevant literature include plastic waste management practices, separation at source, waste collection, and extended producer responsibility which were either used singly or in combination. Quotation marks "" were also used to restrict the literature pool.

3. Review of Waste Management Practices

In 2015, the authors in [24] conducted a study to assess innovations in plastic waste management in Kenya. The author found out that; there were no incentives for innovators, recycling guidelines were lacking and working conditions were poor. With regards to plastic waste management practices, landfilling, illegal dumping and littering were prevalent in the country. The authors recommended the drafting of a plastic recycling framework by all stakeholders including those in the informal sector. This framework would include compulsory recycling targets, guidelines on quality of plastic products and training of informal sector workers.

In 2017, the authors in [25] assessed the management of plastic waste in Bangladesh with the aim of recommending the best way forward. The authors noted that hindrances to effective plastic waste management in the country included poor infrastructure, lack of recycling technologies and inadequate funds to advance waste management services. In addition, recycling and reuse were found to be minimal; with only 20% plastic waste collected while landfilling, open and indiscriminate dumping were the predominant disposal methods. They recommended the use of alternative plastic waste management technologies such as pyrolysis, bitumen production and use of plastic waste as solid refuse fuel in cement kilns in addition to recycling.

In the same year of 2017, the authors in [26] assessed current solid waste management practices and policies in Malaysia. Their findings were that there was limited separation of waste at source in the country hence poor recycling. In addition, the authors noted the lack of commitment by the public to participate in the initiative as well as the unavailability of accurate documented data. Open dumping and landfills were observed to be prevalent with 95% of waste being disposed of through these methods and the balance being recycled/treated or illegally dumped. In [26], the authors recommended; regularisation of informal waste pickers, updating of waste management policies and mandatory separation of waste at source to increase recycling rates and reduce illegal dumping incidences.

In 2018, the authors in [27] conducted a sustainability impact assessment of three scenarios for plastic waste management in Sweden. According to the authors, the country's dependence on incineration hampers recycling efforts. Plastic waste is also not prioritised in the different waste management policies of the country. The authors concluded that increasing recycling rates and phasing out incineration will be the most sustainable pathway to manage plastic waste in the country.

A study in Austria on plastic packaging waste management conducted in 2018, found that in the year 2013:

- 300,000 ± 3% tonnes of packaging waste was produced at 35 kg per capita,
- recycling rate was 26% ± 7%,
- use of plastic packaging waste in the cement industry was 32% ± 6% and,
- waste to energy plants (WtE) used 40% ± 3% packaging waste [28].

The authors recommended that recycling rates should be calculated using the output rate and not the recycling input rate to improve accuracy of results.

In 2019, the authors in [29] reviewed plastic waste management strategies in Nigeria and concluded that lack of funds and mismanagement of this limited resource and poor infrastructure hampered waste management in the country. Furthermore, over 50% of generated plastic wastes are either indiscriminately dumped or in drains and waterways. The authors recommended; educating the public on separation of waste at source; institution of fines for unsorted waste, setting up of collection centres, WtE plants, and establishing frameworks that enhance sufficient record keeping.

In the same year of 2019, a study in Singapore was conducted to determine the best environmentally performing plastic waste management scenario [30]. The author highlighted the dominance of WtE plants in the country due to shortage of space for landfilling. WtE plants result in a reduction in volume of waste by 90%. Approximately 634 kWh of energy is released from a tonne of mixed plastic waste. Residues (fly and bottom ash) from WtE plants are landfilled offshore. In addition, recycling is minimal in the country. In [30], the author recommended the use of pyrolysis in addition to WtE plants and mechanical recycling.

In [31], the author assessed plastic waste management practices in the United Kingdom (UK). According to the study's findings, of the 3.3 million tonnes of plastic waste produced in 2013, 2.26 million tonnes was packaging. The author also highlighted the need for attention to be directed to other plastics such as nurdles, synthetic fibres and microbeads and not only on plastic waste packaging. The commonly practiced waste management methods in UK are WtE plants followed by recycling and landfilling while some waste is also exported. Some of the recommendations given based on findings included enforcement of bans on fishing residues, enforcement of Operation Clean Sweep to prevent leakages of nurdles, extending ban of microbeads to all products and installation of sand filters in wastewater treatment facilities for the removal of plastic fibres.

In [32], the authors also conducted a 2020 study to assess plastic waste management practices in the Kingdom of Eswatini rural households. The amount of plastic waste generated per household was found to be 15.9 g/day. Common practices employed to manage waste were open burning, burying, indiscriminate dumping, use of pits at the backyard, reuse, upcycling, and recycling. The authors suggested the roll out of waste collection services in rural areas as well as educating consumers on plastic pollution effects and various ways of managing plastic waste sustainably.

In 2020, the authors in [33] assessed plastic packaging waste management in South Korea. Three million tonnes of plastic packaging waste were generated. The packaging waste was managed as follows; use as solid refuse fuel (39.3%), incineration without energy recovery (33.4%), recycling (13%) and the balance landfilled. The authors noted that only 22.3% of plastic packaging waste was under the EPR scheme and therefore they advocated for the list of plastic items covered by EPR to be expanded to reduce plastic pollution.

In 2021, the authors in [34] investigated SUP waste management in Hanoi, Vietnam and described the secondary use of plastic shopping bags as bin liners after single use. In addition, although plastic bags should be taxed, implementation is lacking. Landfilling was found to be prevalent while recycling is limited. According to the authors, there are gaps in drafted waste management policies which need to be addressed.

Waste management practices from five countries namely, Belgium (Europe), Australia, South Korea (Asia), Canada (North America) and South Africa (Africa) were also studied in this paper and these are highlighted in the ensuing sections.

3.1. Europe

Many countries in the European Union (EU) are trying to avoid landfilling, simultaneously improving their recycling rates of plastic waste. The EU must be admired for its unified approach as a region in trying to address waste challenges although countries like Malta, Greece, Romania and Cyprus still have a long way to go in reducing their rates of landfilled plastic waste from the current 70-80%. In Europe, a recycling fund is included on purchases which is subsequently reimbursed on returning bottles. This compels the public to recycle. Denmark, Germany, Austria, Sweden and Belgium have the lowest disposal rates; with less than 3% of MSW generated heading to landfills, Figure 3 [35-37].

Compared to the 50% target of recycled household waste by 2020, as outlined in the Waste Framework Directive of 2008, these 5 countries have indeed set the bar extremely high for other countries in the region. From the beginning of 2006 to 2016, the amount of plastic waste recycled in the region increased by 79%, an increase of 61% was observed for energy production while plastic waste landfilling decreased by 43%. During the same period, the recycling of plastic packaging increased by nearly 75% [36].

Extended Producer Responsibility, where manufacturers of products are responsible for their products throughout their entire life cycles [7, 38] is another scheme that is being advocated for within the European Union [35].

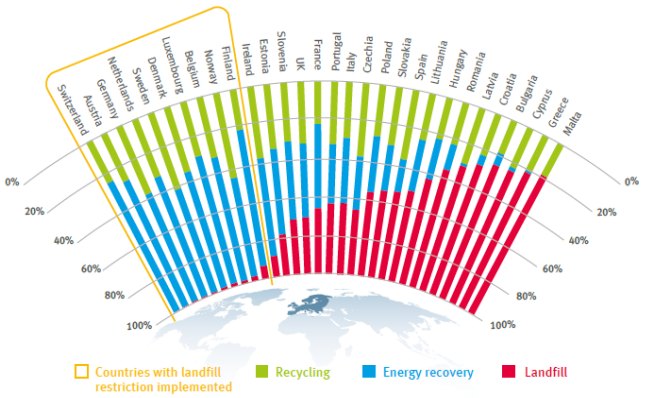


Figure 3: Recycling, Landfilling and Energy Recovery rates in Europe, 2016 [36, 37]

3.1.1 Belgium's Case

Belgium has a population of 11.6 million [39] and is comprised of three regions; namely Brussels Capital Region, Flanders and Wallonia. The enactment of Article 11(2), of the Waste Directive of 2008 has been instrumental in the way Belgium manages its waste. In the directive the various waste types such as glass, paper, plastic and garden waste are separated at source. The Pay as You Throw (PAYT) schemes have also been fundamental in compelling the public to adhere to the sorting of waste regulations. Residents are given about 4 bags to sort their waste in their homes. The bag of waste meant for disposal, costs more than the bags for recyclable waste. In 2013, Belgium was the leader in waste management across Europe [40] and profound lessons can be learnt by other governments by emulating Belgium. The amount of waste either disposed off or incinerated was a low 197 kgs per capita compared to its other counterparts in Europe as a result of reduced waste generation and increased rates of recycling [40].

According to the authors in [35], in 2015, US$2.45 was charged per bag of waste to be disposed in Northern Belgium. This was approximately 5 times more than the rest of the bags and discourages waste dumping. In South Africa this would have been equivalent to R35.80 per bag versus R7.16 for recyclables. However, credit should also be given to the Belgians for playing their part in sorting waste and not resisting change in attitudes [35].

The company, Fost Plus, responsible for financing and handling the collection, sorting and recycling of waste has also been impeccable in its operations [41]. This has had an enormous positive impact on Belgium's waste statistics. Residents are also given a waste collection calendar yearly or they can download the Recycle mobile application since collections for different coloured bags maybe carried out on different days. Residents are also fined if they do not sort or leave an improperly sorted bag in the street. In a report written in 2013 by authors in [42], the region, Brussels-Capital was apparently penalising residents as much as €625. Training is also offered to the public including children in schools, emphasizing the importance of waste sorting and its benefits [41]. Belgium is also a signatory to the conventions outlined in Figure 4.



Figure 4: Conventions and Bodies Supported by Belgium

3.2. Australia's case

Australia has a population of 25.6 million [43] and is made up of six states namely New South Wales, Queensland, South Australia, Tasmania, Victoria, and Western Australia as well as three internal territories and seven external territories each with its own government except the Jervis Bay Territory [44]. Australia set itself some ambitious targets in its 2018 National Waste Policy, among which was the ban after June 2020 on plastic, tyres, glass and paper exports. The country has realised that these materials often regarded as waste are a resource that can be used to; generate valuable products, create employment, boost the economy, protect the environment and health of its people. Approximately 9.2 jobs are created from 10 000 tonnes of recycled material in comparison to 2.8 jobs from landfilling [45]. The policy also reflects the banning of SUPs and the need for recording (and sharing) of accurate waste data including imports almost on a real time basis to assist with decision making coupled with unsophisticated online systems that are easy for everyone to understand. Other highlighted targets included a recovery rate of 80% for recyclables in MSW, purchasing of recycled materials by government and industries in order to increase the demand for these, and drive innovation by funding upcoming plastic recycling and waste prevention solutions. The underlying principle in Australia's Waste policy is the circular economy approach; a concept discussed in greater detail in an earlier study in [37].

The 2019 National Action Plan was then drafted to aid in the implementation of the National Waste Policy. Some of the documented strategies and targets given are shown in Table 1 [46]:

Table 1: 2019 National Action Plan Strategies

Strategy	Target
Drafting of legislation to avoid landfilling of recyclable material	2022
Sourcing new markets for recycled products	ongoing
Funding industries that make recycled products	2020
Increasing kerbside recycling rates through education and use of the Australasian Recycling Label (ARL)	ongoing
Launching a recycled products online market where buyers and sellers can connect	2021
Development of an application that helps the public to minimise contamination of recyclables in MSW	2020

Determining factors contributing to contamination of recyclables in kerbside collections and implement remedial actions	2020
Instituting an investment fund for Product Stewardship of oil containers made of plastic	(2020)
Development or adoption of standards that extend the life of a product including standards for use of recycled products in roads (Reconophalt) [42] and rail projects	2020
Develop national standards for collections	2022

Figure 5 shows the targets in relation to packaging and plastic as detailed in the 2019 action plan in [46].

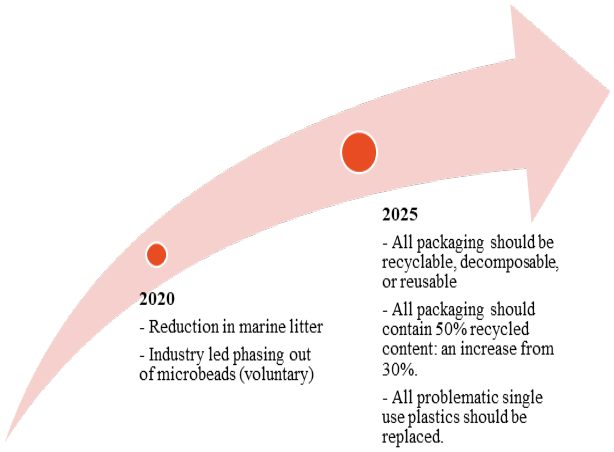


Figure 5: Targets in relation to packaging and plastic

In [47], the authors also mentioned programmes such as Do the Right Thing, Neat Streets, Don't be a Tosser and Bin your Butts which all help to make the public aware of the need to prevent littering. Another programme; Keep Australia Beautiful (KAB) outlined in [48] is also linked to initiatives such as Tidy Towns, Clean Beaches, Sustainable Cities, National Litter Index (NLI), Beverage Container Recycling Grants, Eco-Schools, Adopt a Patch, and KAB Week which occurs annually in August [48]. The NLI measures the degree of littering across the country annually.

Australia also formulated a framework known as the Threat Abatement Plan for the impact of marine debris on the vertebrate wildlife of Australia's coasts and oceans (2018), to protect marine animals from injury or death caused by marine debris. Australia also has two national plans namely, the Marine Turtle Recovery Plan (implemented in 2003 and reviewed in 2013) and the Grey Nurse Recovery Plan (initially adopted in 2002 and a new plan initiated in 2014) which are meant to boost the numbers of these species [49].

Australia is also actively involved in the United Nations General Assembly and the United Nations Environment Programme (UNEP) which seeks to protect the marine environment from land-based pollution. Australia, therefore, is in full support of the UN's Sustainable Development Goal 12 which encourages sustainable consumption and production. The country participates in regional initiatives such as the Coral Triangle Initiative, the Coordinating Body on the Seas of East Asia (COBSEA) and the Marine Resources Conservation Working Group of Asia Pacific Economic Cooperation (APEC) [49] and is launching in 2021, the Australia, New Zealand and Pacific Island nations (ANZPAC) Plastics Pact; a collaboration between Australia, New Zealand and other Pacific island countries which will be part of Ellen MacArthur's Plastics Pact. The pact is led by APCO and The Waste and Resources Action Programme (WRAP), based in the United Kingdom (UK) is offering support. WRAP also manages the UK Plastics Pact [50].

In addition to this, the country is a member of other various international conventions and agreements on waste control, as shown in Figure 6. This ensures that the country is accountable, a trait that can only be of benefit to the country. Australia has managed to meet its international commitments by instituting regulations.

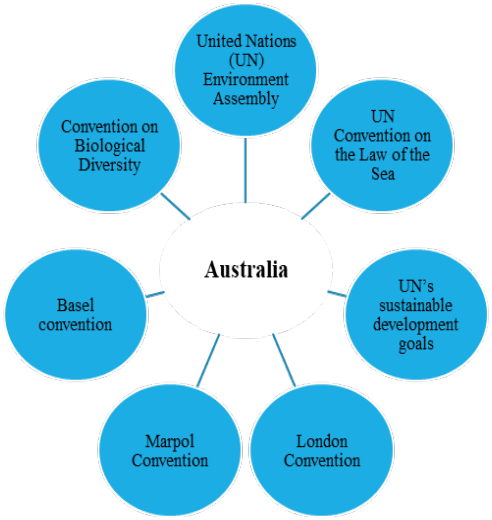


Figure 6: Conventions and Bodies Supported by Australia

3.3. South Korea's case

South Korea has a population of about 51.3 million [51] and the country is divided into 9 provinces namely North Chungcheong, South Chungcheong, Gangwon, Gyeonggi, North Gyeongsang, South Gyeongsang, North Jeolla, South Jeolla, and Jeju Special Self-Governing Province [52]. The Ministry of Environment in South Korea oversees the waste disposal policy referred to as "jongnyangje". The policy emphasises mandatory separation of waste into various fractions namely, recyclables, organics, large waste and landfill wastes. Fines are instituted for failure to abide by the policy guidelines [53,54]. Rewards are also given as an incentive to people who report non-compliance [54]. This explains why South Korea placed second out of 37 countries in the Organisation for Economic Co-operation and Development (OECD) for achieving a recycling rate of 59% in 2013 [53,55] and placed fifth globally in 2018 with a recycling rate of 53.7% [56]. On produced plastic products, over 60% of these are recycled [57].

The collection of waste is done at municipal level where revenue for this service is generated from the sale of differently coloured garbage bags. The purchase of the bin bags is compulsory and colour codes also differ per district. While organic wastes should be dried before disposal, recyclables should be flattened, and large wastes such as televisions should bear large object disposal stickers that are purchased from district offices [58]. The government of South Korea has been lauded for financially supporting the 48 plastic recycling businesses in the country which were struggling as a result of the ban imposed by China on importation of recyclable wastes which had resulted in piles of plastic waste in the country [58].

Figure 7 shows the progression of waste management acts and schemes implemented in the country and these are described in more detail in Table 2.

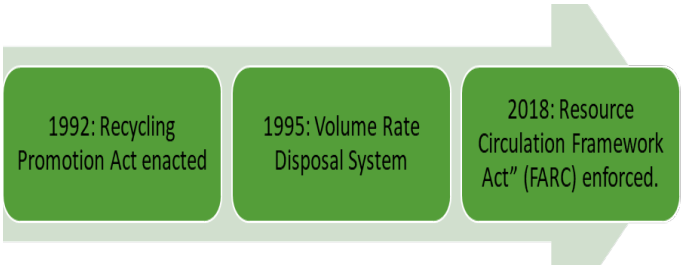


Figure 7: Progression of Waste Management Acts/Schemes [58]

Table 2: Description of Waste Management Acts/Schemes in South Korea [58]

Act/Policy	About the Act/Scheme
Recycling Promotion Act enacted	The aim was to promote recycling through introducing once-off use product laws, waste deposit and fee schemes and establishment of recycling industries.
Volume Rate Disposal System	Established on the premise of the "Pay as you throw scheme" similar to Belgium's case. The aim was to discourage waste generation and increase recycling. This was a deviation from the fixed charges which were previously imposed regardless of the amount of waste a household would generate.
Resource Circulation Framework Act" (FARC)	The aim of this framework was to shift from the linear "take-make-waste" model to a circular economy approach. Hence the country is now working towards a "Zero Waste" policy. The country expects to yield socio-economic and environmental benefits from the framework as a result of reduced pollution and better managed resources.

Between 1994 and 2013, household waste per capita reduced from 1.3 kg to 0.94 kg (47,940 tonnes per day in total), buried waste decreased from 81.2% to 9.6%, burned waste decreased from 15.3% to 6.4% while recycled waste increased from 15.3% to 83.2%, Figure 8 [59]. These outstanding figures brought rapid economic growth into the country [55,60].



Figure 8: South Korea Waste Management Successes

The government of South Korea aims to reduce plastic waste by 50% as well as increase recycling rate from 34% to 70% by 2030. In 2020, coloured plastic bottles were also banned as they are difficult to recycle and polyvinyl chloride products may also follow suit. A target has also been set to eliminate disposal cups and straws by 2027 while cafés can be fined as much as US$1800 for using plastic cups for indoor sitting [60].

South Korea is also a signatory to the international conventions and agreements shown in Figure 9.



Figure 9: Conventions and Agreements Supported by South Korea

3.4. Canada's case

Canada's population is around 38 million [61] and the country has ten provinces namely Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, and Saskatchewan and three territories namely, Northwest Territories, Nunavut, and Yukon [62]. The Environment and Climate Change Canada (ECCC) is the government entity mandated to coordinate environmental policies to protect the environment and the public [63]. In 2018, the Aspirational Canada-wide Waste Reduction Goal was approved with the aim of promoting waste reduction and assessing progress in the country. The goal is to reduce the annual waste generation rate per person from 706 kg (2014) to 490 kg per person in 9 years (2030) and eventually 350 kg per person in 19 years (2040) [64].

On the international stage, during the tenure of the country's G7 Presidency in 2018, Canada launched the Ocean Plastics Charter aimed at protecting the marine environment and was adopted by over 20 countries and more than 50 organisations around the globe [65-67]. The charter outlines:

- the prevention of plastic waste mismanagement
- redesign of plastic products to ensure ease of recovery and recyclability.
- recycling and recycled content targets
- commitment to reduce plastic usage and waste generation.
- importance of seeing plastic as a valuable and not trash [66].

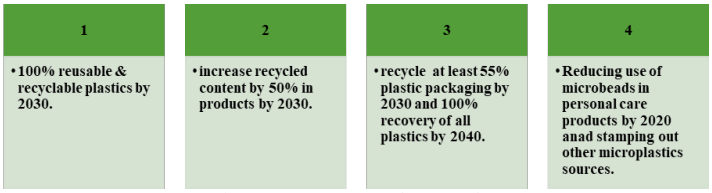


Figure 10: Targets set by Canada

Figure 10 shows the targets set which require collaborations between governments and industry [67].

Canada also pledged $100 million towards the improvement of waste management services in developing countries [65] and is a signatory to the conventions and agreements in Figure 11.

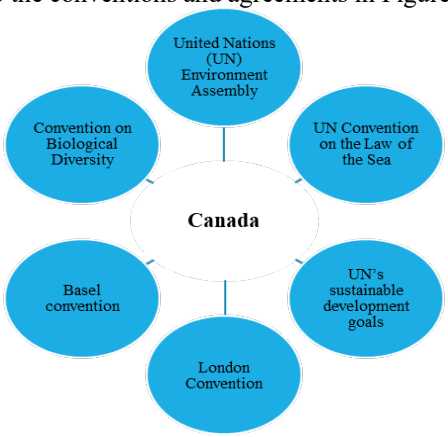


Figure 11: Conventions and Agreements Supported by Canada

Building on the Charter, Canada drafted its zero plastic waste strategy and its success hinges on enforcement of regulations, voluntary industry led initiatives, partnerships with various environment organisations and the community as well as EPR initiatives. In 2019, the Canadian Council of Ministers of the Environment (CCME) drafted an action plan to meet the goals of the zero-waste strategy [67]. The six broad action items outlined in the plan are outlined in Figure 12 below.



Figure 12: Action Areas of the Canada-Wide Action Plan on Zero Plastic Waste Phase 1



Figure 13: Action Areas of the Canada-Wide Action Plan on Zero Plastic Waste Phase 2

In 2020, the CCME launched Phase 2 of the Action Plan whose main focus areas are highlighted in Figure 13 [68].

Figure 14 shows plastic waste management practices in Canada in 2016 which reveal the linearity of the country's waste management techniques. According to the report by authors in [69] Canada lost around US$6 billion through failure to recycle plastic waste and this will increase to US$8.7 billion with a "business as usual" approach [69]. However, with the zero-waste strategy that has been commissioned in the country, this will avert the problems associated with a linear waste management model.

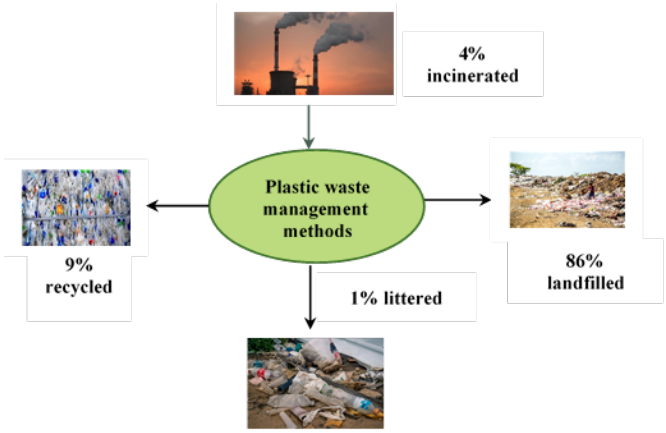


Figure 14: Plastic Waste Management Practices in Canada in 2016 [69]

Although Canada's track record in managing plastic waste still has a long way to go, it has begun to move on a positive trajectory with all the strategies that are being implemented. There are lessons that can be learnt from these and the wealth of research that has been conducted in the country. The success of the zero-waste strategy will result in the following benefits for the country:

- annual cost saving of about US$400 million
- creation of 42,000 jobs
- Greenhouse gas savings of 1.8Mt of CO2e [69].

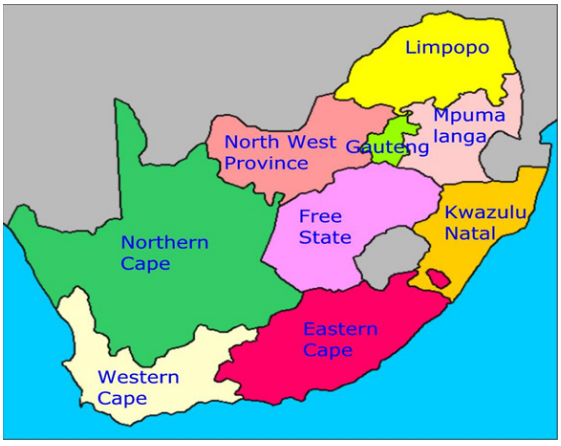


Figure 15: Map of South Africa [70]

3.5. South Africa's case

South Africa has nine provinces, Figure 15 [70] and eight metropolitan municipalities [71], namely; Buffalo City (East London), City of Cape Town (Western Cape), Ekurhuleni Metropolitan Municipality (East Rand, Gauteng), City of

eThekwini (Durban, KwaZulu Natal), City of Johannesburg (Gauteng)), Mangaung Municipality (Free State), Nelson Mandela Metropolitan Municipality (Eastern Cape) and City of Tshwane (Gauteng).

The annual plastic consumption per capita in South Africa is 30-50kg [72] which gives a total consumption of approximately 1.8 to 3 million tonnes. Figure 16 shows a comparison of the daily generation of plastic waste per person for Australia, Belgium, South Korea, Canada and South Africa [73].

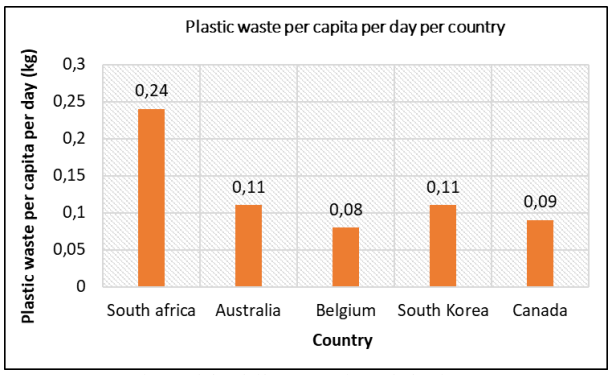


Figure 16: Comparison of the daily generation of plastic waste per person for South Africa and previously mentioned countries Data sourced from [73,74].

However, in [75] the authors argue that the data used in [74] was inaccurate and as a result they determined a significantly lower daily plastic waste generation rate of 0.053kgs/per capita. The authors in [75] used data from the Department of Environmental Affairs (DEA) in [76] in their calculations although the authors in [76] highlighted in their report that there was under-reporting for waste streams such as plastic, organic waste, glass and tyres due to lack of weigh bridges for example at landfill sites as well as the classification of these streams under MSW. The inconsistencies in literature on the waste management statistics reported for South Africa are also highlighted by various authors in [77-79]. Inaccurate data results in the understating of the environmental impact of plastic waste [78].

Furthermore, approximately 19.7 million people (34.1%) in South Africa do not have access to waste collection services and consequently resort to either burning or illegally dumping their waste, in which case valuable recyclable materials leak into the environment and are also not accounted for [7,80]. Availability of correct data ensures that costly problems of over-designing and under-designing equipment for example in an incineration plant, are avoided, as well as allow for effective and appropriate mitigation strategies to be employed. Lack of "buy in" from stakeholders on proposed projects citing exorbitant costs can be a negative consequence that arises from over-designing.

Figure 17 shows the disparity in waste collection services in South Africa's provinces [76]. Provinces with a higher urban population have higher collection rates compared to those with a higher rural population. As the population of informal settlements residents who do not pay rates in cities continues to increase due to rural-urban migration, problems associated with waste management in these cities will intensify [79].

In [81], the authors reported that Europe and other developed countries are 20 to 30 years ahead of South Africa when it comes to waste management. Unlike South Africa, these countries have diverted from the use of landfills and adopted a culture of prevention, reuse, recycling, and recovery [81]. In 2018, Plastics South Africa, classified the country as a "mechanical recycling champion" for having recycled 46.3% of all the plastic waste generated taking into consideration locally manufactured products only [82] against 31.1% for Europe [80]. In the same year, 2018, the highest figure of US$354.4 million was recorded for imports of Plastic & Rubber Articles in South Africa [83], while in 2019, the country imported plastics and plastic articles worth US$2.5 billion, representing a proportion of 2.8% of the overall value of products imported [84]. It is very crucial for recycling statistics reported to consider all imported plastic products for a clearer picture on plastic waste management to be ascertained.

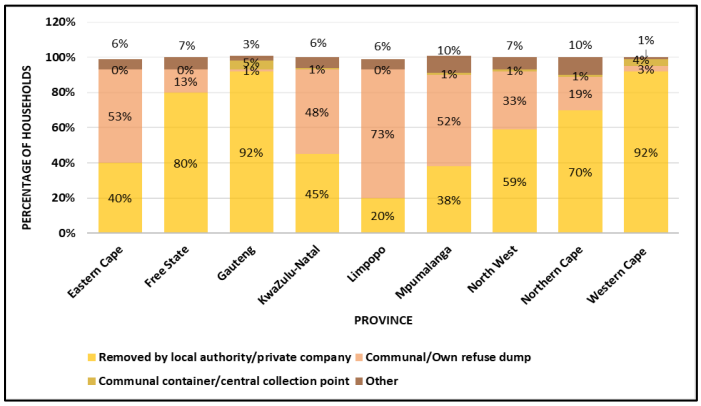


Figure 17: Waste Collection Services in South Africa Per Province (data sourced from [76])

In [85], the authors also argue that since countries employ various approaches when determining recycling rates, comparisons per country are difficult to undertake. In addition, it is also worthwhile to note that, in Europe, the balance of plastic waste remaining after recycling, was either used for energy generation in WtE plants (41.6%) or landfilled (27.3%) while in South Africa's case, the balance of 53.7% was landfilled.

In South Africa, plastic is the material which has the lowest recycling rate among other recyclables such as paper, glass, tyres and metals. In a State of Waste Report released in 2018, plastic had a recycling rate of 43.7% in 2017, while glass, metals (ferrous and non-ferrous) and tyres were at 78.4%, 75% and 100%, respectively, Figure 18.

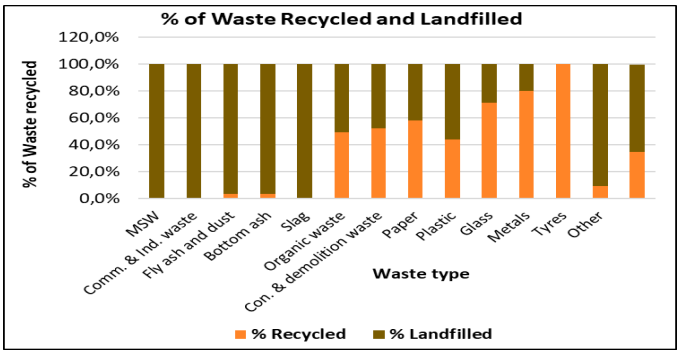


Figure 18: Percentage of Waste Recycled and Landfilled Per Waste Type (data sourced from [76])

The balance of 56.3% plastic was disposed off at landfills [76] an indication of significant plastic leakage.

Further accumulation of plastics in landfills, which are running low on landfill space as well and reaching maximum acceptable height should be prevented [86]. The lack of innovative ideas and research and development inadequacies have been cited as a limitation in South Africa's progression in the Plastics sector [87]. The recycling rate of 100% for tyres is disputable though, since it is common to see tyres indiscriminately dumped.

The World Wide Fund for Nature (WWF), however put the plastics recycling rate for South Africa at 16% with 84% being dumped in waste bins and ending up in landfills while some plastic wastes are transported by wind and end up in the marine environment through storm water drains [72]. It is not clearly defined in [72], whether the mentioned figures are from the input or output recycling rate.

Nonetheless, South Africa still needs to be complimented for the strides that it has made in managing plastic waste. Voluntary producer responsibility organisations (PROs) such as the PET Recycling Company (PETCO) for polyethylene terephthalate (PET) [7], Polystyrene Association of South Africa for all variants of polystyrene, Southern African Vinyls Association (SAVA) for plastics such as polyvinyl chloride and Polyco responsible for polypropylene, high and low density polyethylene and other types [88] have been formed to assist with the recycling of the various plastic waste streams thereby preventing the leakage of plastic into the environment as well as attaining value from post-consumer and landfilled plastics. These companies all fall under the umbrella body Plastics South Africa.



Figure 19a: Shrink wrap before baling



Figure 19b: Baled mixed PET bottles



Figure 19c: Baled plastic bags



Figure 19d: Separately baled coloured PET bottles

Figure 19: The recyclable waste

At the heart of South Africa's recycling industry are about 60 000 waste pickers [89]. However, according to the authors in [81], this figure may be conservative, and they estimated the number to be around 215 000 in 2017 [81]. Waste pickers rummage through bins or are based at landfills where they remove recyclables from MSW destined for the landfill such as plastics which they take to buy-back centres and get paid to earn a living. Approximately 80 to 90% of packaging is recovered by waste pickers [79]. The presence of waste pickers resulted in a landfill cost saving of around $US21.3-US$51.5 million (R309.2-R748.8 million) for municipalities in 2014 [89]. Buy back centres such as Remade Recycling have also been key in South Africa's plastic waste management, receiving all types of plastic waste from waste pickers and collectors, sorting, and baling the recyclable waste before transporting it to convertors as feedstock (Figure 19a-d).

3.5.1 The Role of Government

The National Waste Management Strategy (NWMS) is a statutory requirement of the National Environmental Management Act: Waste Act, 2008 (Act No. 59 of 2008) referred to as the "Waste Act". This strategy aims to ensure that the objectives of the Waste Act are fulfilled [90]. The NWMS of 2011 stressed the importance of re-using, recycling, and recovering waste (3Rs), Figure 20. In [91], targets to have waste collection in 95% of households in urban areas and 75% in the rural areas by 2016 are given. To date, in urban areas and rural areas, 64.7% of households and 75.1% households respectively now have access to waste collection services.

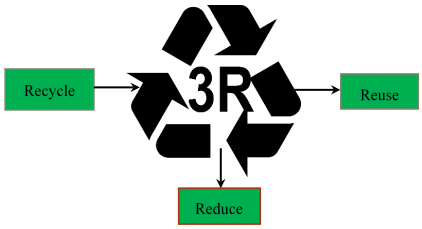


Figure 20: The 3 Rs

On separation of waste at source, the 2011 strategy failed to meet its target of ensuring that households in different municipalities would be separating their waste at source by 2016 [91]. Prior to the drafting of the NWMS, the Polokwane Declaration which envisioned zero waste to landfill for South Africa was signed in September 2001. In the declaration, South Africa set ambitious targets which are summarised in Figure 21.

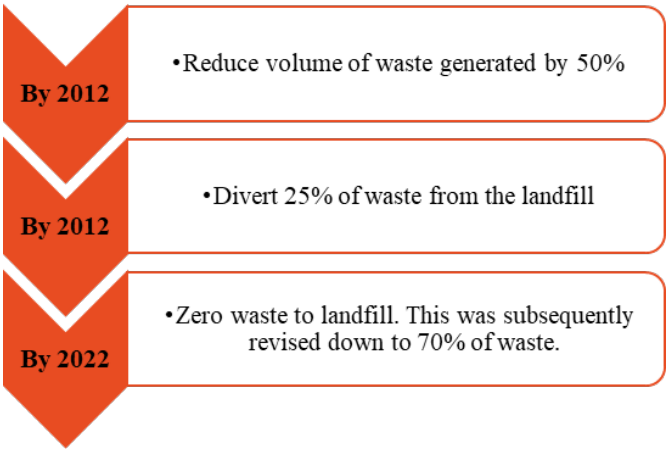


Figure 21: Polokwane Declaration Targets [92,93].

In 2020, the government released a new NWMS which now advocates not only for the 3Rs but also a circular economy approach coupled with EPR, which encourages waste prevention and product regeneration at the end of its lifecycle [37,86] whose waste reduction to landfill pathway is shown in Figure 22. The failure of the Polokwane Declaration to achieve zero waste to landfill has resulted in a new target being set in the 2020 NWMS which goes beyond 2035.

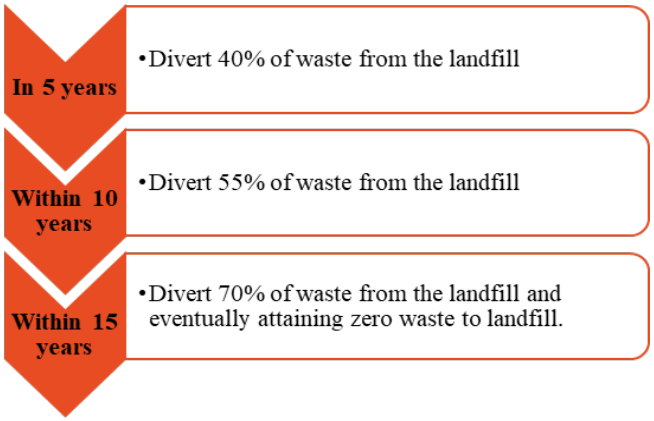


Figure 22: Reduction of waste to landfill pathway

The current NWMS of 2020 also reinforces the need for the continuation of awareness campaigns [86]; which was first outlined in the 2011 NWMS. The objective was to have 80% of municipalities and 80% of schools conducting such campaigns by 2016 [91] and to date 60% of municipalities are conducting these awareness programmes while all schools now teach about waste as part of the curriculum [86].

Minimisation of illegal dumping of waste, littering and the use of single use plastics such as disposable cups and straws is also outlined in the 2020 NWMS. One of the main drivers for this strategy is the need to prevent landfilling, as previously reflected in the 2011 strategy [91]. It remains to be seen whether the latest NWMS strategy will succeed in the areas where the 2011 NWMS strategy has not with the new target year for waste collection in 95% of households set for 2024.

The prevalence landfilling in South Africa has been attributed to the absence of alternatives as well as the belief that this method of waste management, despite being the least recommended in the waste hierarchy, is less costly to implement and yet its environmental impacts are not accounted for [86]. According to the authors in [94], recycling reduces the environmental impact of a product, therefore, increasing recycling rates in South Africa, can only benefit the country.

3.5.1.1. Introduction of other Policy Initiatives

In May 2003, South Africa instituted plastic bag regulations meant to stop the production and importation of sub-standard plastics that could neither be reused nor recycled. This meant that only bags with a minimum thickness of 30 µm were permissible [7,95]. This regulation was accompanied by a plastic levy of 3 cents (R0.03) in 2004, which latter increased to 6 cents (R0.06) and 25 cents (R0.25) in 2017 and 2020, respectively. However, this policy intervention has not yielded the expected results of curtailing plastic bag consumption [96,97].

The Department of Environmental Affairs (DEA) has also begun discussions with businesses to ban microbeads in the cosmetic industry. It is not clear at what stage these consultations are now at [98]. Other initiatives launched by the government to change the public's attitudes and perceptions and lessen the impact of plastic waste while creating employment at the same time include Operation Phakisa (2014) [99], the Recycling Enterprise Support Programme (RESP) (2016) [100] , and the Good Green Deeds programme (2019) [101].

On the international stage, South Africa has pledged to support several Conventions on the protection of the environment, Figure 23.

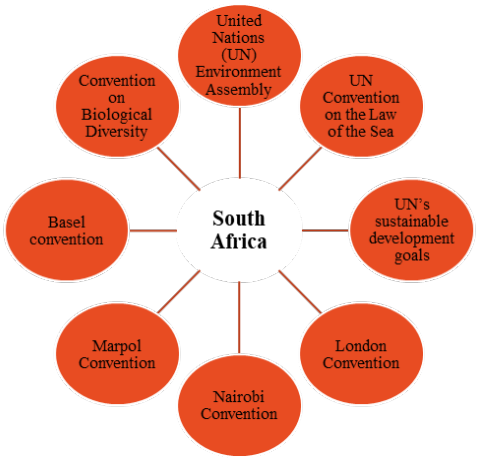


Figure 23: Conventions and Bodies supported by South Africa.

3.5.2 Waste Management Services: A Case Study of Pikitup

Pikitup, which provides waste management services to the City of Johannesburg in the Gauteng Province rolled out a voluntary separation of waste at source programme in some of its suburbs in 2009 [102]. This programme has been moving at a snail's pace because more than a decade later, it has not moved into the rest of Johannesburg. The amount of dry waste collected in the financial year 2016/17 was reportedly 4.5 kgs of dry recyclables per household per month out of an expected 13 kgs per household because not all households separated their waste. On 1 July 2018, Pikitup implemented mandatory waste separation in the communities where the programme had been rolled out, to increase the recycling rates [102]. Therefore, this mandatory call covered approximately 26% of all the households in Johannesburg [102]. Despite the mandatory call, there are no fines instituted for households that do not separate the waste, opting instead to incentivise residents whilst still collecting mixed waste. This creates a point of plastic waste leakage.

The company has mentioned that the roll out to all of Gauteng will happen around 2021 and they expect that this campaign will eventually be implemented not only in Johannesburg but the rest of South Africa. Households where this initiative has not been commenced are encouraged to separate their waste and drop off at nearby drop off centres. Pikitup also mention that dirty clamshell containers, sweet wrappers, detergent bags, potato crisps and sauce packaging are not recyclable [102].

This information and the motivation behind it, needs to be communicated to all households for the programme to be a success. We are now in 2021, so the complete rollout of the separation of waste at source programme appears not achievable. The positive however, is that it shows that the country is aware of what needs to be done but a shift in gears is required so that targets can be met within reasonable and set timeframes. It is undeniable that if this programme is to move like a well-oiled machine, its success hinges on the availability of financial resources as well as willingness by the public to separate their waste. The public also needs to be aware of the various plastic types and the recycling codes associated with them, as shown in Table 3.

Table 3: Plastic Recycling Codes [103]

Recycling Code	Plastic Type	Application
1	Polyethylene terephthalate (PETE)	Soft drink bottles, mineral water, fruit juice container, cooking oil
2	High density polyethylene (HDPE)	Milk jugs, detergents, shampoo bottles
3	Poly vinyl chloride (PVC)	Trays for sweets, fruit, plastic packaging (bubble foil), food wrapping foils
4	Low density polyethylene (LDPE)	Crushed bottles, shopping bags, wrappings
5	Polypropylene (PP)	Furniture, toys, car bumpers
6	Polystyrene	Hard packaging, cosmetic bags, CD cases
7	Other	Acrylic, polycarbonate, fibres, nylon

Figure 24 shows the household separation levels in South Africa's nine provinces, with the Western Cape Province having the highest level of 20% followed by Gauteng Province with 13%.

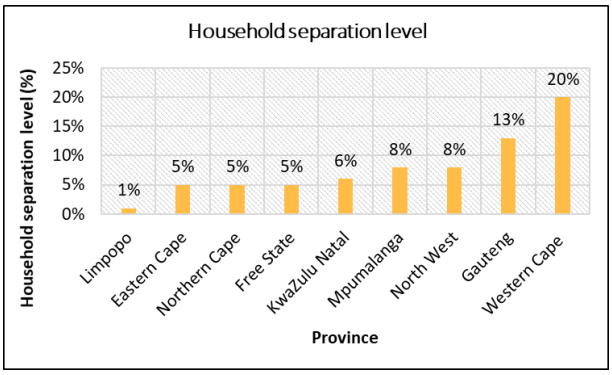


Figure 24: Household separation level per province in South Africa [104]

3.5.3 Other Waste Campaigns

There are also other campaigns that are happening in South Africa. However, it is not clear to what extent the public is aware about them and a national survey would need to be conducted to assess public awareness and knowledge levels. These include the annual Clean-up and Recycle SA Week, the International Coastal Clean-up Day (ICC) [105] and the weekly #KleenaJoburg campaign launched in December 2019 [106]. The Two Oceans Aquarium in Cape Town also runs a campaign called "Rethink the Bag", which educates people on the harm that SUP bags can cause in the marine environment. They have also launched a petition to urge the government to ban single use plastic bags [107]. In some municipalities, beach clean-up campaigns are also done.

4. EPR Reviewed Studies

Various studies on EPR have been conducted across the globe; some of which have been reviewed in this study under this section. In [108] (2010), the author compared mandatory and voluntary EPR schemes in South Africa and found that the latter appeared more effective after comparing the Plastic bag regulations instituted by the government against the PETCO initiative for PET recycling. However, the study also highlights that the differences could also be attributed to the fact that unlike PET, plastic bags are not easily recyclable, are of lower value, recycled bags have limited areas of application and their use as plastic bag liners leads to leakages into the environment. The need to avoid government regulations thought to be harsh may also have possibly fuelled voluntary EPR's success.

In [109], the authors conducted a study in Serbia which revealed several problems. First, the quantity of recyclables recovered from MSW is low due to lack of recycling infrastructure, programmes, and funding. Further to this, landfill taxes are not applied to all municipalities, consequently there is no motivation to remove recyclables from MSW. However, if landfill taxes are exorbitant, this can result in an increase in illegal dumping. In addition, the exclusion of waste pickers from projects involving municipalities and PROs was also observed to be an impediment in the collection of recyclables with waste pickers vandalising infrastructure put in place to take out the recyclables. Lack of expertise during policy formulation was also highlighted. For example, after providing colour coded bins to residents, a single truck would be used to collect the separated waste, which then demotivated the residents. Setting low recycling targets has also affected recovery of materials from MSW. For example, PROs can meet these targets by collecting waste from industrial and commercial wastes such that there is no motivation to collect from residential areas. PROs may also prioritise funding towards one type of recyclable and not others while some may not contribute to infrastructure development. In addition the authors in [109] also mention that there needs to be monitoring of operations of PROs, a minimum target that PROs should contribute towards recovery of recyclables from MSW and these organisations should submit reports to the government annually detailing expenditure for the sake of transparency.

In [110], the authors conducted a study in Colombia (2018), and recommended having multiple PROs for a product as well as delegating enforcement of laws to more than one organisation.

In [111], the authors conducted a study in Europe and found that strict enforcement of EPR related regulations was deficient. Further to this, the lack of incentives for companies that comply resulted in companies not participating in the programme. Despite these limitations, the study concluded that EPR programmes can boost recycling levels and this increase in plastic recyclate, will in turn require effective collection, sorting and treatment of the plastic waste. Therefore, regional and local authorities together with PROs should ensure that collection services are efficient as well as educate the public on how to properly sort waste. Revenue from EPR schemes can be used for these initiatives. The study also highlights the importance of EPR in achieving a circular economy and recommended implantation of Deposit Refund Schemes (DPR) across the European Union. In addition, the need for the opinions of manufacturers, packers, fillers and retailers to be considered when implementing DPR as well as the importance of the public advocating for the use of biodegradable plastics were highlighted. Success of EPR will depend on its confluence with other initiatives such as, labelling, procurement policies, pay-as-you-throw schemes, recycling goals, prohibitions, goods and waste taxes, non-mandatory agreements and public consciousness with no seclusion of any type of plastic.

In [112], the authors refer to the EU Waste Framework Directive, which has assisted waste management in Europe to remain financially viable and independent. The study recommend the roll out of a global EPR as well as ensuring that plastic product

designs are standardised at a global level such that products that do not meet the specifications are banned or taxed. However, the authors also highlight the possibility of resistance from industry regarding the Polluter Pays principle. Deposit return schemes as well as consumer pay-as-you-throw programmes which discourage unsustainable consumption patterns are recommended.

In [113], the authors revealed that EPR in China is hindered by low levels of recycling including lack of recycling technology and therefore, producers should be encouraged to develop sorting and recycling technologies. The study mentions the importance of citizenry involvement, in assisting the government to identify industries not complying with the EPR regulations.

In [33], the authors explained that the reliance of the EPR scheme in South Korea on money paid by producers is a drawback. The study found that some recycled products were of poor quality and therefore, investing in modern recycling equipment is crucial. In South Korea, producers are given an annual target of plastic waste to recycle and this can be done directly or through a PRO, which they pay a fee to. The PRO then pays a subsidy to the Korea Resource Circulation Service Agency (KORA) which is responsible for collection and recycling. Recycling subsidies and private investments have sustained the EPR scheme. Producers that fail to meet their recycling targets are fined a fee that is higher than the recycling fee. PRO and KORA submit reports on the performance of the producers under the EPR scheme to the Korea Environment Corporation. The authors recommended the inclusion of more plastic products under the EPR scheme to improve recycling rates. The study also highlights that since recycling plants operate on a small-scale basis, they are affected by domestic and global trends in the recycling market.

In [114], the authors revealed a number of hindrances to the successful implementation of EPR schemes in Brazil (São Paulo). First, lack of enforcement of the law for those not practicing in EPR schemes as well as absence of incentives could demotivate companies that are practicing. Incentives include offering tax rebates to companies in the scheme. In addition, consumers do not always return the waste, for example if drop off centres are far, while some would rather keep the waste, for example in the case of cell phones and resell. Discounts given on raw materials also do not promote the use of recycled products. Collection of waste may also be a problem if the distances to be covered are long and the waste is little. The authors recommended industry led EPR schemes and that retailers should also form part of the schemes as they are involved in waste collection from consumers. Integration of municipalities into the scheme although necessary, could pose a challenge when determining how much compensation they should get if they are part of the scheme. It was noted that many cases have gone to court because of these disputes and as a result binding agreement must be drafted to avoid such disputes. The importance of awareness campaigns which should be funded by retailers and producers was also noted. Other recommendations include;

- banning the sale of goods subject to EPR but without the scheme in place
- the need for collaborations between municipalities and businesses to also build infrastructure and to stamp out non-compliance
- that all new companies should have a proposed EPR scheme before they can be given permits to operate.

5. Discussion and Conclusions

5.1. Setting of Targets

Reviewing findings from the authors in [74]; of the five countries studied in this paper, South Africa has the highest plastic waste generation rates per capita. There are vital lessons to learn from Belgium, Australia, South Korea and Canada. South Africa has many right elements in its policies to become a pace setter in the continent when it comes to plastic waste management. However, implementation seems to be lacking as set target dates continue to be shifted forward without tangible results being realised in some cases; an example being the "zero waste to landfill target" which was reduced from 100% to 70% waste diversion from landfill by 2022 [93]. Further to this, the target year of 2022 has now been moved to 2035 as highlighted in the 2020 NWMS [86]. Continuous resetting of goals will be detrimental to the country's success in managing its plastic waste as it does not yield benefits except shifting responsibility from one actor to another and this can persist for decades to come. Moreover, if 9.2 jobs are being created for every 10 000 tonnes of waste recycled compared to 2.8 jobs with landfilling [45], then South Africa is losing a significant number of jobs in the waste sector by not stamping out the practice of landfilling which is the main waste management method. Despite these challenges, South Africa has made some headways in trying to manage plastic pollution.

5.2. Waste Collection Services

Ensuring that the entire population of the country has access to waste collection services followed by compulsory separation of waste at source and dedicated collection points for small plastic items are key elements to successful waste management. Currently, with a single bin of mixed waste in most households, 34.1% of the population lacks waste management services [80] and at times refuse collections are not always on schedule as a result of backlogs caused by breakdowns. Therefore, unless there are other role players to assist with the collection of sorted recyclable bins or bags the separation of waste at source programme will be unsuccessful.

5.3. Data Accuracy

This study noted the problem of data inconsistencies in the country making it difficult to ascertain with confidence, the country's waste management record as has been highlighted by other authors in [76-79]. Perhaps, the accuracy of data can be improved by employing real time data logging at recycling centres that receive post-consumer plastic as highlighted in Australia's National Action Plan, installing and properly maintaining weigh bridges at landfills as well as stamping out illegal dumping and burning of plastic waste which both contribute to the load of unaccounted waste.

5.4. Public's Attitudes

As the government and various stakeholders drive the agenda of zero waste to landfill, the consumers should not be left behind because they are also vital in achieving a future where plastic waste is not dumped but repurposed (circular economy). Behaviour

changes go hand in glove with knowledge and if the public is not made aware of the importance or need to separate waste, then the amount of waste sent to the landfills will not decrease to levels observed in Belgium. A change in the public's attitudes has been highlighted as one of the reasons why waste management in Belgium has been a success [35]. Investment in awareness campaigns is necessary and these must reach as far as informal settlements and rural areas. It was noted that rural households are less likely to recycle as compared to urban households [115] and this could be attributed to South Africa's National Domestic Waste Collection Standards which put more emphasis on waste separation in urban areas. Furthermore, young adults and families were also found less likely to recycle [115], and therefore, campaigns like Europe's Generation Awake, which encouraged sustainable living for young adults (25-40 years old) and children, can play a very important role in educating this group. This campaign involved the release of short films such as The Awakeners and Water Maniac Walter [116], which were comical, but got the message through [117]. The Generation Awake campaign was a success as more than 1 million people reportedly accessed their website, their videos were watched over 10 million times, they reached 140 000 followers on Facebook and over 2000 articles were written regarding the campaign [117]. Incentives such as the PAYT scheme practised in Belgium will also encourage reciprocation [118] as seen in Belgium; while enforcing this, will strengthen the regulation.

Moreover, the public should be educated on the different plastic recycling codes. However, a major challenge that is likely to be encountered is that some plastic products in the country do not have recycling codes. Therefore, regulations should be put in place to ensure that all plastic products including imports have these codes. In addition, some of the printed codes are quite small to detect and therefore, increasing their size on the product will also make it easier for the consumer to separate or recycle. A National Littering Index (NLI) similar to the one employed in Australia can be used as one of the tools to assess change in the attitudes and behaviours of people post anti-littering campaigns.

5.5. Lessons from the Covid-19 Pandemic

The Covid 19 phase exposed a major weakness in our waste management system. When the government declared a Level 5 lockdown of the country from the 27th of March 2020 to 30 April 2020 (35 days), where only people regarded as essential services were allowed to work, waste pickers were not on that list, but the municipalities still carried out their mandate of collecting MSW. With no one to remove recyclables such as plastic, this meant that plastic in the MSW was landfilled. It could be argued by some, that since over 70% of South Africa's plastic waste that is recycled is from landfills; the plastic that leaked into these landfills during the lockdown will eventually be reclaimed. However, plastic from landfills is contaminated and therefore not only does this restrict its uses but it also requires washing and is sometimes rejected by recyclers. South Africa, being a water scarce country, with several municipalities under water restrictions means diverting water towards washing of plastic from landfills will only strain this resource further [119]. Therefore, it is necessary to develop a framework to regularise and integrate informal waste pickers in the waste management value chain of all municipalities across the country as they have an important role to play in the prevention of leakages of recyclables.

5.6. Alternatives to mechanical recycling

Other innovative projects such as the construction of a section of a road in Jeffreys Bay with plastic infused tar need to be expanded. About 1.8 million plastic bags can be used for a 1 km stretch of road [120]. Similarly, to Australia as highlighted in [46] South Africa should consider also developing its own standards for use of plastic waste and other recyclables in roads and rail projects.

In Ivory Coast, the country in partnership with the United Nations Children's Emergency Fund (UNICEF) has begun building schools using plastic bricks that are also resistant to fire and are waterproof. The bricks also do not require cement and sand and only require a hammer to lay them [121]. UNICEF has projected that a total of 500 schools; enough to accommodate 25 000 students, will be built by 2021. This project has also empowered women in the process as they are the ones collecting the plastic for recycling into bricks. UNICEF also built a plastic brick making factory in the country to enable all the manufacturing to be done in Ivory Coast. The bricks were initially made by Conceptos Plasticos, a Colombian company. Prior to this initiative, only 5% of Ivory Coast's waste was recycled, so this figure is set to rise as more people begin to realise the value of plastic that is in their own "backyard" [121]. There is a significant amount of plastic that can be reclaimed from South Africa's landfills and used in such projects where contamination may not be a problem.

5.7. Lessons for South Africa from the Reviewed EPR Studies

To ensure EPR's success in South Africa, Multi-National Companies should disclose how much plastic they are putting on the market as well as how they are managing it post-consumer use [6]. Furthermore, producers must meet the recycling targets set by the government otherwise they should be fined fees that are higher than the recycling fees [33]. As industries in South Africa work on producing integrated waste management plans as requested by the government, informal waste pickers, retailers, municipalities and consumers should not be left out. As evidenced in the Serbia study, the exclusion of waste pickers led them to damage infrastructure in order to access the collected recyclables [109].

The government should also be wary no to set low recycling targets, as this can result in depressed recoveries of recyclables [109]. Plastic producers must continue to pay their recycling fees to PROs, which is currently being done in SA for voluntary PROs and plastic importers should not be absolved from paying [110]. The government should also periodically review and ascertain if the fees that are being contributed by producers to PROs are enough to sustain the sector.

Where a producer decides to take the responsibility, and by-pass the PRO, although the producer may submit performance reports, honesty and auditing are also required in this case. Furthermore, the quality of recyclables should also be of high standard so that the recycling industry remains viable up to a point where it can sustain itself without requiring external investments. A high quality will enable high value products to be made as well as increasing the product range that can be made. In that regard, the public should be aware [111,114] of how to sort waste, avoiding contamination of recyclables and to clean any

contaminated plastic waste preferably with used water (grey water).

On the other hand, retailers also play a crucial role in the collection of post-consumer plastic [111,114], for example through deposit refund schemes (DRS). Further to this, municipalities currently own the existing infrastructure for waste collection and therefore may only need assistance from industry to increase capacity [114]. The inclusion of consumers will also be critical as they need to ensure that waste is sorted correctly as well as be prepared to return the waste to the retailers for the DRS to work. In addition, consumers may also play a role by reporting industries flouting the EPR laws [111,113,114].

Enforcement of the law where compliance is lacking which has also been a concern in South Africa [77,79] should also be prioritised in order to avoid frustrating companies that are observing the law [111,114]. Other aspects to be considered are; whether it will be advantageous to have multiple PROs to make them competitive [110], whether it would be beneficial to ban products whose companies are not implementing their EPR schemes [114], and whether permits to operate should only be given to industries who have a proposed EPR scheme in place on registration [114]. There are many lessons on EPR implementation that South Africa can learn as it prepares its own blueprint on EPR and these should be enumerated and deliberated by the government and all concerned stakeholders.

6. Road Map for Future Actions in Plastic Waste Management in South Africa

Figure 25 shows a proposed road map that South Africa can implement moving forward.

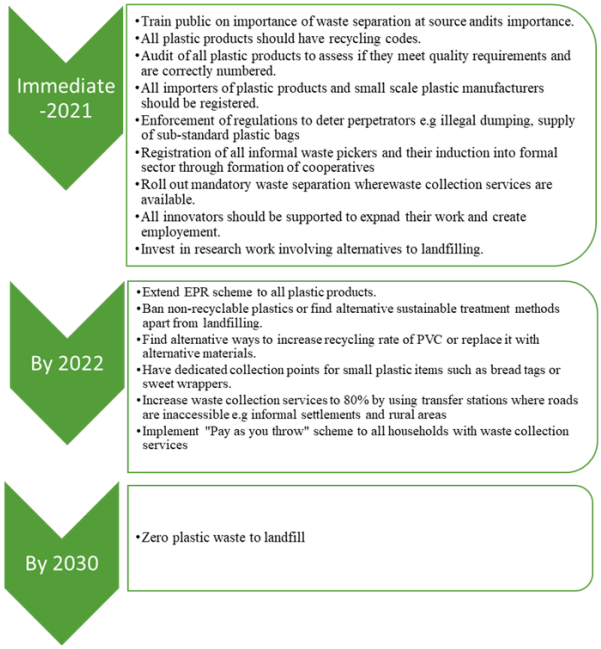


Figure 25: Road Map for South Africa

This paper has highlighted South Africa's strengths, weaknesses, and opportunities in waste management as well as lessons from Belgium, Australia, South Korea and Canada that can be adopted by the country in order to make it a role model for Africa in plastic waste management. There exists an untapped opportunity for improvement of plastic recycling rates to levels comparable to other recyclables in the country through compulsory separation of waste at source. Hence an enabling environment should be created to encourage this practice. Since this will require a fully functional waste management infrastructure, collection services should expand to cover rural areas and informal settlements while industries can assist municipalities to upgrade infrastructure through the EPR scheme. In addition, there is potential for more jobs to be created in the waste sector through recycling than landfilling, thus urgent attention is needed to divert 100% waste from the landfill. Finally, the integration of informal waste pickers into the waste management should be prioritised. If the above-mentioned action points in the road map are succinctly applied, then zero plastic waste to landfill can be achieved earlier (2030) rather than aiming for a 70% reduction in plastic waste to landfill in the same year.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors are grateful to the University of Johannesburg and the Botswana International University of Science and Technology for the financial and technical support.

References

[1] EPA, Municipal Solid Waste, Environmental Protection Agency, 2016.
[2] P. Magutu, C.O. Onsongo, "Operationalising municipal solid waste management," IntechOpen, **2**, 2011, doi:10.5772/16457.
[3] J. A. Nathanson, Solid-waste management. Definition, Methods, Importance, & Facts, Encyclopaedia Britannica, 2020.
[4] DEAT, Guideline on Recycling of Solid Waste, Department of Environmental Affairs and Tourism of South Africa, 2019.
[5] D. Hoornweg, P. Bhada-Tata, What a Waste: A Global Review of Solid Waste Management. Urban development series; knowledge papers no. 15. World Bank, 2012.
[6] M. Williams, R. Gower, J. Green, E. Whitebread, Z. Lenkiewicz, P. Schröder, No Time to Waste. Tackling the Plastic Pollution Crisis before It's Too Late, Tearfund, 2019.
[7] Z.S. Mazhandu, E. Muzenda, T.A. Mamvura, M. Belaid, T. Nhubu, "Integrated and consolidated review of plastic waste management and bio-based biodegradable plastics: challenges and opportunities," Sustainability, **12**(20), 8360, 2020, doi:10.3390/su12208360.
[8] K. Naidoo, An Analysis of Municipal Solid Waste Management in South Africa using the Msunduzi Municipality as a Case Study, M.Sc Thesis, University of KwaZulu-Natal, 2009.
[9] Z.Z. Rasmeni, D.M. Madyira, "A review of the current municipal solid waste management practices in Johannesburg city townships," Procedia Manufacturing, **35**, 1025–1031, 2019, doi:10.1016/j.promfg.2019.06.052.
[10] Worldometer, South Africa Population 2020, 2020.
[11] FACS, Bread - Food Facts, Food Advisory Consumer Service, 2019.
[12] Bread Tags for Wheelchairs, Bread Tags for Wheelchairs – Helping Others Through Top and Tag Recycling, 2020.
[13] S.D. Anuar Sharuddin, F. Abnisa, W.M.A. Wan Daud, M.K. Aroua, "A review on pyrolysis of plastic wastes," Energy Conversion and Management, **115**, 308–326, 2016, doi:10.1016/j.enconman.2016.02.037.
[14] B. Kunwar, H.N. Cheng, S.R. Chandrashekaran, B.K. Sharma, "Plastics to fuel: a review," Renewable and Sustainable Energy Reviews, **54**, 421–428, 2016, doi:10.1016/j.rser.2015.10.015.
[15] S.M. Fakhrhoseini, M. Dastanian, "Predicting pyrolysis products of PE, PP, and PET using NRTL activity coefficient model," Journal of Chemistry, **2013**, 7–9, 2013, doi:10.1155/2013/487676.
[16] M.P. Aznar, M.A. Caballero, J.A. Sancho, E. Francés, "Plastic waste elimination by co-gasification with coal and biomass in fluidized bed with

air in pilot plant," Fuel Processing Technology, **87**(5), 409–420, 2006, doi:10.1016/j.fuproc.2005.09.006.
[17] A. Sivan, New perspectives in plastic biodegradation, Current Opinion in Biotechnology, **22**(3), 422–426, 2011, doi:10.1016/j.copbio.2011.01.013.
[18] O. Gulseven, S. Ashkanani, S. Abdullah, H. Ismaeil, H. Alkandari, M. Baroun, "A sustainable model for enhancing road quality with recycled plastic bags," Kuwait Journal of Science, **46**(4), 112–119, 2019.
[19] A.K. Jassim, "Recycling of Polyethylene Waste to Produce Plastic Cement," Procedia Manufacturing, **8**, 635–642, 2017, doi:10.1016/j.promfg.2017.02.081.
[20] N. Saikia, J. De Brito, "Use of plastic waste as aggregate in cement mortar and concrete preparation: A review," Construction and Building Materials, **34**, 385–401, 2012, doi:10.1016/j.conbuildmat.2012.02.066.
[21] Averda, South Africa Can Turn Plastic Waste into Building Blocks, 2019.
[22] B.L. Hartley, R.C. Thompson, S. Pahl, "Marine litter education boosts children's understanding and self-reported actions," Marine Pollution Bulletin, **90**(1–2), 209–217, 2015, doi:10.1016/j.marpolbul.2014.10.049.
[23] C.M. Rochman, M.A. Browne, B.S. Halpern, B.T. Hentschel, E. Hoh, H.K. Karapanagioti, L.M. Rios-Mendoza, H. Takada, S. Teh, R.C. Thompson, "Policy: Classify plastic waste as hazardous," Nature, **494**(7436), 169–171, 2013, doi:10.1038/494169a.
[24] L. Oyake-Ombis, B.J.M. van Vliet, A.P.J. Mol, "Managing plastic waste in East Africa: Niche innovations in plastic production and solid waste," Habitat International, **48**, 188–197, 2015, doi:10.1016/j.habitatint.2015.03.019.
[25] M. Mourshed, M.H. Masud, F. Rashid, M.U.H. Joardder, Towards the effective plastic waste management in Bangladesh: a review, Environmental Science and Pollution Research, **24**(35), 27021–27046, 2017, doi:10.1007/s11356-017-0429-9.
[26] Y.C. Moh, L. Abd Manaf, "Solid waste management transformation and future challenges of source separation and recycling practice in Malaysia," Resources, Conservation and Recycling, **116**(2017), 1–14, 2017, doi:10.1016/j.resconrec.2016.09.012.
[27] L. Milios, A.E. Davani, Y. Yu, "Sustainability impact assessment of increased plastic recycling and future pathways of plastic waste management in Sweden," Recycling, **3**(3), 2018, doi:10.3390/recycling3030033.
[28] E. Van Eygen, D. Laner, J. Fellner, "Circular economy of plastic packaging: Current practice and perspectives in Austria," Waste Management, **72**, 55–64, 2018, doi:10.1016/j.wasman.2017.11.040.
[29] R.U. Duru, E.E. Ikpeama, J.A. Ibekwe, "Challenges and prospects of plastic waste management in Nigeria," Waste Disposal & Sustainable Energy, **1**(2), 117–126, 2019, doi:10.1007/s42768-019-00010-2.
[30] H.H. Khoo, "LCA of plastic waste recovery into recycled materials, energy and fuels in Singapore," Resources, Conservation and Recycling, **145**, 67–77, 2019, doi:10.1016/j.resconrec.2019.02.010.
[31] L. Peake, Plastic waste in the United Kingdom, Academic Press.
[32] S.M. Nxumalo, S.D. Mabaso, S.F. Mamba, S.S. Singwane, "Plastic waste management practices in the rural areas of Eswatini," Social Sciences & Humanities Open, **2**(1), 100066, 2020, doi:10.1016/j.ssaho.2020.100066.
[33] Y.C. Jang, G. Lee, Y. Kwon, J. hong Lim, J. hyun Jeong, "Recycling and management practices of plastic packaging waste towards a circular economy in South Korea," Resources, Conservation and Recycling, **158**, 104798, 2020, doi:10.1016/j.resconrec.2020.104798.
[34] C. Liu, T. Thang Nguyen, Y. Ishimura, "Current situation and key challenges on the use of single-use plastic in Hanoi," Waste Management, **121**, 422–431, 2021, doi:10.1016/j.wasman.2020.12.033.
[35] M. E. Hall, Belgium takes out the trash, EURACTIV, 2015.
[36] Plastics Europe, Plastics – the Facts, 2018.
[37] Z.S. Mazhandu, E. Muzenda, M. Belaid, T.A. Mamvura, T. Nhubu, "Incineration as a potential solution to Africa's plastic waste challenges: A narrative review," in 2020 International Conference on Energy, Environment and Storage of Energy (ICEESEN), Kayseri, Turkey, 2020.
[38] European Commission, New Waste Rules Will Make EU Global Front-Runner in Waste Management and Recycling, 2018.
[39] Worldometer, Belgium Population 2020, 2020.
[40] J. M. Simon, And the European Waste Champion is… Belgium!, Zero Waste Europe, 2015.
[41] Fostplus, Organisation, 2020.
[42] E. C. Gentil, Municipal Waste Management in Belgium, EEA, 2013.
[43] Worldometer, Australia Population 2020, 2020.
[44] Wikipedia, States and Territories of Australia, 2020.
[45] Department of the Environment, Water, Heritage and the Arts, Waste and Recycling Publications: Employment in Waste Management and Recycling, Australian Government, 2009.
[46] Department of the Environment and Energy, National Waste Policy Action Plan 2019, Australian Government, 2019.
[47] K. Willis, C. Maureaud, C. Wilcox, B.D. Hardesty, "How successful are waste abatement campaigns and government policies at reducing plastic waste into the marine environment?," Marine Policy, **96**, 243–249, 2018, doi:10.1016/j.marpol.2017.11.037.
[48] KAB, Keep Australia Beautiful, 2020.
[49] Department of Agriculture, Marine Debris Including Plastic Pollution, Australian Government, 2020.
[50] APCO, ANZPAC Plastic Pact, Australian Packaging Covenant Organisation, 2020.
[51] Worldometer, South Korea Population 2021, 2021.
[52] Wikipedia, Provinces of South Korea, 2021.
[53] Wikipedia, Recycling in South Korea, 2021.
[54] Anglo Info, Waste Disposal and Recycling in South Korea, 2021.
[55] N. McCarthy, Chart: The Countries Winning the Recycling Race, Statista, 2016.
[56] T. Parker, From Australia to Wales: The Five Best Recycling Countries in The World, NS Packaging, 2020.
[57] Earthsquad, Plastic recycling in South Korea, 2021.
[58] Wastemetasia, Waste Management Policies in South Korea, 2020.
[59] Y. Jun-suk, Breakdown of Korea's Recycling Process, Korea Herald, 2015.
[60] D. Ng, In South Korea, a Lesson to Be Learned from A Plastic Waste Crisis, CNA Insider, 2018.
[61] Worldometer, Canada Population 2021, 2021.
[62] Trip Savvy, Guide to Canadian Provinces and Territories, 2020.
[63] Wikipedia, Environment and Climate Change Canada, 2021
[64] CCME, Aspirational Canada-wide Waste Reduction Goal, Canadian Council of Ministers and of the Environment, 2018.
[65] CCME, Strategy on Zero Plastic Waste, Canadian Council of Ministers and of the Environment, 2018.
[66] IUCN, IUCN Endorses Global 'Oceans Plastics Charter, International Union for Conservation of Nature, 2019.
[67] CCME, Canada-wide Action Plan on Zero Plastic Waste Phase 1, Canadian Council of Ministers of the Environment, 2019.
[68] CCME, Canada-wide Action Plan on Zero Plastic Waste Phase 2, Canadian Council of Ministers of the Environment, 2020.
[69] Deloitte and Cheminfo Services Inc., Economic Study of The Canadian Plastic Industry, Markets and Waste, 2019.
[70] South Africa Travel Online, South Africa's Provinces, 2017.
[71] South African Government, Local Government, 2019.
[72] WWF, Plastic, World Wide Fund for Nature, 2020.
[73] H. Ritchie, M. Roser, Plastic Pollution, Our World in Data, 2018.
[74] J.R. Jambeck, R. Geyer, C. Wilcox, T.R. Siegler, M. Perryman, A. Andrady, R. Narayan, K. Law, "Plastic Waste Inputs from Land into The Ocean," Science, **347**(6223), 768–771, 2015, doi:10.1126/science.1260352.
[75] C. Verster, H. Bouwman, "Land-based sources and pathways of marine plastics in a South African context," South African Journal of Science, **116**(5–6), 1–9, 2020, doi:10.17159/sajs.2020/7700.
[76] DEA, South Africa-State of Waste Report, Department of Environmental Affairs, 2018.
[77] WWF, Legal Framework Study of Extended Producer Responsibility, World Wide Fund for Nature, 2019.
[78] C. Rodseth, P. Notten, H. von Blottnitz, "A revised approach for estimating informally disposed domestic waste in rural versus urban South Africa and implications for waste management," South African Journal of Science, **116**(1–2), 2020, doi:10.17159/sajs.2020/5635.
[79] APWC, South Africa – Waste Classification Report, The Commonwealth Litter Programme, Asia Pacific Waste Consultants, 2020.
[80] Plastics SA, Key Statistics from Latest Recycling Report, 2019.
[81] L. Godfrey, S. Oelofse, "Historical review of waste management and recycling in South Africa," Resources, **6**(4), 57, 2017, doi:10.3390/resources6040057.

[82] SAPRO, About Us, South African Plastics Recycling Organisation, 2020.
[83] Trading Economics, South Africa Imports of Plastic & Rubber Articles, 2020.
[84] D. Workman, South Africa's Top 10 Imports 2019, World's Top Exports, 2019.
[85] BPF, Plastic Recycling, British Plastics Federation, 2020.
[86] DEA, National Waste Management Strategy, Department of Environmental Affairs, 2020.
[87] C. Steyn, Briefing on the Progress of the South African Plastics Sector, DTI, 2015.
[88] Plastics SA, Plastics Industry Shows Its Support for Waste Pickers During Lockdown, 2020.
[89] SAPRO, Why Recycle, South African Plastics Recycling Organisation, 2020.
[90] DEA, National Waste Management Strategy, Department of Environmental Affairs, 2019.
[91] DEA, National Waste Management Strategy 2011, Department of Environmental Affairs, 2011.
[92] DEA, Working on Waste, Department of Environmental Affairs, 2019.
[93] O. Taiwo, F. Otieno, C. Venter, "Towards attaining the Polokwane waste reduction goals – where are we?," Town and Regional Planning, **2008** (53), 25-31, 2008, doi:10.10520/EJC108684.
[94] D.A.R. George, B.C. Lin, Y. Chen, "A circular economy model of economic growth," Environmental Modelling and Software, **73**, 60-63, 2015, doi:10.1016/j.envsoft.2015.06.014.
[95] UNEP, Single-Use Plastics: A Roadmap for Sustainability, United Nations Environment Programme, 2018.
[96] DEA, Department of Environmental Affairs Strives to Improve Plastic Bag Recycling in South Africa, Department of Environmental Affairs, 2017.
[97] DEA, Minister of Environment, Forestry and Fisheries, Ms Barbara Creecy Hosts Plastic Colloquium, Department of Environmental Affairs, 2019.
[98] DEA, Minister Edna Molewa Calls on South Africans to Join Hands to Eradicate Plastic Pollution, Department of Environmental Affairs, 2018.
[99] Operation Phakisa, Chemical and Waste Economy, Department of Planning, Monitoring and Evaluation, 2017.
[100] DEFF, Minister Drives Thuma Mina Call for a Cleaner SA, Department of Environment, Forestry and Fisheries, 2018.
[101] South African Government, Good Green Deeds, 2019.
[102] Pikitup, Mandatory Separation at Source Frequently Asked Questions, 2020.
[103] K. Wilson-Harris, Recycling Codes - Understanding the Plastic Recycling Codes, Proficient Packaging, 2021.
[104] GreenCape, Waste: 2020 Market Intelligence Report, 2020.
[105] DEA, South Africa commemorates International Coastal Clean-Up Day, Department of Environmental Affairs, 2018.
[106] City of Johannesburg, Mayor Makhubo takes Kleena Joburg Campaign to Alexandra, 2020.
[107] Two Oceans Aquarium, Rethink the Bag for a Single-Use Plastic Shopping Bag Free South Africa, 2020.
[108] A. Nahman, "Extended producer responsibility for packaging waste in South Africa: Current approaches and lessons learned," Resources, Conservation and Recycling, **54**(3), 155–162, 2010, doi:10.1016/j.resconrec.2009.07.006.
[109] V. Mrkajić, N. Stanisavljevic, X. Wang, L. Tomas, P. Haro, "Efficiency of packaging waste management in a European Union candidate country," Resources, Conservation and Recycling, **136**, 130–141, 2018, doi:10.1016/j.resconrec.2018.04.008.
[110] J. Park, N. Díaz-Posada, S. Mejía-Dugand, "Challenges in implementing the extended producer responsibility in an emerging economy: The end-of-life tire management in Colombia," Journal of Cleaner Production, **189**, 754–762, 2018, doi:10.1016/j.jclepro.2018.04.058.
[111] W. Leal Filho, U. Saari, M. Fedoruk, A. Iital, H. Moora, M. Klöga, V. Voronova, "An overview of the problems posed by plastic products and the role of extended producer responsibility in Europe," Journal of Cleaner Production, **214**, 550–558, 2019, doi:10.1016/j.jclepro.2018.12.256.
[112] K. Raubenheimer, N. Urho, "Rethinking global governance of plastics – The role of industry," Marine Policy, **113**, 2020, doi:10.1016/j.marpol.2019.103802.
[113] H. Shan, J. Yang, "Promoting the implementation of extended producer responsibility systems in China: A behavioral game perspective," Journal of Cleaner Production, **250**, 119446, 2020, doi:10.1016/j.jclepro.2019.119446.
[114] F. de M. Ribeiro, I. Kruglianskas, "Critical factors for environmental regulation change management: Evidences from an extended producer responsibility case study," Journal of Cleaner Production, **246**, 119013, 2020, doi:10.1016/j.jclepro.2019.119013.
[115] StatsSA, Only 10% of waste recycled in South Africa, Statistics South Africa, 2020.
[116] European Commission, Generation Awake Resource Efficiency Campaign Video, 2011.
[117] European Commission, Generation Awake, 2011.
[118] W. Allen, Sustainability: Six Ways to Influence Behavior Change, Learning for Sustainability, 2016.
[119] DEA, South Africa: A Water Scarce Country, Department of Environmental Affairs, 2011.
[120] N. Manyana, Pilot Plastic Road Completed in Jeffreys Bay, Infrastructure News, 2019.
[121] UNICEF, UNICEF breaks ground on Africa's First-of-its-Kind Recycled Plastic Brick Factory in Côte d'Ivoire, United Nations International Children's Emergency Fund, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 1036-1041 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Social skills and Resilience in Adolescent of Secondary Level of a public Educational Institution in Puente Piedra Lima - 2020

Niurka Jacome-Olacua[1], Joselyne Rodríguez-Paucar[1], Prhitty Marin-Garcia[1], Brian Meneses-Claudio[2,*], Hernan Solis-Matta[1], Eduardo Matta-Solis[3]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[3]*Health Sciences, Instituto Peruano de Salud Familiar, 15304, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 28 September, 2020
Online: 20 October, 2020

Keywords:
Social skills
Resilience
Teenagers
Relationship with the opposite sex
Spirituality

ABSTRACT

Social skills and resilience are very important aspects for mental health; therefore, it is necessary to take into account the positive contribution in the development of adolescents, they are in a vulnerable stage, adapting to physical, mental, emotional, etc. changes. Adolescents, not knowing how to handle the difficulties that may arise, if they do not have the ability to face it and show their positive qualities, can easily fall into depression, student desertion, early pregnancy, addictions to toxic substances (alcohol, tobacco and drugs). That is why in this study to measure social skills, the Elena Gismero Scale of Social Skills questionnaire data collection instrument was used, it has 33 items, 28 refer to the lack of assertion or social skills deficit, 5 of them refers to a positive sense. For resilience, the Connor-Davidson Resilience Scale will be used, it has 25 items. The results obtained with respect to social skills show that in the dimension positive interactions with the opposite sex shows the low level 14.3%, which is equivalent to 45 students, they are presenting difficulties in relating to the opposite sex, regarding resilience in the dimension of spirituality, the low level of 21.7%, indicates that 65 students do not have positive attitudes to fulfill their purposes. The minimum age was 11 years, the maximum age was 17 years, the mean being 13.39 in terms of sex, males predominate with a number of 154 students and 146 females.

1. Introduction

According to the World Health Organization and the Pan American Health Organization, they consider that in recent years adolescents from 13 to 16 years old have suffered from some mental health problem such as stress, depression or apprehension, the use of technology, cyberbullying as well as the condition that adolescents live, whether in areas where violence is used [1].

In Peru, it is estimated that there are 295 thousand people who have limitations to relate either by conduct or their way of thinking, feeling and managing their emotions, showing that 67.8% of them, which is equivalent to 200 thousand people, have moderate and severe limitations [2].

A study carried out on the main problem adolescents face reveals that pregnancy is one of the consequences of not using contraceptive methods, due to lack of information or the myths they believe, followed by that is the state of how they live and the economic income that each family has, the physical abuse that the adolescent faces, these problems leads to the adolescent tending to suffer mental health problems such as stress, depression or apprehension, the use of technology, cyberbullying.

On the other hand, a study on depression in Peruvian adolescents was carried out in a report prepared by the Instituto Nacional de Salud Mental del Perú (INSM) with the result that 7% of the constitutional province of Callao Lima - Peru, suffers from depression disorder, being the most frequent. In the Peruvian Amazon jungle, they reported a prevalence of 1.9%, most relevant in women with 2.2% being adolescents from Lima [3].

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & bmeneses@uch.edu.pe

A study of 352 adolescents on alcohol consumption in the city of Cajamarca revealed that 5.5% of adolescents have alcohol problems. In addition, a comparison was made in which men have a high risk of alcohol consumption with 9.5%, while women have a 3.8% high risk of alcohol consumption [5].

A study was carried out to determine the consumption of toxic and alcohol in adolescents in which 55% of its population were women between 13 and 14 years of age, 70.9% of adolescents consume alcohol. On the other hand, it was evident that 26.4% use tobacco and 14.2% cannabis. However, people who did not consume both products presented better mood and a good family relationship, therefore a good school environment [5].

However, a study was conducted to determine the relationship between resilience and self-concept with social support in a group of adolescents, resulting in a positive correlation of 26.6% in family support and support from friends, women showed a high level of self-concept with 25.4%, as opposed to men who obtained 21.1% [6].

This study will show the most relevant signs that the adolescent presents in order to intervene and therefore avoid future consequences such as stress, depression, alcoholism and drug addiction. In addition, it has been identified through a bibliographic review of the National and International scientific articles that there is a lack of research in mental health in adolescents [6]. The objective of this study is to determine the relationship of social skills and resilience in adolescents of secondary level in a Public Educational Institution of Puente Piedra.

The Elena Gismero Scale of Social Skills questionnaire will be used to measure social skills, it has 33 items. The answers consist of 4 alternatives that are “I do not identify myself at all; most of the time I do not act or I would not "as well as" I very much agree and I would feel or act like this in most cases" [7]. On the other hand, the Connor-Davidson resilience instrument will be used to propose 25 items, which are distributed in 5 dimensions that are personal competence, confidence in one's own instincts, tolerance of adversity, positive acceptance of change, control and spiritual influence¨ [8]. The data from this research will be entered into the statistical software SPSS 24.0, since it is a correlational design study, the Pearson's Chi-square correlation test will be used for data analysis to evaluate the general objective of the study and a descriptive analysis to see the behavior of the main variables of the study.

The following research work is structured as follows: In section II, the methodology that was carried out to apply the survey collection technique will be presented. Section III will deal with the results showing the sociodemographic data of Social Skills and Resilience in adolescents, indicating the high, medium and low level according to its dimensions. In section IV, it will present the discussions of the research work. Finally, in section V, the conclusions of the research work.

2. Methodology

In this section, the steps for completing the questionnaire will be explained, in addition to showing the sociodemographic data.

2.1. Research Focus and Design

The present research is of quantitative approach, as for the methodological design, it is a cross-sectional correlational research.

2.2. Population and Sample

The population is made up of 300 students from the María de Los Ángeles 3070 Educational Institution which involved 154 men and 146 women, the mean age being 13.39.

2.3. Inclusion Criteria

- Students who are in the first to fifth year of high school in the Puente Piedra district.
- Students whose Parents have signed the informed consent to participate in the study.
- Students who agree to participate in the research study.

2.4. Study Variable

The present study presents two main variables that are social skills, these are a group of strategic behaviors that help developing a capacity for adolescents to contribute to solving a social situation in a way that is acceptable to both the person and the environment social, resilience is the ability that it develops people or group, to deal with difficult adversity such as trauma or a situation. Social skills and resilience in adolescents at the secondary level of a public educational institution.

2.5. Data Collection Technique

The technique to be used in this research work will be a questionnaire. As they are 2 variables, it presents Elena Gismero's Social Skills questionnaire consisting of 33 items, divided into 6 dimensions in which self-expression in social situations, the defense of its own rights as a consumer, the expression of anger or disagreement, being able to say no and cut interactions, make a request, and initiate positive interactions with the opposite sex. 28 of the items refer to the lack of assertion or deficit in social skills, and the remaining 5 refer to the positive sense. The responses are dichotomous for each of them, consisting of 4 response options, “I do not identify myself in absolute; most of the time it does not happen to me or it would not "as well as the positive responses" I would feel or act like this in most cases; strongly agree". In the sum of the results, it obtains high global scores, the adolescent would be expressing its social abilities and in different contexts its desertion capacity [7]. Regarding resilience, the Connor-Davidson Resilience Scale was used, it has 24 items divided into 5 dimensions that will assess personal competence, confidence in one's instincts, tolerance of adversity, positive acceptance of change and control, and spiritual influence [8].

2.6. Procedure for Data Collection

Arrangements will be made to the authorities to access the educational institution. First, it processed the letter of presentation of the university, the team went to the educational institution and presented ourselves to the director of the current campus. The team initiates activities such as classroom accounting. Regarding the student census, the assistant principal provided with an updated list, which allowed to identify the number of students available between 1st and 5th grade of secondary school. When identifying the number of students, coordinated with the assistant principal to carry out the surveys in an average time of one month, The time it took to do the surveys was approximately 15 to 20

minutes per classroom, the team surveyed 4 classrooms per day, two days a week, at the time of conducting the surveys, it was quiet since the teachers collaborated so that everything is carried out correctly, as future health personnel, by applying the surveys the team has been allowed to see more closely the means of how they socialize, to know the capacities that some develop in their environment, among the needs, etc. The team also make sure that the filling and process is correct and detect in a timely manner some erroneous process that may have been presented.



Figure 1: Carrying out the Survey on adolescents in secondary education at the María de los Ángeles Educational Institution N° 3070, in the Puente Piedra district - Lima, 2019.

3. Results

The tables are showing the research results and the relationship between social skills and resilience.

Table 1: Sociodemographic Data on Social Kkills and Resilience in adolescents of Secondary Education of the María de Los Angeles N° 3070 Educational Institution, District of Puent Piedra - Lima, 2020 (n = 300)

Participant Information	**Total**	
	N	**%**
Total	300	100
Gender		
Female	146	48,7
Male	154	51,3
Who do adolescents live with?		
Parents	191	63,7
Other relatives	108	36,0
Alone	1	0,3
Types of Family		
Nuclear	184	61,3
Extended	115	38,3
Aggrandized	1	0,3
Study Grade		
First	72	24,0
Second	87	29,0
Third	40	13,3
Fourth	70	23,3
Fifth	31	10,3

In Table 1, it has the sociodemographic data of the study participants, there were 300 adolescents from secondary education. The minimum age is 11 years old; the maximum age is 17 years old; the average is 13.39 years old.

Regarding gender, the male sex predominated with 154 students representing 51.3%, followed by the female with 146 students representing 48.7%.

Regarding the dimension; Who do students live with? the highest percentage was that they live with their parents with a population of 191, which is equivalent to 63.7%, followed by other relatives, 108 (36.0) and only 1 (0.3).

The type of family nuclear families predominates with 184 (61.3%) cases, extended family with 115 (38.3%) cases and aggrandized with 1 (0.3%) cases. Regarding the study grade, the distribution is heterogeneous, the students of the First grade 72 (24.0%), the second grade 87 (29.0%) and the fourth grade 70 (23.3%) exceed in proportion to those of third grade 40 (13.3%) and fifth grade 31 (10.3%) of high school.

Table 2: Social Skills and Resilience in adolescents of Secondary Education of the María de Los Angeles N ° 3070 Educational Institution, of the Puente Piedra District - Lima, 2019 (N = 300)

Resilience			
Spearman's Rho	**Social skills**	**Correlation coefficient**	0,086
		Sig. (Unilateral)	0,068
		N	300

Test correlation coefficient is + 0.086 ($p<0.05$). Regarding the previous results, there is no statistical evidence to confirm that there is a relationship between the variables Social Skills and Resilience.

Table 3: Social Skills in Adolescents of Secondary Education of The María De Los Angeles N ° 3070 Educational Institution, of the Puente Piedra District - Lima, 2019 (N = 300)

		N	**%**
Social Skills	**Low Level**	0	0,0
	Medium Level	292	97,3
	High Level	8	2,7
Total		300	100,0

In Table 3, it can see the social skills of secondary school adolescents from the María de los Ángeles No. 3070 Educational Institution, in the Puente Piedra District - Lima where a population of 300 adolescents, 292 Students representing the 97.3 have medium level when making requests, 8 adolescents representing 2.7% have a high level when saying no and cutting

off interactions and 0.0% do not present a low level when initiating positive interactions with the opposite sex.

Table 4: Resilience, in adolescents of Secondary Education of the María de Los Angeles N ° 3070 Educational Institution, of the Puente Piedra District - Lima, 2019 (N = 300)

		N	%
Resilience	**Low Level**	1	0,3
	Medium Level	174	58,0
	High Level	125	41,7
Total		300	100,0

In Table 4, the resilience in adolescents of secondary education of the María de los Ángeles N ° 3070 Educational Institution of the Puente Piedra district - Lima 2019, where 174 adolescents representing 58% have a medium level of adaptation and capacity to recover, 125 adolescents representing 41.7% have a high level of control and purpose and finally 1 adolescent representing 0.3% has a low level of spirituality.

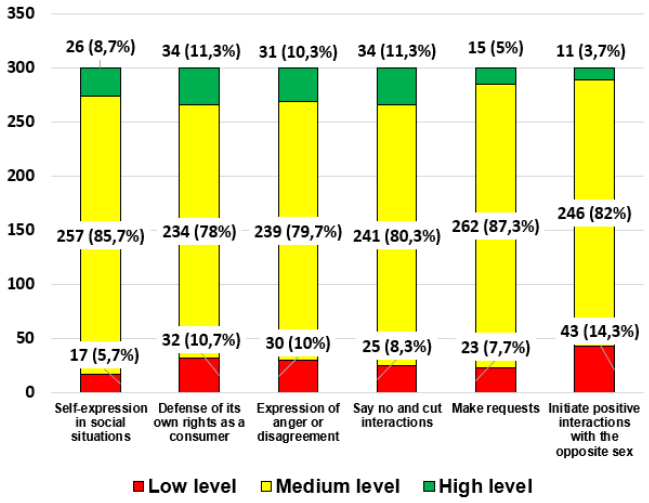


Figure 2: Social skills according to their dimensions in adolescents of Secondary Education of the María de Los Angeles N ° 3070 Educational Institution, of the Puente Piedra District - Lima, 2019 (N = 300)

In Figure 2, it can observe social skills according to their 6 dimensions, where the most affected dimension with low level was to initiate positive interactions with the opposite sex with 43 representing 14.3% adolescents followed by the defense dimension of the its own rights as a consumer 32 that represents 10.7% of the population, the dimension of expression of anger or disagreement with 30 that represents 10%, the dimension of saying no and cutting interactions with 25 which is 8.3%, the dimension of making requests with 23 representing 7.7%, the self-expression dimension with 17 which is 5.7%. The medium level dimension is to make requests with 262 adolescents, which is equivalent to a percentage of 87.3%, the self-expression dimension with 257 represents 85.7% of the population, the dimension of initiating positive interactions with the opposite sex 246 students represent 82%, the dimension saying no and cutting interactions with 241 represents 80.3%, expression of anger or disagreement 239 which represents a percentage of 79.7%, the dimension of defense of its own rights as a consumer with 234 which represents 78%, the dimension of high level that is say no and cut interactions with 34 which represents 11.3%.

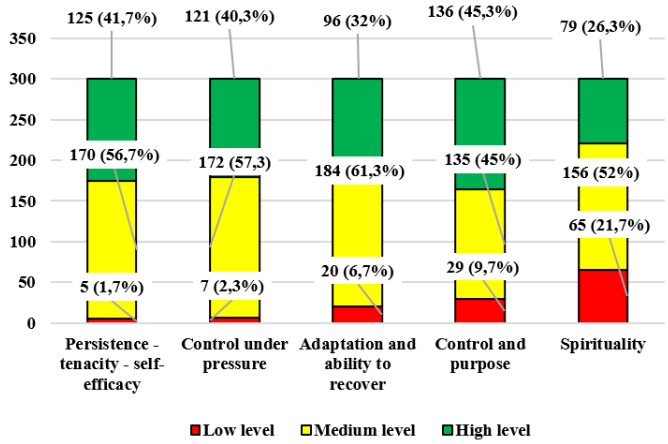


Figure 3: Resilience according to its dimensions in adolescents of Secondary Education of the María de Los Angeles N ° 3070 Educational Institution, of the Puente Piedra District - Lima, 2019 (N = 300)

In Figure 3, it can see resilience according to its dimensions, where the most affected dimension was spirituality with 65 which represents 21.7% of adolescents, the control and purpose dimension with 29 which represents 9.7% of the population, the adaptation dimension and ability to recover with 20 which represents a percentage of 6.7%, followed by the control dimension under pressure with 7 which represents a percentage of 2.3% of adolescents, the persistence-tenacity-self-efficacy dimension with 5 represents a 1.7%. The medium level dimension is adaptation and ability to recover with 184 which represents 61.3%, followed by the control dimension under pressure with 172 which represents 53.3% of adolescents, the persistence-tenacity-self-efficacy dimension with 170 represents 56.7%, spirituality with 156 which represents a percentage of 52%, control and purpose with 135 represents 45%. The high level dimension is control and purpose with 136 which represents a percentage of 45.3% of adolescents.

The importance of knowing the data from this study will contribute to the school authorities are going to detect any problem that may be harmful to students, such as student desertion, dependence on toxic substances such as alcohol, tobacco, drugs, depression, among others. It will also contribute to the health level in which it allows to directly identify the needs and deficits that adolescents may present, this stage is important since they can develop mentally as appropriate as possible and thus be able to focus with more attention to health Mental that is at the limit of risk in what is social skills and resilience.

4. Discussion

In the present research work, it shows that 58.0% of the students are at a medium level with respect to resilience, presenting higher percentages in the dimensions of spirituality, a percentage of 52.0% and in control under pressure shows the 61.3%. Corroborating data on resilience in adolescence is that in the first grades there is a significant decrease in resilience, but it also shows that in intermediate grades, such as 3rd grade, it is still low, but partially recovering in later grades. as 4th and 5th grades [9].

A study was carried out in which the effects of behavior in a social skills program in young people between 14 and 18 years of age in vulnerable situations were identified, for which two interventions were carried out, which they called pre-test and post-test, giving as result in the first one an average of 28.40% and in the post, result 86.64%. This research work also obtain similar results which reinforces the risk that adolescents are in [10].

A research work was carried out in an Educational Institution regarding the well-being in adolescence in which they looked for the relationship between resilience and subjective well-being in adolescents between 12 and 15 years old, resulting in the relation between the subjective well-being and resilience in which have an average of 74.80% in their total sample, making women the most affected [11].

On the other hand, a research on the resilience implied regarding sex and educational level was carried out, it is observed that the students who participated in the research, have a high level of resilience in what is behavioral involvement, 20.29%, which is important that make more emphasis on the emotions of adolescents in schools [12].

Compared with the reviewed studies, we reinforce that the results obtained in this study show that adolescents are at a risk level. Regarding resilience, which is worrisome since if this is not treated in time, adolescents can present difficulties to develop and easily fall into consequences such as depression, student desertion, etc. It was also evidenced that students with fewer grades represent a lower level of resilience, but that would recover progressively since in the higher grades they represent a higher level. In the following study of resilience, it indicates that women would be the most affected regarding to well-being in adolescence and resilience, even though our predominant population are men with a percentage of 51.3%, it is vitally important to be aware of warning signs. Regarding to social skills, in the analyzed study, we have obtained similar results, which reinforces the importance of this study since it is necessary for the authorities to pay attention to the possible problems that the students may have, and to act as soon as possible with the support of mental health professionals and thus the quality of development of these adolescents can be improved, reinforcing their self-esteem, ability to face adversities and improve their social skills, being able to be an important contribution to society. What we want is demonstrate that at this stage of life, it is necessary to be aware because it is very easy for social problems to appear and it is also a stage where if the qualities and mental health are reinforced, these adolescents can be positively motivated resulting in good citizens.

5. Conclusion

Regarding Social Skills and Resilience in adolescents at the secondary level of a Public Educational Institution in Puente Piedra District, it was found that there is no significant relationship between these 2 variables.

Regarding social skills in its self-expression dimension in social situation, the medium level predominated with 85.5%, followed by the high level 8.7%, low level 5.7%, the defense of its own rights as a consumer dimension, the medium level predominated 78%, followed by the high level 11.3%, low level 10.7%, in its dimension expression of anger or disagreement, the medium level predominated 79.7% followed by the high level 10.3%, low level 10%, in its dimension saying no and cutting interactions predominated the medium level with 80.3%, followed by the high level 11.3%, low level 8.3%, in its dimension making requests predominated the medium level with 87.3%, followed by low level 7.7% the high level with 5%, in its dimension to initiate positive interactions with the opposite sex, the medium level predominated with 82%, followed by the low level with 14.3%, high level 3.7%.

It can conclude that the social skills in this study indicate that 6 dimensions that the measurement scale has, 5 of them are at a medium level, followed by high level and low level. The 6th dimension must be taken into consideration, initiate positive interactions with the opposite sex, which has an medium level of 82%, but followed by this is the Low level with 14.3% and the high level obtains 3.7 %; this would represent that students are at risk, the 6th dimension is the most worrisome since adolescents could have difficulties in relating with the opposite sex, although it is true that it considers that they are in a stage of attraction and infatuation typical of age, for this reason, it is worrying that this dimension may decrease, the authorities of this institution should have professional advice such as psychologists who can detect the reason why this interaction is decreasing, which may be an alarm sign such as low self-esteem, upbringing, family history (such as dysfunctional families, among others), leading to difficulties in developing this stage normally. Using this study, we demonstrate and emphasize that it is necessary to take mental health into account at this stage of life, since these adolescents will be the future of the country and society. Take into account how necessary is to carry out a timely control, schools or colleges and especially families add mental health as one of their priorities.

Regarding resilience in its persistence-tenacity-self-efficacy dimension, the medium level predominates with 56.7%, followed by the high level with 41.7%, low level 1.7%; in its control dimension under pressure, the medium level predominates with 57.3%, followed by the high level 40.3%, low level 2.3%; in its adaptation and ability to recover dimension, the medium level predominated with 184 which represents 61.3%, followed by the high level 32%, low level 6.7%; in its control and purpose dimension, a high level predominated with 45.3%, followed by a medium level 45%, low level 9.7%; in its spirituality dimension, a medium level predominated with 52%, followed by a high level 26, 3%, low level 21.7%.

It can conclude that the Resilience in the present study showed that, in the 5 dimensions of the measurement scale, they are all at a medium level, followed by the high level and finally the low level. But in the spirituality dimension, at the low level it is 21.7% which is equivalent to 65 students, this amount is very high, comparing it with the results obtained in the low levels of the other dimensions, this is worrying since that dimension refers to the conformist attitude indicator, this means that 65 students do not have positive attitudes to be able to set a goal and finish it, which is very likely that they do not want to improve their quality of life, this could be considered an alarm sign and is necessary to seek the intervention of medic professionals, psychologists among others, in order to detect the reason for this attitude, they will be able to make a more complete observation. The authorities of the institution must act as

soon as possible so that this is improved and good mental health can be guaranteed in these adolescents, and in this way enhance their capacities to face the difficulties that arise.

Conflicts of Interest

The authors declare no conflict of interest**.**

References

[1]. W.H.O. Pan American Health Organization, Los jóvenes y la salud mental en un mundo en transformación, PAHO, 2018.

[2]. Ministerio de Salud, Prioridades de Investigación en Salud, Instituto Nacional de Salud, 2016.

[3]. L. Donatus, D.J. Sama, J.M. Tsoka-Gwegweni, S.N. Cumber, "Factors associated with adolescent school girl's pregnancy in kumbo east health district north west region Cameroon," Pan African Medical Journal, **31**(3), 1–11, 2018, doi:10.11604/pamj.2018.31.138.16888.

[4]. J.S. Navarro-Loli, M. Moscoso, G. Calderón-De La Cruz, "Research on depression in adolescents in Peru: a systematic review," Liberabit: Peruvian Psychology Journal, **23**(1), 57–74, 2017, doi:10.24265/liberabit.2017.v23n1.04.

[5]. R. Leal, L. Vásquez, "INFLUENCIA DE LOS ESTILOS DE CRIANZA Y LA RESISTENCIA A LA PRESIÓN DE GRUPO SOBRE EL CONSUMO DE ALCOHOL EN ADOLESCENTES DE LA CIUDAD DE CAJAMARCA," Perspective Journal, **17**(1), 33–45, 2016.

[6]. A. Rodríguez-Fernández, E. Ramos-Díaz, I. Ros, A. Fernández-Zabala, "Relaciones de la resiliencia con el autoconcepto y el Apoyo Social Percibido en una muestra de adolescentes," Psychology Action, **12**(2), 1, 2015, doi:10.5944/ap.12.2.14903.

[7]. E. Gismero, EHS. Escala de Habilidades Sociales, 3rd ed., TEA Ediciones, Madrid-España, 2010, doi:10.5901/ajis.2016.v5n3s1p36.

[8]. M.I. Soler Sánchez, M. Meseguer de Pedro, M. García Izquierdo, "Propiedades psicométricas de la versión española de la escala de resiliencia de 10 ítems de Connor-Davidson (CD-RISC 10) en una muestra multiocupacional," Latin American Journal of Psychology, **48**(3), 159–166, 2015, doi:10.1016/j.rlp.2015.09.002.

[9]. I. Axpe, I. Ros, E. Ramos, "Resiliencia y bienestar subjetivo en estudiantes de secundaria en funcion de variables sociopersonales," Proceedings of 3rd International Congress of Educational Sciences and Development, **26**(4), 15–19, 2018.

[10]. J. Redondo-Pachecho, J.S. Parra-Moreno, M. Luzardo-Briceño, "Efectos comportamentales de un programa de habilidades sociales en jóvenes de 14 a 18 años en situación de vulnerabilidad," Pensando Psicología, **11**(18), 45–58, 2015, doi:10.16925/pe.v11i18.1003.

[11]. A. Rodríguez-Fernández, E. Ramos-Díaz, I. Ros, A. Fernández-Zabala, L. Revuelta, "Bienestar subjetivo en la adolescencia: El papel de la resiliencia, El autoconcepto y el apoyo social percibido," Psychology Sum, **23**(1), 60–69, 2016, doi:10.1016/j.sumpsi.2016.02.002.

[12]. A. Rodríguez-Fernández, E. Ramos-Díaz, I. Ros, A. Fernández-Zabala, L. Revuelta, "Resiliencia e implicación escolar en función del sexo y del nivel educativo en educación secundaria," Open School, **44**(2), 1–6, 2015, doi:10.1016/j.aula.2015.09.001.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1345-1349 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Social skills and Resilience in Adolescents of Secondary Education of the Kumamoto I 3092 Educational Institution, of the Puente Piedra District – Lima 2019

Evelyn Roncal-Cespedes[1], Gloria Castillo-Laban[1], Brian Meneses-Claudio[2,*], Hernan Matta-Solis[1], Lourdes Matta-Zamudio[1], Eduardo Matta-Solis[3]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[3]*Health Sciences, Instituto Peruano de Salud Familiar, 15304, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 05 December, 2020
Online: 16 December, 2020

Keywords:
Social skills
Psychological resilience
Adolescents
Students

ABSTRACT

Social skills are the behaviors that the individual adopts on a regular basis at specific moments in everyday life. Resilience is the way in which the individual faces the adversity that life imposes on them, as well as manages to overcome it. In adolescence, problems are perceived because important changes are generated in an accelerated way, many adolescents develop psychological disorders or alterations that could compromise their own lives, if not detected in time, therefore, the objective of this study is to determine the social skills and resilience in adolescents of secondary education of the Kumamoto Educational Institution I 3092, of the Puente Piedra district - Lima, 2019. Applying a quantitative, non-experimental, descriptive, and correlational study, with a population of 626 adolescents from the Kumamoto I 3092 educational institution in the Puente Piedra district, who answered a questionnaire with sociodemographic data and the instruments of the social skills scale and the Connor-Davidson questionnaire for resilience. In the results, the predominance of the medium level in both variables, in social skills with 87.5%, in resilience with 58.3%, concluding the need to carry out interventions in adolescent students, to prevent the inadequate development of their personality.

1. Introduction

The social skill or assertive behavior is a set of expressions that is manifested in a verbal and non-verbal way, where the person externalizes in an interpersonal way all his deficiencies, emotions, ideas, favoritism, rights, avoiding anxiety and hostility [1], [2]**.** Resilience is the ability that a subject or a group obtains that projects itself into the future despite difficult situations and on some occasions can generate trauma and serious consequences in its own life [3].

In adolescence, internal problems such as anxiety disorders, depression, post-traumatic disorders, and even obsessive-compulsive disorders can be perceived. They may also perceive external problems associated with abrupt behavior, conflicts with their environment and aggressive behavior [4]. Family problems can have a strong impact on the adolescent's life, which is why it is important that the family manages its problems correctly during this period, since otherwise this problem will affect the adolescent, diminishing his or her abilities in the academic aspect in society and in a personal way [5].

Worldwide, social skills are specifically oriented to prevent risk behavior and promote a healthy lifestyle [6]. Skills education has emerged since the increase of the problems that affects the community, but directly to adolescents as there are several events such as diseases generated through sex, drug addiction, bullying and even suicide [7].

According to the Pan American Health Organization, it promotes healthy lifestyle and the development of the adolescent stage in the Caribbean and Latin America. Regarding resilience, it establishes support for adolescents based on strategies to reduce various risk behaviors. This process will be continuous, where

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & bmeneses@uch.edu.pe

adolescents will use adaptation skills to seek adequate development within society, the family, school and culture [8].

PAHO establishes that social skills require the adoption of attitudes that are indispensable for optimal behavior to achieve success in the various challenges that life imposes. This favors healthy physical, social and mental behaviors, where adolescents have various possibilities for satisfying their needs, developing their aspirations and coping their environment [9].

In [10], un studio entre 170 adolescents between 15 to 17 years old participated in this study, to whom the social skills questionnaire was applied, where the predominance of the medium level is seen with 50.6%, continuing the low level with 38.2%, then the high level with 11.2%. In the explanation of the dimensions, it is observed that the self-expression factors in social situations predominate the low level with 40.6%, making requests 59.4%, while the high level corresponds to the expression of anger and disagreement with 45.3%. Regarding its study, it mentions the importance of developing social skills in an adequate way to reduce the risk of developing antisocial behaviors.

In [11], it developed the study with a population of 49 adolescents, where the prevalence of medium level resilience was found with 51%, then the low level with 43% and a high level of 6%. In its study, it was shown that adolescents have a medium capacity of resistance to remake and try to escape complex situations.

In [12], the present study had a population of 119 adolescents, where the results show that 71.75% is equivalent to the medium level, in their study, they determined to resort to development programs for teachers in order to include social skills topics in the training of adolescents.

In [13], they indicate that their total population was 160 adolescents, of whom the prevalence of medium-level resilience was obtained with 52.5%, continuing the high level with 46.9% and the low level of 0.6%. Regarding its control and purpose dimensions, they obtained a high level of 61.9%, while control under pressure had a medium level with 59.4% and spirituality a low level of 8.1%. Concluding that the present work will be use as reference to suggest precise interventions where the beneficiary will be the adolescent.

The main objective of this research study is to determine the relationship of social skills and resilience in adolescents because there is a deficit in research on this enigmatic topic in adolescence.

This study presents a quantitative approach, it will be carried out without manipulating the variable, which is why it is a non-experimental and descriptive cross-sectional design. The technique to be used in the research work will be the survey, since it is a measurement instrument which the data will be collected appropriately in a reliable, objective, and valid way.

The Elena Gismero Gonzales, E.H.S questionnaire includes 33 items, 28 of which are written in the sense of lack of assertion or deficit in social skills and 5 of them in the positive sense. The Connor –Davidson Questionnaire (CD-RISC) This test was created in 2003 and consists of two fundamental versions, one of 25 items and the other of 10, in which to carry out the research work the 25-item version was used. All the data collected was entered into the matrix designed by the statistical software SPSS 24.0., This procedure was carried out cautiously to not have errors and avoid values being lost.

This study is important because it will show us the capacity that adolescents have in the face of stimuli of changes in their personality and in their environment, in which it allows the adolescent to improve their capacity for understanding and adaptation both on a social, personal, and family level.

2. Methodology

Starting by making a request and receiving the authorization of the director of the educational institution, it was always established that ethical aspects were applied during the process, respectively, the students voluntarily answered the questionnaires. It is worth mentioning that these questionnaires were provided by the researchers, in the different classrooms, in the presence of the teachers from each grade. The descriptive cross-sectional and correlational methodological design was chosen. Once the database was assembled, the descriptive analysis began, where the percentages achieved, as well as its dimensions, are emphasized by the contingency table that will later be reflected in the bar charts. The stages of data collection are shown in Figure 1 where the processes by which the students were part are shown.

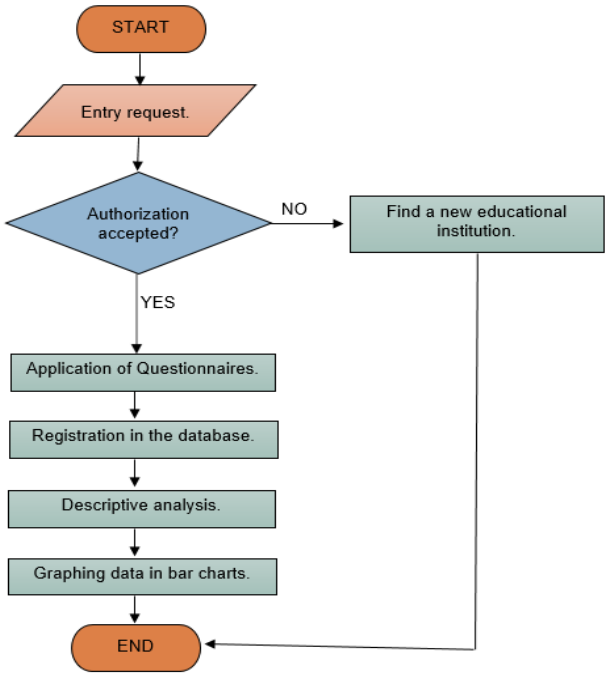


Figure 1: Flowchart for data collection

1.1. Population and sample

For this research, a vulnerable population was selected from a state school in the Puente Piedra district, where 626 students participated, they are the total population, ranging in age from 11 to 18 years old, who are in 1st to 5th grade of secondary education.

1.2. Type of study

In the present research work, due to its characteristics, data collection and the measurement of the variables involved is a quantitative approach. Regarding the methodological design, it is a non-experimental, descriptive, and correlational study.

Table 1: Distribution of classrooms at the secondary level of the Kumamoto I 3092 educational institution

Grade	Shifts	Classrooms	Total Students
1st	Afternoon	5	145
2nd	Afternoon	5	135
3rd	Afternoon	6	133
4th	Afternoon	4	103
5th	Afternoon	4	110
TOTAL		24	626

1.3. Inclusion and exclusion criteria

Inclusion:

- Adolescents who study in 1st to 5th grade of high school.
- Adolescents who have signed the informed consent.
- Parents who have signed the informed consent for the application of the survey in the adolescent.

Exclusion:

- Adolescents and parents who have missed classes and who have not signed the informed consent.
- Parents who have not signed the informed consent.

1.4. Analysis of the variable

The following study has two main variables, social skills, and resilience in adolescents, both variables according to their nature, are qualitative or categorical, and their measurement scale is ordinal.

The technique for data collection in the educational institution was the survey, which is very useful to collect data appropriately in a reliable, objective and valid way [14].

To assess social skills, the instrument of Elena Gismero Gonzales E.H.S was used, which comprises 33 items, 28 of which are written in the sense of lack of assertion or deficit in social skills and 5 of them in the positive sense. The final factor analysis has revealed 6 factors: self-expression in social situations, defense of their own rights as a consumer, expression of anger or disagreement, saying no and cutting interactions, making requests and initiating positive interactions with the opposite sex [15]. This instrument has been valid in different Latin American countries and in other countries, as well as in Peru that has construct validity, the meaning attributed to the measured construct (assertion) is correct. Regarding statistical validity, it was obtained using the Kaiser-Meyer-Olkin (KMO) obtained a sample adequacy coefficient of 0.923 and the Bartlett's Sphericity Test of 0,000 (x^2 = 4036,183; g.l. = 300; $p < 0,05$). Regarding the reliability of the instrument, it was determined based on the Cronbach's Alpha statistic. For the Social Skills Scale instrument, it was determined that it presents an internal consistency index of 0.805 ($\alpha>0.6$).

The Connor –Davidson questionnaire (CD-RISC) was used to assess resilience. This test was created in 2003 and consists of two fundamental versions, one of 25 items and the other of 10 in which the 25-item version was used to carry out the research work. The construct consists of five factors: persistence-tenacity-self-efficacy (items: 10-12, 16, 17, 23-25); control under pressure (6, 7, 14, 15, 18, 19, 20); adaptability and support networks (1, 2, 4, 5, 8); control and purpose (13, 21, 22) and spirituality (3, 9). The scores for each item are added together and it is interpreted that the higher the score in each dimension, the more indicators of resilience the individual shows. Measurement scale 0 - 33 Low resilience, 34 - 66 Medium resilience, 67 - 100 High resilience [16][17]. The Spanish translation was carried out by two members of the research team and supervised by a bilingual person with training in Psychology. Regarding the statistical reliability of the data collection instrument, a coefficient of 0.882 ($\alpha>0.6$) was obtained.

All the collected data will be entered into the matrix designed by the statistical software SPSS 24.0., this procedure was carried out cautiously to not have errors and avoid the values being lost.

1.5. Place and Application of the Instrument

To start the process, the followed steps were taken, making a request document attaching the letter of presentation provided by the nursing professional school of the Universidad de Ciencias y Humanidades, addressed to the director of the educational institution Kumamoto I 3092, located in the department de Lima, Puente Piedra district, in the AA. HH Laderas de Chillón, to access the permit and be able to apply the questionnaires formally, also it coordinated a meeting with the director, in the interview the authors explained in detail about the importance of conducting such research, which was accepted for the director and allowed the study to be carried out.

On July 19, 2019, the informed consent and assent was delivered to all students from the 1st to the 5th grade of secondary education, on July 22 the documents were collected verifying the correct filling, then the authors proceeded with the delivery of the surveys. The time elapsed in each classroom was 25 minutes. They were informed of the importance of their participation, in a brief explanation of the instrument, also it was mentioned that the confidentiality of their responses would be respected, the doubts they presented were also clarified. The students voluntarily answered the questionnaire provided. At the end of the application of the instrument, each of the surveys was verified with the correct completion and they were thanked for their participation in the research.

3. Results

After applying the surveys, the data was processed in a statistical program, resulting in the summary tables presented below.

Table 2 shows the results of the gender of the high school adolescents, where 306 (48.9%) of the adolescents are female and 320 (51.5%) of the adolescents are male.

Table 3 shows the mean age of secondary school adolescents, where the minimum age of adolescents is 11 years and the

maximum age of adolescents is 18 years, therefore the average age of adolescents in primary education is 14 years old.

Table 2. Gender of secondary education adolescents from the Kumamoto I 3092 educational institution, from the Puente Piedra district - Lima, 2019 (N = 626)

Gender of the adolescents			
		Frequency	**Percentage**
Valid	Female	306	48,9
	Male	320	51,1
	Total	626	100,0

Table 3. Age of secondary education adolescents from the Kumamoto I 3092 educational institution, from the Puente Piedra district - Lima, 2019 (N = 626)

Adolescents Age					
	N	**Minimum**	**Maximum**	**Mean**	**SD**
Adolescent Age	626	11	18	14,33	1,485
Valid N (per list)	626				

Table 4: Social Skills in adolescents of Secondary Education of The Kumamoto I 3092 Educational Institution, of The Puente Piedra District - Lima, 2019 (N = 626)

		N	**%**
Social Skills	**Low Level**	76	12,1
	Medium Level	548	87,5
	High Level	2	0,3
Total		626	100,0

In Table 4, the social skills in the adolescents of secondary education of the Kumamoto I 3092Educational Institution, of the Puente Piedra district - Lima, 548 adolescents that represent 87.5% have a medium level, followed by 76 adolescents representing 12.1% have a low level and finally 2 adolescents representing 0.3% have a high level.

Table 5: Resilience in adolescents of Secondary Education of The Kumamoto I 3092 Educational Institution, of The Puente Piedra District - Lima, 2019 (N = 626)

		N	**%**
Resilience	**Low Level**	24	3,8
	Medium Level	365	58,3
	High Level	237	37,9
Total		626	100,0

In Table 5, it can see the resilience of adolescents in secondary education at the Kumamoto I 3092 Educational Institution, in the Puente Piedra district - Lima, where 365 adolescents representing 58.3% have a medium level, 237 adolescents representing 37.9% have a high level and finally 24 adolescents representing 3.8% have a low level.

Table 6: Social Skills and Resilience in adolescents of Secondary Education of The Kumamoto I 3092 Educational Institution, of The Puente Piedra District - Lima, 2019 (N = 626)

			Social Skills			**Total**
			Low Level	**Medium Level**	**High Level**	
Resilience	**Low level**	Count	4	20	0	24
		% Resilience	16,7%	83,3%	0,0%	100,0%
	Medium Level	Count	34	330	1	365
		% Resilience	9,3%	90,4%	0,3%	100,0%
	High Level	Count	38	198	1	237
		% Resilience	16,0%	83,5%	0,4%	100,0%
Total		Count	76	548	2	626
		% Resilience	12,1%	87,5%	0,3%	100,0%
Pearson's chi-square			X^2 Value	g.l.		Sig. (p)
			6,767	4		0,149

In Table 6, the relationship between the variables Social Skills and Resilience was determined based on the Pearson's Chi square (X^2) statistical test. The significance level of the test obtained a value of 0.149 ($p>0.05$) ($X^2 = 6.767$; g.l. = 4). Therefore, the dissociation hypothesis cannot be rejected, for which there is no statistical evidence of a significant relationship between social skills and resilience of a sample of students from the Kumamoto I 3092 Educational Institution, in the Puente Piedra District - Lima. 2019.

4. Discussion

In this research study, the issue of social skills and resilience was raised, promoting adolescent mental health, therefore, it seeks to collaborate with educational institutions that are linked to education as part of the global training of adolescents, where they must incorporate plans or programs that allow the adolescent to strengthen and improve all the capacity to interrelate and be able to face better in daily life and social aspect.

Regarding social skills in adolescents at the secondary level of the Kumamoto I 3092 Educational Institution, in the Puente Piedra district - Lima, the medium level stands out, then the low and finally the high. This study is similar to [10], in which they found medium scores of 50.6% as the predominant average, however, it differs in terms of dimensions, the lowest level most affected is making requests with 59.4% and in our study shows a low score in the expression of anger or disagreement with 34.3%. At the same time, in [12], it presents a percentage of 44.54%, being the predominant medium level, for this reason we agree with the proposal of resorting to development programs for educators and this allows covering topics of social skills within the training of students.

Regarding resilience, it was evident that the medium level predominates, a similar value was found in the study [11], where a greater proportion of the medium level of resilience is evident, represented by 51% followed by 43% of low level and 6% of the high level. At the same time, it shows the average capacity of adolescents to cope with complicated situations. Similarly, in [13], the author determine that the predominant level of resilience was

the medium, and also proposes to carry out interventions with a multidisciplinary vision to comprehensively address adolescents, in the same way, it is evidenced in the study [18], who argue that the resilience in the participants was 17.5% low resilience, 67.1% medium resilience and 15.4% high resilience.

The research study shows that the results obtained indicate the need to carry out an intervention to the adolescent population with social skills and medium-level resilience, otherwise they would present an inadequate development in their personality.

5. Conclusions

Eighty-seven-point five percent of the study participants reported to have medium level of social skills, twelve point one, medium level and cero point three, high level. Regarding social skills according to their dimensions, where the most affected dimension was the expression of anger or disagreement with 215 representing thirty-four-point three percent of adolescents. All students from 1st to 6th grade of secondary level presented medium level resilience with fifty-eight-point three percent, high level with thirty-seven-point nine percent and low level with three-point eight percent.

Regarding resilience according to its dimensions, the most affected dimension was spirituality with 132 representing twenty-one-point one percent of adolescents.

As future work, it is expected that health professionals will have a greater interest in conducting more research, in vulnerable locations, that cover the adolescent stage in a holistic way, physically, emotionally, and socially, where adolescents can develop in an intrapersonal or interpersonal context.

The limitation in the present research work is that some adolescents were not present at the time of conducting the surveys also the interpretation of some questions, so nurse professionals were taking the survey to help school students understanding the questions.

Likewise, educators must assume the responsibility they have regarding the adolescent's physical, affective, psychological-cognitive, social development, guiding them to increase their activities to contribute to the management of good self-esteem, social skills, resilience, among others.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] E. Gismero, Escala de Habilidades Sociales - EHS. Manual. 3a ed. Madrid - España: TEA- Ediciones; 2010. 56 p.

[2] E. Flores, M. Garcia, W. Calsina, A. Yapuchura, "Las Habilidades sociales y la comunicación interpersonal de los estudiantes de la Universidad Nacional del Altiplano – Puno.," Comunicacción, **7**(2), 2016.

[3] J. García del Castillo, Á. García del Castillo, C. López, P. Dias, "Conceptualización teórica de la resiliencia psicosocial y su relación con la salud.," Health and Addictions/Salud y Drogas, **16**(1), 59–68, 2016, doi:10.21134/haaj.v16i1.263.

[4] R. Vera, F. Raquel, L. Tobar, V. María, V. Delgado, "Habilidades Sociales.," Revista Salud y Ciencias, **1**(2), 08–15, 2017.

[5] C. Salavera, P. Usán, P. Teruel, "Contextual problems, emotional intelligence and social skills in Secondary Education students. Gender differences.," Annales Médico-Psychologiques, Revue Psychiatrique., **177**(3), 223–230, 2019.

[6] B. Gunn, E. Feil, J. Seeley, H. Severson, H. Walker, "Promoting School Success: Developing Social Skills and Early Literacy in Head Start Classrooms," NHSA Dialog, **9**(1), 1–11, 2006, doi:10.1207/s19309325nhsa0901_2.

[7] M. Aparicio, Habilidades para la vida. Manual de conceptos básicos para facilitadores y educadores., CEDRO, 1–38, 2013.

[8] Organización Panamerica de la Salud, Manual de identificación y promoción de la resiliencia en niños y adolescentes., 1–90, 1998.

[9] Ministerio de Salud, "Guía técnica de gestión de promoción de la Salud en Instituciones educativas para el Desarrollo Sostenible," 46, 2011.

[10] M. Quispe, "Habilidades sociales y consumo de alcohol en adolescentes de un Colegio Nacional de Lima.," Revista Ciencia y Arte de Enfermería, **2**(2), 43–49, 2017, doi:10.24314/rcae.2017.v2n2.08.

[11] L. Badaracco, "Factores de resiliencia y riesgo suicida en adolescentes en estado de abandono.," Temática Psicológica, **9**(9), 35–42, 2013, doi:10.33539/tematpsicol.2013.n9.838.

[12] A. Gonzales, P. Quispe, "Habilidades sociales y rendimiento Académico de los estudiantes de la Facultad de Ciencias de la Educación - Universidad Nacional del Altiplano - Puno Perú.," Revista de Investigaciones Altoandinas - Journal of High Andean Research, **18**(3), 331–336, 2016, doi:10.18271/ria.2016.222.

[13] E. Choque, H. Matta, "Nivel de resiliencia y funcionalidad familiar en adolescentes de un albergue de Ventanilla , Perú.," Ágora Rev. Cient., **5**(2), 1–7, 2018.

[14] C. Fernández, P. Baptista, Metodología de la Investigación. 6ta ed. México: Mc Graw-Hill/Interamericana., 2015.

[15] H. As, E. Huaire, "Desarrollo de habilidades sociales en contextos universitarios," Horizonte de La Ciencia, **8**(14), 123, 2018, doi:10.26490/uncp.horizonteciencia.2018.14.430.

[16] M. Serrano, M. Garrido, B. Notario, R. Bartolomé, M. Solera, V. Martínez, "Validity of the Connor-Davidson resilience scale (10 items) in a population of elderly | Validez de la escala de resiliencia de Connor-Davidson(10 ítems) en una población de mayores no institucionalizados.," Enfermeria Clinica, **23**(1), 2020, 2013, doi:10.1016/j.enfcli.2012.11.006.

[17] M. Cresp, V. Fernandez, C. Soberon, "Adaptación española de la escala de resiliencia de connor-davidson (cd-risc) en situaciones de estrés crónico," Behavioral Psychology / Psicología Conductual, **22**(2), 219–238, 2014.

[18] R. Singh, S. Mahato, B. Singh, J. Thapa, D. Gartland, "Resilience in nepalese adolescents: Socio- demographic factors associated with low resilience.," Journal of Multidisciplinary Healthcare, **12**, 893–902, 2019, doi:10.2147/JMDH.S226011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1350-1355 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Social Skills and Resilience in Adolescents of an Educational Institution in North Lima, 2019

Lili Sifuentes-Gomez[1], Doris Vega-Davila[1], Betty Flores-Paz[1], Brian Meneses-Claudio[2,*], Hernan Matta-Solis[1], Eduardo Matta-Solis[3]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[3]*Research and Intellectual Creativity Direction, Universidad María Auxiliadora, 15408, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 25 November, 2020
Online: 16 December, 2020

Keywords:
Social Skills
Resilience
Educational institution
Daily Life
Coping

ABSTRACT

Social skills are a set of partially independent and situationally specific verbal or nonverbal responses, in terms of resilience, it is the set of qualities, resources or strengths that favor adolescents to face the adversities of life, therefore the objective of the research work is to determine the social skills and resilience in adolescents of an educational institution in Lima - North. It is a non-experimental study, descriptive of a quantitative approach, it is a cross-sectional correlational research, with a population of 706 students of the secondary level, also with demographic data, and the Social Skills Scale (SSS), and the scale of Resilience (CD-RISC) that will measure the level of resilience of adolescents. The relationship between the variables Social Skills and Resilience. It was determined based on the Pearson's Chi square test (X^2). The significance level of the test obtained a value of 0.000 ($p<0.05$) ($X^2 = 28.626$; g.l. = 4). So, the dissociation hypothesis is rejected and it can be affirmed with statistical evidence that there is a significant relationship between social skills and resilience. The predominant level of social skills is medium with 648 (91.8%) participants and in terms of resilience, the predominant level is high with 366 (51.8%) of participants. Social skills and resilience are mutually related since both allow the adolescent to face situations of daily life looking the ability to develop their coping capacity.

1. Introduction

Social skills or assertive behaviors are understood as a set of partially independent and situationally specific verbal and nonverbal responses, through which adolescents express in an interpersonal their needs, feelings, preferences, right opinions without excessive anxiety and in a non-aversive way, expressing itself in self-expression of social situations, defense of its own rights as a consumer, expression of anger or disagreement, saying no and cutting interactions, initiating positive interactions with the opposite sex and making requests. Social skills are involved in many areas of a person's life, such as adaptation to the environment in which they live and health protection, both physical and mental. It helps to solve various situations such as appropriateness in behavior, decision making, facing difficult situations, assigned essential responsibilities in the development of learning [1].

Regarding resilience, it is the set of qualities, resources or strengths that favor adolescents to face life's adversities, overcome them or even be transformed by it, which is characterized by presenting traits of persistence-tenacity-self-efficiency, control under pressure, adaptation and support environments, control and purpose and spirituality [2].

Social skills and resilience can prevent problems that may affect adolescents in their daily life, where risky conditions such as family problems, drug abuse or abandonment make it evident that resilient factors are looking for a way to excel in risky conditions in an adolescent-friendly way [3], [4].

Therefore, social skills and resilience play an important role in the adolescent since it will allow them to develop on a social and emotional level so that they can adapt to their environment and at the same time they can face adverse situations in their daily life [5].

In Peru, according to the report of the Instituto Nacional de Estadística e Informática, 32.9% of households are made up of

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & bmeneses@uch.edu.pe

girls and boys and/or adolescents, of whom there is at least one child under 18 with a caloric deficit, being higher in rural areas (33.5%) than in urban areas (33.2%) and Metropolitan Lima (32.0%). On the other hand, 41.4% of children under 18 years of age had a problem that affected their health such as a new disease, a relapse or an accident. Health problems occurred more frequently in children and adolescents in the urban area, not including Metropolitan Lima (42.5%), Metropolitan Lima (41.4%) and rural areas (39.8%) [6].

In [7], the authors presented a research work carried out in Colombia, significant differences were shown in emotional intelligence. In the high social acceptance group, low scores were found in the social skills evaluated, being the ability to make requests lower. Likewise, statistically significant correlations were found between several of the factors of the Baron test and those of the Social Skills Scale.

In [5], the authors described a study carried out in Spain, in the sample made up of 844 participants of both sexes, it shows that the most resilient adolescents have better results in all dimensions of quality of life and the effect of resilience is greater in the dimensions related to mental health and in all those that measure social relationships.

In [3], the authors described a study carried out in Colombia, a study was presented regarding the ability to listen, the ability to start a conversation, the ability to hold a conversation and the ability to ask for help, it is evident that most of people obtained scores between high and medium level. On the other hand, in the expression subscale of feelings, it is stated that most men have a medium level, while women are between the medium level and a high level.

The objective of the study is to determine the social skills and resilience in adolescents of an educational institution in North Lima in which it will allow to give important data about the social skills and resilience of adolescent students at the secondary level because nowadays, adolescents tend to have little communication with others in which this behavior cannot help them in a social situation effectively.

The study applied the data collection instrument Social Skills Scale (SSS), it was created by Elena Gismero Gonzales and the Connor-Davidson Resilience Scale (CD-RISC), which has been useful to assess social skills and resilience of adolescents at the secondary level of an educational institution.

The data collection was processed through the survey of adolescents at the secondary level, the data was entered in a data matrix that will be designed in the statistical program SPSS (Statistical Package for the Social Sciences) in its version 24.0 , in which it will allow a better data processing to make statistical tables and graphs so that they can be described and interpreted in results and discussions, respectively.

The following research work is structured as follows: In section II, the development of data collection processing for each adolescent at the secondary level will be presented, as well as the guidelines to consider so that they are included in the research work. In section III, the results will show the social skills and resilience of adolescents at the secondary level of the educational institution according to the specified dimensions of the instruments in the measurement of the variables. In section IV, it presents the discussions of the research work, in section V, the conclusions and in section VI the recommendations as well as the future work that is intended to be reached with the research work.

2. Methodology

In this section, the type and design of the research will be evidenced, as well as the population and sample that will be carried out in the research work, in addition to the inclusion and exclusion criteria and finally the technique and the instrument of data collection.

2.1. Research Type and Design

The present research is non-experimental, descriptive of a quantitative approach, it is a cross-sectional correlational research.

2.2. Population and sample

The population is made up of 706 students from 1st to 5th secondary level of the educational institution 3049 Imperio Del Tahuantinsuyo.

Inclusion Criteria

- Students who normally attend the educational institution.
- Students who have informed consent and assent.
- Students who have the authorization signed by their parents and agree to voluntarily join the study.

Exclusion Criteria

- Students who do not regularly attend to the institution.
- Students who do not have the authorization of their parents.

2.3. Technique and instrument

The technique used is the survey, using the questionnaire or instrument for collecting data on social skills and resilience, which aims to measure social skills and resilience in adolescents of an educational institution in North Lima.

The Social Skills Scale (SSS) was created by Elena Gismero Gonzales, from the Universidad Pontificia de Comillas (Madrid) in 1997, its application is given individually and collectively, with an approximate duration of 10 to 15 minutes, it is focused in adults and adolescents. The social skills scale consists of 33 items grouped into 6 dimensions, self-expression in social situations (8 items), defense of one's rights as a consumer (5 items), expression of anger or disagreement (4 items), saying do not cut off interactions (6 items), make requests (5 items) and initiate positive interactions with the opposite sex (5 items) that are evaluated by a Likert-type scale where “1 = I do not identify myself, in most times it does not happen to me or it does not I would do ”,“ 2 = it has nothing to do with me, even if it happens to me sometime ”,“ 3 = describes me roughly, even if I don't always act that way or feel that way ”and“ 4 = strongly agree, I would feel that way or act that way in most cases". The scale range varies from 33 to 132, the higher the score, the higher the adolescent's social skills [8].

The Connor-Davidson Resilience Scale (CD-RISC) was built by Connor and Davidson in the United States in 2003.

Its application is focus for the adolescent and adult population from 11 to 80 years old, the resilience scale consists of 25 items grouped into 5 dimensions, persistence-tenacity-self-efficacy (8 items), control under pressure (7 items), adaptation and ability to recover (5 items), control and purpose (3 items) and spirituality (2 items) that are evaluated with a Likert-type scale where "0 = never", "1 = rarely", "2 = sometimes", "3 = often" and "4 = almost always". The range of the scale varies from 0 to 100, the higher the score, the higher the resilience in the adolescent [9].

2.4. Place and Application of the Instrument

The survey to measure social skills and resilience in adolescents was carried out at the 3049 educational institution of Imperio del Tahuantinsuyo in North Lima.

The survey was carried out in two consecutive shifts, morning and afternoon according to the schedule of the students of the secondary level, the questionnaire was taken to the secondary level students with a time of 20 minutes to each selected section (1st from high school to 5th grade according to the inclusion criteria), it was concluded with a good satisfaction at the time of collecting the surveys since support from high school students was presented to carry out this research work.

It is very important to emphasize the presence of classroom teachers at the time of completing the surveys, because this helps students to relay on making questions using the sign of upper hand. In addition to the presence of nursing professionals since the mental health of adolescents is very important, because they are the future where the ways of communication tend to be high and that allow them to generate self-confidence and be able to perform effectively to any situation that is happening.



Figure 1: Application of the Survey to the students of the Educational Institution



Figure 2: Nursing staff is supporting in carrying out the survey

In Figure 1, the resolution of survey is shown in a classroom of the educational institution at the secondary level by the students, the collaboration of the students can be seen since this allows interaction with them.

In Figure 2, the reader can see the direct interaction with the adolescent student for the research process, it should be noted that the students gave confidence when explaining about the completion of the surveys.

3. Results

The table 1 is a summary of the surveys carried out following the guidelines corresponding to the research work:

Table 1: Social Skills and Resilience in secondary level students of Independencia Educational Institution – North Lima, 2019 (N = 706)

			Social Skills			Total
			Low level	Mediu m level	High level	
Resilience	Low level	Count	0	16	0	16
		% within Resilienc e	0,0%	100,0%	0,0%	100,0 %
	Mediu m level	Count	6	312	6	324
		% within Resilienc e	1,9%	96,3%	1,9%	100,0 %
	High level	Count	3	320	43	366
		% within Resilienc e	0,8%	87,4%	11,7 %	100,0 %
Total		Count	9	648	49	706
		% within Resilienc e	1,3%	91,8%	6,9%	100,0 %
Pearson's Chi-square			Value X^2	g.l.	Sig. (p)	
			28,626	4	0,000	

The significance level of the test obtained a value of 0.000 ($p < 0.05$) ($X^2 = 28.626$; g.l. = 4). This indicates that there is a significant relationship between social skills and resilience.

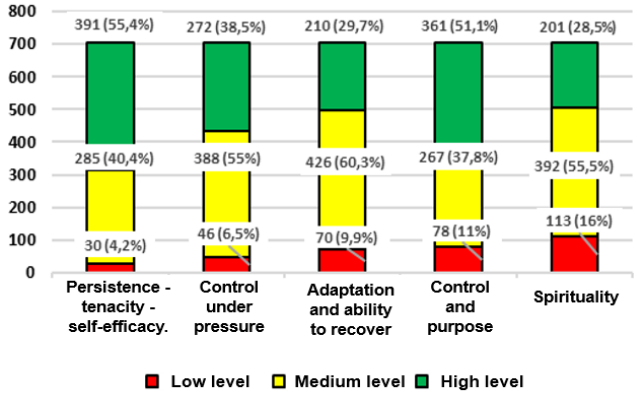


Figure 3: Social skills according to their dimensions in secondary level students in an Educational Institution of Independencia – North Lima, 2019 (N = 706)

In Table 2, it can see the social skills in secondary level students in an Educational Institution of Independencia – North Lima, where the middle level predominates with 648 (91.8%) students.

Table 2: Social Skills in secondary level students in an Educational Institution of Independencia – North Lima, 2019 (N = 706)

		N	%
Social Skills	**Low Level**	9	1,3
	Medium Level	648	91,8
	High Level	49	6,9
Total		706	100,0

In Figure 3, it can observe social skills according to their dimensions, where the most affected dimension was the expression of anger or disagreement with 97 (13.7%) students.

Table 3: Resilience in secondary level students in an Educational Institution of Independencia – North Lima, 2019 (N = 706)

		N	%
Resilience	**Low Level**	16	2,3
	Medium Level	324	45,9
	High Level	366	51,8
Total		706	100,0

In Table 3, it can see the resilience of secondary level students in an Educational Institution of Independence, where the high level predominated with 366 (51.8%) students.

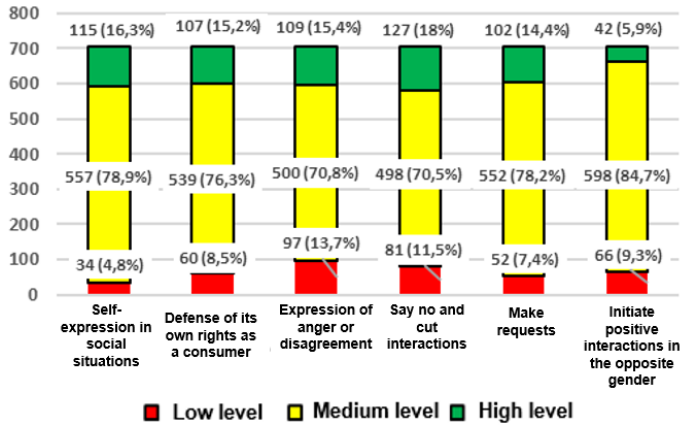


Figure 4: Resilience in secondary level students in an Educational Institution of Independence - North Lima, 2019 (N = 706)

In Figure 4, it can see the resilience according to its dimensions in secondary level students in an Educational Institution of Independence, where the most affected dimension was spirituality with 113 (16%) students.

Both studies can be interpreted that social skills and resilience in adolescents are linked to the promotion of mental and physical health of adolescents, knowing that resilience is relevant because it will allow the person to respond appropriately to negative events or situations. It also focuses on social skills because it is an essential element to achieve the socio-emotional development of the human being and favors school learning. It is vitally important to develop social skills and have a high level of resilience, but unfortunately there is a knowledge gap, reviewing the scientific information related to this topic.

These results will help to the educational institution since it will provide them with information that will allow to evaluate more concisely the adolescents of the secondary level, giving importance to the mental health of the student, in which it is beneficial for them and so they can communicate in an appropriate way that allows to have opportunities in the future.

4. Discussion

In the present study, the issue of social skills and resilience is raised from the point of view of promoting adolescent health, which seeks to contribute to institutions that are in the educational field as part of comprehensive adolescent training, including programs that allow adolescents to enhance their ability to interact, be extrovert, and cope with problems to improve their social life.

Based on the social skills in adolescents of the educational institution 3049 Imperio del Tahuantinsuyo, it prevailed in medium level followed by high and low. These results are based on two purposes: the first, if there are no interventions in the group with a medium level of social skills, these will not have a full development their personality. On the other hand, making interventions can raise the level of social skills, which would be beneficial for adolescent development. In [10], the authors showed that adolescents are normally receptive to new ideas, and are eager to take full advantage of their growing capacity to make decisions; their curiosity and interest lead to a great openness to new possibilities and among them could be considered relating to their peers. Engaging in positive and constructive activities offers opportunities to build relationships with adults and peers. The reason for these findings is due to the fact that in the institution there are no teachers prepared for the tutoring course, they are the teachers of different courses who are in charge of some sections and the participation of parents in the meetings is required.

By implementing well-designed social skills programs, the level of social skills can be raised for the benefit of adolescents and mental health. In these processes of strengthening social skills, the family must also be involved, since this constitutes an important Socialization space and parents play an important role as Trainers and counselors. Conditions must be implemented in schools such as not saturating student classrooms and having the necessary infrastructure for the work of the teacher and tutor; favoring the work of professionals for the benefit of adolescents [11].

Regarding social skills in its self-expression dimension in social situations in adolescents of the educational institution 3049 Imperio del Tahuantinsuyo-Independencia 2019, the medium level prevailed followed by the high and low. It reflects the ability of each individual to express themselves spontaneously in different social and work events. We agree with what was stated by Pérez, who points out in its study that an institutional intervention program can significantly improve social skills in adolescents.

Regarding to social skills in their defense of their rights as a consumer dimension in adolescents of the educational institution 3049 Imperio del Tahuantinsuyo-Independencia 2019, the medium level prevailed followed by the high and the low. Obtaining a high score reflects the expression of assertive behavior towards strangers in defense of their own rights in situations of consumption. Regarding social skills in their dimension defending their own rights as consumers in adolescents from a pre-university center in Lima Norte, the medium level predominated, followed by the low and high. This dimension is related to being an active participant in social life that allows to exercise full citizenship. In [12], the author deduces in his study that the development of social skills and their dimensions, play a protective role for the adolescent, away from risky doings and behaviors such as marijuana use.

Regarding to social skills in its expression of anger or disagreement dimension in adolescents from the educational institution 3049 Imperio del Tahuantinsuyo-Independencia. 2019, the medium level prevailed followed by the high and low. It rates the individual's ability to handle their conflicts with another person in a negative or positive way. In [13], the authors point out from the findings of their study, that the training institution must assume a leading role in promoting social skills; teachers play an important role, so they must be permanently trained to join and contribute to the promotion of social skills in them.

In relation to social skills in it say no and cut interactions dimension in adolescents of the Educational Institution 3049 Imperio del Tahuantinsuyo-Independencia 2019, the medium level prevailed followed by the low and high. This dimension reflects the ability of the individual to cut off interactions that they do not want to maintain, as well as refusing to lend something when dislike doing so, it was found that the vulnerable adolescents who are victims of bullying do not know how to respond to these situations and they have a hard time saying “enough” or “not”.

Respecting of social skills in its making requests dimension in adolescents from the educational institution 3049 Imperio del Tahuantinsuyo-Independencia 2019, the medium level prevailed followed by the low and high. This dimension reflects the ability of the individual to make a request without difficulties, it should be noted that in cooperative learning processes they promote the development of social skills and can enhance the ability to make requests in an appropriate and timely manner.

Regarding social skills in its dimension, initiating positive interactions with the opposite sex in adolescents from the educational institution 3049 Imperio del Tahuantinsuyo-Independencia 2019, the medium level prevailed followed by the low and high [10]. This dimension qualifies the ability of the individual to relate, interact with the opposite sex. This dimension presented a higher percentage in the low level of social skills.

Concerning the level of resilience in adolescents of the educational institution 3049 Imperio del Tahuantinsuyo-Independencia 2019, the high level predominated, followed by a minimal difference from the medium and low level. The results would be explained in different ways, one of the advantages is that the highest percentage lives with their parents, they share family experiences, which favors them as a protective factor so that they can overcome adversities and get out with great ease. The adolescence plays a very important role in resilience since at this stage of life, it goes through various changes such as: biological, psychological and social. This is where most conflicts, rebellions are generated, testing their ability to adapt [12].

In relation to its spiritual dimension, it has a higher percentage in the low level of resilience. The role of spirituality in improving an adolescent's ability to cope with stressful life situations is often under-appreciated, especially in the context of mental health services within a juvenile detention environment.

5. Conclusions

We can conclude that social skills and resilience are mutually related since both allow adolescents to develop functions that allow them to have resources or strengths to cope with situations of daily life.

It is concluded that the social skills are very important for adolescents since it is fundamental for their daily life because they can get success, have proper communication with other people and not only of verbally as well as affectively and emotionally.

It is concluded that resilience is very favorable for adolescents since, within their social environment, they will have the capacity to solve stressful situations and will make them vulnerable to any situation, since resilience will allow adolescents to adapt to these situations and be able to solve them effectively.

6. Recommendations

It is recommended in this institution to have programs that help improve communication between the students themselves so that they can improve their communication with other people.

It is recommended in this institution that it be carried out annually with a control system on social skills and resilience since these are essential for the adolescents so this will allow them to communicate with society and also with their environment.

It is also recommended that health professionals attend educational institutions where they can observe certain deficiencies in social skills and resilience in order to allow them to carry out psycho-affective programs where the adolescent can interact with their social environment and also develop both verbal and non-verbal skills.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] R. Estrada, E. Escoffié, R. García, “Habilidades sociales en adolescentes institucionalizados: Una intervención con juego grupal.,” Psicumex, **6**(2), 27–38, 2016, doi:10.36793/psicumex.v6i2.285.

[2] A. Rodríguez, E. Ramos, I. Ros, A. Fernández, L. Revuelta, “Bienestar subjetivo en la adolescencia: el papel de la resiliencia, el autoconcepto y el apoyo social percibido.,” Suma Psicologica, **23**(1), 60–69, 2016, doi:10.1016/j.sumpsi.2016.02.002.

[3] S. Betancourth, C. Zambrano, A. Ceballos, V. Benavides, N. Villota, “Habilidades sociales relacionadas con el proceso de comunicación en una muestra de adolescentes.,” Psicoespacios, **11**(18), 133, 2017, doi:10.25057/21452776.898.

[4] C. Huang, Y. Chen, S. Cheung, L. Greene, S. Lu, “Resilience, emotional problems, and behavioural problems of adolescents in China: Roles of mindfulness and life skills.,” Health and Social Care in the Community, 1–

9, 2019, doi:10.1111/hsc.12753.
[5] M. Simón, R. Fuentes, M. Garrido, M. Serrano, E. Larrañaga, S. Yubero, "Influencia de la resiliencia sobre la calidad de vida relacionada con la salud en adolescentes.," Enfermería Clínica, **28**(5), 283–291, 2018, doi:10.1016/j.enfcli.2018.06.003.
[6] Instituto Nacional de Estadistica e Informatica, Estado de la Niñez y Adolescencia., 2018.
[7] C. Coronel, M. Levin, S. Mejail, "Las habilidades sociales en adolescentes tempranos de diferentes contextos socioeconómicos.," Electronic Journal of Research in Educational Psychology, **9**(1), 241–261, 2011.
[8] E. Miranda, E. Riquelme, H. Cifuentes, "Revista Latinoamericana de Psicología Análisis factorial confirmatorio de la Escala de habilidades sociales en universitarios chilenos.," Child Psychiatry & Human Development, **41**(2), 73–82, 2010.
[9] G. Guihard, L. Deumier, B. Alliot-Licht, L. Bouton-Kelly, C. Michaut, F. Quilliot, "Psychometric validation of the French version of the Connor-Davidson Resilience Scale.," Encephale, **44**(1), 40–45, 2018, doi:10.1016/j.encep.2017.06.002.
[10] M. Valera, L. Buil, E. Rigo, A. Casero, E. Aguilar, "Habilidades sociales en preadolescentes con trastorno específico del lenguaje.," Revista de Logopedia, Foniatria y Audiologia, **36**(2), 55–63, 2016, doi:10.1016/j.rlfa.2015.03.002.
[11] E. Maguiña, R. Mino, Facultad De Ciencias De La Salud Escuela Profesional De Enfermería., 2018.
[12] F. Morales, "Relaciones entre afrontamiento del estrés cotidiano, autoconcepto, habilidades sociales e inteligencia emocional.," European Journal of Education and Psychology, **10**, 41–48, 2017, doi:10.1016/j.ejeps.2017.04.001.
[13] A. Gonzales, P. Quispe, "Habilidades sociales y rendimiento Académico de los estudiantes de la Facultad de Ciencias de la Educación - Universidad Nacional del Altiplano - Puno Perú.," Revista de Investigaciones Altoandinas - Journal of High Andean Research, **18**(3), 331–336, 2016, doi:10.18271/ria.2016.222.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 1038-1045 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Closed Loop Capacitive Accelerometer Model using Simple Regression Test for Linearity

Mamudu Hamidu[1,*], Jerry John Kponyo[2]

[1]*Kumasi Technical University, Electrical/Electronic Engineering, Kumasi, 00233, Ghana*

[2]*Kwame Nkrumah University of Science and Technology, Electrical/Electronic Engineering, Institute, Kumasi, 00233, Ghana*

ARTICLE INFO

Article history:
Received: 02 September, 2020
Accepted: 02 March, 2021
Online: 22 April, 2021

Keywords:
Capacitive accelerometer
PVA model
PI-controller
Modelling
Closed loop

ABSTRACT

This article extends novelty of modeling capacitive accelerometer with PID controller to provide PI controller for better tuning and statistical test to determine the linear validation characteristics of closed capacitive accelerometer. Capacitive accelerometer is a sensor which uses the dynamic law of physics model Position-Velocity-Accelerator (PVA) by the movement of an electrode coupled to mass proof sandwich between parallel plates to detect vehicle/object displacement. The modelling of closed loop system helps to mitigate steady state error accumulation of measurements in open loop model. The accelerometer gives linear time dependence on output displacement after an input step-like function of acceleration. The closed model can predict the desired output signal. The linearity of the model is tested statistically using simple regression of 120 dataset. This shows a p-value of 2e-16 indicating that at any time, the acceleration predicts the displacement/position of vehicle/object.

1. Introduction

Accelerometer is a microelectromechanical sensor (MEMS) which is generally used in the measurement of moving body's acceleration[1]. MEMS is widely used across the current settings of interconnected world in our daily lives [2]. The obtained acceleration is measured in gravity (g) [3]. The measurements of g can be transformed into velocity and position/displacement.

In the application of accelerometer, it provides a wide range of services in a lot of technologies aiding the prediction of most human and no-human activities. In [4], accelerometer is used on a bridge in Nottingham Wilford in the UK with the main purpose of determining the dynamic characteristics. This provided real-time information on suspended deck movement. It is applied using triaxial accelerometer and Real-time Kinematic Global position system (RTK-GPS). The inclusion of the RTK-GPS is applied to measure the low-frequency vibrations of medium span suspension bridge. The reason for this inclusion is to aid the accelerometer which measures based on high-frequency rate. Similar bridge technique is used but in this domain it is applied in bridge management system and land safety [5]. The only component which changed is the replacement of RTK-GPS with three Leica-GPS.

In the determination of human kinematics, accelerometer places an exceptional role as well. According to [6], the authors presented triaxial accelerometer with a specific value of CDXL04M3 is applied in the recognition of classification of problem during people activities. Also in [7], the article provides a description in the development real-time classification persons activities such as: walking, standing and running using wearable device. The only differentiation this application is the system composition where a two-accelerometer (ADXL202JE) layer model that integrates multiple component Gaussian mixture model with Markov models is applied. This is to accurately classify the zone of user state.

Accelerometer application in human context [8, 9] applied triaxial accelerometer (ActiGraph GT3X+) to validate the process of determining everyday human physical activities. The device is placed at the thigh and hip to identify: cycling, walking, climbing of stairs, running, sitting, and no movements. In [10], authors worked on the determination and understanding of physical activities and key behavioral changes among children. This led to the implemented four different accelerometers namely: ActiGraph, Actical, Actiwatch, and RT3 Triaxial devices. This produces a regression model. According to [11] and [12], triaxial accelerometer and Wireless Wearable Multisensory Integrated Measurement System (WIMS) were implemented together to measure activities such as human heartbeats, hip motion, and other sensors. Additionally, in[13], the author developed a wearable

*Corresponding Author: Mamudu Hamidu, Email: drsennet@gmail.com

https://dx.doi.org/10.25046/aj0602118

wireless base displacement system which utilizes inertial measurement unit (IMU) having triaxial accelerometer onboard. This system monitors the daily indoor activities of humans in buildings such as health statuses and locations.

In other applications such as vehicular domain, the triaxial has received enormous integrated application. Especially, the use of tracking of vehicle velocity and position coupled with other MEMS sensors [14].

There is another form of accelerometer called uniaxial or single axial which used in predicting the behavior of person in motion [15]-[20]. In [19], the authors presented ideas of applying 2 to 3 uniaxial sensors mounted on human body for behavioral investigations. All uniaxial accelerometers were in 2-dimensions. According to [20], the author comparison between uniaxial and triaxial accelerometer application in human activities, showed no difference in the results obtained.

However, in this paper the research design looks at the implementation of the triaxial accelerometer since it is in a 3D form associated in measuring X, Y and Z of the acceleration of vehicle dynamics.

Velocity and position are computed mathematically from acceleration. The position represents the displacement from the mechanical model. This displacement measurement very small. To achieve a better measurement, and electronic readout unit is attached. There are various types of mechanical models. These include: capacitive [21]-[31], piezo-resistive [32], and thermal [33], [34].

The choice of capacitive accelerometer in vehicle tracking is mainly based on its advantages such as:

- Zero static biasing with high sensitivity ability, and better thermal stability. According to [35], these characteristics make capacitive accelerometer model preferred to others.
- It has the edge over other due to size, less expensive, and more flexible in interfacing with other circuitry [36].
- Low temperature and better linearity [37]

However, the capacitive accelerometer has the challenge of high capacitance variation sometimes due the type of mass proof, low damping factor, high sensitivity providing higher resolution that give higher frequencies [38].

A triaxial capacitive accelerometer IC module is integrated with GPS module used in determining vehicle position. The integrations provides a uniform sensitivity which is distributed in the 3D of the vehicle position in the x-y-z domain. [39], [40]. The accelerometer used is made up of glass-silicon which serves as a mass proof. This glass is suspended to the chassis of the vehicle with four locks [41]. This detects parallel motion in z-acceleration and measurements are made in X or Y axis acceleration during tilting of the proof mass. Displacement is detected when electrodes connected to the mass proof which is then sandwich between parallel plates. By this mechanism, the acceleration is measured in the X, Y, and Z axis (a_x, a_y, a_z) [9], [3].

It is based on this that the velocity and position cand be obtained using Position-Velocity-Acceleration (PVA) model which Obeys Newton's law of physical object in motion and that of Hooke's law [31]. the principle underlining PVA is the first and second integral of acceleration model to produce velocity and position respectively [42]. Therefore, for any measurements of acceleration should provide a proportional outputs of velocity and position of a body in motion with an accelerometer [40], [43], [44].

This article looks at modelling the linear characteristics of capacitive accelerometer using PI-controller to determine the linear dependence relationship between acceleration and position. It also verifies the model for statistical significance using simple regression model. The model is based on mathematical modelling simulations using MATLAB/Simulink.

2. Method of analysis

The method of design is by mathematical modelling of system dynamics of a moving mass proof using law of physics. Figure 1 shows an eight stage methodological view of analyzing and simulating of the capacitive accelerometer.

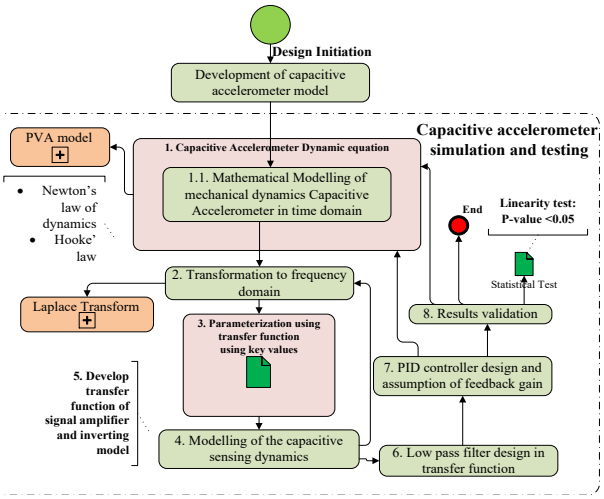


Figure 1: Methodological view of modelling a capacitive accelerometer

- Sensing dynamics: is the model for the change in capacitance during the movement of object (vehicle) [30].
- Model for capacitive sensing dynamics: is the model for the change in capacitance during the movement of object (vehicle) [30].
- Develop transfer function of signal amplifier-inverting model: the output from 4 is amplified using OpAmp and inverted to obtain the original signal. These are all handled in frequency domain for easy computation.
- Low pass filter transfer function: this filters high frequency measurements associated with acceleration measurements. LPF will allow only the low frequency for better measurements of position variables.
- PI controller design and assumption of feedback gain: this stage is done by developing a controller gain matrix and a feedback gain.

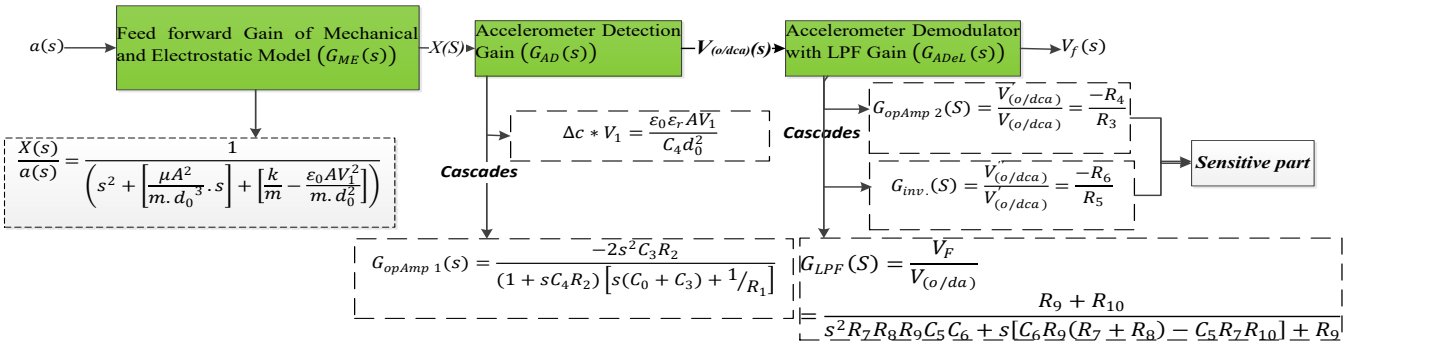


Figure 2: open loop accelerometer system with readout components[50].

- Results validation: to validate the results from the model, data collected is tested using simple regression model for input to output linear dependency using the P-value as a criteria of validity.

3. Open loop accelerometer system

The linear time variant (LTV) model of the mechanical model consisting of mass, damper and spring is represented in (1) [45]-[48]:

$$\left.\begin{aligned} m\frac{d^2x}{dt^2} &= kx + b\frac{dx}{dt} + m\frac{dx^2}{dt^2} \\ b &= \frac{1}{2}\mu A^2\left[\frac{1}{(d_0-d_x)^3}+\frac{1}{(d_0+d_x)^3}\right] \\ \frac{d^2y}{dt^2} &= \frac{d^2y}{dt^2} - \left[\frac{k}{m}x + \frac{1}{2}\mu A^2\left[\frac{1}{(d_0-x)^3}+\frac{1}{(d_0+x)^3}\right].\frac{dx}{dt}\right] + \frac{\varepsilon_0 AV_1^2}{4m}.x \\ a &= \frac{d^2y}{dt^2} \\ a &= \frac{d^2x}{dt^2} + \left[\frac{\mu A^2}{{d_0}^3}.\frac{dx}{dt}\right] + x.\left[\frac{k}{m} - \frac{\varepsilon_0 AV_1^2}{4m}\right] \end{aligned}\right\} (1)$$

where:

b : damping coefficient.
μ: Air gap displacement
A : Mobile plaque area
$(d_0,\ d_x)$: Distance of electrode displacement between parallel plates.

The Laplace transfer of (1) is provided in (2):

$$\left.\begin{aligned} s^2.X(s) + \left[\frac{\mu A^2}{{d_0}^3}.s\right].X(s) + X(s).\left[k - \frac{\varepsilon_0 AV_1^2}{d_0^2}\right] &= a(s).m \\ G_{ME}(s) = \frac{X(s)}{a(s)} &= \frac{1}{\left(s^2+\left[\frac{\mu A^2}{m.{d_0}^3}.s\right]+\left[\frac{k}{m}-\frac{\varepsilon_0 AV_1^2}{m.d_0^2}\right]\right)} \end{aligned}\right\} (2)$$

$G_{ME}(s)$: Gain of Mechanical module

In figure 2, block diagrams of various units of the accelerometer and their models for open loop system is represented. the diagram has the mechanical unit, electronic readout circuit which consist of gain detection the demodulation unit with low pass filter (LPF) [19], [45]-[47], [49].

The following represent key nomenclature in Figure 2:
(G_{AD}) : Gain Detection
$G_{opAmp1}(s)$: signal Detection unit
$G_{ADeL}(s)$: Demodulation and LPF Gain
$G_{opAmp2}(S)$: Electronic Read-out Operational Amplifier Gain
$G_{inv.}(S)$: Electronic Read-out inverter Gain
$G_{LPF}(S)$: Electronic Read-out Low Pass Filter Gain
$V_F(s)$: output displacement of x

Accelerometer detection block is the block which electronically detects the displacement of the mass proof with a rod to dangle between the parallel plate. This unit produces a gain, $G_{AD}(s)$. This gain consists of the electrostatic force of parallel plate and the operational amplifier. The operational amplifier helps boast the electrical signal detected from the mechanical displacement for processing [34].

$V_{(o/dca)}$ in Figure 2 is passed through a Accelerometer Demodulation unit to produce $V_f(s)$.

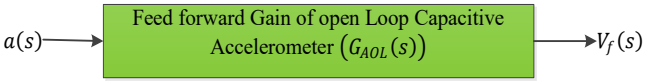


Figure 3: The open loop gain $(G_{AOL}(s))$ block diagram

Figure 3 represents overall open loo transfer function, $(G_{AOL}(s))$ which consists of the cascade of $(G_{ME}(s)), (G_{AD}(s))$ and $(G_{ADeL}(s))$.

$$G_{AOL}(s) = (G_{ME}(s)) * (G_{AD}(s)) * (G_{ADeL}(s)) \quad (3)$$

The open loop function is represented in (4) using the models from Figure 2:

$$\Rightarrow G_{AOL}(s) = (V_f(s))/\ a(s);\ \ V_f(s) = G_{AOL}(s)a(s) \quad (4)$$

By inputting all functions into (4), (5) is obtained:

$$\left.\begin{aligned} V_F(s) \equiv V_{(o/dca)}(s)G_{ADeL}(s) &= \left(V_{(o/dca)}(s)\right) * \\ &\left(G_{inv.}(s)G_{opAmp2}(s)\right)(R_9+R_{10}) \\ V_{(o/dca)} &= \frac{1}{T}\left[\int_0^{\frac{T}{2}}(V_1(t))dt - \int_{\frac{T}{2}}^{T}(V_1(t))dt\right][\Delta Cx] \\ V_1(t) &= V_1\sin(\omega t) \\ = \frac{1}{T}\left[\int_0^{\frac{T}{2}}(V_1\sin(\omega t))dt - \int_{\frac{T}{2}}^{T}(V_1\sin(\omega t))dt\right]&\left[\frac{\varepsilon_0\varepsilon_r A}{C_4 d_0^2}x\right] \\ solving, V_{(o/dca)} &\gg \frac{4V_1}{T\omega}\left[\frac{\varepsilon_0\varepsilon_r A}{C_4 d_0^2}x\right] \\ \therefore V_f(s) = \frac{4V_1}{T\omega}\left[\frac{\varepsilon_0\varepsilon_r A}{C_4 d_0^2}x\right]&\left[\left(G_{inv.}(s)G_{opAmp2}(s)\right) * (R_9+R_{10})\right] \\ Finally;\ V_F(s) &= \\ \frac{4V_1}{\pi}\left[\frac{\varepsilon_0\varepsilon_r A}{C_4 d_0^2}x\right]&\left[\left(G_{inv.}(s)G_{opAmp2}(s)\right)(R_9+R_{10})\right] \end{aligned}\right\} (5)$$

The open loop of the capacitive accelerometer in (5) is now simplified into two main function blocks, thus the mechanical unit and the electronic readout unit in Figure 4.

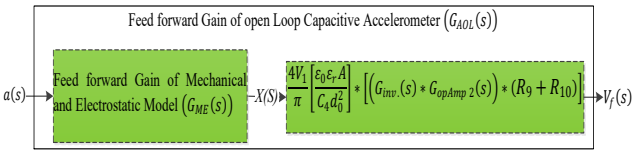


Figure 4: Mechanical and electronic readout of accelerometer

The modeling of the two blocks is represented in (6):

$$\left.\begin{aligned} G_{AOL}(s) &= \frac{G_{ocf}{}^{1}}{\left(s^2 + \left[\frac{\mu A^2}{m.\,d_0{}^3}.s\right] + \left[\frac{k}{m} - \frac{\varepsilon_0 A V_1^2}{m.\,d_0^2}\right]\right)} \\ G_{AOL}(s) &= \frac{7.436e10}{s^2 + 316.1\,s + 1.016e0\,4} \end{aligned}\right\} \quad (6)$$

The modeling of (6) into a closed loop form is discussed in the next section using Proportional-Integral (PI) controller.

4. Closed loop accelerometer modeling

The design of a closed loop accelerometer is considered in this section by designing the feedback and control gain matrices of PI controller models.

4.1. Designing of feedback model

Feedback model is designed to help compare the acceleration (reference signal) to the position (output signal). Base on the feedback a suitable controller can be tuned to achieve a desirable output.

4.2. Choosing Feedback signal

According to [51], feedback model is based the type of sensor which provides a proportional signal to the dynamic model (mechanical unit). Therefore, in choosing the feedback signal, a balancing force is needed to check the inertial forces mechanical unit of the accelerometer. This balancing force is called the electrostatic force. It is chosen because it establishes the reading of small displacement from the proof mass between the two fixed electrodes in parallel plates[23]. The electrostatic force is used to provide restoration of sensing force to balance acceleration force. Figure 5 shows the representation of the accelerometer closed loop structure.

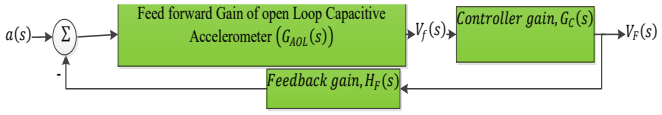


Figure 5: closed loop structure of accelerometer

A bias signal V_b is introduced with excitation signal v_1 to obtain a feedback signal V_F to the two electrodes. This generates the V_{ne} and V_{pe} as negative electrode and positive electrode supplies respectively. This is represented in (7) as:

$$\left.\begin{aligned} V_{ne} &= v_1 - V_b + V_F \\ V_{pe} &= v_2 + V_b + V_F \end{aligned}\right\} \;\&\; \left.\begin{aligned} v_1 &= V_1 \sin\omega t \\ v_2 &= -V_1 \sin\omega t \end{aligned}\right\} \quad (7)$$

The resultant signal gives (8):

$$V_{npe} = V_{ne} - V_{pe} \quad (8)$$

The application of the electrostatic force feedback gives (9):

$$\left.\begin{aligned} F_{el_F} &= F_{el_{ne}} + F_{el_{pe}} \Big\} F_{el_F} \\ &= \frac{\varepsilon_0\varepsilon_r A}{4}.\left[\frac{V_{ne}^2}{(d_0 - d_x)^2} - \frac{V_{pe}^2}{(d_0 + d_x)^2}\right] \\ F_{el_F} &= \frac{\varepsilon_0\varepsilon_r A}{4}.\left[\frac{(V_1 sin\omega t - V_b + V_F)^2}{(d_0 - d_x)^2} - \frac{(-V_1 sin\omega t + V_b + V_F)^2}{(d_0 + d_x)^2}\right] \end{aligned}\right\} \quad (9)$$

The accelerometer having high system frequency and considering the initial position of the proof mass, the electrostatic force feedback mean signal is represented mathematically in (10):

$$F_{el_F} = \frac{1}{T}\int F_{el_F} dt = \frac{2\varepsilon_0\varepsilon_r A}{d_0{}^3}.\int [x(V_1 \sin\omega t - V_b + V_F)^2 - (-V_1 \sin\omega t + V_b + V_F)^2]\, dt \quad (10)$$

Expanding and integrating all terms with respect to t gives (11):

$$\left.\begin{aligned} F_{el_F} &= \frac{2\varepsilon_0\varepsilon_r A}{d_0{}^3}\left[x\left(\frac{V_1^2}{2} + V_b^2 + V_F^2\right) - V_b V_F d_0\right]; \\ &\text{with inertial displacement of } x = 0 \\ &\therefore \text{ the feedback electrostatic force,} \\ F_{el_F} &= -\frac{2\varepsilon_0\varepsilon_r A V_b V_F}{d_0{}^2} \end{aligned}\right\} \quad (11)$$

Rewriting the equations again for all the acting forces on the proof mass the feedback sensor gives (12):

$$\left.\begin{aligned} \vec{F}_i &= \vec{F}_e + \vec{F}_a + \vec{a}_x - F_{el_F} \\ ma &= kx + \frac{1}{2}\mu A^2\left[\frac{1}{d_0{}^3} + \frac{1}{d_0{}^3}\right]\frac{dx}{dt} + m\frac{d^2x}{dt^2} + \frac{2\varepsilon_0\varepsilon_r A V_b V_F}{d_0{}^2} \\ ma &= kx + \left[\frac{\mu A^2}{d_0{}^3}\right]\frac{dx}{dt} + m\frac{d^2x}{dt^2} + \frac{2\varepsilon_0\varepsilon_r A V_b V_F}{d_0{}^2} \\ a &= \frac{k}{m}x + \frac{1}{m}\left[\frac{\mu A^2}{d_0{}^3}\right]\frac{dx}{dt} + \frac{d^2x}{dt^2} + \frac{1}{m}.\frac{2\varepsilon_0\varepsilon_r A V_b V_F}{d_0{}^2} \end{aligned}\right\} \quad (12)$$

The model for feedforward including the controller in figure 5 represented in (13):

$$V_F(s) = V_f(s) * G_C(s)$$

[1] G_{ocf}: open loop cascade factor $= \left[\frac{4V_1}{\pi}\left(\frac{\varepsilon_0\varepsilon_r A}{C_4 d_0^2}\right)\right] * \left[\left(G_{inv.}(s) * G_{opAmp2}(s)\right) * (R_9 + R_{10})\right]$

$$\Rightarrow V_F(s) = \left(G_{AOL}(s)\right) * \left(G_C(s)\right) \quad (13)$$

∴the close Accelerometer Loop Transfer function $\left(G_{ACL}(s)\right)$ is in (14):

$$\left.\begin{aligned} G_{ACL}(s) &= \frac{V_F(s)}{a(s)} \\ &= \frac{\left(G_{AOL}(s)\right) * \left(G_C(s)\right)}{1 + \left(G_{AOL}(s)\right) * \left(G_C(s)\right) * \left(H_F(s)\right)} \\ G_{ACL}(s) &= \frac{\left[\frac{7.436e10}{s^2 + 316.1\,s + 1.016e0\,4}\right]^2 * \left(G_C(s)\right)}{1 + \left[\frac{7.436e10}{s^2 + 316.1\,s + 1.016e0\,4}\right] * \left(H_F(s)\right)} \\ G_{ACL}(s) &= \\ &\frac{(7.436e10) * \left(G_C(s)\right)}{denum^3} \end{aligned}\right\} (14)$$

5. Closed loop accelerometer simulation and results

5.1. Open loop accelerometer model and characteristics

The open loop model of accelerometer is represented by the mechanical-charge detection model and the electronic demodulator model in Figure 6 and 7, respectively.

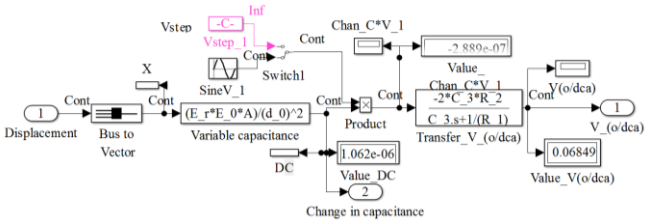


Figure 6: Acceleration Charge Detection unit

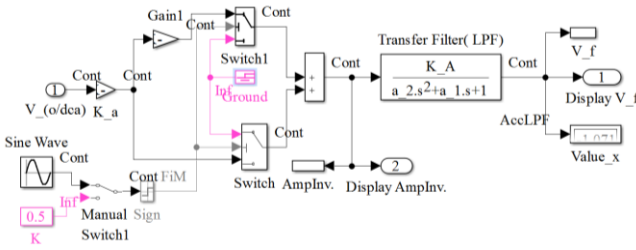


Figure 7: Demodulator Electronic processor

The mechanical, charge detection, and electronic demodulator transfer function is given in (15):

$$G_{opl}(s) = \frac{7.436e10}{7.29e8\,s^4 + 5.13\,s^3 + 11620s^2 + 3.213e6\,s + 1.16e8} \quad (15)$$

The open loop characteristics are shown in Figure 8 is based on Figure 6 and 7. For an input acceleration of 1g the model provides Mechanical displacement response of 1.969×10^{-04} and the respective charge detection unit and low pass filter output is displayed as well.

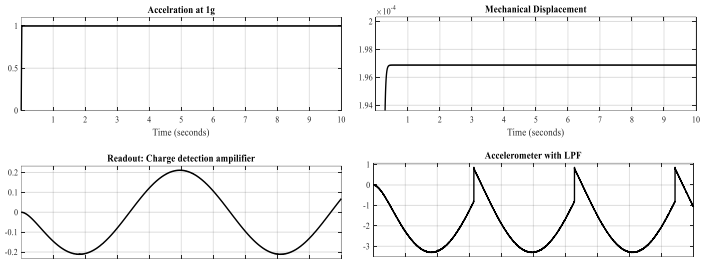


Figure 8: open loop characteristics of accelerometer model.

The overall transfer function of the open loop model in (15) produces an overshoot of about 641 with an input-output response rise time of $0.0528s$ as shown in Figure 9.

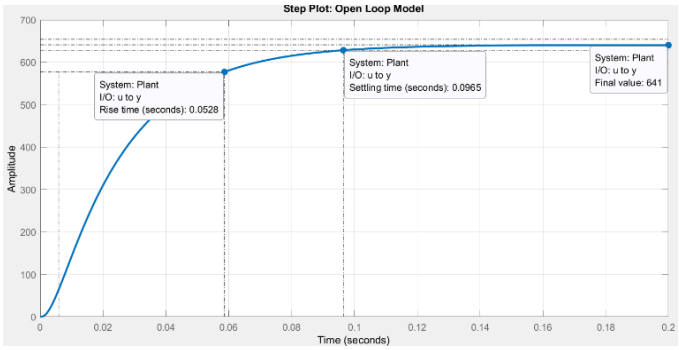


Figure 9: overall open loop transfer function characteristics of accelerometer model.

In Figure 10, the open loop bode plot shows system gain of 56.1 dB which was constant until it reached a frequency response is 791 rad/s ($\omega = 791$).

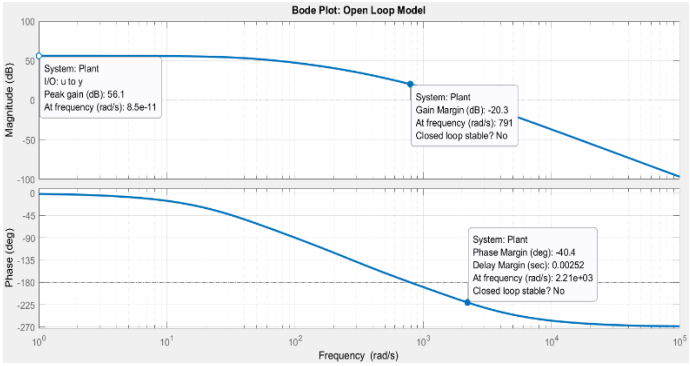


Figure 10: the magnitude and phase of the open loop accelerometer transfer function $G_{opl}(s)$ against angular frequency.

The open loop gain gives the open loop minimum stability margins at $Gain = -20.3\ dB$ and frequency= $791 rad/s$.

[2] Feed forward Gain of open Loop Capacitive Accelerometer $\left(G_{AOL}(s)\right)$with electronic readout circuit.

[3] $denum = [s^2 + 316.1\,s + 1.016e0\,4] + (7.436e10) * \left(G_C(s)\right) * \left(H_F(s)\right)$

For the phase, at low frequency, the phase is 0^o and at high frequency is -268^o. The minimum stability margins are -40.4^0against frequency of $2.21krad/s$.

5.2. Accelerometer closed loop model and characteristics

This section looks at the modelling of the close system of the accelerometer. The system uses proportional integral (PI) controller to establish the stability of the open loop gain. The gain of the controller is represented in (16):

$$G_C(s) = \frac{K_p s + K_i}{s} \tag{16}$$

The control gain matrices are represented by:

$$\begin{bmatrix} K_p \\ K_i \end{bmatrix} = \begin{bmatrix} 0.002369 \\ 0.1613 \end{bmatrix} \tag{17}$$

Considering (15) and (116) with a feedback of 1 gives the closed loop gain in (18):

$$G_{CLp}(s) = \frac{K_1 s + K_{cLp}}{K_2\, s^5 + K_3\, s^4 + K_4\, s^3 + K_5\, s^2 + K_6 s + K_{cLp}} \tag{18}$$

where:
$K_1 = 1.027 \times 10^8$, $K_{cLp} = 2.87 \times 10^9$, $K_2 = 7.29e \times 10^8$, $K_3 = 5.13$, $K_4 = 1.162 \times 10^4$, $K_5 = 3.213 \times 10^6$, and $K_6 = 2.043 \times 10^8$

The overall closed loop transfer function gain in (18) is represented in Figure 11.

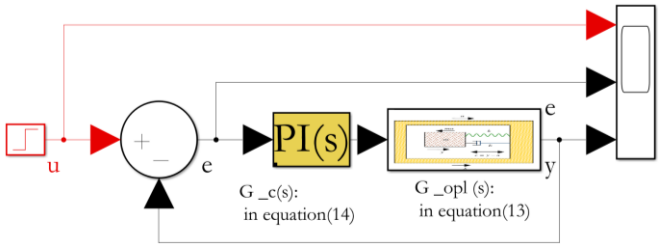


Figure 11: closed loop accelerometer model in MATLAB/SIMULINK.

In Figure 12, the response time is $0.0132s$ at amplitude of 0.9 with an overshoot of 10.4% at $0.04s$ and an amplitude of 1.1 the model shows a settling in at $0.0696s$.

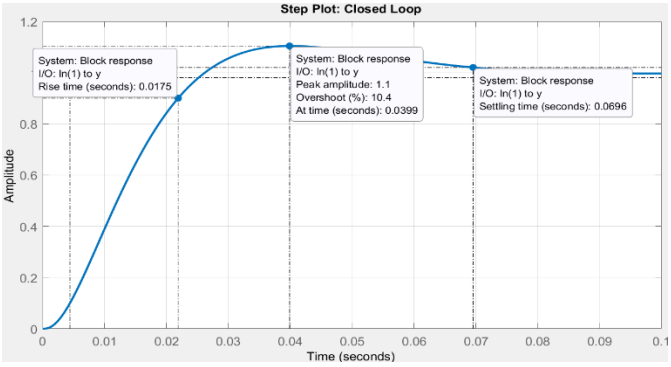


Figure 12: Closed loop accelerometer characteristics with PI controller.

Figure 13 represents the Bode plot for closed loop gain which gives a minimum stability margin at $Gain = -29.5dB$ and $frequency = 689rad/s$.

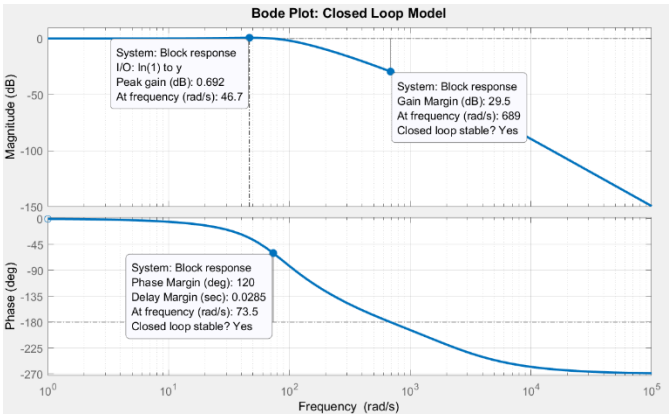


Figure 13: Bode plot for closed loop accelerometer.

The minimum stability margins are -120^0against frequency of $73.5rad/s$ having a delay margin of $0.0285s$.

6. Statistical Test on Accelerometer Linearity

In this section, a linearity test is applied on the acceleration model. The test uses simple regression and p-value to establish linearity of the model.

A simple regression model is developed for the accelerometer model using the algorithm in (19):

$$x = b + w * a \tag{19}$$

where:
Dependent variable: Displacement (x)
Independent variable: Accelerometer (a)
b: $Bias\ term$
w: $Weight$

By building a simple regression model using sample test run of 120 datasets, the model for the accelerometer is obtained statistically, (20) and the figure 14 shows the linear relationship.

$$x = 9.224\text{e} - 08 + 1.00 * a \tag{20}$$

where: $b = 9.224e - 08$, and $w = 1.00$

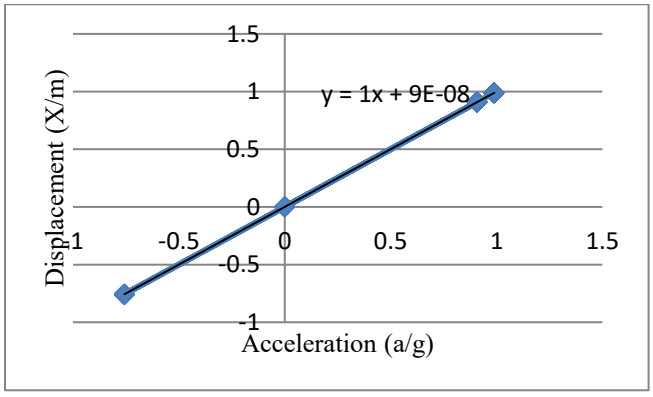


Figure 14: Linearity test of closed loop accelerometer.

In Figure 14, it is shown that the acceleration and the displacement are linearly dependent. The statistical test of p-value is below 0.05. The model's p-value is 641, which means the acceleration can predict the displacement.

7. Conclusion

This paper looks at the modelling and simulation of closed loop capacitive accelerometer. The model applied is the Position-Velocity-Acceleration (PVA) which consist of mechanical unit, charge detection unit, and signal demodulator with filter unit.

The results from the open loop after applying acceleration of $1g$ shows an overshoot to an amplitude of 641. However, the by controlling the accelerometer in in a closed loop form with a Proportional-Integral (PI) control gain, the desirable output is obtained.

The linearity of the model is with simple regression which shows it is predictable with a statistical significance of $2e-16$ as its p-value. However, for better optimization, the accelerometer model must be modelled in a state space domain using Gaussian stochastic model where noise is factored as independent of system memory. This would also provide a better technique for better prediction of position. In another vein, future model can consider temperature as a component of damping coefficient (b).

Conflict of Interest

The authors of this manuscript have seen and agree with the contents. We declare that this work is the original work submitted for publication and has undergone peer review process in no other journal other than this.

References

[1] S. Yan, H. Liu, Q. Xu, D. Liu, M. Zhang, W. Wu, L. Tu, "A method for improving out-of-plane robustness of an area-changed capacitive displacement transducer used in a micro-accelerometer," Sensors and Actuators, A: Physical, **312**(May), 112156, 2020, doi:10.1016/j.sna.2020.112156.

[2] G. Lammel, "The future of MEMS sensors in our connected world," Proceedings of the IEEE International Conference on Micro Electro Mechanical Systems (MEMS), **2015-Febru**(February), 61–64, 2015, doi:10.1109/MEMSYS.2015.7050886.

[3] M. Devi, B.A. Kumar, M.T. Scholar, "Accelerometer Based Direction Controlled Wheelchair Using Gesture Technology," **1070**(3), 1065–1070, 2014.

[4] M. Meo, G. Zumpano, X. Meng, E. Cosser, G. Roberts, A. Dodson, "Mechanical Systems and Signal Processing Measurements of dynamic properties of a medium span suspension bridge by using the wavelet transforms," Mechanical Systems and Signal Processing, **20**, 1112–1133, 2006, doi:10.1016/j.ymssp.2004.09.008.

[5] X. Meng, Real-time Deformation Monitoring of Bridges Using GPS/Accelerometers, Jul. 2020.

[6] N. Ravi, N. Dandekar, P. Mysore, M.L. Littman, "Activity recognition from accelerometer data," in Proceedings of the National Conference on Artificial Intelligence, 1541–1546, 2005.

[7] R.W. DeVaul, S. Dunn, "Real-time motion classification for wearable computing applications," 2001, Project Paper, Http://Www. Media. Mit. Edu/Wearables/Mithril/Realtime. Pdf, 1–14, 2001.

[8] M. Korshøj, J. Kristiansen, C. Hanisch, A. Holtermann, E.U.K. Online, "Key findings," Article in Journal of Physical Activity and Health, (May), 2035, 2010, doi:10.1016/0267-3649(88)90030-1.

[9] M. Korshøj, J. Kristiansen, C. Hanisch, A. Holtermann, "Detection of Physical Activity Types Using Triaxial Accelerometers," Article in Journal of Physical Activity and Health, 2012, doi:10.1123/jpah.2011-0347.

[10] P. Freedson, D. Pober, K.F. Janz, "Calibration of Accelerometer Output for Children," Med. Sci. Sports Exerc, **37**(11), 523–530, 2005, doi:10.1249/01.mss.0000185658.28284.ba.

[11] S.S.M.E. Assistant, B. Birosekhan, G. Chandru, K. Naveen, K. Praveen, "Predicting Human Activities Using Accelerometer And Sensors," **3**(1), 3743–3747, 2014.

[12] D. Dissertation, LOW-POWER FRONT-ENDS FOR CAPACITIVE THREE-AXIS ACCELEROMETERS Doctoral Dissertation, 2010.

[13] M. Menai, V.T. Van Hees, A. Elbaz, M. Kivimaki, A. Singh-Manoux, S. Sabia, "Accelerometer assessed moderate-To-vigorous physical activity and successful ageing: Results from the Whitehall II study," Scientific Reports, **8**, 2017, doi:10.1038/srep45772.

[14] J.P. Merepala, "Tri-axis MEMS Accelerometer for Vehicle Accident Monitoring," 1068–1070, 2016.

[15] P.H. Veltink, H.B.J. Bussmann, W. De Vries, W.L.J. Martens, R.C. Van Lummel, "Detection of static and dynamic activities using uniaxial accelerometers," IEEE Transactions on Rehabilitation Engineering, **4**(4), 375–385, 1996, doi:10.1109/86.547939.

[16] F. Bagalá, V.L. Fuschillo, L. Chiari, A. Cappello, "Calibrated 2D angular kinematics by single-axis accelerometers: From inverted pendulum to N-Link chain," IEEE Sensors Journal, **12**(3), 479–486, 2012, doi:10.1109/JSEN.2011.2107897.

[17] J.C. Eisenmann, S.J. Strath, D. Shadrick, P. Rigsby, N. Hirsch, L. Jacobson, "Validity of uniaxial accelerometry during activities of daily living in children," European Journal of Applied Physiology, **91**(2–3), 259–263, 2004, doi:10.1007/s00421-003-0983-3.

[18] N. Ruch, M. Rumo, U. Mäder, "Recognition of activities in children by two uniaxial accelerometers in free-living conditions," European Journal of Applied Physiology, **111**(8), 1917–1927, 2011, doi:10.1007/s00421-011-1828-0.

[19] S. Toumodge, "Book reviews - Modern control systems," IEEE Control Systems Magazine, **6**(5), 56–56, 1986, doi:10.1109/MCS.1986.1105138.

[20] L.A. Kelly, D. Ge Mcmillan, A. Anderson, M. Fippinger, G. Fillerup, J. Rider, Validity of actigraphs uniaxial and triaxial accelerometers for assessment of physical activity in adults in laboratory conditions, 2013.

[21] Y. Matsumoto, M. Esashi, "Integrated silicon capacitive accelerometer with PLL servo technique," Sensors and Actuators: A. Physical, **39**(3), 209–217, 1993, doi:10.1016/0924-4247(93)80221-2.

[22] W. Kuehnel, S. Sherman, "A surface micromachined silicon accelerometer with on-chip detection circuitry," Sensors and Actuators: A. Physical, **45**(1), 7–16, 1994, doi:10.1016/0924-4247(94)00815-9.

[23] B.E. Boser, R.T. Howe, "urf ace Micromachined Accelerometers," **31**(3), 1996.

[24] C. Hierold, A. Hildebrandt, U. Näher, T. Scheiter, B. Mensching, M. Steger, R. Tielert, "A pure CMOS surface-micromachined integrated accelerometer," Sensors and Actuators, A: Physical, **57**(2), 111–116, 1996, doi:10.1016/S0924-4247(97)80101-X.

[25] N. Yazdi, K. Najafi, "An All-Silicon Single-Wafer Micro-g Accelerometer Process," Journal of Microelectromechanical Systems, Vol. 9, **9**(4), 544–550, 2000.

[26] H. Luo, G. Zhang, L.R. Carley, G.K. Fedder, "A post-CMOS micromachined lateral accelerometer," Journal of Microelectromechanical Systems, **11**(3), 188–195, 2002, doi:10.1109/JMEMS.2002.1007397.

[27] B.V. Amini, F. Ayazi, "A 2.5-V 14-bit ∑Δ CMOS SOI capacitive accelerometer," IEEE Journal of Solid-State Circuits, **39**(12), 2467–2476, 2004, doi:10.1109/JSSC.2004.837025.

[28] P.L. Walter, "The history of the accelerometer," S V Sound and Vibration, **31**(3), 16–22, 1997.

[29] L.M. Roylance, J.B. Angell, "A batch-fabricated silicon accelerometer," IEEE Transactions on Electron Devices, **26**(12), 1911–1917, 1979, doi:10.1109/T-ED.1979.19795.

[30] H. V. Allen, S.C. Terry, D.W. De Bruin, "Accelerometer systems with self-testable features," Sensors and Actuators, **20**(1–2), 153–161, 1989, doi:10.1016/0250-6874(89)87113-6.

[31] K. Yamada, K. Higuchi, H. Tanigawa, "A novel silicon accelerometer with a surrounding mass structure," Sensors and Actuators: A. Physical, **21**(1–3), 308–311, 1990, doi:10.1016/0924-4247(90)85061-8.

[32] C. Song, B. Ha, S. Lee, "Micromachined inertial sensors," IEEE International Conference on Intelligent Robots and Systems, **2**(8), 1049–1056, 1999, doi:10.1109/iros.1999.812819.

[33] R. Hiratsuka, D.C. van Duyn, T. Otaredian, P. de Vries, P.M. Sarro, "Design considerations for the thermal accelerometer," Sensors and Actuators: A. Physical, **32**(1–3), 380–385, 1992, doi:10.1016/0924-4247(92)80016-V.

[34] A.M. Leung, J. Jones, E. Czyzewska, J. Chen, B. Woods, "Micromachined accelerometer based on convection heat transfer," Proceedings of the IEEE Micro Electro Mechanical Systems (MEMS), 627–630, 1998, doi:10.1109/memsys.1998.659830.

[35] A. Kazama, T. Aono, R. Okada, "Stress relaxation mechanism with a ring-shaped beam for a piezoresistive three-axis accelerometer," Journal of Microelectromechanical Systems, **22**(2), 386–394, 2013, doi:10.1109/JMEMS.2012.2227139.

[36] J. Chae, H. Kulah, K. Najafi, "An in-plane high-sensitivity, low-noise micro-g silicon accelerometer with CMOS readout circuitry," Journal of Microelectromechanical Systems, **13**(4), 628–635, 2004, doi:10.1109/JMEMS.2004.832653.

[37] B. Li, D. Lu, W. Wang, "Micromachined accelerometer with area-changed capacitance," Mechatronics, **11**(7), 811–819, 2001, doi:10.1016/S0957-4158(00)00050-7.

[38] K. Najafi, A. Selvakumar, K. Najafi, J. Bernstein, A. Arbor, P. Examiner, H. Williams, P.C. Kushman, "United States Patent (19)," (19), 2000.

[39] T. Mineta, S. Kobayashi, Y. Watanabe, S. Kanauchi, I. Nakagawa, E. Suganuma, M. Esashi, "Three-axis capacitive accelerometer with uniform axial sensitivities," Journal of Micromechanics and Microengineering, **6**(4), 431–435, 1996, doi:10.1088/0960-1317/6/4/010.

[40] T. Lehtonen, J. Thurau, "Monolithic Accelerometer for 3D Measurements," Advanced Microsystems for Automotive Applications 2004, 11–22, 2007, doi:10.1007/978-3-540-76989-7_2.

[41] V. Josselin, P. Touboul, R. Kielbasa, "Capacitive detection scheme for space accelerometers applications," Sensors and Actuators, A: Physical, **78**(2), 92–98, 1999, doi:10.1016/S0924-4247(99)00227-7.

[42] M. Youssef, M.A. Yosef, M. El-Derini, "GAC: Energy-efficient hybrid GPS-accelerometer-compass GSM localization," GLOBECOM - IEEE Global Telecommunications Conference, 2010, doi:10.1109/GLOCOM.2010.5684304.

[43] A.-M. Dinarvand, N. Dinarvand, M. Khaleqi, Q. Joogh, "Current Trends in Technology and Science Behavioral Modeling and Simulation of an Open-loop MEMS Capacitive Accelerometer with the MATLAB/SIMULINK," 2014.

[44] B. Guldimann, P. Dubois, P.A. Clerc, N. De Rooij, "Transducers '01 Eurosensors XV," Transducers '01 Eurosensors XV, (April), 2001, doi:10.1007/978-3-642-59497-7.

[45] P. Chen, J. Bai, S. Lou, Q. Lu, D. Han, X. Jiao, "Design of the closed-loop capacitive microaccelerometer based on PSpice," 1st International Conference on Electronics Instrumentation and Information Systems, EIIS 2017, **2018-Janua**, 1–4, 2018, doi:10.1109/EIIS.2017.8298762.

[46] K.N. Khamil, K.S. Leong, N. Bin Mohamad, N. Soin, "Analysis of MEMS Accelerometer for Optimized Sensitivity," International Journal of Engineering and Technology, **6**(6), 2705–2711, 2014.

[47] M. Kraft, C.P. Lewis, T.G. Hesketh, "Control System Design Study for a Micromachined Accelerometer," IFAC Proceedings Volumes, **30**(21), 139–143, 1997, doi:10.1016/s1474-6670(17)41429-7.

[48] Y. Terzioglu, S.E. Alper, K. Azgin, T. Akin, "A capacitive MEMS accelerometer readout with concurrent detection and feedback using discrete components," Record - IEEE PLANS, Position Location and Navigation Symposium, 12–15, 2014, doi:10.1109/PLANS.2014.6851351.

[49] Y. Chu, J. Dong, B. Chi, Y. Liu, "A novel digital closed loop MEMS accelerometer utilizing a charge pump," Sensors (Switzerland), **16**(3), 1–11, 2016, doi:10.3390/s16030389.

[50] M. Hamidu, J.J. Kponyo, A. Acakpovi, "Innovative prediction of vehicle position based on closed loop modeling of capacitive accelerometer," in Proceedings - 2019 International Conference on Cyber Security and Internet of Things, ICSIoT 2019, 54–61, 2019, doi:10.1109/ICSIoT47925.2019.00017.

[51] P.H. Zipfel, Modeling and Simulation of Aerospace Vehicle Dynamics, Second Edition, 2007, doi:10.2514/4.862182.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1040-1048 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Simulating Get-Understand-Share-Connect Model using Process Mining

Shahrinaz Ismail[*,1], Faes Tumin[2]

[1]*Universiti Kuala Lumpur (UniKL), Process Mining Research Cluster (PMineReC), Malaysian Institute of Information Technology (MIIT), Kuala Lumpur, 50250, Malaysia*

[2]*Permodalan Nasional Berhad, Technology Division, Core Application Systems Department (CASD), Kuala Lumpur, 50400, Malaysia*

ARTICLE INFO

Article history:
Received: 24 October, 2020
Accepted: 09 December, 2020
Online: 16 February, 2021

Keywords:
Process mining
Process discovery
Simulation
Personal knowledge management
GUSC Model

ABSTRACT

This paper presents the method of simulating the personal knowledge management (PKM) processes, based on Get-Understand-Share-Connect (GUSC) Model, using real event logs data from an online learning platform. The method used in here is process discovery and conformance, which are the process mining techniques. Having the model proven at granular level of multi-agent system, this research is found significant in proving that PKM indeed exists in students' online learning behavior and needs to be monitored to ensure that they are managing knowledge in a complete cycle, to support their credibility as future graduates and knowledge workers in organizations. The ideal process starts from Get, then Understand, and followed by Share and Connect, but this study proves that the sequence may vary although the original theory is construed. This depends on the way the online activities being mapped to the Get, Understand, Share and Connect processes during the data processing stage. The results from this paper include the simulation of the GUSC model as discovered from real event logs data.

1. Introduction

Due to the habit and intuitive that they possess, the digital natives, i.e. the generation that grows up with the Internet technologies, are used to independent learning and do not rely much on physical classroom learning. It is a competition of being the most knowledgeable to them, to be able to gain more knowledge and manage their personal knowledge in as many means as they can through the Internet, hence the shift from traditional classrooms to online learning environment in many universities today.

Looking at the need to ensure that learners gained the right knowledge at the right time, and being competitive and marketable graduates, it is necessary to also ensure that they could well manage their personal knowledge. With the current experience in managing classes and teaching materials online, the data is available for analysis on online learning behavior of learners using the university online learning platform. Instead of relying on questionnaire and interview surveys that highly depend on respondents' feedback that could be bias and distorted from truth, it is a better way to approach the case of online learning behavior directly from the data source on the server.

Recent research has proven the existence of personal knowledge management (PKM) in online learning platforms, including social media and learning management system provided by universities. However, these studies were mostly done using surveys. One significant research developed an agent-based simulation to visualize the PKM processes based on those surveys. There is a gap between these two types of previous research, in which this study intends to close this gap with its findings visualized directly from the original source, which is the event logs data from the online learning server. In doing so, this research aims to achieve the following objectives:

- to discover PKM model by visualizing the process flow in online learning environment; and
- to simulate the PKM processes using real case data for further verification on conforming the model.

This paper is an extension of work originally presented in the 6th International Conference On Research & Innovation In Information Systems (ICRIIS2019), Malaysia [1].

[*]Corresponding Author: Shahrinaz Ismail, Universiti Kuala Lumpur City Campus, 1016 Jalan Sultan Ismail, +603-21754435, shahrinaz@unikl.edu.my

https://dx.doi.org/10.25046/aj0601115

2. Related Works

2.1. GUSC Model Simulation

As mentioned in the introduction, a recent research has developed an agent-based simulation on GUSC Model, a model that is proven in quite a number research before this. The abbreviation "GUSC" comes from the words Get, Understand, Share and Connect, which are the processes deem necessary for personal knowledge management (PKM). The GUSC Model was simplified from a handful of previous models on PKM by renowned authors since 2009 [2]-[4]. The only difference between GUSC Model and other models before this is the proven suitability of this model for intelligent agents modelling, in which the PKM processes can be developed at granular level for software agents to perform.

In lieu to the previous models, this PKM model is based on the following processes [4], [5]:

- ***Get or Retrieve knowledge (G)***: This process simulates how learners retrieve knowledge in explicit form, in which that knowledge has been converted from tacit form by the person who shared it;
- ***Understand or Analyze Knowledge (U)***: This process simulates how learners analyze the explicit knowledge that they have retrieved, and convert it to a tacit form in a way that they understand;
- ***Share or Publish Knowledge (S)***: This process demonstrates how learners share or post the knowledge that they have understood in an explicit form, so that others can gain knowledge from it;
- ***Connect to Knowledge Source (C)***: This process is about connecting learners to sources of knowledge, including materials and people, which may involve communication between the two, resulting in a transformation from tacit form of knowledge to another tacit form.

The decision of adopting GUSC Model for this research is due to its credibility in proving the existence of PKM in learning environment [6], smart classroom [7], [8] , social network [9] , and social messaging applications [10], [11]. In addition to that, prototypes have been developed to prove that GUSC Model can be used to design PKM platforms for individuals [12], and a simulation of GUSC Model in learning environment is developed to show and prove PKM processes among learners [13].

It was suggested in the recent research on GUSC simulation [14] to relook into the simulation itself by improving the way the measurement of each PKM process is done. There is a trend discovered from the simulation, in which there are some "learning sessions" that are not used, or termed as "non-usage". This is a significant finding that is found interesting for this research to investigate using real event logs data.

2.2. Process Model Discovery

Derived from business process management domain, process mining helps the analysis of business process based on actual data of event logs. It is often thought as the same as data mining, but it is not. Data mining algorithm is significant to process mining but the latter provides more comprehensive benefits in terms of the organizational system as a whole.

Event logs are analysed in the context of process-aware system [15], in which the full view of the case situation is being understood, hence the importance of the event logs to researchers. Process mining techniques are expected to be non-trivial because it is based on extraction of only useful information from the logs [16], which in turn could provide truthful, real data without biasness. Nevertheless, process mining can show that a real process in a system is more complex than the way it is planned to be in documents. The event logs from e-learning environments might contain huge amounts of fine granular events and process-related data, which consist of different categories that make the whole process messy and prone to producing waste of resources and useless data [17].

There are three different techniques suggested for process mining [18], as follows:

- ***Process Model Discovery***: By using the event logs retrieved from the database, fuzzy miner algorithms are applied by the process mining software to produce a process model. This requires data to be readily available, in a situation where process model is not yet planned or developed in documentation.
- ***Conformance Checking***: When a process model is available (either from the plans and documentation or from the discovery), the retrieved event logs are analysed to conform whether the real activities happening in the system is as how the process is modelled for, or not.
- ***Process Improvement***: The purpose of this technique is to improve the existing process model using the information retrieved from the event logs. It is different than conformance checking because it measures the alignment between the model and the real online activities, hence producing an improved model that fits the condition deemed necessary.

There are several algorithms that can be used to perform process discovery including Fuzzy miner and Heuristic miner [19]. Fluxicon Disco is a recommended tool for automated process model discovery based on fuzzy miner algorithms. The strength of the fuzzy miner is the seamless process simplification and highlighting of frequent activities that are very useful for the research purpose [20]. The process model could be messy and fuzzy mining will come in handy in terms of filtering out the data and adjusting the level of details available in the model, in order to better perceive the process model.

3. Methodology

For this research, process model discovery will be used to initiate the process of discovering the GUSC Model in online learning environment and achieving the first objective. This is followed by conformance checking, to prove that GUSC processes really happen in the online environment, by simulating the processes through animated visualization, to achieve the second objective.

3.1. Case Settings

This research has chosen a case university that implements a Moodle-based online learning environment called "VLE", i.e.

virtual learning environment. The event logs data from year 2016 to 2017 was extracted from the VLE database for the purpose of this study. The event log consists online activities data that the students went through every day in the VLE. The extracted fields from the event logs data are date and time, case ID, and activity.

The students who performed the online learning activities in the selected case are those from semester three to six. They were chosen because they have experience using the VLE for more than one semester.

The selected courses for this study are from software engineering and business technology. These courses are selected to ensure that both technical and less technical contents are covered in the process model discovery, since the lecturers or instructors of both courses have their own way of delivering teaching materials online and assigning activities. For example, business technology may have a more theoretical contents compared to software engineering, hence the instructor may use discussion forum and wiki more than the latter.

Other important matters to be identified during case settings are the activities in the event log and the relevance of data to this research. Irrelevant data of less or non-relevant activities need to be omitted out, for example "register", "delete comment" and "update profile". Since the purpose of this research is to analyze PKM processes in VLE, the activities selected for analysis are those related to PKM activities only, such as "course module viewed", "quiz attempt started" and "quiz attempt summary viewed".

3.2. Process Discovery

In process discovery, the event log was extracted and imported into the chosen process mining tool (i.e. Fluxicon Disco), in which the log is automatically transformed to create a newly discovered process model using fuzzy miner algorithm. There are two ways of process discovery: play-in, i.e. from event log to process model; and play-out, i.e. from process model to event log [21].

The first step in classification of requirement is the extraction of the event log from VLE database and importing the event log into the Fluxicon Disco. Figure 1 shows the interface after the event log is imported into the tool.

From Figure 1, the event log (in *.CSV file format) that is inserted into the process mining tool needs to be classified according to the requirement, i.e. ID, Activity and Timestamp. Due to the difficulty of having process instances scattered over multiple rows (i.e. in *.CSV file), it is found that the Fluxicon Disco is a suitable tool for data extraction compared to a normal spreadsheet.

The basic data requirement for process mining is to look into the historical process data precisely, such as a "process lens". The three requirements are as follows:

i) *Case ID*: A case identifier, else called as process case ID, is important to recognize various executions of a similar process. The correctness of case ID relies on the domain of the process. For this research case, the case ID are *Get*, *Understand*, *Share* and *Connect*, i.e. the GUSC processes.

ii) *Activity*: There ought to be names for various procedure steps or status changes as activities were performed in the process. In the event that the information only appears once or in a single passage (i.e. one row) for each procedure occasion, then it could be concluded that at that point the information is not detailed enough.

iii) *Timestamp*: From the timestamp, the delays between activities can be identified as it could show the duration of each activities in a process. This could be further related to the process of a series of activities and its duration across time.

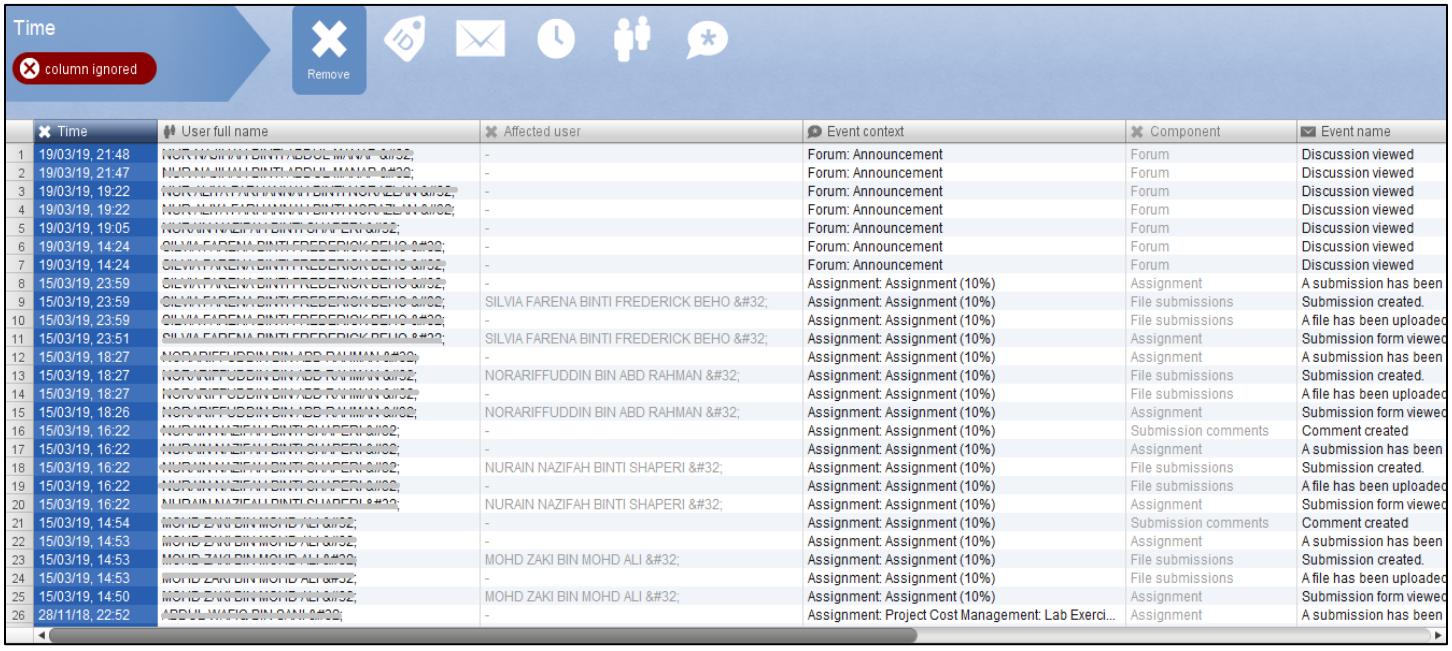


Figure 1: Event log being imported into Fluxicon Disco

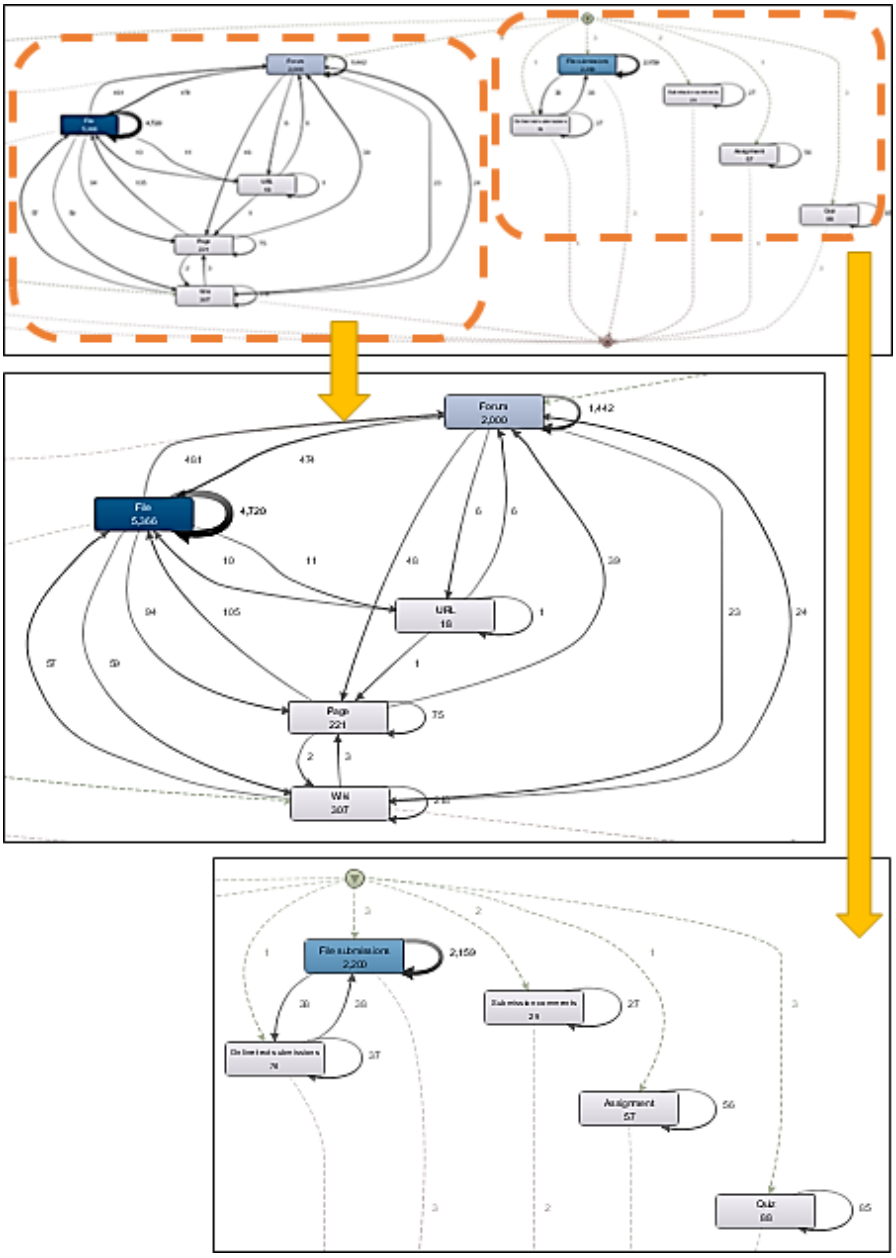


Figure 2: Results after classification of requirement

3.3. Process Mapping

Figure 2 shows the result after classification of requirement has been confirmed. From the classification, a process model is developed based on the mapping that has been assigned in the event log. This process model can be manipulated by adjusting its part and activity frequency. This is to ensure that the process model is better understood, instead of merely showing the full process that involves so many paths and activities that will often be displayed as spaghetti-like process. Adjusting the activities to lower the frequency would not make the process model unreliable, but it could hide some activities that may not seem fit to the process model.

In a glance as shown in Figure 2, it is observed that there are two groups of activities being defined in the overall process. The left part of the process flow shows the learning activities over forum, file, URL, webpage and wiki, whereas the right part of the process flow shows the assessment related activities, mainly file submission, assignment and quiz.

Filtering of data can be easily done in Fluxicon Disco, as compared to using spreadsheet. Data that is identified to be useless for the process model simulation is filtered out. The Timeframe can also be filtered according to year, month and/or day, according to the time an active process happened. For this research, the whole timeframe is taken into account to see active months and years, so that none of the activities are left behind, e.g. student logging in to VLE during mid-semester break or in semester break. Figure 3 shows the interface of filtering process in Fluxicon Disco.

The statistics interface shown in Figure 4 displays an overview of the data, which can be in a huge amount. The statistics table shows which case is frequently active in the process, and the connection between each data. The *Connect* process is identified throughout the process in a more understandable way. This table

generated in Fluxicon Disco also shows the frequency of each activity that consists the number of events data. The duration of each activity can also be viewed here.

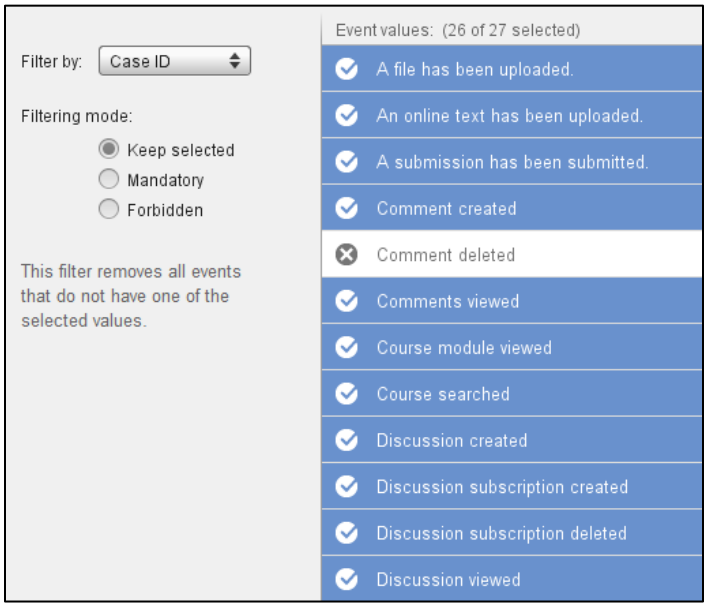


Figure 3: Filtering of data in Fluxicon Disco

activities and their frequencies. The results shown here are presented according to three types, i.e. process model by components, by PKM process variables, and by activities that happened in the VLE.

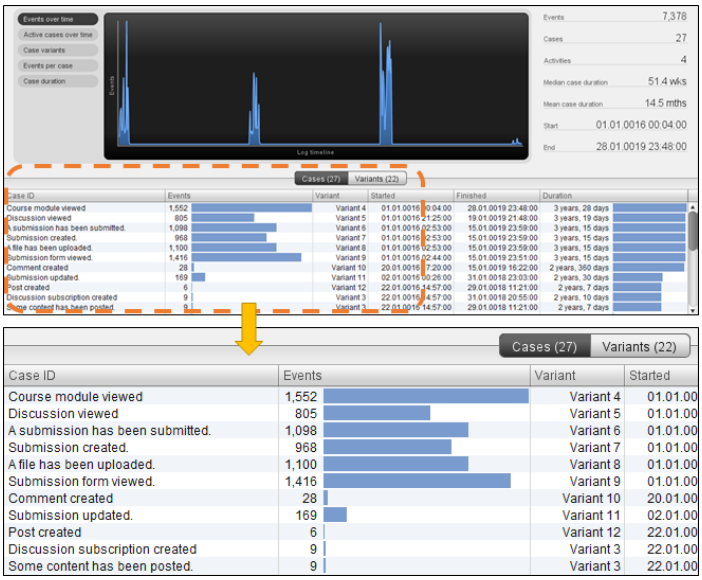


Figure 4: Statistics of activities in Fluxicon Disco

4. GUSC Model Simulation

This section presents the results of the GUSC Model simulation using the fuzzy miner algorithm in Fluxicon Disco. The simulation results are displayed in two modes: static view of GUSC process flow; and animated view of GUSC process flow.

The process model presented here basically shows the actual process that happened in VLE, which consists of the numbers of

4.1. Static View of GUSC Process Flow

Figure 5 shows the process model that was developed according to components from the event log. In general, there are 10 components analyzed for process model discovery simulation: File; File submissions; Submission comments; Quiz; Assignment; Forum; Wiki; Online text submissions; URL; and Page.

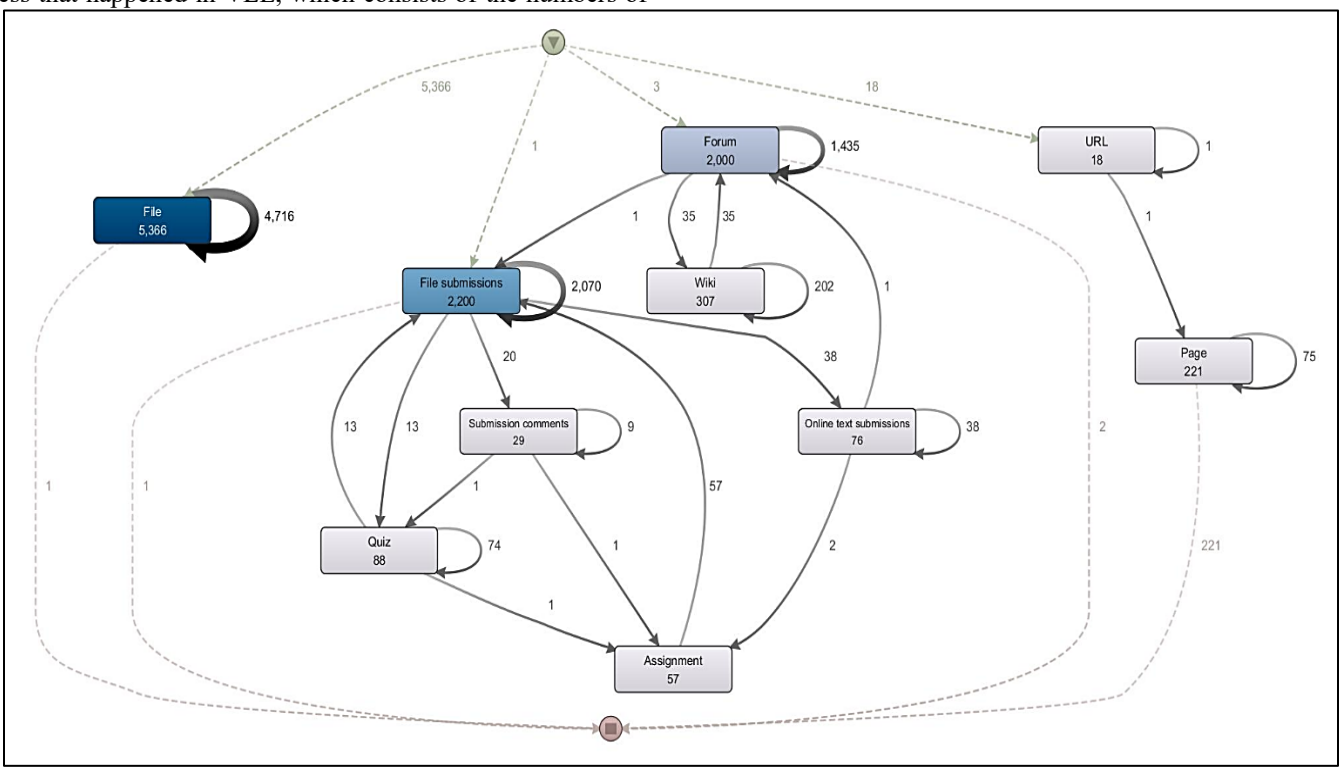


Figure 5: GUSC process model by components

The 'File' component (shown in darkest shade in Figure 5), is the dominant one because the activities that happened on a 'File' is frequently active (i.e. 5,366 occurrences). Students' activities of downloading or uploading files would fall under this 'File' component. Activities like 'downloading notes' are quiet common in the event logs and it happens quite frequently among the students.

'URL' (18 occurrences) is a component that is posted by the lecturers, in which the URL or hyperlink of a website is shared in the VLE for students to click on and be directed to the landing page with materials to learn. A 'Page' (221 occurrences) or webpage serves the same purpose as the 'URL', but instead of directly bringing the users to an external website, it is an internal webpage created by the lecturers, in which the students can view notes and multimedia contents like a page in a website within the VLE.

Each activity has different component and each component may share the same activity name. For example, the component 'File' may have two activities in it that are 'Course module viewed' and 'Wiki page viewed', and these two activities share the same component with 'Forum', 'Page', 'URL', and 'Wiki'. This is due to the way the 'File' component is used within the other activities.

Table 1 shows the number of occurrences for each component shown in Figure 5. It also presents the number of recurrences for each component, which happened when students return to the same component during a session. Figure 5 also shows the frequencies of inflow and outflow processes, which are not presented in Table 1. However, the total number of recurrences, inflow and outflow processes for each component should be the same as the number of occurrences for each component.

Table 1: Component Occurrences in VLE

Component	Occurrence	Recurrence
File	5,366	4,716
File submissions	2,200	2,070
Submission comments	29	9
Quiz	88	74
Assignment	57	-
Forum	2,000	1,435
Wiki	307	202
Online text submissions	76	38
URL	18	1
Page	221	75

Figure 6 shows the process model that was developed according to PKM processes, based on the GUSC model. The processes are *Get*, *Understand*, *Share* and *Connect*, as stated in Section 2. These processes are mapped to the VLE activities in the imported event log data (i.e. in the *.CSV file), as the mapping was developed according to the PKM process model requirements made prior to this study.

As shown in Figure 6, the number shows the frequencies of the VLE activities that are mapped to the GUSC processes. The *Get* activity is the dominant one as compared to the other processes (shown in darkest shade in Figure 6). This result is as expected because the *Get* activities is common to happen more frequently than *Share* or *Connect* activities. The basic idea of having VLE is to provide a platform for students to *Get* information and knowledge shared by others, especially in the form of learning materials and notes.

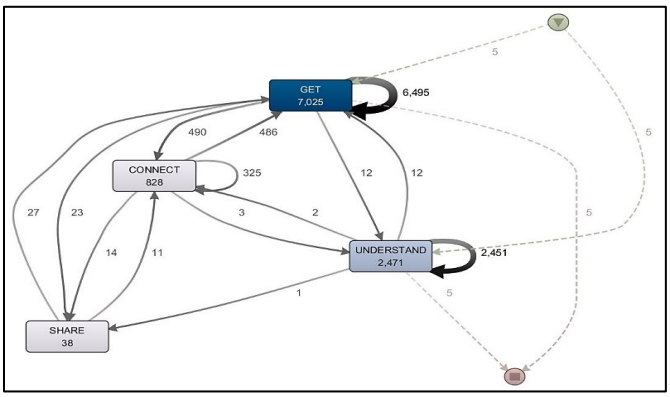


Figure 6: GUSC process model by PKM processes

The looping arrow displayed on *Get* process shows a number of reworks that occurred in the event. An activity that keeps looping like this basically means that students kept downloading the same notes or materials, e.g. for three days in a row. The same pattern is seen on the *Understand* process, in which the process keeps looping on the same activity as the students attempt to understand the information and knowledge in VLE. An example of this is uploading of assignment, in which students are allowed to keep uploading an assignment, or attempt multiple submissions of an assignment, as long as the deadline has not yet past.

Table 2 shows the number of occurrences and recurrences for each GUSC process derived from Figure 6. As described for Table 1, the inflow and outflow processes are not shown here, but the number of occurrences should show the total number of inflow processes, outflow processes and recurrences as a whole.

Table 2: GUSC Process Occurrences in VLE

Process	Occurrence	Recurrence
Get	7,025	6,495
Understand	2,471	2,451
Share	38	-
Connect	828	325

Figure 7 shows the overall view of the GUSC process model according to VLE activities. The most dominant activity is "File – Course module viewed – Get" with 5,366 occurrences (shown in darkest shade in Figure 7). This is followed by "Forum – Course module viewed – Get" (1,167 occurrences) and "File submissions – A file has been uploaded – Understand" (1,100 occurrences). In other words, the *Understand* process does happen in an online learning environment like VLE, even if can only be proven by uploading of files to VLE.

In a glance, the process model presented in Figure 7 looks quite structured. Unlike the previous two views, Figure 7 shows no recurrences of any activity. Recurrences only happened between two activities, e.g. between "Forum – Course module viewed – Get" and "Forum – Discussion viewed – Connect" (470 occurrences to, and 473 occurrences return). Table 3 shows the summary of occurrences according to the activities derived from Figure 7.

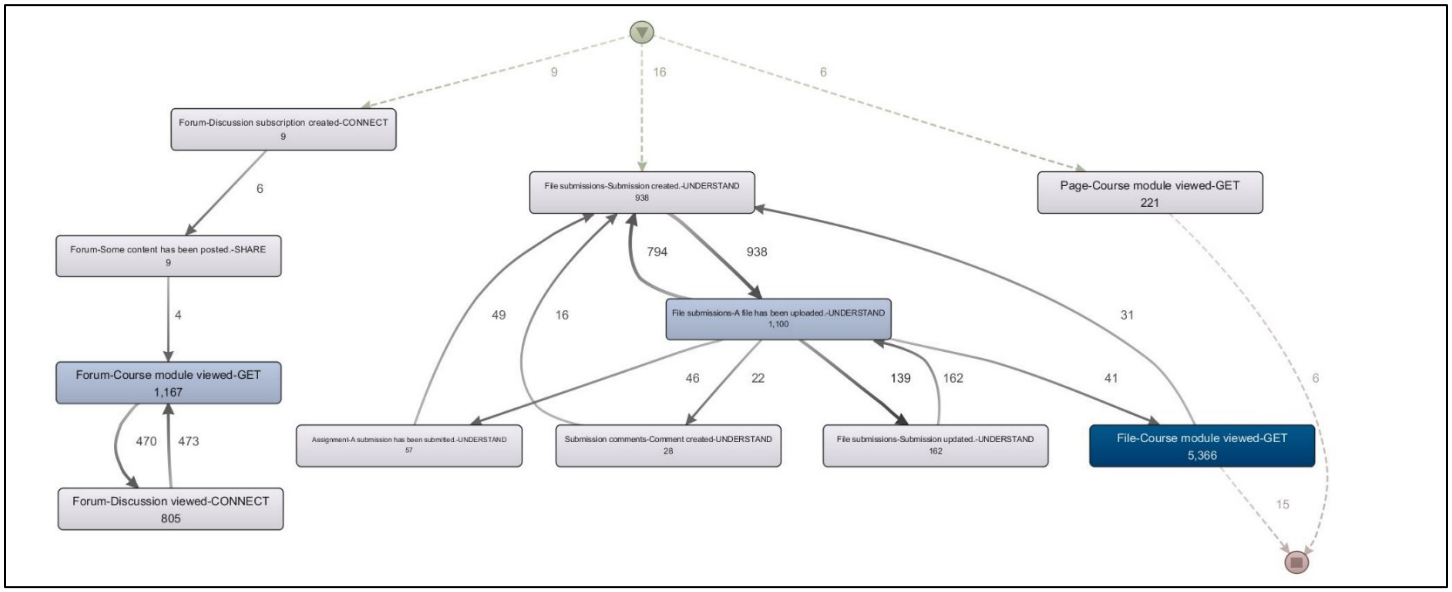


Figure 7: GUSC process model by activities

Table 3: Activity Occurrences in VLE

Activity	Occurrence
Forum – Discussion subscription created – Connect	9
Forum – Some content has been posted – Share	9
Forum – Course module viewed – Get	1,167
Forum – Discussion viewed – Connect	805
File submissions – Submission created – Understand	938
File submissions – A file has been uploaded – Understand	1,100
Assignment – Submission has been submitted – Undestand	17
Submission comments – Comment content – Understand	28
File submissions – Submission updated – Understand	102
Page - Course module viewed – Get	221
File - Course module viewed – Get	5,366

4.2. Animated View of GUSC Process Flow

The GUSC model simulation shows the movement of the processes in 'blobbing' shapes. The movement can be seen with different rhythm and speed depending on the time duration and frequency of the running processes. The thicker the line movement of the process travels, the frequent the process is. Figure 8 shows the snapshot of this animation of GUSC process model, in which the thicker line in red is where a process travels, and the yellow 'blobbing' shape shows the high volume per time unit.

As shown in Figure 8, the *Understand* process does not travel to other variables except for one way towards *Share* process, while *Get*, *Connect* and *Share* processes travel to each other. This can be explained that *Get*, *Connect* and *Share* processes happen before the *Understand* process can happen. The situation supports the findings of previous research that proved the *Get*, *Connect* and *Share* are connected to each other and they are the main processes of PKM [6]. This was statistical proven to justify that an *Understand* process can only happen when a *Share* process happens (the one-way line shown from *Understand* to *Share* in Figure 8), and *Share* process happens when *Get* and *Connect* processes happen (shown in darker lines in Figure 8).

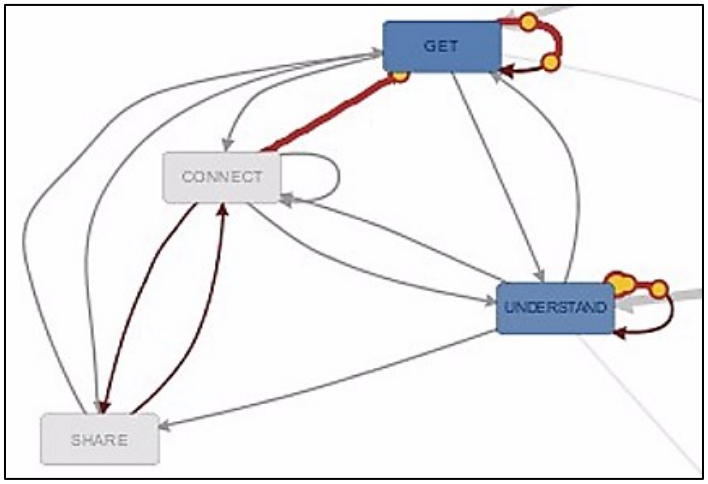


Figure 8: Animated view of GUSC process model

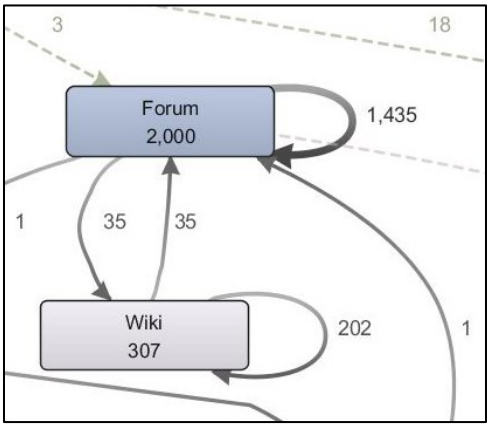


Figure 9: Frequency of occurrences and recurrences on Forum and Wiki

	Time	User full name	Event context	Component	Event name
2	24/03/19, 19:54	[redacted]	Forum: Announcement	Forum	Course module viewed
3	23/03/19, 15:02	[redacted]	Forum: Announcement	Forum	Course module viewed
4	23/03/19, 11:26	[redacted]	Forum: Announcement	Forum	Course module viewed
5	22/03/19, 23:44	[redacted]	Forum: Announcement	Forum	Course module viewed
6	22/03/19, 21:14	[redacted]	Forum: Announcement	Forum	Course module viewed
7	22/03/19, 12:33	[redacted]	Forum: Announcement	Forum	Course module viewed
8	22/03/19, 10:18	[redacted]	Forum: Announcement	Forum	Course module viewed
9	19/03/19, 21:48	[redacted]	Forum: Announcement	Forum	Discussion viewed
10	19/03/19, 21:47	[redacted]	Forum: Announcement	Forum	Discussion viewed
11	19/03/19, 21:46	[redacted]	Forum: Announcement	Forum	Course module viewed
12	19/03/19, 19:22	[redacted]	Forum: Announcement	Forum	Discussion viewed
13	19/03/19, 19:22	[redacted]	Forum: Announcement	Forum	Discussion viewed
14	19/03/19, 19:22	[redacted]	Forum: Announcement	Forum	Course module viewed
15	19/03/19, 19:05	[redacted]	Forum: Announcement	Forum	Discussion viewed
16	19/03/19, 19:05	[redacted]	Forum: Announcement	Forum	Course module viewed
17	19/03/19, 15:37	[redacted]	File: Chapter 3: The PM Process Groups	File	Course module viewed
18	19/03/19, 15:34	[redacted]	File: Chapter 3: The PM Process Groups	File	Course module viewed

Figure 10: Timestamp in event log data

5. Discussions

Among the features available in VLE, forum is a powerful platform that can facilitate students in performing all PKM processes. It has been mapped to all four PKM processes, namely GUSC processes, and the results have shown that Forum receives high occurrences and recurrences in certain areas of PKM. Although Wiki has similar strength in providing the same PKM facilities, the exposure on its usage is still low, in which students do not find that Wiki could help in enhancing their learning capacity through collaboration and knowledge sharing. Both of these features or components require high commitment and interaction among the students, and thus the processes of *Connect* and *Understand* could be well performed on top of *Get* and *Share*.

The process model discovered through this simulation research has shown the real scenario of processes that occurred in VLE, but it is far from the expectation this research earlier perceived. A simple activity as downloading lecture notes is actually an important activity in VLE to achieve the process of *Get* knowledge, but it happens as a looping activity. In business process view, a looping activity is not good because it shows that the first time of performing the activity is not done properly. As a result, it can be concluded that the students are not managing their knowledge very well because they keep on downloading the same notes several times.

This research shows that using Forum and Wiki in VLE will not only let the students interact with each other but it can also boost the full potential of the VLE itself. Not many users know all the features in VLE that can provide the full potential benefits of managing knowledge, hence time is wasted on available precious resources. Figure 9 shows the frequency of the activities happened on Forum and Wiki, in which the process travels a lot between these two components that all the PKM processes happened in this loop simultaneously. In contrast, components like Assignment only fall under *Understand* process, as students only tend to submit their assignments once they have understood on how to complete them.

Figure 10 shows the event log data when it was first retrieved from the database. It does not show the start and end time of a process, but only the overall time when a process happened. The unavailability of both start and end timestamp in the VLE event log data has caused this limitation. This limitation has caused some difficulties in analyzing further on students' activities, especially when the duration of each activity can produce significant measurement and findings for this research, in which the duration can be derived from having two timestamps.

Overall, the use of process mining has helped in proving a theoretical model like personal knowledge management processes, without biasness of respondents' feedback. This research has proven that the same event log data can be used to analyze different theoretical models, as the same data was used in a research on self-regulated learning (SRL) model prior to this [22].

6. Conclusion

In a nutshell, this research has achieved its objectives of discovering PKM model by visualizing the process flow in online learning environment, and simulating the PKM processes using real case data for further verification on conforming the model. With the simulation views (i.e. both static and animated views), it is expected that online learning environment users can benefit in terms of knowing their status of managing knowledge. It should benefit both the lecturers (who can use the simulation to gauge students' learning behavior as well as improving teaching initiatives) and students (who can know where they can improve in terms of responding to the learning system for own future benefits). Nevertheless, the process model discovered in this research is highly dependent on how the online learning environment is used in the case organization, in which the features used by both lecturers and students reflect whether the PKM model is fully complied or not. Missing features or activities may affect the learners' capability of managing personal knowledge, as they depend on the features to exist for them to have more choices and fully utilize as part of their learning sessions.

It is recommended that the future work could improve the way this research is conducted, in terms of activity and component mapping to the PKM processes (i.e. GUSC), as well as identifying the start and end time for each activity in an online learning system to better analyze the overall PKM processes. Other opportunities

include adopting suitable techniques to perform the process mapping to activities and extending this research on other courses and case settings.

Acknowledgment

This journal publication is funded by Centre of Research & Innovation, Universiti Kuala Lumpur, Malaysia.

References

[1] S. Ismail, F. Tumin, "Analysis on online learning environment using process mining technique for personal knowledge management mapping," in International Conference on Research and Innovation in Information Systems, ICRIIS, 2019, doi:10.1109/ICRIIS48246.2019.9073269.

[2] L. Razmerita, K. Kirchner, F. Sudzina, "Personal knowledge management: The role of Web 2.0 tools for managing knowledge at individual and organisational levels," Online Information Review, **33**(6), 1021–1039, 2009, doi:10.1108/14684520911010981.

[3] H. Jarche, From observation to breakthrough, Adapting to Perpetual Beta Blog, 2012.

[4] S. Ismail, M.S. Ahmad, Z. Hassan, "Emerging personal intelligence in collective goals: Data analysis on the bottom-up approach from PKM to OKM," Journal of Knowledge Management, **17**(6), 2013, doi:10.1108/JKM-08-2013-0313.

[5] S. Ismail, T.D. Nguyen, M.S. Ahmad, "A multi-agent knowledge expert locating system: A software agent simulation on personal knowledge management (PKM) model," in International Conference on Intelligent Systems Design and Applications, ISDA, 2014, doi:10.1109/ISDA.2013.6920705.

[6] S. Ismail, A. Othman, M.S. Ahmad, "Knowledge Management in Learning Environment: A Case Study of Students' Coursework Coordination," in Knowledge Management International Conference (KMICe), 2014.

[7] S. Ismail, S.F.M. Suhaimi, M.S. Ahmad, "The GUSC model in smart notification system: The quantitative analysis and conceptual model," in 2013 8th International Conference on Information Technology in Asia - Smart Devices Trend: Technologising Future Lifestyle, Proceedings of CITA 2013, 2013, doi:10.1109/CITA.2013.6637581.

[8] S. Ismail, M.S. Ahmad, "Knowledge Management in Agents of Things: A case study of smart classroom management," in International Conference on Research and Innovation in Information Systems, ICRIIS, 2013, doi:10.1109/ICRIIS.2013.6716685.

[9] S. Ismail, Z. Mohammed, N.W. Yusof, M.S. Ahmad, "Personal Knowledge Management among Adult Learners: Behind the Scene of Social Network," International Journal of Humanities and Social Sciences, 2013.

[10] S. Ismail, M.S. Ahmad, "Personal Knowledge Management among Managers: Mobile Apps for Collective Decision Making," Journal of Information Systems Research and Innovation, 2015.

[11] S. Ismail, N. Jamaludin, "Managing knowledge over social messaging application: The case of an event management project group," in 2016 3rd International Conference on Computer and Information Sciences, ICCOINS 2016 - Proceedings, 2016, doi:10.1109/ICCOINS.2016.7783184.

[12] S. Ismail, M.F. Ghazali, "Mobile coursework coordination deploying the concept of agent-mediated personal knowledge management in learning environment," in ICICTM 2016 - Proceedings of the 1st International Conference on Information and Communication Technology, 2017, doi:10.1109/ICICTM.2016.7890797.

[13] S. Ismail, "Simulating Get-Understand-Share-Connect model for personal knowledge management in learning environment," in 2nd International Symposium on Agent, Multi-Agent Systems and Robotics, ISAMSR 2016, 2017, doi:10.1109/ISAMSR.2016.7809995.

[14] S. Ismail, "Simulating Get-Understand-Share-Connect model for personal knowledge management in learning environment," in 2nd International Symposium on Agent, Multi-Agent Systems and Robotics, ISAMSR 2016, 2017, doi:10.1109/ISAMSR.2016.7809995.

[15] W. Van Der Aalst, T. Weijters, L. Maruster, "Workflow mining: Discovering process models from event logs," IEEE Transactions on Knowledge and Data Engineering, 2004, doi:10.1109/TKDE.2004.47.

[16] W.M.P. Van Der Aalst, "Process-aware information systems: Lessons to be learned from process mining," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2009, doi:10.1007/978-3-642-00899-3_1.

[17] A.H. Cairns, B. Gueni, M. Fhima, A. Cairns, S. David, N. Khelfa, "Process Mining in the Education Domain," International Journal on Advances in Intelligent Systems, 2015.

[18] W. Van der Aalst, Process mining: Data science in action, 2016, doi:10.1007/978-3-662-49851-4.

[19] K. Rattanathavorn, W. Premchaiswadi, "Analysis of customer behavior in a call center using fuzzy miner," in International Conference on ICT and Knowledge Engineering, 2015, doi:10.1109/ICTKE.2015.7368485.

[20] V.A. Rubin, A.A. Mitsyuk, I.A. Lomazova, W.M.P. Van Der Aalst, "Process mining can be applied to software too!," in International Symposium on Empirical Software Engineering and Measurement, 2014, doi:10.1145/2652524.2652583.

[21] W.M.P. Van Der Aalst, "Relating process models and event logs 21 conformance propositions," in CEUR Workshop Proceedings, 2018.

[22] M.H.B.A. Bakar, S. Ismail, S.H.S. Ali, "A process mining approach to understand self regulated-learning in moodle environment," International Journal of Advanced Trends in Computer Science and Engineering, 2019, doi:10.30534/ijatcse/2019/1581.32019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 5, 1042-1049 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Cavity Structure based Flexible Piezoelectric for Low-Frequency Vibration Energy Harvesting

Khairul Azman Ahmad[*,1], Siti Noraini Sulaiman[1], Noramalina Abdullah[2], Muhammad Khusairi Osman[1]

[1]*Faculty of Electrical Engineering, Universiti Teknologi MARA, Cawangan Pulau Pinang, Permatang Pauh, Pulau Pinang, 13500, Malaysia*

[2]*School Of Electrical & Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, Seberang Perai Selatan, Nibong Tebal Pulau Pinang, 14300, Malaysia*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 28 September, 2020
Online: 20 October, 2020

Keywords:
Flexible piezoelectric energy harvester
Low-frequency vibration energy
Interdigitated electrode circuit

ABSTRACT

Piezoelectric energy harvesters (PEH) can be used in many areas of application, including human walking, railways, pavements and bridges. Piezoelectric energy harvesters are currently based on two types of external forces, namely pressure load and mechanical oscillation or vibration. A vibration energy harvesting (VEH) is a mechanical oscillation in a piezoelectric energy harvester that harvested electric energy. In the market, there is available energy harvesting device in good electric energy harvesting and very sensitivity. However, the price is too high and the fabrication process is too complex. Furthermore, one of the aimed of the research is to install the energy harvesting device at rotary compressor machine which has noise vibration frequency at 1 kHz to 10 kHz. This paper presented a cavity structure-based flexible piezoelectric vibration energy harvester (FPVEH) based on an IDE circuit for low-frequency vibration applications. A cavity structure (IDE circuity) combine with the flexible circuit (polyimide) and flexible membrane (polyvinylidene fluoride, PVDF) will increase the electric energy harvesting and sensitivity of the device. Therefore, the four designs (Design A to D) are used to investigate the effect of the electrode finger width and the gap between the electrode fingers (to investigate the cavity structure applying in the design). All designs have been characterized by FEA simulation using COMSOL Multiphysics 5.0 and experimental work using a sieve shaker vibration machine. A sieve shaker machine is worked as vibration frequency calibrator. However, the sieve machine can operate at 5 kHz and 7 kHz. Since these two vibration frequencies are in targeted vibration frequency. It is used as vibration frequency calibrator in this experimental work. The results from the FEA simulation and experimental work show the Design D has the highest electric energy harvesting compare to other designs. It has electric energy harvesting at 27.3 V for 1 minutes period. Design D has a wide electrode finger width and the wide gap between electrodes compare to other designs. The vibration frequency was also given the impact to energy harvesting whereby the vibration frequency at 5 kHz has the highest electric energy harvesting compare to vibration frequency at 7 kHz.

1. Introduction

Previously, the research [1] has been looked up into a vibration effect towards PMUT design and then, it has been energised and stored in a capacitor. This paper is an extension of work originally presented in 9th International Conference on Control System, Computing and Engineering (ICCSCE2019) [1]. The paper will look into the flexible membrane (PVDF) and cavity design (the gap between electrodes) in PMUT for energy improvement.

Recently, numerous studies have been investigated the use of piezoelectric energy harvesters (PEH) in various fields of application, such as human walking [2], railways [3], pavements [4–6], and bridges [7]. Several applications have emerged based on frequency and vibration studies [8]-[10], while others have been

*Corresponding Author: Khairul Azman Ahmad, azman062@uitm.edu.my

based on load or pressure acting on an energy harvesting device [4] and [11].

Flexible piezoelectric energy harvesting (FPEH) systems use elastomer pillars based on a piezoelectric polymer, whereby the stored energy has been converted into electrical energy in the clamped boundary on the piezoelectric layer [12]. Double-sided tribological layers, including polydimethylsiloxane (PDMS) and PDMS/multiwall carbon nanotubes with flexible polyvinylidene fluoride (PVDF), have been used as a triboelectric mechanism to enhance the performance of electric harvesters at various frequencies [13]. A flexible piezoelectric energy harvester with a sandwich structure using PIN-PMN-PT/epoxy has also been designed for wearable applications. The sandwich structure and the interdigitated electrode (IDE) film increase the flexural modulus of the membrane, resulting in a high electrical output [11]. The flexible piezoelectric energy harvester has been further developed using lead-free (Na0.5Bi0.5) TiO3-BaTiO3 piezoelectric nanofibers, which exhibit a high peak voltage output with a high load resistance [14]. Polymer piezoelectric materials, in collaboration with a beating mechanism, have been shown to generate a large force during low-frequency vibration and thus generate large output [15]. In one study, a piezoelectric PZT film was clamped to one end of a fixed wall and a load was mounted onto its other end [3]. This is called a cantilever piezoelectric beam model, and it has since been used in many vibration-type piezoelectric energy harvesters [3], [9], [16], [17].

Many researchers are examining the relationship between vibration and a resonance frequency that is capable of increasing the output power [8]. A vibration energy harvester with multiple nonlinear techniques can broaden the bandwidth of energy harvesting systems; however, there is a need to tune the resonance frequency [18], [19]. This design uses a parallel-plate structure consisting of a suspended spring-plate, which is a very complicated design and extra work is needed to tune the frequency [18]. A resonant vibration energy harvester was developed based on the cantilever model that can achieve a high strain and thus maximize the output power [20]. Another study on broadening the bandwidth in piezoelectric energy harvesters used a bi-stable composite laminate. This bi-stable piezoelectric energy harvester has a high geometrical nonlinear response that can broaden the frequency bandwidth during vibration [21]. Besides, a bimorph piezoelectric vibration energy harvester with flexible piezoelectric material was designed to improve the output power while lowering the resonance frequency. The bimorph-type piezoelectric increases the deflection, and this increases the strain inside the piezoelectric material and thereby improves the output power [9].

This paper presents a flexible piezoelectric vibration energy harvester (FPVEH) for use in low-frequency vibration energy harvesting applications. The energy harvesting device that available in the market can produce low energy harvester. There is an available energy harvesting device that can produce high electric energy harvester and sensitive to vibration frequency in the market. However, the price is a too high and complicated fabrication. The aimed of this research is to design and characterize a cavity structure-based flexible piezoelectric energy harvesting for low-frequency vibration applications. Beside of that, this design is to improve the quality of electric energy harvester and sensitive to vibration frequency. The target vibration frequencies are in the range 1 kHz to 10 kHz because the device will be installed at the rotary compressor that has noise vibration frequency from 1 kHz to 10 kHz. However, the available vibration frequency calibrator that is sieve shaker machine only can support 5 kHz and 7 kHz frequency. Since these two frequencies are in the range of target vibration frequencies, this sieve shaker has been used as vibration frequency calibrator in the experiment.

The four designs of FPVEH involving different in the gap between electrodes and different in the width of electrode fingers have been simulated and tested and demonstrated a significant performance in the result. All of these will be explained in the next section. The contribution of this research is a cavity design which is the gap between electrode fingers and a flexible circuit that has significant performance in electric energy harvesting. The investigation of different electrode fingers gap will be shown the significant of cavity design in FPVEH. The next section presents the design and methodology for the piezoelectric vibration energy harvester, while the finite element analysis (FEA) simulation method and experimental method are also described. The FEA simulation results and experimental results are discussed in the results and discussion section, and the last section presents the conclusion.

2. Design and Method

The PMUT design and method are presented in this section. A cavity-based flexible piezoelectric energy harvesting has been elaborated in terms of structure and mechanical properties that related to energy harvesting. Furthermore, a procedure and method are presented in this section.

2.1. IDE Circuits Structure

Four different IDE circuits have been designed to examine the effect of the different gap between electrodes of FPVEH and the electrode finger width of the FPVEH that harvested the electrical energy. Hereby, the electrode finger width and electrode finger gap were the two parameters that were investigated in low-frequency vibration. The design parameters are shown in Table 1. The schematic diagram of the IDE circuit is shown in Figure 1 and is referred to as Design A.

Table 1: IDE circuit design parameters

Parameter	Design A(mm)	Design B(mm)	Design C(mm)	Design D(mm)
IDE width	0.5	0.5	1.0	1.0
IDE gap	0.5	1.0	0.5	1.0
IDE length	20.5	26.0	25.0	29.0

As presented for Design A in Figure 1, the electrode finger gap between the positive and negative electrodes was 0.5 mm, the electrode finger width was 0.5 mm, and the IDE circuit length was 20.5 mm.

Polyvinylidene fluoride (PVDF) was attached to the top of the IDE circuit. The IDE circuit was made of polyimide (PI) as a dielectric and flexible membrane, and copper (Cu) was used for the finger electrode of the IDE circuit. The 3M tape was attached

on top of PVDF and IDE circuit and then worked together as a new structure flexible membrane in harvesting electric potential difference of low-frequency vibration. The schematic diagram of the flexible piezoelectric vibration energy harvester (FPVEH) designed in this study is shown in Figure 2.

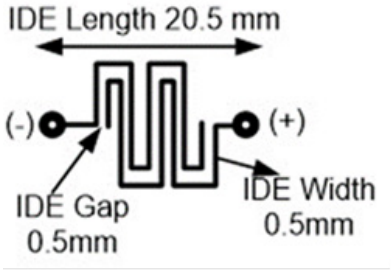
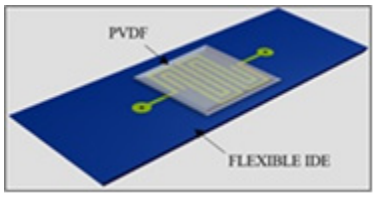
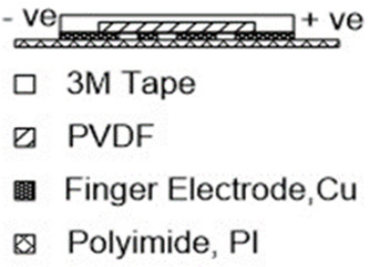


Figure 1: Schematic diagram of IDE circuit Design A.



(a)



(b)

Figure 2: Schematic diagram of the flexible piezoelectric vibration energy harvester; (a) top view and (b) side view.

The thicknesses of the 3M tape, PVDF, Cu and PI are shown in Table 2. The contribution in this design of FPVEH is a cavity added between PVDF and PI and clamped by electrode fingers at both ends left and right. It gives more flexure to the membrane and oscillates smoothly if there an external force acting on both top and bottom membrane. The four IDE circuit designs were created using Proteus 8 software; these are illustrated in Figure 3. The fabricated IDE circuit is depicted in Figure 4, whereby the yellow material is a coverlay made from PI and the shiny circuit is made from Cu. The completed fabrication of the FPVEH is shown in Figure 5.

Table 2: The thickness of material in flexible piezoelectric vibration energy harvester design.

Material name	Thickness (mm)
3M tape	0.88
PVDF	0.11
Finger electrode, Cu	0.035
Polyimide, PI	0.025

The readout circuitry for the energy harvester system is presented in Figure 6, and it is also called a full-bridge rectifier circuit, consisting of four diodes, a capacitor and a resistor. This full-bridge rectifier circuit converted the AC electric potential difference from the FPVEH into DC voltage and stored it in a capacitor. The value of the capacitor was 47 μF and it was used as a filter to reduce the ripple voltage while also acting temporary storage. The resistor had a value of 1 kΩ and was used as a load. A 1N4007 diode was considered suitable for use in low-power consumption because its peak-to-peak voltage was around 70 V.

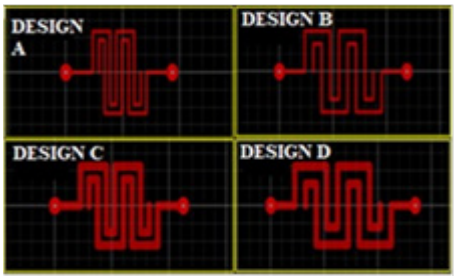


Figure 3: The four IDE circuit design in Proteus software.



Figure 4: A fabricated IDE circuit.

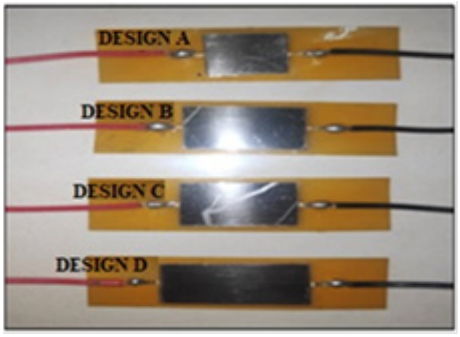


Figure 5: The completed fabrication of the FPVEH.

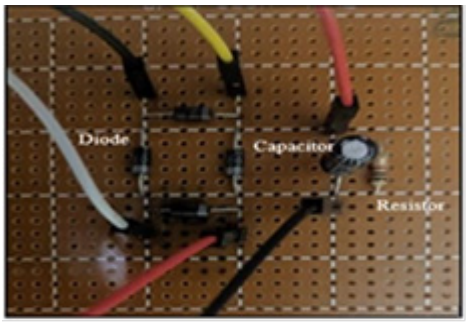


Figure 6: A full-bridge rectifier circuit.

2.2. Simulation Method

In the FEA simulation method, COMSOL Multiphysics 5.0 presented the direct piezoelectric effect on the piezoelectric material and elastic material. Firstly, defined the gravitational acceleration as amount 9.764 ms^{-2} in parameter definitions. A mass is represented of a pressure or a load per unit area with 0.5 MPa also has been defined in parameter definitions.

Then, set up the geometry in the geometry column. The flexible circuit consists of polyimide and copper, Cu is shown in Figure 7. Then, the geometric has been set up until the geometric structure shown in Figure 8. It consisted of 3M tape at the top layer followed by PVDF, copper and polyimide.

Then, set the material properties in the material properties column. The material properties of all materials are shown in table 3.

Table 3: Mechanical properties of all material

Material name	Mechanical properties		
	Young's modulus (GPa)	Poisson ratio (na)	Density (kg/m^3)
PVDF	8.3	0.18	1780
Copper, Cu	128	0.36	8920
Polyimide	7.5	0.35	1420
3M Tape	0.0045	0.499	980

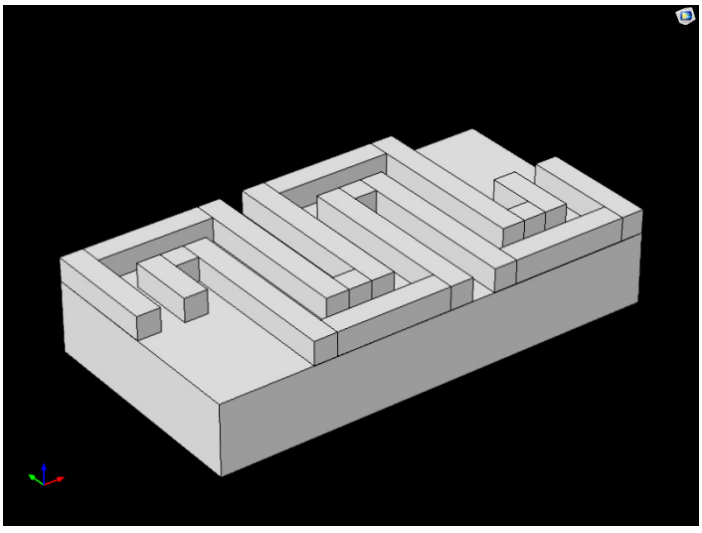


Figure 7:. Top view of set up flexible circuit in the geometric column in COMSOL Multiphysics

Next, select the suitable physics studies such as solid mechanics, electrostatics and Multiphysics. Selected the proper domain and put in physics studies column. For example, all materials are solid, then all domain belongs to solid mechanics. In solid mechanics, both ends at the left and right of the FPVEH are clamped. These clamped made the FPVEH oscillated freely when there was an external force acted on the surface of the membrane from the top or bottom. Then, select PVDF as an electrostatic domain. PVDF is a material that can convert mechanical property to electrical property and put the domain it into electrostatics study.

Then, select the suitable mesh, for instance, used size normal in the range 0.342 mm to 1.9 mm, used swept and the distribution, and select all domain put into distribution column. The mesh of geometry is shown in Figure 9.

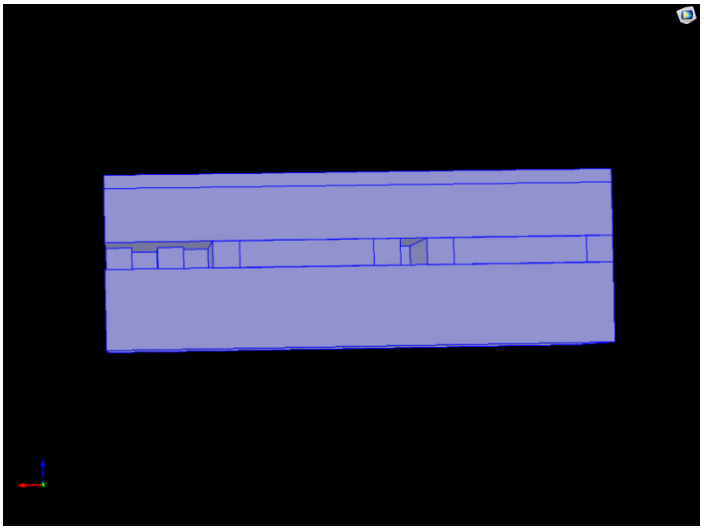


Figure 8: Side view of completed geometric setup in COMSOL Multiphysics

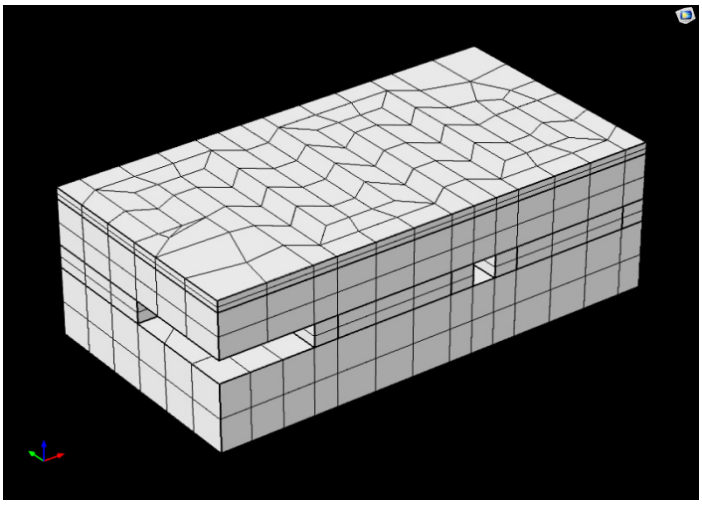


Figure 9: Mesh of geometry

Then, select the frequency domain as the study method. Furthermore, selected frequency range from 0 to 10 kHz with step 0.5 kHz as frequency input and put into the frequency domain. Lastly, in the result, the electric potential and electrical displacement were selected for plotted the graph. In electric potential result, the graph total voltage against frequency input has been plotted whereby in electrical displacement, the graph membrane displacement against frequency input has been plotted.

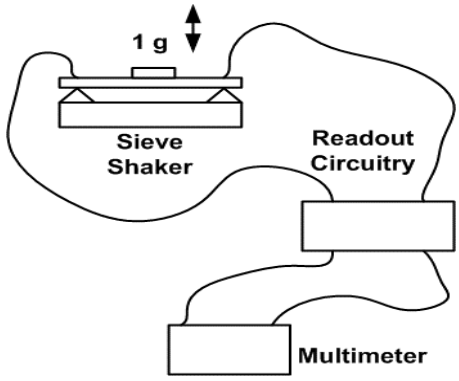
2.3. Experimental method

In this experimental, two parameters have been studied such as the effects of different electrode finger width of FPVEH and the effect of the different gap between electrodes of FPVEH during vibration. The electrical output voltage has been used as a benchmark of different studies to indicate the performance of the energy harvesting device. The four designs of FPVEH with different electrode finger and the different gap between electrodes have been tested using a sieve shaker to characterize their performances.

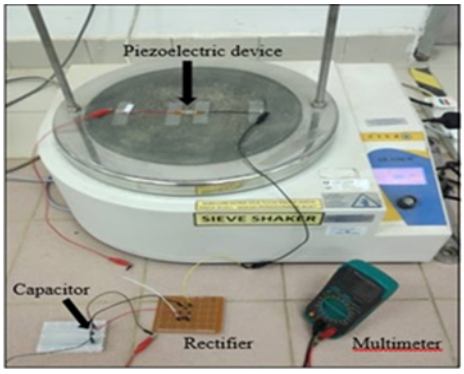
In experimental, a sieve shaker machine has been used to apply a vibration effect to the FPVEH. The FPVEH was placed on top of sieve shaker with both ends clamped using foam tape. Foam tape was worked as clamped at both ends as shown as a triangle in Figure 10 (a). The gravitational acceleration as amount 9.764 ms^{-2} has been set up to the sieve shaker machine with 0.5 MPa mass load. As one of the aimed of this research is to obtain electric harvesting energy from the rotary compressor within the range of frequency noise 1 to 10 kHz. In the laboratory, there is no calibrator machine can support 1 kHz to 10 kHz vibration. However, this sieve shaker machine can be operated at 5 kHz and 7 kHz only. Since the operation frequency, 5 kHz and 7 kHz are still in the range of rotary compressor frequency noise (1 kHz to 10 kHz), this sieve shaker machine has been used as a calibrator vibration machine.

The sieve shaker has a function to set time duration. The time duration can be used for investigating the energy stored in the period of time. The sieve shaker can support the period of 1 minute, 5 minutes, 10 minutes, 15 minutes and 20 minutes.

The output measurements comprised the electric field charge collected in the capacitor. Each of four designs (Design A to D) energy harvesting device was placed on top of the sieve shaker machine, and its electrodes were connected to the readout circuitry. The output of readout circuitry was measured by multimeter, and then it was observed and recorded the measurements. The experiment utilized the sieve shaker vibration machine to characterize the performance of four designs of different electrode finger width of FPVEH and different of the gap between electrodes of FPVEH that caused by vibration were shown in Figure 10.



(a)



(b)

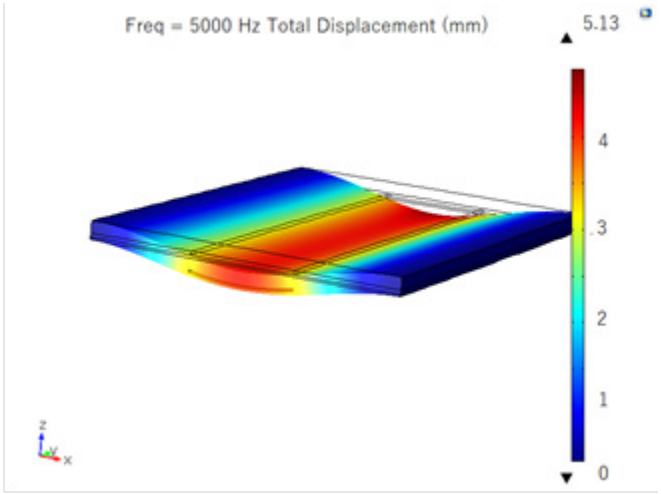
Figure 10: The experimental method to characterize the four designs with a different electrode finger width of FPVEH and the different gap between electrodes of FPVEH caused by vibration, (a) the schematic diagram, and (b) test field.

3. Results and Discussion

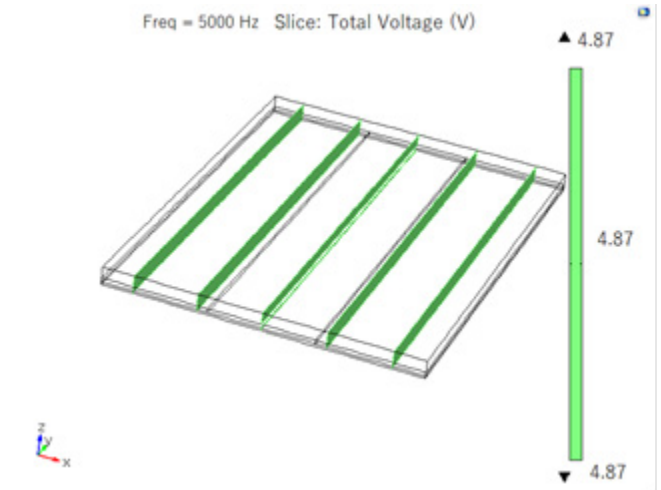
This section is divided into two parts that, respectively, present the simulation results and the experimental results. In the simulation results, two investigations were conducted to examine the effect of different IDE electrode finger widths and different IDE electrode gaps on the displacement of the membrane and the subsequently harvested an electric field accordingly to frequency input given. In the experimental result, the investigations were done on four designs with different electrode fingers width and different electrode gap (Design A to D) that have been affected on potential different, V by the energy harvesting device.

3.1. Simulation Results

FEA simulation is performed to characterize the total displacement of the membrane and the total electric potential difference harvested caused by the different of IDE electrode finger width and IDE electrode accordingly to the frequency input given. Figure 11 presents the simulation results for an IDE electrode finger width and gap of 1.0 mm and 1.0 mm, respectively for input frequency, 5000 Hz.



(a)



(b)

Figure 11: Simulation results for an IDE electrode finger width of 1.0 mm and an IDE electrode finger gap of 1.0 mm according to 5000 Hz; (a) displacement of the membrane (mm), (b) and total electric potential difference (V).

Figure 11(a) shows that the centre of the FPVEH membrane between the electrodes gap has been displaced by about 5.13 mm towards the negative z-axis at a frequency of 5000 Hz. The gap between the electrode fingers created a cavity between the PVDF and the polyimide. Therefore, the displacement of the flexure membrane increased when pressure was exerted on the surface of the FPVEH. Figure 11(b) shows that the FPVEH has a total electric potential difference of 4.87 volts for a frequency of 5000 Hz. Hence, Figure 11 was demonstrated the flexural membrane that has relation with total electric potential difference harvested by the FPVEH.

The simulation results for the four different electrode finger widths and gaps (Design A to D) are shown in Figure 12. The results show Design D with electrode finger width 1.0 mm and the gap between electrodes 1.0 mm has the highest electrical energy harvesting value compare to all design. Furthermore, Design D and Design C that has the same electrode finger width are the higher electric energy harvester compare to design B and Design A which are the smaller size of electrodes fingers width.

Next, Design D has higher electric energy harvesting compare to Design C whereby the Design D has a higher gap between electrodes compare to Design C. The Design A and Design B which are smaller in electrode fingers width and the gap between electrodes has lower electric energy harvesting. Subsequently, the vibration frequency has given the impact on harvested electric energy at low-frequency vibration applications whereby in the results shown the total voltage is highest at 5 kHz vibration frequency. It can show that the resonance frequency of Design D is at 5 kHz.

3.2. Experimental Results

In this section, the experiment results have been demonstrated. The investigation on the different electrode finger width of FPVEH and the different gap between electrodes of FPVEH were demonstrated. As explained in methodology, all four FPVEH designs (Design A to D) with different electrode finger and the different gap between electrodes were measured using a multimeter and sieve shaker as a vibrator calibrator. Table 4 is tabulated the experimental results of Design A, Design B, Design C and Design D for the different lengths of the periods.

The tabulated results are plotted in Figure 13. Figure 13 shows that the electric potential difference harvested by all four designs (Design A to D) increase relative to the increase in the length of time. Design D has an electrode finger width of 1.0 mm and a gap between the electrode fingers of 1.0 mm, and it has the highest electric potential difference value compared to the others. In addition to that, Design D and Design C with an electrode finger width of 1.0 mm have higher electric potential difference value compare to Design A and Design B that have lower an electrode finger width at 0.5 mm. Therefore, the wide electrode finger width has affected the electric potential difference value to increase.

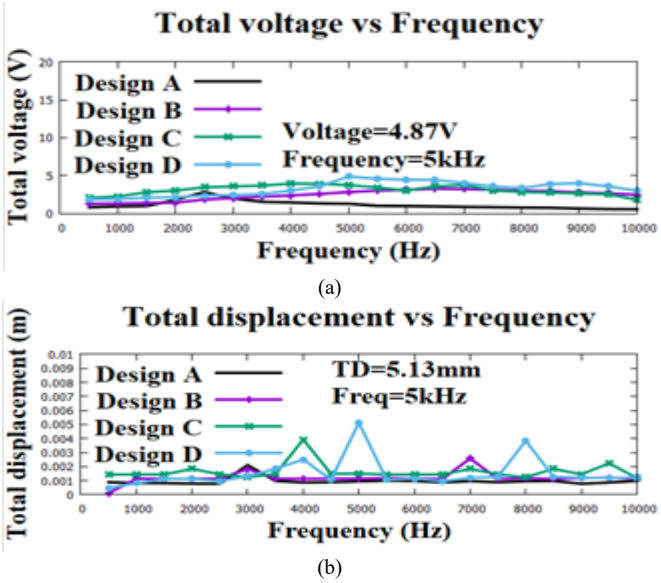


Figure 12: Simulation results for the four different electrode finger widths and gaps based on different frequency input; (a) total voltage (V) vs frequency (Hz), and (b) total displacement (mm) vs frequency (Hz).

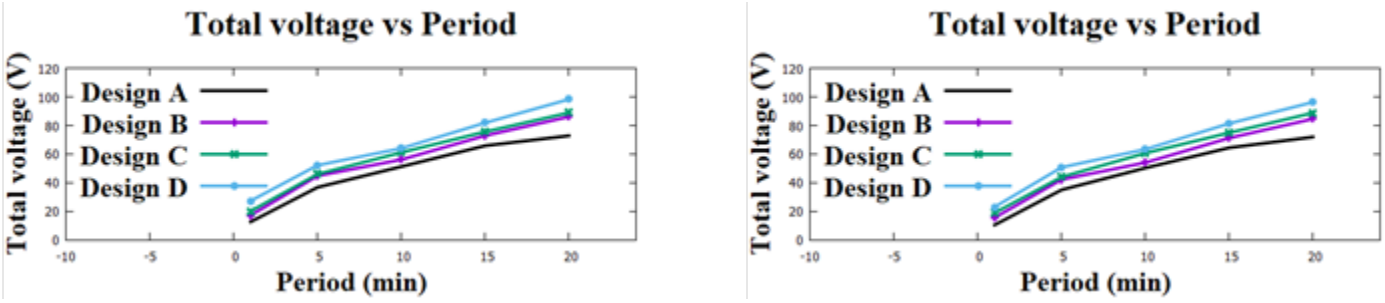
Furthermore, Design D has higher electric potential difference value compare to Design C which indicates the gap between electrode fingers also affected an increase of electric potential difference value. Therefore, the increase of the electrode fingers gap has higher electric potential difference value.

Next, it can be seen that the electric potential difference harvesting in Figure 13(a) has a higher value compared to the value in Figure 13(b). The electric potential difference harvested in the experiment using the sieve shaker machine at 5 kHz was higher than at 7 kHz. Hence, the frequency vibration will affect electric energy harvesting by FPVEH.

In Figure 13 (a), Design D with an electrode finger width of 1.0 mm and electrode finger gap of 1.0 mm has the highest electric potential difference harvesting at 98.4 volts. The frequency of the sieve shaker machine is 5 kHz.

Table 4: Tabulated data for Design A, Design B, Design C, and Design C

Period (min)	Total Voltage stored (V)							
	Design A		Design B		Design C		Design D	
	Amplitude 5 kHz	Amplitude 7 kHz	Amplitude 5 kHz	Amplitude 7 kHz	Amplitude 5 kHz	Amplitude 7 kHz	Amplitude 5 kHz	Amplitude 7 kHz
1	12.6	10.4	17.1	15.5	20.1	18.9	27.3	23.0
5	36.8	35.0	44.8	42.2	46.0	44.0	52.2	50.8
10	51.3	50.3	56.2	54.1	61.2	60.7	64.1	63.4
15	65.9	64.5	73.0	71.3	75.7	75.0	82.0	81.5
20	72.9	72.0	86.2	84.7	89.0	88.9	98.4	96.4



(a) (b)

Figure 13: Total electric potential difference by experimental vibration results for the different designs and different lengths of time; (a) 5 kHz vibration and (b) 7 kHz vibration.

Table 5 shows the summary of the previous study that designed the vibration piezoelectric energy harvesting and their performances. The highest voltage has been harvested by [21] at 12 V with a resonance frequency of 19.1 Hz. The design can energise high voltage however, it needs to tune their frequencies until it reached the high electrical potential differences. The previous design was used PCB as substrate and it was hard to bend, thus the flexible circuit has been used to replace PCB as a substrate that can bend together with PVDF substrate. Therefore, more electrical energy has been produced. In this work has shown the significant performance in increasing the electric energy harvesting.

Table 5: Summary of previous study and their achievement.

Author	Method	Achievement
[22]	Design PEH using clamped-clamped beam with proof mass.	Voltage: 4.5V at 1 g Resonance Frequency: 42.8 Hz Displacement of plate: 1.56 mm Stress: 8Mpa
[23]	Design multi-layer piezoelectric energy harvester (MPEH).	Voltage: 3.79V at 1 g Resonance frequency: 91 Hz
[24]	Design PEH based macro-fibre composite	Voltage: 2.8 V in 2.12 s Resonance frequency: 8 Hz
[25]	Design metamaterial beam with double-layer resonators.	Voltage: 0.48 V Resonance frequency: 348 Hz
[21]	Design PEH based cantilever tunable beam.	Voltage: 12 V at 0.7 g Resonance frequency: 120 rad/s = 19.1 Hz
[1]	Previous work, design interdigitated electrodes on the PCB board with PVDF membrane and 1.0 mm finger width.	Voltage = 322 mV at 1 ($9.764\ m^{-2}$) g and 1 minutes, Voltage = 0.005 V at 1 s. Frequency = 50 Hz
This work.	Design flexible piezoelectric based on PVDF and polyimide substrate.	Voltage = 27.3 V at 1 g ($9.764\ m^{-2}$) and 1 minutes, Voltage = 0.455 V at 1s. Frequency = 5 kHz.

4. Conclusion

This paper presented an investigation into the effect of electrode finger width as well as the gap between the electrode fingers and vibration frequency on the potential difference harvested by an FPVEH. The novel of FPVEH presented here was successfully designed and fabricated using IDEs, a flexible printed circuit, PVDF and 3M tape. The new structure flexible membrane and novel cavity design in FPVEH were successfully presented in this paper.

The experiment showed that the electrode finger gap had a greater effect on the electric potential difference harvesting by the FPVEH than the width between the electrode fingers. In this experiment with different setting frequencies at vibration machine was shown that the electric potential difference harvesting by all four designs using vibration machine at 5 kHz have higher than the electric potential difference harvesting by all four designs using vibration machine at 7 kHz. This indicates the vibration frequency has affected the electric potential difference has been harvested. Design D is the best design that harvesting the highest electric potential difference of 98.4 volts with a frequency input of 5 kHz. These concluded that the increases of electrode finger gap of FPVEH, and the wide of electrode finger width of FPVEH can increase the potential difference that harvesting by FPVEH. Moreover, the vibration frequency was also gave affected to energy harvesting when the results show that the vibration of 5 kHz has higher electric potential difference compare to vibration of 7 kHz that been harvested by FPVH.

Furthermore, a cavity that is a gap between electrodes has caused the increases of electrical energy harvesting by FPVEH. As a recommendation for future research, we propose the use of a power bank as a form of storage that can keep the electrical energy harvesting.

Conflict of Interest

The authors would like to declare that it is no conflict of interest in this paper publication.

Acknowledgement

The authors would like to acknowledge and express their highest appreciation to Research Management Unit (RMU), Universiti Teknologi MARA, Cawangan Pulau Pinang, Kampus Permatang Pauh for funding this fee of the paper. The authors also would like to express appreciation to Adavances Control System and Computing Research Group (ACSCRG), Universiti Teknologi MARA, Cawangan Pulau Pinang, Kampus Permatang Pauh for their contributions to this research.

References

[1] K.A. Ahmad, M.F. Abdullah, N. Abdullah, "Design and Characterization of an Interdigitated Electrode PVDF based Energy Harvesting Device," in 2019 9th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 172–177, 2019, doi:10.1109/ICCSCE47578.2019.9068549.

[2] A.C. Turkmen, C. Celik, "Energy harvesting with the piezoelectric material integrated shoe," Energy, **150**, 556–564, 2018, doi:10.1016/j.energy.2017.12.159.

[3] M.Y. Gao, P. Wang, Y. Cao, R. Chen, C. Liu, A rail-borne piezoelectric transducer for energy harvesting of railway vibration, Journal of Vibroengineering, **18**(7), 4647–4663, 2016, doi:10.21595/jve.2016.16938.

[4] C. Wang, S. Wang, Z. Gao, X. Wang, "Applicability evaluation of embedded piezoelectric energy harvester applied in pavement structures," Applied Energy, **251**, 113383, 2019, doi:10.1016/j.apenergy.2019.113383.

[5] X. Xu, D. Cao, H. Yang, M. He, "Application of piezoelectric transducer in energy harvesting in pavement," International Journal of Pavement Research and Technology, **11**(4), 388–395, 2018, doi:10.1016/j.ijprt.2017.09.011.

[6] X. Zhao, H. Xiang, Z. Shi, "Piezoelectric energy harvesting from vehicles induced bending deformation in pavements considering the arrangement of harvesters," Applied Mathematical Modelling, **77**, 327–340, 2020, doi:10.1016/j.apm.2019.07.048.

[7] G. Yesner, A. Jasim, H. Wang, B. Basily, A. Maher, A. Safari, "Energy harvesting and evaluation of a novel piezoelectric bridge transducer," Sensors and Actuators A: Physical, **285**, 348–354, 2019, doi:10.1016/j.sna.2018.11.013.

[8] L. Lamprecht, R. Ehrenpfordt, C.K. Lim, A. Zimmermann, "A 500 Hz-wide kinetic energy harvester: Outperforming macroscopic electrodynamic arrays with piezoelectric arrays," Mechanical Systems and Signal Processing, **119**, 222–243, 2019, doi:10.1016/j.ymssp.2018.09.025.

[9] T. Tsukamoto, Y. Umino, S. Shiomi, K. Yamada, T. Suzuki, "Bimorph piezoelectric vibration energy harvester with flexible 3D meshed-core structure for low frequency vibration," Science and Technology of Advanced Materials, **19**(1), 660–668, 2018, doi:10.1080/14686996.2018.1508985.

[10] X. Wang, H. Xiao, "Dimensionless Analysis and Optimization of Piezoelectric Vibration Energy Harvester," International Review of Mechanical Engineering (IREME), **7**(4), 607-624–624, 2013, doi:10.15866/ireme.v7i4.3812.

[11] Z. Zeng, L. Gai, A. Petitpas, Y. Li, H. Luo, D. Wang, X. Zhao, "A flexible, sandwich structure piezoelectric energy harvester using PIN-PMN-PT/epoxy 2-2 composite flake for wearable application," Sensors and Actuators A: Physical, **265**, 62–69, 2017, doi:10.1016/j.sna.2017.07.059.

[12] H.G. Çetin, B. Sümer, "A Flexible Piezoelectric Energy Harvesting System for Broadband and Low-frequency Vibrations," Procedia Engineering, **120**, 345–348, 2015, doi:10.1016/j.proeng.2015.08.631.

[13] Y. Zhu, B. Yang, J. Liu, X. Wang, X. Chen, C. Yang, "An Integrated Flexible Harvester Coupled Triboelectric and Piezoelectric Mechanisms Using PDMS/MWCNT and PVDF," Journal of Microelectromechanical Systems, **24**(3), 513–515, 2015, doi:10.1109/JMEMS.2015.2404037.

[14] B. Liu, B. Lu, X. Chen, X. Wu, S. Shi, L. Xu, Y. Liu, F. Wang, X. Zhao, W. Shi, "A high-performance flexible piezoelectric energy harvester based on lead-free (Na0.5Bi0.5)TiO3–BaTiO3 piezoelectric nanofibers," Journal of Materials Chemistry A, **5**(45), 23634–23640, 2017, doi:10.1039/C7TA07570G.

[15] H.-H. Huang, K.-S. Chen, "Design, analysis, and experimental studies of a novel PVDF-based piezoelectric energy harvester with beating mechanisms," Sensors and Actuators A: Physical, **238**, 317–328, 2016, doi:10.1016/j.sna.2015.11.036.

[16] Y. Kuang, M. Zhu, "Design study of a mechanically plucked piezoelectric energy harvester using validated finite element modelling," Sensors and Actuators A: Physical, **263**, 510–520, 2017, doi:10.1016/j.sna.2017.07.009.

[17] J. Fu, Y. Hou, M. Zheng, M. Zhu, "Flexible Piezoelectric Energy Harvester with Extremely High Power Generation Capability by Sandwich Structure Design Strategy," ACS Applied Materials & Interfaces, **12**(8), 9766–9774, 2020, doi:10.1021/acsami.9b21201.

[18] X. Wang, C. Chen, N. Wang, H. San, Y. Yu, E. Halvorsen, X. Chen, "A frequency and bandwidth tunable piezoelectric vibration energy harvester using multiple nonlinear techniques," Applied Energy, **190**, 368–375, 2017, doi:10.1016/j.apenergy.2016.12.168.

[19] M. Yuan, Z. Cao, J. Luo, J. Zhang, C. Chang, "An efficient low-frequency acoustic energy harvester," Sensors and Actuators A: Physical, **264**, 84–89, 2017, doi:10.1016/j.sna.2017.07.051.

[20] S. Du, Y. Jia, S.-T. Chen, C. Zhao, B. Sun, E. Arroyo, A.A. Seshia, "A new electrode design method in piezoelectric vibration energy harvesters to maximize output power," Sensors and Actuators A: Physical, **263**, 693–701, 2017, doi:10.1016/j.sna.2017.06.026.

[21] C. Lihua, X. Jiangtao, P. Shiqing, C. Liqi, "Study on cantilever piezoelectric energy harvester with tunable function," Smart Materials and Structures, **29**(7), 075001, 2020, doi:10.1088/1361-665X/ab859f.

[22] A. Damya, E. Abbaspour Sani, G. Rezazadeh, An innovative piezoelectric energy harvester using clamped–clamped beam with proof mass for WSN applications, Microsystem Technologies, 9, 2018, doi:10.1007/s00542-018-3890-6.

[23] Q. Lu, L. Liu, F. Scarpa, J. Leng, Y. Liu, "A novel composite multi-layer piezoelectric energy harvester," Composite Structures, **201**, 121–130, 2018, doi:10.1016/j.compstruct.2018.06.024.

[24] S. Balguvhar, S. Bhalla, "Evaluation of power extraction circuits on piezo-transducers operating under low-frequency vibration-induced strains in bridges," Strain, **55**(3), 2019, doi:10.1111/str.12303.

[25] J.S. Chen, W.J. Su, Y. Cheng, W.-C. Li, C.-Y. Lin, "A metamaterial structure capable of wave attenuation and concurrent energy harvesting," Journal of Intelligent Material Systems and Structures, **30**(20), 2973–2981, 2019, doi:10.1177/1045389X19880023.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1049-1057 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Formal Proof of Properties of a Syntax-Oriented Editor of Robotic Missions Plans

Laurent Nana*, François Monin, Sophie Gire

Univ Brest, Lab-STICC, CNRS, UMR 6285, F-29200 Brest, France

ARTICLE INFO

Article history:
Received: 23 October, 2020
Accepted: 28 January, 2021
Online: 16 February, 2021

Keywords:
Missions programming
Robotics
Modeling
Verification
Formal proof

ABSTRACT

This article copes with the formal verification of properties of the missions building module of PILOT's software. PILOT is a language dedicated to remote control of robots. An incremental syntax-oriented editor was built in order to increase the dependability of PILOT's missions and we showed that, under a maximum size of plan, this editor allows building only all plans that are syntactically correct. The limitation in size was due to state space explosion problem inherent to the Model-checking approach used for the proof. In order to extend the proof to all plans without any limitation in size, we investigated the theorem-proving approach, and especially PVS (Prototype Verification System). This paper therefore focuses more on modeling of PILOT plans and related building operations and the use of PVS to verify properties of the built models, in view of proving the aforementioned properties of PILOT software's missions building module.

1. Introduction

This paper extends the work originally presented in [1]. In the prolongation of preceding works aiming at enhancing the dependability of robotic applications [2-12], it targets the use of verification systems, and especially PVS (Prototype Verification System), for properties verification of a syntax-oriented editor of missions plans conceived for PILOT, a programming language devoted to the control of robots from remote. PILOT is developed within the Laboratory of Sciences and Techniques of Information, Communication and Knowledge (Lab-STICC). Proof systems rest on methods of formal proof or verification that can be described as approaches enabling to define systems properties and check their correctness using mathematics techniques and inference rules. In an earlier work, with the help of Prolog and University of Amsterdam's SWI-Prolog tool, we proved that the syntax-oriented editor allows building all and only plans having a correct syntax. Nevertheless, we could only do this proof for plans whose size was under a maximum limit, due to the explosion of state-space inherent to the model-checking approach implemented using the SWI-Prolog tool. Theorem proving approaches and related tools such PVS should make it possible to get rid of the constraint on plans size, and to extend the proof to all plans whatever their size. PVS has been chosen because of our experiment in its use for protocol verification [13, 14]. The formal verification of properties of PILOT's syntax-oriented editor involves a proper formalization of the language syntax and of the editor's working. This article is dedicated to the modeling of PILOT plans and building operations of plans, and the use of PVS to verify properties of the built models, in view of proving the correctness of PILOT's incremental syntax-oriented editor.

The rest of this paper is structured as follows. A state of the art of robotic systems' formal verification is presented in the second section. The third section addresses preliminaries on formal proof methods and gives an overview of PVS. The language PILOT as well as PILOT plans construction and checking method are presented in the fourth section. The fifth section deals with verification of properties with the help of SWI-Prolog tool. The sixth section is dedicated to the modeling of plans and building operations of plans, and verification of properties on the models using PVS. Conclusions are presented in the seventh and last section.

2. Related works

In [15], the authors reminded, firstly that formal methods of verification were only recently introduced in the control community to assist developers in the construction of complex robotic systems' control architectures, and secondly that the hybrid nature of such systems (combination of discrete and continuous behaviors) makes their formal treatment hard and necessitates new methods that are both operational and efficient.

In [16], the authors tackle issues and outlooks in robotics. One of the issues they pinpoint is long-term mission execution

*Corresponding Author: Laurent NANA, Univ. of Brest, Computer Science Department, 20 Avenue Le Gorgeu, 29238 Brest, +33298052283, nana@univ-brest.fr

capability. Despite the fact that dependability, formal verification and fault tolerance methods are not directly referred, they are needed to reach that objective.

In [17], the authors address the inclusion of operations into the control architecture of a UUV (unmanned underwater vehicle) for the mapping and the watching of the ocean. The proposed control system integrates safety mechanisms to lessen hazards at the Mission level and necessitates to use formal verification approaches.

In [18], the authors portray an approach for the use of formal tools to verify the controllers of unmanned autonomous vehicles (UAV) working in congested areas. They mention that the design of controllers of such systems is difficult, in consequence of the required speed and responsiveness. They also notify that, since it is costly to test and compare distinct UAV controllers, formal verification is the number one phase for their performance and efficiency optimization.

In [19], the authors propound a model-checking tool for user assistance in properties checking of systems having complex behaviors. The system modeling is done with the help of a finite-state automaton. The latter is very large and its manual verification is therefore very difficult. The model-checking tool helps the user by automating part of the verification process.

In [20], the authors propose a review of research works related to the application of formal specification and verification methods to the autonomous robotic domain. They pinpoint the insufficiency of testing on real deployment or simulation of autonomous robots. Indeed, the complexity of these robotic systems and their use in safety critical applications do not allow to test some of their behaviors. The authors claim the necessity to use formal methods to guarantee the exactness of such systems.

In [21], the authors propose an integration of the modeling language Timed Rebeca in ROS, a middleware for mobile robots' program development. A conceptual model of ROS programs in Timed Rebeca is proposed and used to verify the correctness of properties defined by users on ROS programs. Timed Rebeca is based on a model checking approach.

In [22], the authors deal with the formal proof of a robotic system. Model checking approach is applied as well as theorem proving approach of higher-order logic. The first one requires the discretization of the differential equations describing the continuous dynamics of the system and therefore limits the model to an abstracted view of the system, whereas the second one allows using it in its true form and makes it possible to take into account all the real possibilities of the dynamics of the system. Different techniques used for the analysis of the robotic system are compared, based on expressiveness, accuracy and automation.

Above works show the usefulness of formal verification approaches in robotics. They also show that solutions and tools are needed to facilitate the use of formal verification methods in the robotic domain. Analysis of state of the art reveals that the majority of works on formal verification of robotic systems use model checking techniques. In [20], only 3 between 49 works related to formal verification of robotic systems use theorem proving, despite their ability to avoid the state-space explosion issue that sometimes occur when using model checking. It is necessary to develop the use of theorem provers for robotic systems, given their complexity that makes their verification with model checkers prone to the state-space explosion problem.

In the next section, preliminaries on formal proof methods are first introduced, then the theorem prover PVS is presented.

3. Preliminaries on formal proof methods and Prototype Verification System

3.1. Formal proof methods

Two main formal proof methods exist: model-checking and demonstration also called theorem proving.

In the model checking approaches (illustration in figure 1), system's state-space model is built and its properties are first specified. Both are then input to the model checking system that goes through the state space exhaustively and checks if the system satisfies the given properties. When a property is not satisfied, the model checker generates error information.

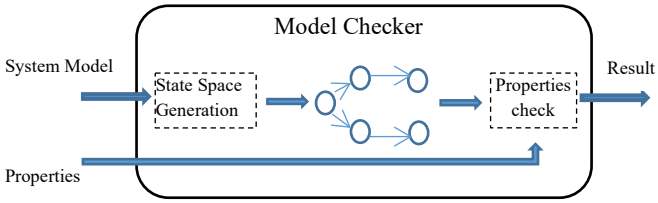


Figure 1: Model checking

In the theorem proving approaches (illustration in figure 2), system's mathematical model is built and its properties are formalized using a well-defined logic. Both are then input to the theorem proving system as theorems. Using axioms, hypothesis and deductive reasoning (inference rules), the theorem prover helps the user to develop his proof and to verify if the system satisfies the given properties. Inference rules together with theorems that have already been verified make it possible to prove new theorems. When the theorem proving relies on a decidable (propositional) logic, it can be automatic. When it is based on an un-decidable (higher-order) logic, it is interactive.

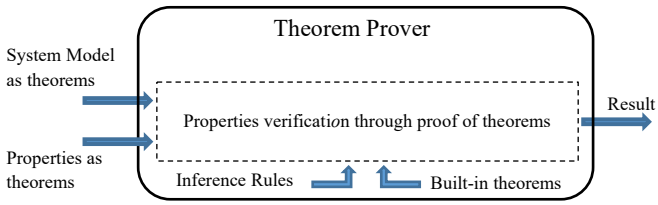


Figure 2: Theorem proving

In comparison to theorem proving, the advantage of model-checking is its entire automation. Its disadvantage is that it is subject to state-space size explosion that could cause problems due to limited computer memory space and limited computational resources.

3.2. Prototype Verification System

PVS [23] is a deductive verification system. It manages a tree of proofs and helps the user to build a full proof tree, i.e. a tree whose terminal nodes are all admitted as being true. Each non-terminal node is a goal from which children nodes are obtained by

a step of the proof. Goals are sequents having each the following shape:

$$p_1, p_2, \ldots, p_n \vdash q_1, q_2, \ldots, q_m$$

where the p_i are the antecedents and the q_i are the consequents.

PVS system uses backward reasoning: each proof step results in sequents that are at least as strong as previous ones. The root of the proof is the sequent $\vdash q$ where q is the theorem to prove.

PVS system furnishes a specification language that is strongly typed [24] as well as an interface that makes it possible to specify systems under Emacs, to formally specify properties on those systems and to prove them with the help of the proof system of PVS [25].

PVS's specification language is founded upon a logic of higher order. It provides various types and subtypes, including elementary ones (strings, numbers, predicates, etc.), abstract data and compound types (records, union, etc.). PVS system offers several useful functions for the process of formal verification. The main commands implementing the decision procedures are the following:

- *typecheck:* it makes it possible to analyze the file containing the specification of the system and to detect semantic errors.
- *prove:* it makes it possible to start the proof of properties that are non-trivial for the system.
- *flatten*: enables to flatten the structure of the current goal
- *split*: enables to separate a goal into subgoals
- *inst*: makes it possible to instantiate variables with given terms
- *expand*: enables to expand/develop a definition or an expression
- *skolem*: it makes it possible to "skolemize" quantified variables (quantifiers are removed and quantified variables are replaced by skolem constants).
- *grind/ground*: enable to launch the process of decision/simplification of a rule.
- *induct*: makes it possible to perform an induction on a variable.
- *undo*: enables to go back in the proof.
- *lemma*: makes it possible to add a lemma to the assumptions.
- More elaborated commands, such as *skosimp* that iterates the application of *skolem* and *flatten* commands on the current rule.

The next section presents the language PILOT and the method of construction and checking of plans.

4. Language PILOT and method of incremental construction and checking of missions' plans

4.1. The language PILOT

PILOT [5, 26, 27] is founded on the action concept. An action is composed of an order that the robot can execute, a rule of precondition and rules of supervising each having a tied treatment. An action is either elementary or continuous. Elementary actions end by themselves, in general once they reach their goal, while continuous actions' termination is provoked by a parallel or a preemption primitive. Whether elementary or continuous, an action only executes if its precondition is true. Similarly, when a supervision rule of an action becomes true during its execution, the corresponding treatment is launched. For each supervising rule, the default treatment consists in stopping the related action. Precondition and supervision rules are usually conditions expressed on values of sensors. Figure 3 shows the graphical representation of elementary and continuous actions.

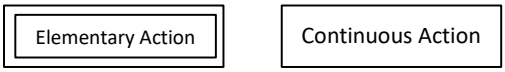


Figure 3: Elementary action and continuous action

In PILOT language, the following control primitives are available for missions' plans programming:

- Sequentiality: it starts by a "sequence beginning" and ends by a "sequence end". It enables to define an order of execution on other primitives of the language (actions and control primitives). A graphical symbol named "intersequence" is used to connect the primitives of the sequence. Figure 4 illustrates a sequence comprising 2 elementary actions.

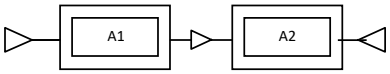


Figure 4: Sequentiality

- The conditional: it is formed by branches, each composed of a Boolean expression and a sequence following it. Starting from top, the first sequence whose condition is true is the only one to be executed. Figure 5 shows an example of conditional primitive with 2 alternatives.

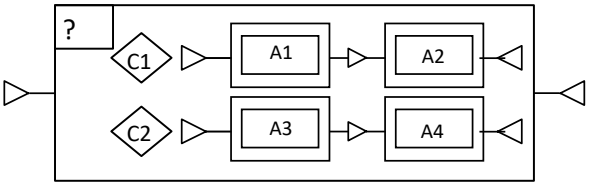


Figure 5: Conditional

- Iteration: it is constituted of a criterion of continuation and a sequence following it. Depending on the continuation criterion, the iteration is said to be fixed or to be conditional. In the first case, the criterion is a number of loops. In the second one, the criterion is a boolean condition. Figure 6 shows and example of fixed iteration on the left and an example of conditional iteration on the right. In the example of fixed iteration, elementary action A1 is executed 3 times. In the example of conditional iteration, elementary action A2 is executed while condition s>2 is true.

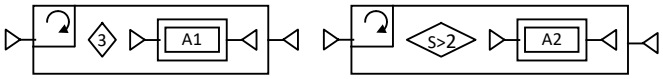


Figure 6: Fixed and conditional iterations

- Parallelism: it is composed of sequences executed in parallel. It terminates its execution when all its sequences have ended theirs. Figure 7 illustrates an example of parallel primitive. In this

example, elementary actions A1 and A2 are executed in parallel and the parallel execution ends when both actions reach the end of their execution.

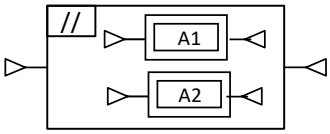


Figure 7: Parallelism

• Preemption: it is composed of sequences executing in parallel, but unlike parallelism structure, the termination of its execution occurs as soon as one of the sequences ends. Figure 8 illustrates an example of preemption with 2 sequences. In this example, once one of the sequences ends, it causes the termination of the second one and leads to the end of the parallel execution.

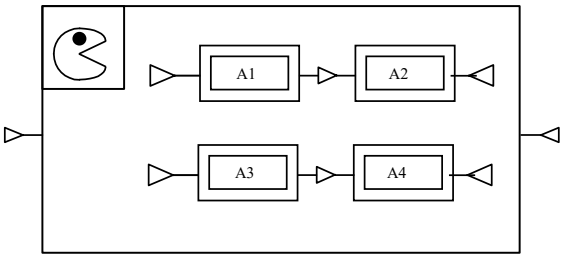


Figure 8: Preemption

After the above presentation of the language PILOT, the next subsection describes the incremental syntax-oriented building approach of missions' plans.

4.2. Method of construction and checking of PILOT missions' plans

The syntax-oriented edition of PILOT plans guarantees that the plan is syntactically correct at each step of plan building. When the programmer starts building a mission plan, he obtains a sequence that is empty. In order to continue the construction of the plan, he uses operations *insert*, *erase* or *modify*. The *modify* operation makes it possible to change elements such as the Boolean expression of a conditional primitive, etc. Whenever an operation is applied by the user, the editor checks the syntactic validity of the plan obtained and only takes the modification into account in the positive case. When the syntax is incorrect, the editor informs the user through a message. For the purpose of ensuring the syntactic correctness of the plan whenever insertion operation is performed, the following default primitives are associated with the structures of PILOT language:

- The default conditional structure which contains a unique alternative consisting in a condition set to "false" followed by an empty sequence.
- The default iteration structure that contains a continuation criterion set to 0 followed by an empty sequence.
- The default parallelism structure which contains a unique empty sequence.
- The default preemption structure that contains a unique empty sequence.

These default structures are illustrated in the example of plan shown in Figure 11.

Assuming that:

- *plan* is the current mission plan,
- *elt* is the element to insert in the plan,
- *sel* is the element of the plan selected by the user to indicate where to insert the new element (i.e. he wants *elt* to be inserted just before *sel*),
- *cont* represents the immediate encapsulating structure containing *sel*.
- Type *(*<param>*)* represents the type of <param>. Type values are BS, ES, BE, NL, EA, CA, CP, IP, PP. They respectively correspond to "Beginning of Sequence", "End of Sequence", Boolean Expression", "Number of Loops", "Elementary Action", "Continuous Action", "Conditional Primitive", "Iteration Primitive", and "Parallel Primitive". Here, parallelism and preemption primitives are considered of type PP.
- PredecessorOf (<param>) represents the element preceding <param> in the plan.
- SetOfSequencesOf (<param>) defines the set of sequences of <param>. In this case <param> is supposed to be a parallel primitive.
- IsContinuousActionSequence (<param>) is true if <param> is a sequence made of a beginning of sequence, followed by a continuous action followed by an end of sequence. Otherwise, it is false.

The precondition of the implemented insertion operation can be represented as follows:

$$
\begin{aligned}
&\exists\ elt \wedge \\
&(\neg\exists\ sel \Rightarrow (\exists\ cont \Rightarrow (\text{Type}\ (cont) \neq \text{IP} \wedge \\
&\quad (\text{Type}\ (cont) = \text{PP} \Rightarrow \text{Type}\ (elt) = \text{BS}) \wedge \\
&\quad (\text{Type}\ (cont) = \text{CP} \Rightarrow \text{Type}\ (elt) = \text{BE})))) \wedge \\
&(\exists\ sel \Rightarrow ((\text{Type}\ (sel) = \text{BS} \Rightarrow \\
&\quad (\exists\ cont \wedge \text{Type}\ (cont) = \text{PP} \wedge \text{Type}\ (elt) = \text{BS})) \wedge \\
&\quad (\text{Type}\ (sel) = \text{BE} \Rightarrow \\
&\quad (\exists\ cont \wedge \text{Type}\ (cont) = \text{CP} \wedge \text{Type}\ (elt) = \text{BE})) \wedge (\text{Type}\ (sel) \neq \text{NL}))) \wedge \\
&(\text{Type}\ (elt) = \text{CA} \Rightarrow \\
&\quad ((\exists\ cont \wedge \text{Type}\ (cont) = \text{PP}) \wedge \\
&\quad (\exists\ sel \wedge \text{Type}\ (sel) = \text{ES} \wedge\ \text{Type}\ (\text{PredecessorOf}\ (sel)) = \text{BS}) \wedge \\
&\quad (\exists\ \text{seq} / \text{seq} \in \text{SetOfSequencesOf}(cont) \wedge \text{sel} \notin \text{seq} \wedge \\
&\quad \neg\text{IsContinuousActionSequence}\ (\text{seq}))))
\end{aligned}
$$

After this presentation of PILOT and the description of the method of construction and checking of PILOT missions' plans, the next section deals with properties' proof of the latter, based on Prolog and an associated tool.

5. Verification of the method of construction and checking of PILOT missions' plans using Prolog and an associated tool

It is necessary demonstrating that the implementation of the syntax-oriented editor enables the construction of all but solely plans whose syntax is valid. For the sake of simplicity, only the

case of plans building using insertion operations is considered. In this case, the approach used for the validation can be illustrated by the colored Petri net of Figure 9. In colored Petri nets [28], colors are represented by associating values to tokens.

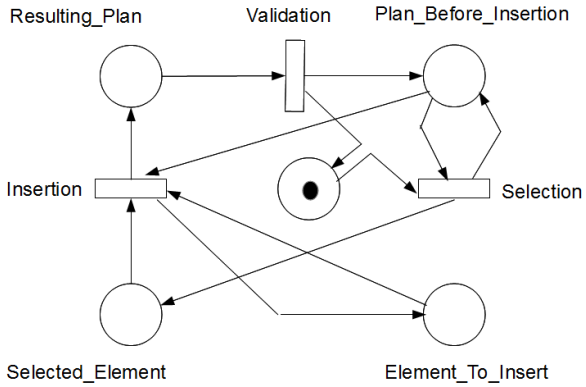


Figure 9: Validation approach

The initial marking of places "Plan_Before_Insertion" and "Element_To_Insert" are respectively {[beg_seq, end_seq]} and {beg_seq, end_seq, cont_act, elem_act, par, pre, iter, cond, bool_exp, loops_num}. Initially, place "Selected_Element" has an empty marking and so is place "Resulting_Plan".

The transition "Selection" (respectively "Insertion") models an element selection (respectively insertion) of (respectively in) the plan. As far as transition "Validation" is concerned, it represents the formal syntax analyzer of the language PILOT. The mechanism implemented for the syntax-oriented building of plans is free of errors if the Petri net of Figure 9 does not contain any deadlock.

For the application of this certification method to the syntax-oriented editor of PILOT plans, a syntax-checker has been built in PROLOG based on the approach suggested in [29]. Thereafter, the plan construction has been modeled and the properties hereafter have been checked:

- Are there insertions that lead to plans whose syntax is not correct?
- Are there plans with correct syntax whose construction is not possible with the proposed insertion model?

The first property aims at ensuring that the proposed syntax-oriented building mechanism disallows the construction of plans whose syntax is not correct. Regarding the second property, the goal is to make sure that the mechanism does not disallow the building of correct plans. Indeed, the checking mechanism may be too restrictive and lead to the rejection of plans that are syntactically correct.

Due to the working of PROLOG, it was necessary to limit the size of the set of plans built. If not, the PROLOG tool would have tried to generate all syntactically correct plans and this would have led to memory space problems. PROLOG parser and plan construction models were therefore modified to solely generate plans of size under a threshold. Here, the size is that of the list modeling the plan. Figure 10 shows the translation of the above properties in PROLOG, taking the size constraint into account. In this PROLOG code, *convert* converts a model representing the graphic plan into a model which the syntax analyzer can process. *convert_set* works similarly, but applies to set of models.

```
inserion_issue (Size_Max, NG_Plan, G_Plan):-
 automated_insertions (Size_Max, G_Plans_List),
 member (G_Plan, G_Plans_List),
 convert (G_Plan, NG_Plan),
 not valid_syntax (NG_Plan,[]).
anomalous_rejection (Size_Max, NG_Plan):-
 automated_insertions (Size_Max, G_Plans_Set),
 convert_set (G_Plans_Set, NG_Plans_Set),
 valid_syntax_plans (Size_Max,Valid_NG_Plans_Set),
 member (NG_Plan, Valid_NG_Plans_Set),
 not member (NG_Plan, NG_Plans_Set).
```

Figure 10: Syntax analyzer' properties validity translation in Prolog

SWI-Prolog tool of University of Amsterdam were used for the programming. The tests performed for plans of size under 15 showed no insertion problem and no anomalous rejection. Consequently, for plans whose size is under 15, the proposed syntax-oriented building mechanism enables constructing only but all plans whose syntax is valid. For greater sizes, an exception related to lack of memory space is raised. The lack of space is caused by the exponential increase of the set of plans.

The next section presents the modeling of PILOT plans' properties as well as the modeling of plans' building operations, for their verification using Prototype Verification System. As indicated in the third section, Prototype Verification System offers a theorem proving system which enables to avoid the aforementioned memory space problem.

6. Plans and operations modeling for properties verification using Prototype Verification System

6.1. Prolog approach model analysis

In the approach used for PILOT's syntax-oriented editor properties' proof using SWI-Prolog, plan's model is a numbered elements list, so as to be able to locate components of the plan as in its graphical version. Indeed, for operations such as insertion, it is necessary identifying where insertion is wished. In PILOT's Graphical User Interface (Figure 11), in order to insert an element in a plan, the user selects the element among the operators list on the left, then he clicks on the location of the plan before which he wants to insert the element.

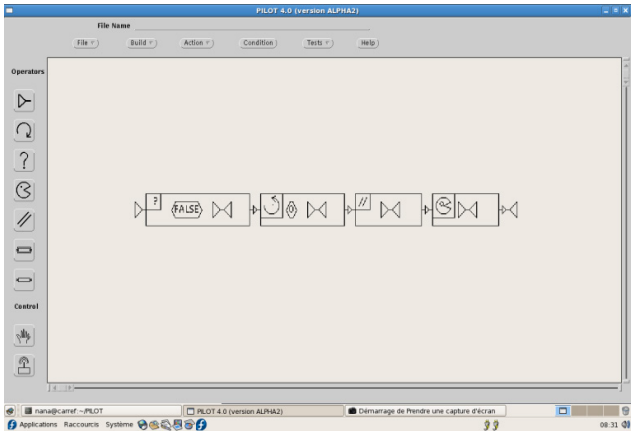


Figure 11: Graphical User Interface of PILOT

The use of numbers to locate the components of the plan is suitable to produce all the plans buildable by insertion of a unique primitive of the language within a plan. Nonetheless, it becomes an obstacle to the generalization needed for theorem proving with the help of Prototype Verification System, the reasoning being done independently of the size of the plan. One of the main difficulties was therefore to find a model of representation of the plan, not based on the numbering of its elements, and making it possible to uniquely locate its elements.

Building and syntax-checking operations necessitate identification of each plan element's container. Indeed, the behavior of operations such as insertion is container-dependent. Mechanisms such as that provided by SWI-Prolog through the notation "par: number: List" make it easy to isolate blocks and locate the containers. PVS offers no similar mechanism. Therefore, the design of a solution for containers' identification is needed for PVS.

6.2. Models proposal for the proof of properties with PVS

Proving PILOT editor's properties with PVS, necessitates defining various models notably for PILOT plan, selection's container and operation "selection" that specifies where to perform the insertion in the plan. The models proposed for these different entities are presented hereafter.

a) Plan model

The plan is represented as a list of elements belonging to the set {*bs, es, e, c, exp, par, pre, condi, iter*} (see example on Figure 12). These elements refer to PILOT' syntax terminals. *bs* denotes sequence beginning, *es* sequence end, *e* (respectively *c*) elementary (respectively continuous) action, *exp* expression of conditional. *pre, par*, *condi* and *iter* respectively denote preemption, parallelism, conditional and iteration primitives.

Plan	**PVS model**
Empty sequence	cons (bs, cons (es, null))
Sequence with a single conditional primitive	cons (bs, cons (condi, cons (exp, cons (bs, cons (es, cons (es, null))))))
Sequence with a single elementary action	cons (bs, cons (e, cons (es, null)))

Figure 12: Examples of plans modeled in PVS

b) Modeling properties of syntactically correct plans

In the proposed approach, each rule of the grammar of the language is represented using a two parameters function whose first parameter is the part of the plan submitted for parsing and the second is the remainder expected after retrieving the pattern matching the rule. The output of the function is true when the real and expected remainders are identical. Otherwise, it is false. In Prolog's case, this approach enables to automatically generate the program that recognizes the language, as shown in Table 1.

Using the same approach under PVS, the program recognizing the language PILOT has the shape shown in Figure 13.

It turns out that this solution cannot be used as such under PVS. Indeed, in PVS's description language, a function call can only be done after the definition of the function. So, in figure 13, the call to seq_base in the definition of validplan isn't concretely allowed. An approach to solve this problem could be just to specify (declare) the function before the call without defining it completely, but it is disallowed in Prototype Verification System.

A first method adopted to overpower these drawbacks has been to elaborate PILOT syntax rules' dependencies graph, then using it to define the functions in the order of dependency, beginning by the functions not involving cross recursive calls, and then using the passing of functions as parameters to define functions involved in calls with cross recursion. The principle is shown hereafter through a definition of even and odd parities:

Table 1: Description of PILOT's syntax with Prolog

PILOT Syntax rules		Description in Prolog	
S	: SEQ_BASE	validplan (A,B)	:- seq_base (A,B).
SEQ_BASE	: bs L_PRIMI_BASE es	seq_base ([bs|A],B)	:- l_primi_base (A, [es,B]).
L_PRIMI_BASE	: ε |	l_primi_base (A,A).	
	PRIMI_BASE L_PRIMI_BASE	l_primi_base (A,B)	:- primi_base (A,C), l_primi_base (C,B).
PRIMI_BASE	: PRIMI_PARALLEL |	primi_base (A,B)	:- primi_parallel (A,B).
	PRIMI_PREEMPTION |	primi_base (A,B)	:- primi_preemption (A,B).
	PRIMI_CONDITIONAL |	primi_base (A,B)	:- primi_conditional (A,B).
	PRIMI_ITERATION |	primi_base (A,B)	:- primi_iteration (A,B).
	PRIMI_ACT_ELEM	primi_base (A,B)	:- primi_act_elem (A,B).
PRIMI_PARALLEL	: par '(' LIST_SEQ ')'	primi_parallel ([par,'('|A],B)	:- list_seq (A, [')'| B]).
LIST_SEQ	: LIST_SEQ_A SEQ_BASE LIST_SEQ_A	list_seq (A,B)	:- list_seq_a (A,C), seq_base(C,D), list_seq_a (D, B).
LIST_SEQ_A	: ε |	list_seq_a (A,A).	
	SEQ_LIST_SEQ_A	list_seq_a (A,B)	:- seq (A, C), list_seq_a (C,B).
SEQ	: SEQ_BASE |	seq (A,B)	:- seq_base (A,B).
	SEQ_SPECIFIC	seq (A,B)	:- seq_specific (A,B).
SEQ_SPECIFIC	: bs PRIMI_ACT_CONT es	seq_specific ([bs|A],B)	:- primi_act_cont (A,[es|B]).
PRIMI_ACT_CONT	: c	primi_act_cont ([c|A],A).	
PRIMI_PREEMPTION	: pre '(' LIST_SEQ ')'	primi_preemption ([pre,'('|A],B)	:- list_seq (A, [')'| B]).
PRIMI_CONDITIONAL	: condi '(' LIST_CONDITIONAL ')'	primi_conditional ([cond,'('|A],B)	:- list_conditional (A,[')',B]).
LIST_CONDITIONAL	: CONDITIONAL LIST_CONDITIONAL |	list_conditional (A,B)	:- conditional (A,C), list_conditional (C,B).
	CONDITIONAL	list_conditional (A,B)	:- conditional (A,B).
CONDITIONAL	: exp SEQ_BASE	conditional ([exp|A],B)	:- seq_base (A,B).
PRIMI_ITERATION	: iter '(' SPEC_ITER ')'	primi_iteration ([iter,'('|A],B)	:- spec_iter (A, [')'|B]).
SPEC_ITER	: CONDITIONAL |	spec_iter (A,B)	:- conditional (A,B).
	nb SEQ_BASE	spec_iter ([nb|A],B)	:- seq_base (A, B).
PRIMI_ACT_ELEM	: e	primi_act_elem ([e|A],A).	

```
% definition of types and declaration of variables
primi: type = {bs, es, …}
t: type = list [primi]
a, b, c, x, y: t
% representation of rules
validplan (a, b): bool = seq_base (a, b)
seq_base (a, b): bool = (car (a) = bs) and l_primi_base (cdr (a),
    cons (es, b))
l_primi_base (a, b): recursive bool = ((a = b) or (a /= b and a /=
    null and exists c: (primi_base (a, c) and l_primi_base (c, b))))
    measure lambda (x, y): length (x)
primi_base (a, b): bool = primi_parallel (a, b) or
    primi_preemption (a, b) or primi_conditional (a, b) or
    primi_iteration (a, b) or primi_act_elem (a, b)
```

Figure 13: Specification of PILOT syntax in PVS with same method as Prolog

Even (n: N): Boolean = n=0 or (n ≠0 and odd (n-1))

Odd (n: N): Boolean = n=1 or (n ≠1 and even (n-1))

This definition of even and odd functions can be transformed in PVS as shown in figure 14.

```
Even_1 (n: N, od: N -> Boolean): Boolean  = n = 0 or (n ≠ 1
   and od (n-1))
Odd (n : N) : Boolean = n=1 or (n ≠ 1 and Even_1 (n-1,
   Odd))
Even (n: N) : Boolean = Even_1 (n, Odd)
```

Figure 14: Representation of even and odd functions for PVS with resolution of recursive crossed calls

Nevertheless, the syntax of PILOT contains several complex crossed calls involving more than 2 crossed calls. For example, “seq_base” refers to “l_primi_base” that refers to “primi_base” which refers in turn to “primi_parallel” that also refers to “seq_base”.

Under PVS, defining a function that is recursive also necessitates adding a function called *measure* which decreases in the course of recursive calls and having a down side limit, so as to ensure recursive calls ending.

Figure 15 illustrates the representation proposed in order to solve the problem related to recursive cross calls for the example mentioned above.

A representation of PILOT’s syntax under PVS has been generated with the help of this approach. For its validation, the obligations of proofs produced by Prototype Verification System were proven and a few plans’ definition theorems, among which those in Figure 16, have been defined and proved with the help of PVS.

a) Modeling of plan element selection

The intention is to design a model representing the point of insertion in the PILOT plan. The following two components can be used for characterizing the point of insertion: the clicked element and its container. The proposed model of representation of the selected element is the couple (Lbe, Laf) where Lbe corresponds to the head of the plan list ending just before the selected element, and Laf is the remaining of the plan starting from the selected element. This solution makes it possible to take advantage of the properties of concatenation of lists in the design of the models of operations that apply to the selected element. It also provides a good consistency with the syntactic rules modeling where two parts of the list representing the plan have to be distinguished, namely the part recognized by the rule and the remainder.

```
primi_parallel_i (a, b, lsq): bool = (a/= null and b /= null and
   car (a) = par and lsq (cdr (a), b))
primi_base_i (a, b, lsq): bool = primi_act_elem (a, b) or
   primi_parallel_i (a, b, lsq) or …
l_primi_base_i (lsq) (a, b): recursive bool = (a = b) or (a /=
   b and exists c: (primi_base_i (a, c, lsq) and
   l_primi_base_i (lsq) (c, b)) measure lambda (lsqf) (x, y):
   length (x)
seq_specific (a, b): bool = a = cons (bs, cons (c, cons (es,
   b)))
seq_base_i (a, b, lsq): bool = a /= null and car (a) = bs and
   l_primi_base_i (lsq) (cdr(a), cons (es, b))
seq_i (a, b, lsq): bool = seq_base_i (a, b, lsq) or seq_specific
   (a, b)
l_seq_a (a, b): recursive bool = a = b or (a /= b and exists c:
   seq_i (a, c, l_seq_a) and l_seq_a (c, b))
measure lambda (x, y): length (x)
l_seq (a, b): bool = exists c, d: (l_seq_a (a, c) and
   seq_base_i (c, d, l_seq_a) and l_seq_a (d, b))
seq_base (a, b): bool = seq_base_i (a, b, l_seq)
validplan (a): bool = seq_base (a, null)
```

Figure 15: Excerpt of PILOT’s syntax description in PVS with cross recursive calls solving

In the proposed modeling, the selected element’s container is frequently split in two parts, one belonging to Lbe and the other to Laf. Plan element selection model is shown in Figure 17. In this model, Lcon is the tail of the plan that starts from the selected element’s container.

An extract of the PVS model of the selection operation is given in Figure 18. It shows the specification of the selection of an outermost plan’s element (i.e. an element of the main sequence of the plan), as well as the case of a selection within a parallel box. In the first case, there is no container, whereas in the second case, the parallel box is the container.

```
P01: theorem validplan (cons (bs, cons (es, null)))
P02: theorem validplan (cons (bs, cons (e, cons (es, null))))
P03: theorem validplan (cons (bs, cons (e, cons (e, cons (es,
     null)))))
P04: theorem validplan (cons (bs, cons (par, cons (bs, cons (es,
     cons (es, null))))))
P05: theorem validplan (cons (bs, cons (par, cons (bs, cons (es,
     cons (e, cons (es, null)))))))
P06: theorem validplan (cons (bs, cons (par, cons (bs, cons (e,
     cons (es, cons (bs, cons (es, cons (es, null)))))))))
P07: theorem validplan (cons (bs, cons (condi, cons (exp, cons
     (bs, cons (es, cons (es, null)))))))
P08: theorem validplan (cons (bs, cons (e, cons (condi, cons
     (exp, cons (bs, cons (e, cons (es, cons (e, cons (es,
     null)))))))))))
P09: theorem validplan (cons (bs, cons (iter, cons (exp, cons (bs,
     cons (es, cons (es, null)))))))
P10: theorem validplan (cons (bs, cons (iter, cons (exp, cons (bs,
     cons (e, cons (es, cons (e, cons (es, null)))))))))
```

Figure 16: Samples of PVS demonstrated theorems

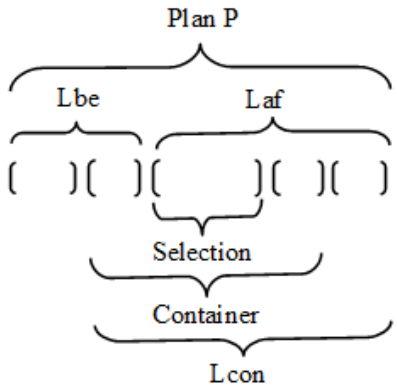


Figure 17: Selection model of a plan element

In this selection operation, “fusion” is an operation that performs the fusion of two lists and “seq” is such that its first parameter is a list L starting by a sequence S followed by a list LS2 of primitives.

7. Conclusion

As shown in the state of the art part of this paper, it is necessary to develop the use of theorem provers for robotic systems, given their complexity that makes their verification with model checkers prone to the state-space explosion problem. This work is a step forward in this direction, through a case study showing the limits of model-checking for the proof of properties of a robotic system, namely the syntax-oriented editor of a robotic missions programming language (PILOT), and investigating the use of the PVS theorem proving system to overcome the problem encountered with the model-checking approach.

Another goal of the work was to check out the appropriateness of Prototype Verification System for accomplishing complicated demonstrations that involve the design of models for different kinds of entities frequently encountered in robotics and more generally in control-command systems (language syntax, operations ...).

```
selection (P, Lbe, Laf, Lcon): recursive bool =
 P = fusion (Lbe, Laf) and
 ( % sequence element selected
  ( Lcon = null and (Laf = cons (es, null) or
   % sequence beginning selected
   Lbe = null or
   % another element of sequence selected
   (l_primi_base (Laf, cons (es, null))))) or
  % selection into a parallel box of the sequence
  % w: parallel box, x: list of primitives after the parallel box
  ( exists Lpar, Lbepar, w, x, y :
   ( P = fusion (Lbepar, Lpar) and primi_par (Lpar, w, x) and
    l_primi_base (x, y, cons (es, null)) and
    ( ( Laf = Lpar and Lcon = Lpar and Lbe = Lbepar) or
     % S: sequence of the parallel box
     % Ls1: list of sequences before S
     % Ls2: list of primitives after S
     % Lcon1: list from the container to the end of S
     ( Exists S, Ls1, Ls2, L, Lbe1, Laf1, Lcon1:
      ( list_seq_a (cdr (Lpar), Ls1, L) and
       seq (L, S, Ls2) and
       selection (S, Lbe, Laf1, Lcon1) and
       ( % element at the top level of S selected
        ( Lcon1 = null and Lcon = Lpar) or
        % selection of an element within a box of the
       sequence S
        ( Lcon1 /= null and Lcon = fusion (Lcon1, Ls2))
       ) and
       ( Lbe = fusion (Lbepar, cons (par, fusion (Ls1, Lbe1)))
      and
        Laf = fusion (Laf1, Ls2))))))
 ) or
 % code for the other encapsulating primitives
 …
)
measure lambda (s, s1, s2, s3): length (s)
```

Figure 18: Extract of selection operation’s PVS model

In order to reach these goals, PVS functions have been used to represent PILOT’s syntax. The proposed model has been validated notably by verifying the obligations of proofs produced by Prototype Verification System and proving some theorems defined on PILOT plans. The modeling of recursive cross-calls has been the major issue faced during PILOT’s syntax specification in PVS.

As matter of fact, contrary to Prolog, the language of specification provided by PVS disallows to use a function before its full definition. Consequently, a PVS model without cross recursivity was proposed. The specification in PVS of the selection operation of the syntax-oriented editor of PILOT has also been presented in this paper.

The results obtained at this step of the work (the representation model of the syntax rules of the language PILOT under PVS and its validation, the modeling of operations of the syntax-oriented editor of PILOT), enables to conclude in a good capacity of PVS for the modeling of aforementioned entities of control-command systems and for the achievement of further formal proofs.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by Lab-STICC (UMR CNRS 6285) and University of Brest.

References

[1] L. Nana, F. Monin, S. Gire, "Proof of properties of a syntax analyzer of robotic mission plans" in Proceedings of 4th International Conference and Workshops on Recent Advances and Innovations in Engineering –ICRAIE 2019, IEEE, Kedah, Malaysia, 2019.

[2] M. Barbier, J. F. Gabard, D. Vizcaino, O Bonnet-Torrès, "ProCoSA: a software package for autonomous systems supervision" in Proceedings of 1st National Workshop on Control Architectures of Robots: software approaches and issues, Montpellier, France, 2006.

[3] C. Barrouil, J. Lemaire, "Advanced Real-Time Mission Management for an AUV" in Proceedings of SCI NATO RESTRICTED Symposium on Advanced Mission Management and System Integration Technologies for improved Tactical Operations, Florence, Italy, 1999.

[4] L. Laouamer, A. Benhocine, L. Nana, A Pascu, "Motion JPEG Video Authentication based on Quantization Matrix Watermarking: Application in Robotics" Int. Journal Computer Application, **47**(24), 1-5, 2012.

[5] L. Nana, "Investigating safety mechanisms for robotics applications" IPSI BGD Transaction Internet Res., **3**(1), 45-50, 2006.

[6] L. Nana, L. Marcé, J. Opderbecke, M. Perrier, V. Rigaud, "Investigation of safety mechanisms for oceanographic AUV missions programming" in Proceedings of the IEEE OCEANS'05 Europe Conference, Brest, France, 2005.

[7] L. Nana, F. Singhoff, J. Legrand, J. Vareille, P. Le Parc, F. Monin, D. Massé, L. Marcé, J. Opderbecke, M. Perrier, V. Rigaud, "Embedded intelligent supervision and piloting for oceanic AUV" in Proceedings of the IEEE OCEANS'05 Europe Conference, Brest, France, 2005.

[8] L. Nana, J. Legrand, F. Singhoff, L. Marcé, "Modelling and Testing of PILOT Plans Interpretation Algorithms" in Proceedings of Multi-conference on Computational Engineering in Systems Applications, CESA'03, IEEE, Lille, France, 2003.

[9] L. Nana Tchamnda, V-A. Nicolas, L. Marcé, "Towards a formal approach for the regeneration of PILOT control system" in Proceedings of 6th World Multiconference on Systemics, Cybernetics and Informatics, SCI'2002, IEEE Venezuela, Orlando, Florida, USA, 2002.

[10] F. Py, F. Ingrand, "Dependable Execution Control for Autonomous Robot" in Proceedings of IROS 2004 (IEEE/RSJ International Conference on Intelligent Robots and Systems), Sendai, Japan, 2004.

[11] E. Rutten, "A framework for using discrete control synthesis in safe robotic programming", Research report, INRIA, 2000.

[12] N. Turro N, "MaestRo: Une approche formelle pour la programmation d'applications robotiques", PhD Thesis, Université de Nice, Sophia Antipolis, 1999.

[13] J. F. Groote, F. Monin, J. C. Van de Pol, "Checking verifications of protocols and distributed systems by computer" in Proceedings of Concur'98, Sophia Antipolis, France, 1998.

[14] J.F. Groote, F. Monin, J. Springintveld, "A computer checked algebraic verification of a distributed summation algorithm". Form. Asp. Comput. **17**(1), 2005, 19–37. DOI:https://doi.org/10.1007/s00165-004-0052-7.

[15] D. Bresolin, L. Di Guglielmo, L. Geretti, R. Muradore, P. Fiorini, T. Villa, "Open problems in verification and refinement of autonomous robotic systems" in Proceedings of the 15th Euromicro Conference on Digital System Design, pp. 469-476, Cesme, Izmir, Turkey, 2012.

[16] E. Zereik, M. Bibuli, N. Miskovic, P. Ridao, A. Pascoal, "Challenges and future trends in marine robotics" Annu. Rev. Control, **46**, 2018, 350-368, ISSN 1367-5788. https://doi.org/10.1016/j.arcontrol.2018.10.002.

[17] M. Ludvigsen, A. J. Sørensen, "Towards integrated autonomous underwater operations for ocean mapping and monitoring". Annu. Rev. Control, **42**, 145-157, 2016.

[18] A. J. Barry, A. Majumdar, R. Tedrake, "Safety verification of reactive controllers for UAV flight in cluttered environments using barrier certificates" in Proceedings of 2012 IEEE International Conference on Robotics and Automation, Saint Paul, MN, USA, 484-490, 2012. doi: 10.1109/ICRA.2012.6225351

[19] C. Armbrust, L. Kiekbusch, T. Ropertz, K. Berns, "Tool-assisted verification of behavior networks" In proceedings of 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 1813-1820, 2013. doi: 10.1109/ICRA.2013.6630816

[20] M. Luckcuck, M. Farrell, L. A. Dennis, C. Dixon, M. Fisher, "Formal Specification and Verification of Autonomous Robotic Systems: A Survey" ACM Comput. Surv. **52**(5), 2019, DOI: https://doi.org/10.1145/3342355

[21] S. Dehnavi, A. Sedaghatbaf, B. Salmani, M. Sirjani, M. Kargahi, E. Khamespanah. "Towards an Actor-based Approach to Design Verified ROS-based Robotic Programs using Rebeca" Procedia Computer Science, **155**, 59-68, 2019, https://doi.org/10.1016/j.procs.2019.08.012.

[22] A. Rashid, O. Hasan, I. T. Bhatti, Formal Verification of Robotic Cell Injection Systems, Editor(s): Ahmad Taher Azar, Control Systems Design of Bio-Robotics and Bio-mechatronics with advanced applications, Academic Press, 2020.

[23] S. Owre N. Shankar, J. M. Rushby, D. W. J. Stringer-Calvert, "PVS System Guide", Technical Report, SRI International, Menlo Park, CA, 1999.

[24] S. Owre N. Shankar, J. M. Rushby, D. W. J. Stringer-Calvert, "PVS Language Reference", Technical Report, SRI International, Menlo Park, CA, 2001.

[25] N. Shankar S. Owre, J. M. Rushby, and D. W. J. Stringer-Calvert, "PVS Prover Guide", Technical Report, SRI International, Menlo Park, CA, 1999.

[26] J.-L. Fleureau, "Vers une méthodologie de programmation d'un système de télérobotique : comparaison des approches PILOT et Grafcet", PhD Thesis, Université de Rennes 1, 1998.

[27] E. Le Rest, "PILOT : un langage pour la télérobotique", PhD Thesis, Université de Rennes 1, 1996.

[28] K. Jensen, Coloured Petri Nets - Basic Concepts, Analysis Methods and Practical Use, Springer-Verlag Berlin Heidelberg, 1997

[29] F. Giannesini, H. Kanoui, R. Pasero, M. Van Caneghem, Prolog, InterEditions, 1985.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 2, 1065-1073 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Gripper Finger Design for Automatic Bottle Opener

Suchada Sitjongsataporn[*,1], Kornika Moolpho[2] and Sethakarn Prongnuch[3]

[1]*Department of Electronic Engineering, Mahanakorn Institute of Innovation (MII), Faculty of Engineering and Technology, Mahanakorn University of Technology, Bangkok, 10530, Thailand*

[2]*Department of Mechatronic Engineering and Automation, Mahanakorn Institute of Innovation (MII), Faculty of Engineering and Technology, Mahanakorn University of Technology, Bangkok, 10530, Thailand*

[3]*Department of Computer Engineering, Faculty of Industrial Technology, Suan Sunandha Rajabhat University, Bangkok, 10300, Thailand*

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 12 April, 2021
Online: 28 April, 2021

Keywords:
Parallel Gripper Finger
Robotic Dual-Arm
Automatic Bottle Opener

ABSTRACT

This paper presents a design of parallel gripper finger for robotic dual-arm working with an automatic push-down bottle crown cork cap opener on ABB's Yumi collaborative bartender robot. A safe gripper finger is made from ABS plastic by 3D printing for human-based interaction design for grasping and holding a glass bottle. Rack design and proposed automatic push-down bottle cap opener using pneumatics are presented to support a gripper finger. Experimental tests as bartender environment with 4-different types of carbonated soft drink with crown cork cap show that can be achieved effectively with average of 91.5% percentage of successful cap opener.

1 Introduction

In recent years, many types of safe collaborative robot with single arm and dual-arm intended for direct collaboration with human-based workers in the industrial automation with incrementally superior capacities have been invented to handle bimanual tasks in the product assembly [1]- [4]. Robot can assist humans in many tasks [5], [6]. Performance of picking and placing objects using robotics technology is an important objective of industrial automation for picking and packing boxes [3], [4] with computer-based vision [7]. Robot Arm is being used the manipulation of static objects that are fast and proven to be reliable. There are many types of robotic arm used. A picking robot arm is scheduled to pick a task in the specific working space to enhance the stability and movement with a specific trajectory points [8]. In [9], the authors have presented a cable-driven underactuated robotic gripper, which is designed for adaptable picking objects in different shapes, weights, sizes, and textures. ABB's Yumi robot has been reported many successful efforts to improve dual-arm operations for synchronized programming in industrial tasks detailed in [3]. Kinematic plan for conventional industrial robotic arms are capable of moving the static object manipulation. However, there is a trouble when moving object manipulation such as a smooth picking up a bottle and a grasping motion [10]. Our inspiration of this research came from the Japanese news that the Japanese owner of restaurants and shops are struggling to hire staff in an aging society. Then in 2020, the Japan's first robot bartender [11] has begun serving up drinks in a Tokyo pub with the industrial robot. In order to avoid jerking the object and figuring out the reasonable time in safety, we present a design of gripper for industrial ABB's Yumi robotic dual-arm.

To enhance the gripping performance, the grippers perform a pick-place process by using the force to compensate the gravitational force of objects in motion [12].

Moreover, the grippers have to be mated with objects to ensure that the objects do not either fall off or get damage from the grippers. The 2-finger parallel gripper was chose because it is the most flexible design and is able to carry out most percentage of applications. The ABB's YuMi is offered gripper options. The basic function of the option is to grasp parts using a parallel grip [13]. There are some inventions related to a crown cap bottle opener [14],[15]. A conventional bottle opener is a specialised lever inserted beneath the pleated metalwork pulling it off when upward force is applied to the handle end of the opener. Autonomous bottle opener robot [14] has presented with a simple bottle opener to assist the bartender. Without suffering from wrist motion disorder while opening, a push bottle opener [15] includes a mechanism configured to remove a

[*]Corresponding Author: Suchada Sitjongsataporn, Mahanakorn University of Technology, ssuchada@mut.ac.th

bottle cap when the opener is introduced over a bottle and pushed downwardly.

To our best knowledge, a design of gripper finger on ABB's Yumi collaborative robot is presented for human-based tasks performing automatic manipulation collaboration with an automatic push-down glass bottle opener using pneumatics in order to operate on grasped objects.

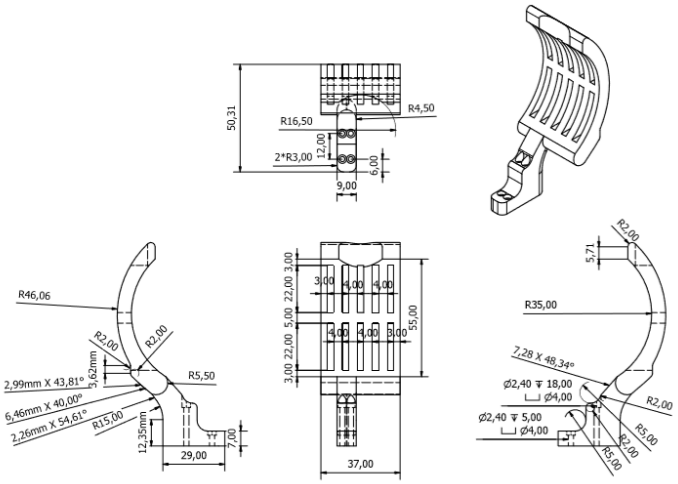


Figure 1: Gripper finger design for grasping a bottle.



Figure 2: Prototype of gripper on ABB's Yumi robot.

In this paper, a design of parallel gripper finger of dual arm robot working by holding safely a glass bottle is proposed for carbonated beverage and opening the crown cork cap opener with the automatic push-down bottle opener using the pneumatics.

The contribution of this research are listed as follows.

- The learning human motion pattern on the left-hand and right-hand side gestures of ABB's Yumi collaborative bartender robot working with the automatic push-down bottle opener will be proposed for gasping a glass, preparing a glass and pouring a beverage, which will be able to replace humans in the future.

- To design a safe gripper finger for holding a glass while movement based on the dual-arm industrial robot.

- To propose an automatic push-down crown cap bottle opener based on the dual-arm industrial robot.

2 Proposed Gripper Finger of Robotic Dual-Arm

In this section, the gripper finger of robotic dual-arm is proposed to apply with an automatic bottle opener. The design of automatic bottle opener controlled by the solenoid circuit are presented as follows.

2.1 Proposed Gripper Finger

In this paper, gripper finger design is to optimise the collaborative robot for picking and placing a bottle with a specific process. The robotic parts handler that physically interacts with the working environment. This leads to increase throughput, improve system reliability and compensate for robot inaccuracy.

The objective is to gasp and hold safely a glass bottle with a parallel gripper while using pneumatic for opening a bottle crown cork cap with an automatic push-down bottle opener. The dimension of carbonated soft drink in 250 ml glass bottle is of $11 \times 20 \times 13$ cm with the 25mm crown cork bottle cap. The design of safe gripper finger is made by 3D-printing ABS (Acrylonitrile Butadiene Styrene) plastic for grasping and holding a glass bottle and human-friendly interaction is shown in Figure 1.

Depending on the material used, a stress limit of designed gripper finger is tested on the stress analysis by the Autodesk Inventor. Prototype of gripper on ABB's Yumi robot is shown in Figure 2.

2.2 Proposed Automatic Bottle opener

The concept of an automatic bottle cap opener mechanism is to effortlessly open and remove the bottle cap in one push-down by single handed robot. Rack design for proposed automatic bottle opener implementation using pneumatics is being worked reliably with the collaborative robot.

Mechanism of proposed automatic bottle opener consists of aluminium profile, bracket, a bottle stand, an automatic push-down bottle cap opener at the top of rack, a bottle opener stand, a pneumatic air cylinder, and a stroke adjustment sensor. Installation of an automatic bottle crown cap opener is shown in Figure 3. The dimension of rack design is $15.1 \times 17.4 \times 42$ cm connected with the pneumatic systems.

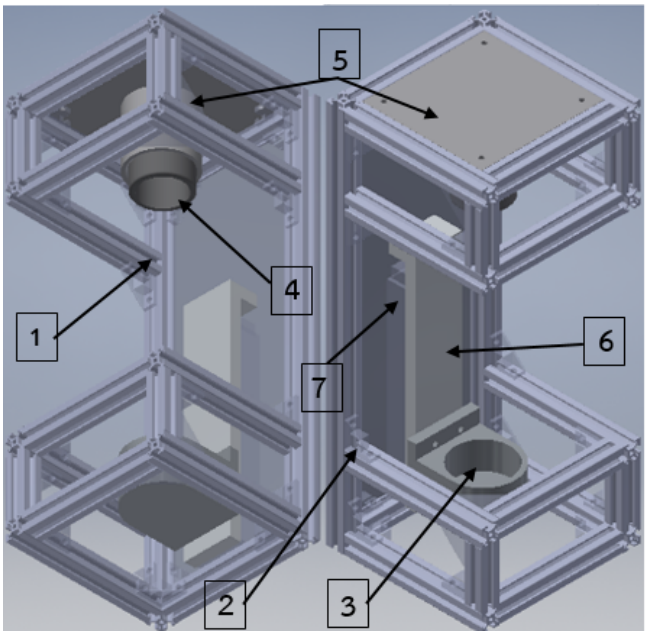


Figure 3: Mechanism of proposed automatic push-down bottle opener: 1) Aluminium profile; 2) Bracket; 3) Bottle stand; 4) Automatic push-down bottle cap opener; 5) Bottle opener stand; 6) Pneumatic cylinder; 7) Stroke adjustment sensor.

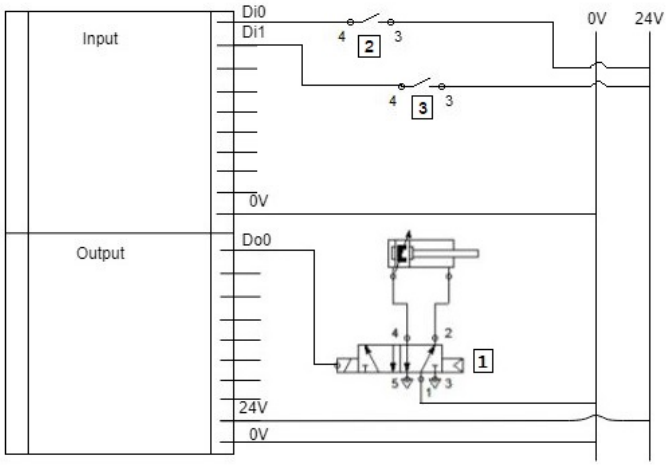


Figure 4: Diagram of solenoid circuit and I/O robot interface: 1) Solenoid circuit for automatic bottle opener; 2) Stroke adjustment sensor circuit at the upper of bottle opener; 3) Stroke adjustment sensor circuit at the lower of bottle opener.

The process of an automatic bottle cap opener starts with the right-hand side of robot arm released a bottle into a bottle stand after selecting and picking a beverage. Then an automatic bottle opener moves vertically by the pneumatic air cylinder in order to remove a crown cap by an automatic push-down bottle cap opener installed at the top of rack in one push down motion. A stroke adjust sensor is referred as a switch for controlling the outstroke of air cylinder while operation.

The relationship on outstroke between force, radius and pressure that can given from

$$F_r = A_e \cdot P \,, \tag{1}$$

where F_r is the resultant force, P is the pressure on the surface and A_e is the effective cross-sectional area of the piston surface.

Solenoid circuit is connected with I/O robot interface as shown in Figure 4. The output signal voltage of robot is used at 24 Vdc while operating. Prototype of automatic push-down bottle opener using pneumatics including with the rack for safety is shown in Figure 5.

3 Coordinate systems of collaborative robot

The mechanical arm of ABB's Yumi robot is divided into an arm, a wrist and an end-manipulator. All the manipulators have marked for four reference points used during process to posture the tool in the workspace with a given orientation. The robot orientation is described with three-dimensional rotation using an order set of four numbers named *quaternions*. The ABB collaborative robots are be able to control manually by using the flex pendant [16], which is a hand held controller connected to the robot.

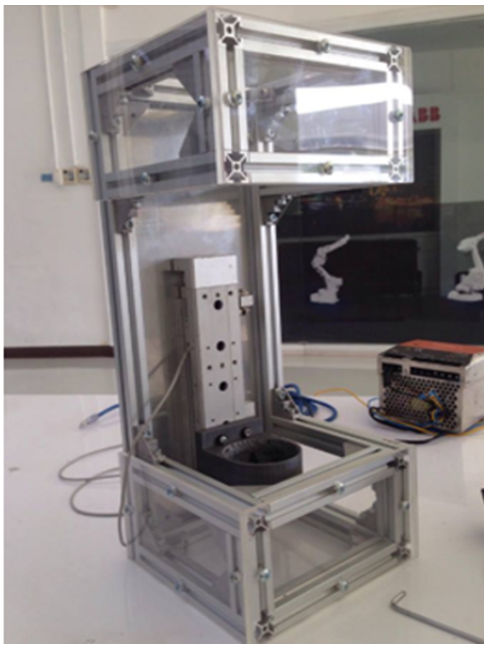


Figure 5: Prototype of automatic push-down bottle opener using pneumatics.

The robot movement is programmed by teach pedant, which are relative to the Tool Center Point (TCP). Normally, TCP is defined as the active point of the tool as one point for each tool at a given time. When the robot is programmed to move along a given path following by TCP expressed in relation to the coordinate system.

Let $\mathbf{t}_{co}$ be the translation vector from the original coordinate system of tool flange to TCP as [17]

$$\mathbf{t}_{co} = [\, t_x \;\; t_y \;\; t_z \,]^T \,. \tag{2}$$

Consider $\mathbb{M}$ be the translation matrix from the tool flange to

TCP as

$$\mathbb{M} = \begin{bmatrix} \mathbf{I} & \mathbf{t}_{co} \\ \mathbf{0} & 1 \end{bmatrix}, \tag{3}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{0}$ is a zero vector.

TCP is assumed at the same coordinate independent of N robot positions that is determined by

$$\mathbb{P}_{1,i}\,\mathbb{M} = \mathbb{P}_{1,j}\,\mathbb{M}\,, \tag{4}$$

where $i, j \in [1 \ldots N]$ and $i \neq j$.

Consider that

$$\mathbb{P}_{1,i}\,\mathbb{M} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbb{B}_a & \tilde{\mathbf{t}}_a \\ \mathbf{0} & 1 \end{bmatrix}, \tag{5}$$

and

$$\mathbb{P}_{1,j}\,\mathbb{M} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbb{B}_b & \tilde{\mathbf{t}}_b \\ \mathbf{0} & 1 \end{bmatrix}. \tag{6}$$

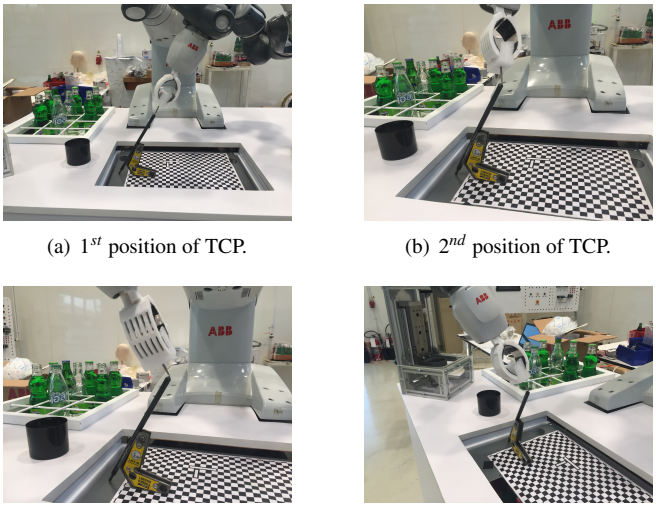
Therefore, the linear equation of system can be expressed as

$$[\mathbb{B}_a - \mathbb{B}_b] \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} + \left[\tilde{\mathbf{t}}_b - \tilde{\mathbf{t}}_b\right] = \mathbf{0}\,. \tag{7}$$

It is seen that all information determining the coordinate of TCP is achieved following (7).

4 Experimental Setup

Simulation setup for ABB's Yumi bartender on specific environment is illustrated with the four different type of beverages. Experimental setup of ABB's Yumi collaborative robot bartender consists of TCP calibration and workobject using the ABB Teach Pedant controller on chessboard pattern [18] at the base frame, including with the proposed bartender algorithm.



(a) 1^{st} position of TCP. (b) 2^{nd} position of TCP.

(c) 3^{rd} position of TCP. (d) 4^{th} position of TCP.

Figure 6: TCP calibration on chessboard pattern.

Algorithm 1 ABB's Yumi Bartender algorithm

```
Initial Parameters: GUI, Status1, 2, , . . . , 4, Drink#1, 2, . . . , 4,
Sensor Top, Sensor Bottom;
Initial Functions:
PickupDrink() {
  Input Position of Robot from CAMERA
  Input Kinematic of Robot
  Position and Kinematic Calculation()
  Output Robot move to pick bottle
  Output Robot move to open bottle cap and pick glass
  SolenoidON ()
  Output Robot move to pick bottle
  Output Robot move to pour
  IF (Drink# == Drink#1) {
    Output Pour Drink#1
  Else
    Output Pour Drink#2
  }
  Output Keep bottle and serve
}
SolenoidON () {
  IF (Sensor Top == ON) {
  SolenoidOFF()
  Else NO }
SolenoidOFF () {
  IF (Sensor Bottom == ON) {
  YES
  Else NO }
Main () {
Switch (GUI == Drink#?) {
  Case Drink#1 :
  Status1 = PickupDrink ()
  Case Drink#2 :
  Status2 = PickupDrink ()
  Case Drink#3 :
  Status3 = PickupDrink ()
  Case Drink#4 :
  Status4 = PickupDrink () }
```



Figure 7: Demonstration of ABB's Yumi bartender robot with gripper finger dual-arm.

4.1 TCP calibration

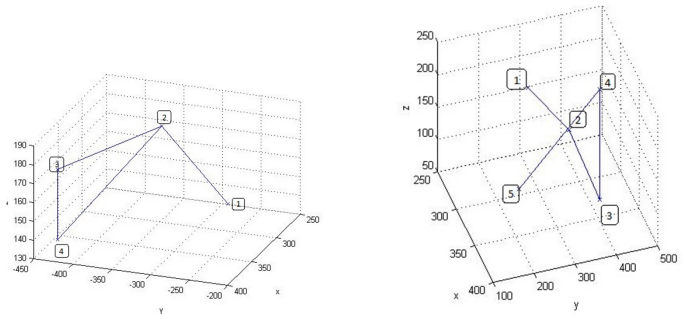
TCP calibration is modified on the chessboard pattern by determining the TCP 4-reference position with different type of grasping gesture object using ABB Teach Pendant as shown in Figure 6.

4.2 Workobject

A workobject [19] is a coordinate system used to describe the position of a work piece. The workobject consists of two frames as a user frame and an object frame. All positions will be related to the object frame, which is related to the user frame and world coordinate system. Working movement is modified on the chessboard pattern by marking the 3-reference point on x-axis and y-axis using ABB Teach Pedant.

4.3 Bartender algorithm

The proposed ABB's Yumi bartender algorithm is as shown in Algorithm 1. The process is starting from taking order from user via GUI interface on web application, then select one beverage and open the cap with an automatic push-down cap opener using the right-hand side of robot. Finally, to pick a glass and then pour the selected beverage into a glass by the left-hand side of robot. Demonstration of ABB's Yumi bartender robot with gripper finger dual-arm working with the automatic push-down bottle opener is depicted in Figure 7.



(a) Path of the right-hand side gesture for opening a bottle.

(b) Path of the left-hand side gesture for preparing a glass.

Figure 8: Path of ABB robot gripper gesture for opening a bottle on the right-hand side and preparing a glass on the left-hand side of robot.

5 Experimental Results

Experimental results show that ABB dual-arm robot performing gesture-based robotic gripper capable of being human-based recognised consist of four paths as two paths for opening a bottle cap using an automatic push-down bottle cap opener by pneumatics and preparing a glass. Next two paths for pouring a selected drink by the right-hand side into a glass holding by the left-hand side of robot gripper as follows.



(a) Right gripper involved taking up a bottle.



(b) Right gripper involved moving a bottle to bottle opener stand.



(c) Left gripper involved rotating a glass right-side up, while cap was being opened.



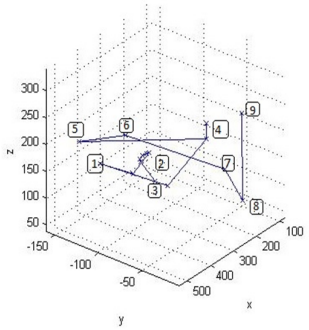
(d) Right gripper involved taking an uncapped bottle, while glass was prepared.

Figure 9: Grasping and movement of opening a bottle cap along Route 1 and Route 2.

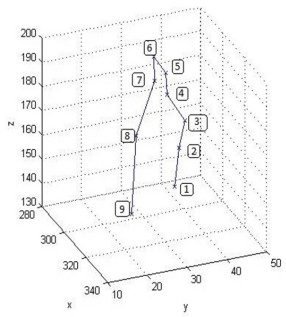
5.1 Path for opening a bottle and preparing a glass

Figure 8 shows the paths of gripper gesture controlled the right-hand side robot for opening a bottle and for preparing a glass on the left-hand side of robot. Experimental route for opening a bottle cap by a right-hand side gripper were the No. 1-4 shown in Figure8(a). Route 1 started at location No. 1, involved taking up a bottle from location No. 2 and moving it to replace on the bottle opener stand at location No. 3 and waiting for opening a bottle cap at location No. 4, then taking up a bottle to location No. 2 again waiting for pouring in the next route. Route for preparing a glass by a left-hand side gripper were the No. 1-5 shown in Figure8(b). Route 2 started at location No. 1, involved picking up a glass from location No. 2 and moving it to location No. 3 and then rotating a glass right-side up to location No. 4, then standby for the next route at the location No. 5.

Figure 9 demonstrates the step-by-step of gesture movement of ABB's Yumi bartender collaborative robot while using an automatic push-down bottle cap opener by pneumatics at the right-hand and holding a glass prepared by a left-hand side of robot gripper followed Route 1 and Route 2. Figure 9(a) shows the right gripper was taking a bottle to the rack of automatic push-down bottle opener, while the left gripper prepared to pick up a glass depicted in Figure 9(b). Meanwhile the bottle was being uncapped by automatic push-down opener using pneumatic, the right gripper was prepared into standby mode and left gripper was rotating a glass right-side up as shown in Figure 9(c). After that, the right gripper was taking out an uncapped bottle presented in Fig 9(d) and left gripper was waiting for pouring in the next route.



(a) Path of the right-hand side gesture for pouring a beverage.



(b) Path of the left-hand side gesture for gasping a glass.

Figure 10: Path of ABB robot gesture movement for pouring an uncapped beverage.



(a) Right gripper involved holding an uncapped bottle and left gripper prepared holding a glass.



(b) Right gripper involved pouring a beverage slowly into a glass, while left gripper tilted a glass at a 45 degree angle.



(c) Right gripper involved shaking a bottle softly, while left gripper straightened holding a glass.



(d) Right gripper involved taking an empty bottle into a bottle stand, while left gripper involved moving a full glass ready to serve.

Figure 11: Grasping and movement of pouring an uncapped beverage into a glass along Route 3 and Route 4.

5.2 Path for pouring a beverage into a glass

Figure 10 shows the paths of gripper gesture controlled the right-hand side robot for pouring a selected beverage without cap into a glass holded by the left-hand side of robot. Experimental route for pouring a uncapped bottle by a right-hand side gripper were the No. 1-9 shown in Figure 10(a). Route 3 started at location No. 1, involved holding an uncapped bottle to location No. 2 and moving it at the front of a glass holding by a left-hand side gripper to location No. 3, pouring slowly from location No. 5 to No. 6 and shaking softly from location No. 7 to No. 8, and then taking up an empty bottle to location No. 9 at the end route.

Now that a glass has been prepared by a left-hand side gripper and ready to pour. Route for pouring a beverage were the No. 1-9 shown in Figure10(b). Route 4 started at location No. 1, moving a glass prepared from location No. 2 to No. 3, tilted a glass at a 45 degree angle from the location No. 4 to No. 5 and then poured a beverage slowly by a right-hand side gripper so that the liquid landed directly in the middle side of a glass holding by a left-hand side gripper of robot. Once a robot poured about half of beverage into a glass at the location No. 6, then straightened a glass and poured a rest of beverage directly into the center of glass from the location No. 7 to No. 8, and then moving a full glass with beverage to location No. 9 at the end route.

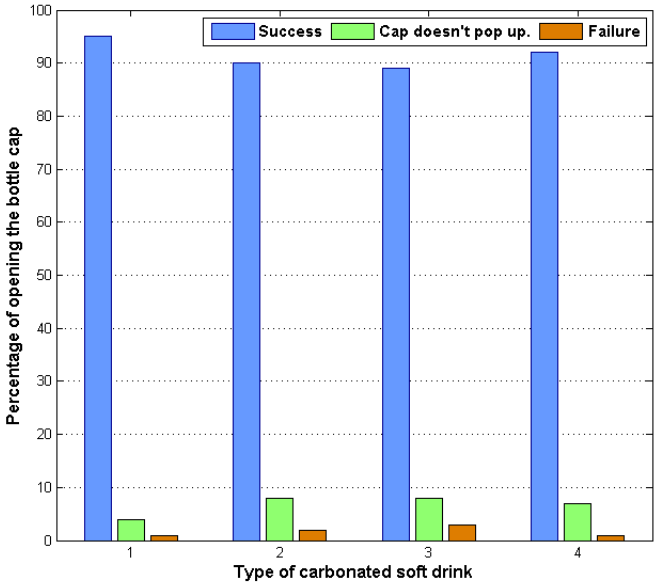


Figure 12: Percentage of opening the bottle cap by automatic push-down bottle opener using pneumatics.

Figure 11 demonstrates the step-by-step of gesture movement of ABB's Yumi bartender collaborative robot while pouring a beverage using a right-hand side gripper into a glass holding by a left-hand side of robot gripper followed Route 3 and Route 4. Figure 11(a) shows that the right gripper was holding an uncapped bottle and then pouring a beverage slowly into the middle side of glass, while the left gripper tilted a glass at a 45 degree angle depicted in Figure 11(b). Until the bottle was empty, the right gripper was shaking a bottle softly, while the left gripper was holding straighten a glass in Figure 11(c). After that, the right gripper was taking an empty bottle into a bottle stand and left gripper was ready to serve as shown in Figure 11(d).

The average percentage of the results from opening a bottle cap is as the success of opening bottle cap is given for 91.5 %, the bottle cap doesn't pop up the opener for 6.75 % and the failure for 1.75 % presented in Figure 12, which are achieved effectively.

6 Conclusion

In this paper, a design of gripper finger working with an automatic push-down bottle crown cork cap opener has been presented using pneumatics on ABB's Yumi collaborative robot for the bartender environments with the 4-different types of carbonated soft drink. The proposed semi-circular shaped gripper is appropriately designed for grasping and holding a bottle safely and working with the automatic bottle opener. The design of safe parallel gripper finger made by ABS plastic for grasping and holding a glass bottle and human-based interaction has been introduced. A gripper finger design with the industrial specification is tested on the stress analysis depending on the material used by the Autodesk Inventor. Rack design for an automatic push-down bottle cap opener using pneumatics supported a proposed gripper finger has been proposed.

All proposed paths are introduced for safely working on bartender responsibilities including with opening a bottle, pouring a beverage and serving a beverage. The paths of successful gesture for grasping a bottle and hold a glass by proposed gripper finger have been presented that they can apply effectively and safely to work as a bartender controlled through a web application. Experimental results shown that the average percentage of 91.5% of successful cap opener can be achieved effectively.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment ABB Robotics provided a YuMi robot for the co-operative education and valuable support regarding YuMi programming. The authors would like to thank Pongthanakorn Yaibuatong, Department of Mechatronic Engineering and Automation, Faculty of Engineering, Mahanakorn University of Technology, Bangkok, Thailand for his support.

References

[1] T. S. Li, P. Kuo, Y. Ho, G. Liou, "Intelligent Control Strategy for Robotic Arm by using Adaptive Inertial Weight and Acceleration Coefficients Particle Swarm Optimization", IEEE Access, **7**, 126929-126940, 2019, doi: 10.1109/ACCESS.2019.2939050.

[2] J. Borras, R. Heudorfere, S. Rader, P. Kaiser, T. Asfour, "The KIT Swiss Knife Gripper for Disassembly Tasks: A Multi-Functional Gripper for Bimanual Manipulation with a single arm", in 2018 IEEE/RSJ International Conference on Intelligent Robots and Sysems (IROS), 2018, doi: 10.1109/IROS.2018.8593567.

[3] M. Stenmark, E. A. Topp, M. Haage, J. Malec, "Knowledge for Synchronized Dual-Arm Robot Programming", in 2017 Artificial Intelligence for Human-Robot Interaction: AAAI Technical Report, 77-84, 2017.

[4] M. Stenmark, E. A. Topp, "From Demonstrations to Skills for High-Level Pogramming of Industrial Robots", in 2016 Artificial Intelligence for Human-Robot Interaction: AAAI Technical Report, 75-78, 2016.

[5] Y. Xiong, V. Shapaval, A. Kohler, J. Li, P. Johan, "A Fully Automated Robot for the Preparation of Fungal Samples for FTIR Spectroscopy Using Deep Learning", IEEE Access, **7**, 132763-132774, 2019, doi: 10.1109/ACCESS.2019.2941704.

[6] M. Chantharasri, W. Suriwong, S. Pengjeam, S. Prongnuch, "The Development of Suan Sunandha Rajabhat University PR Robot", in 2020 Engineering, Science, Technology and Architecture Conference (ESTACON), Nakhon Ratchasima, Thailand, 2020.

[7] T. Wiangtong, S. Prongnuch, "Computer vision framework for object monitoring", in 2012 IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012, doi: 10.1109/ECTICon.2012.6254291.

[8] Z. Zhiyong, H. Dongjian, T. J. Lei and M. Lingshuai, "Picking Robot Arm Trajectory Planning Method", Sensors & Transducers, **162**(1), 11-20, 2014.

[9] K. Lee , Y. Wang, C. Zheng, "TWISTER Hand: Underactuated Robotic Gripper Inspired by Origami Twisted Tower", IEEE Transactions on Robotics, **36**(2), 2020, doi: 10.1109/TRO.2019.2956870.

[10] A. Menon, B. Cohen and M. Likhachev, "Motion planning for smooth pickup of moving object", in 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, doi: 10.1109/ICRA.2014.6906895.

[11] T. Kelly and A. Tomoshige, "Japan's robot bartenders : Last call for human service?", [Online], Available: https://www.japantimes.co.jp/news/2020/02/06/business/tech/japans-robot-bartenders-last-call-human-service/, Accessed on 03-Apr-2021.

[12] D. Li, "Application for the ABB IRB 14000 YuMi robot using Integrated Vision and 3D printing", Thesis, Department of Mechanical Engineering and Production Technology, Lab University of Applied Sciences Technology, Lappeenranta, 2020.

[13] ABB, "IRB 14050 Single-arm YuMi® Collaborative Robot", [Online], Available: https://new.abb.com/products/robotics/industrial-robots/irb-14050-single-arm-yumi, Accessed on 25-Nov-2020.

[14] J.Ph. Clerc, "Autonomous Bottle Opener Robot - ABOR", University of Florida, Department of Electrical and Computer Engineering, Intelligent Machines Design Laboratory, 2002.

[15] Gregory Arthur Jury, "Bottle opener", US20160145087A1, United States, [Online], Available: https://patents.google.com/patent/US20160145087A1/en, Accessed on 03-Apr-2021.

[16] ABB Robotics, "Application manual FlexPendant SDK", [Online], Available: https://library.e.abb.com/public/7a93a58f26acb6acc1257b67004c6b0b/3HAC036958-en.pdf, Accessed on 10-Dec-2020.

[17] J.Hallenberg, "Robot Tool Center Point Calibration using Computer Vision", Master's Thesis, Department of Electrical Engineering, Linkoping University, 2007.

[18] ABB, "Detailed information for: 3HAC028357-001", [Online], Available:https://new.abb.com/products/3HAC028357-001/teach-pendant, Access on 20-Nov-2020.

[19] ABB, "Operating manual RobotStudio 5.12", Document ID: 3HAC032104-001 Revision: B, [Online], Available: https://library.e.abb.com/public/792b4432c1402c40c1257b4b00521525/3HAC032104-001RevBen.pdf, Accessed on 10-Dec-2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1065-1071 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Modeling and Design of a Compact Metal Mountable Dual-band UHF RFID Tag Antenna with Open Bent Stub Feed for Transport and Logistics Fields

Hajar Bouazza[1, 2,*], Aarti Bansal[3], Mohsine Bouya[2], Azeddine Wahbi[4], Antonio Lazaro[5], Abdelkader Hadjoudja[1]

[1]*Electrical and Energetic Systems laboratory Ibn Tofail University Kenitra, 14000, Morocco*

[2]*TIC laboratory, College of Engineering & Architecture, International University of Rabat, 11100, Morocco*

[3]*Electronics and communication Engineering, Thapar Institute of Engineering and Technology, Patiala, 147001, India*

[4]*Laboratory of Industrial Engineering, Information Processing and Logistic, Faculty of sciences Ain Chock, Hassan II University, Casablanca, 20000, Morocco*

[5]*Department of Electronics, Electrics and automatics Engineering Universitat Rovira I Virgili (URV) Tarragona, 43001, Spain*

ARTICLE INFO

Article history:
Received: 06 November, 2020
Accepted: 28 January, 2021
Online: 16 February, 2021

Keywords:
Tag's read range
Conjugate impedance matching
Metal mountable tags
UHF RFID
Reflection coefficient

ABSTRACT

In this paper, we have modeled and designed a metal mountable tag antenna that is applied to cover two major UHF RFID bands, i.e., European (EU) (865-867 MHz) and U.S. bands (902-928 MHz). It is applied for many applications, especially in the transport and logistics fields. The tag antenna configuration utilizes microstrip configuration with open bent stub feed network to attain conjugate matching w.r.t. Monza R6 chip impedance. The proposed microstrip patch-based tag antenna structure is simple without using any shorting pin/holes, thus making it easy and inexpensive to manufacture. Additionally, the proposed tag's impedance has been easily tuned in order to achieve conjugate matching in regard to the employed chip impedance. The presented tag antenna has been fabricated and experimentally characterized to measure its read-range performance in the desired bands. Further, the differential probe set up is used to measure the designed tag's impedance. Also, the designed tag read-range is measured using a reader setup and is observed to exhibit read-range up to 11 m and 9 m in European and U.S. UHF RFID bands, respectively. The tag exhibits an impedance of 9.7 – j 130 ohms at 866 MHz and 8.7 – j 124 ohms at 915 MHz. The proposed tag antenna design's performance is verified, analyzed, and optimized by CST Studio Suite software. The performances of the designed tag are evaluated and analyzed in terms of conjugate matching, reflection coefficient, and read range measurement. From the results, it is noticed that the designed tag exhibit dual-band behavior with good impedance matching, Reflection coefficient, and high read range.

1. Introduction

Ultra-High Radio Frequency Identification (UHF-RFID) technology has emerged as an efficient technique to track the objects/items/products without needing to be in the line of sight. The major RFID applications are in access control, logistics, healthcare, asset tracking owing to its advantage of long read range and compact size in comparison to Low (L.F.) and High Frequency (H.F.) bands [1]. The UHF RFID band is further divided between two major regions, i.e., European and North America/U.S. covering (865-867 MHz) and (902-928 MHz) bandwidth with a maximum allowed transmitted power of 3.28 W and 4 W, respectively [2], [3]. To operate the tag reliably in both the regions, it is required that the designed tag must exhibit dual-band operation covering European and U.S. UHF RFID bandwidth.

The RFID tag has to be affixed on a different type of conductive and non-conductive object to be tracked. Further, the performance of the tag deteriorates significantly especially when placed on conductive items such as metals. This is attributed to the cancellation of radiations resulting from phase electromagnetic

*Corresponding author: Hajar Bouazza, Email: hajar-bouazza@uir.ac.ma

wave reflections from the underlying metal surfaces [4-6]. To track the metallic objects efficiently, some of the recent work utilizes a microstrip patch antenna with a full ground beneath to overcome the influence of conductive objects such as metal, liquids, etc. as depicted in ref [7] and [8]. Further, the planar inverted-F antenna (PIFA) incorporating vias/ holes for the compact-sized tag is presented in [9], [10]. However, the employed vias/holes contribute to increased fabrication cost and complexity to the tag structure.

In this article, the work presented in [11] is extended further. Here, a dual-band metal mountable tag antenna having a planar structure covering European and North American UHF RFID bands is designed. The proposed tag antenna utilizes a microstrip patch antenna having a bent open stub feed. This bent feed structure offers the benefit of easy impedance tuning of the designed tag considering the complex chip impedance.

In this paper, we propose an RFID tag antenna to be integrated to track items in fundamental applications such as in the supply chain for the transport and logistics field. Effective management of this chain leads to very concrete results. To increase this efficiency, the implementation of an RFID based container-tracking system is necessary. This will contribute to real-time identification and reaching new levels of traceability and control.

To that end, and to test the performances of the UHF RFID tag developed, a real test is planned on the metal containers in airports; this test requires the deployment of a chain of readers and tags attached to the objects.

The container is the basic loading unit, usually made of aluminum alloys, with different sizes, the largest, exclusively on cargo aircraft, and the means are loaded in the holds of passenger aircraft.

For better optimization of airfreight, goods are transported in containers designed for safety reasons and to facilitate loading and unloading operations. These goods can be easily tracked using a metal mountable RFID tag. Thus, in this article, a UHF RFID tag is proposed, designed and tested for metallic objects. The designed tag is able to operate on two frequency bands EU and US to ensure interoperability of the system. This work is an extension of the work presented in [11], In this paper, the designed tag antenna's working performance on the basis of circuit analysis is represented. Further the simulated results of the tag presented in [11] are verified experimentally in terms of differential impedance and read range. Also, the designed tag's read range performance is compared for two different heights of the substrate. It is observed from a comparison table (Table II) that the designed tag has better performance with respect to other recently designed tags covering two major RFID bands at UHF range i.e., ETSI and U.S. Further, the tag is specifically designed for metal mountable applications. Also, the designed tag covers a better-measured read range of 11 m and 9.7 m in ETSI and U.S. band when mounted on metal objects.

The paper is organized as follows. The proposed tag antenna structure is introduced in Section 2. The section 3 discusses the equivalent circuit of the patch antenna. The parametric study to show its tuning capability for conjugate impedance matching is shown in Section 4. Further, the simulation performance of the designed tag antenna is shown in Section 5. The impedance measurements for the employed RFID chip i.e. Monza R6, and the designed tag antenna is carried out in Section 6. Also, Section 7, presents the designed tag's measured read range performance. Finally, Section 8 concludes the work.

2. Proposed design

2.1. Antenna structure

The tag antenna proposed here has been designed using a microstrip patch antenna. The proposed tag antenna geometry and its dimensions are illustrated in Fig.1 and Table 1, respectively as described in [11]. It consists of two simple microstrips, a patch, and a bent feed line [8].

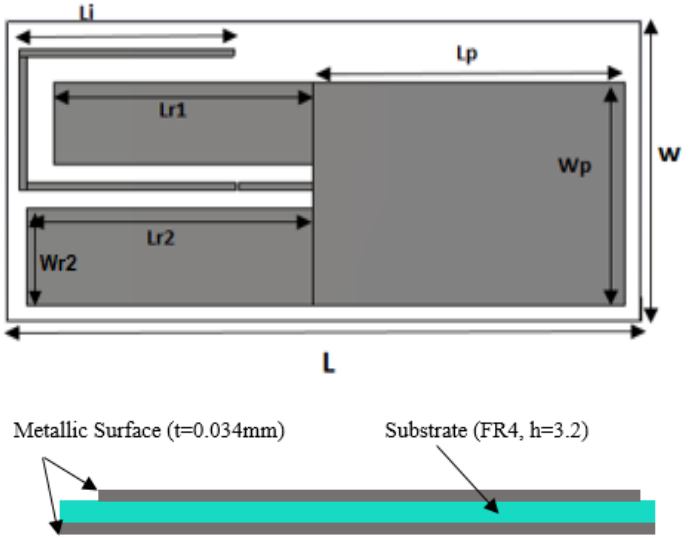


Figure 1: The proposed tag antenna design (Top View and Side view).

Table 1: The proposed antenna dimensions (units in mm)

Par.	L	W	Lp	W_p	L_{r1}	L_i	L_{r2}	W_{r2}
Value	85	40	42	30	30	25	36	11

The tag structure is designed with a full ground plane beneath to be mountable on metallic objects. It is fabricated using FR4 substrate having εr = 4.69, tan δ = 0.02, and a thickness h=3.2 mm, with a total dimension of W*L (40 mm * 85 mm). As the size is a critical factor in designing RFID tags, and in order to reduce tag size, the bent feed line structure has been adopted for feeding the patch. The antenna size can also be reduced by opting for higher permittivity substrates. However, they lead to higher costs. Also, the open stub feed line inserted into the patch decreases the input impedance of the patch. The RFID chip is connected to the feed line and its position is optimized to obtain conjugate impedance matching. The parameters L_1 and L_{r1} are optimized to obtain conjugate matching.

3. Equivalent Circuit Model

Figure 2 illustrates the transmission line model of the designed antenna. From the figure, the global structure of the antenna model has been modeled as two transmission lines connected in series configuration [12] comprising of microstrip patch and the feed line of lengths l_1 and l_2 respectively. The antenna input impedance at the feed port can be calculated as:

$$Z_{in} = Z_{in}^1(l_1) + Z_{in}^2(l_2) \quad (1)$$

where Z_{in}^1 represents the input impedance of the radiation patch, and Z_{in}^2 represents the input impedance of the bent open stub line.

Further, the microstrip patch input impedance is given as in [13]:

$$Z_{in}^1(l_1) = Z_0^1 \frac{R_r + jZ_0^1 \tan(\beta l_1)}{Z_0^1 + jR_r \tan(\beta l_1)} \quad (2)$$

where R_r , Z_0^1 and β are the resistance and characteristic impedance and the propagation constant of the radiation patch.

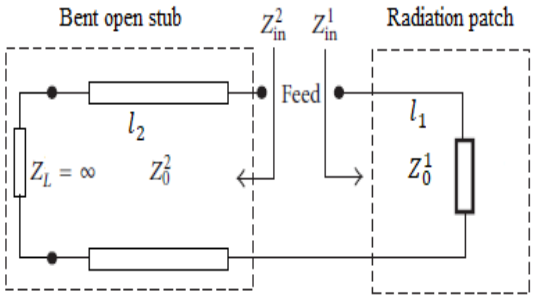


Figure 2: Open stub patch antenna transmission line model representation

Also, the employed open stub feed's input impedance is represented by:

$$Z_{in}^2(l_2) = -jZ_0^2 \frac{1}{\tan(\beta l_2)} \quad (3)$$

where the characteristic impedance of the open stub line is represented by Z_0^2 .

For a preliminary design of the radiating patch antenna, and before proceeding with the optimization, the dimensions were calculated using the classic rectangular patch equations expressed bellow:

$$L = Leff - \Delta L \quad (4)$$

where L is the physical length, ΔL is the extended length due to fringing field effects and $Leff$ is its effective length. The extended length, ΔL is given by [14] as shown in equation (5).

$$\frac{\Delta L}{h} = 0.412 \frac{(\varepsilon_{eff}+0.3)(\frac{w}{h}+0.264)}{(\varepsilon_{eff}-0.258)(\frac{w}{h}+0.8)} \quad (5)$$

where w is the width of the patch, h is the substrate height, and εeff is the effective dielectric constant.

Also, the effective dielectric constant of the patch is represented as follows:

$$\varepsilon_{eff} = \frac{(\varepsilon_r+1)}{2} + \frac{(\varepsilon_r-1)}{2}(1 + \frac{12h}{w})^{-1/2} \quad (6)$$

where, ε_r is the relative dielectric constant of the substrate.

4. Parametric Study

The conjugate impedance matching is optimized by tuning the lengths of the stub, i.e., L_i and L_{r1}, respectively as explained in [11].

Further, the parametric study to investigate the influence of varying the antenna dimensions, i.e., L_{r2} and W_{r2} in order to tune the second resonant frequency, i.e., 915 MHz, is shown in Fig. 3.

It has been observed that the resonance shifts towards the right by varying these parameters simultaneously and keeping the other parameters constant.

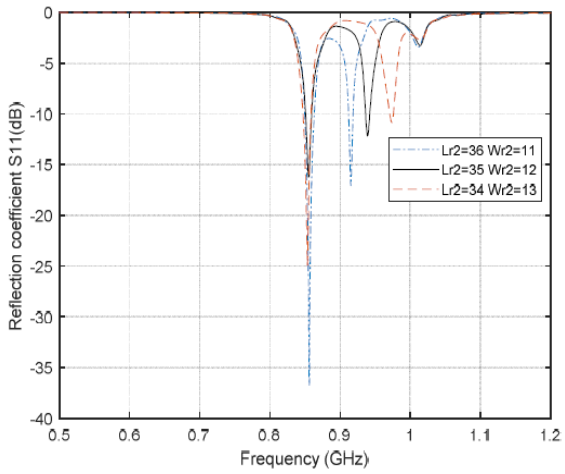


Figure 3: Simulated reflection coefficient versus frequency of the for different values of Lr2 and Wr2.

5. Simulation results

In this context, we have designed, analyzed, simulated, and optimized the proposed tag antenna using CST Studio Suite software.

The tag antenna must exhibit conjugate impedance that matches the employed chip impedance, i.e., Impinj Monza R6 [15]. The simulated and measured S_{11} results for the designed tag antenna are shown in Fig. 4.

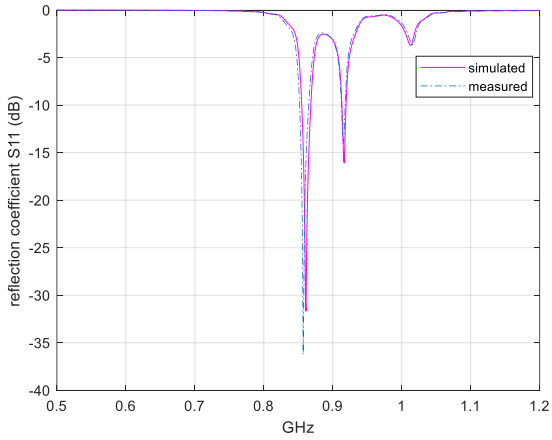


Figure 4: Antenna measured reflection coefficient comparison

6. Impedance Measurement of Designed Tag Antenna and Chip

The Monza R6 chip is employed here at the feeding port of the designed tag antenna and its equivalent circuit is shown in Fig. 3. The chip has a resistance, R_p, of 1200 Ω in parallel to the capacitance, Cp, of 1.43 pF and a chip sensitivity of -20 dBm. The equivalent circuit values for the chip are provided in the datasheet

for the die chip and do not take into account the parasitic impedance that includes substrate, packaging, and input power dependence. The packaged chip, i.e., LXMS21ACMF-183 from Murata is used in the designed tag prototypes. Thus, it is important to consider the chip impedance dependence with respect to frequency.

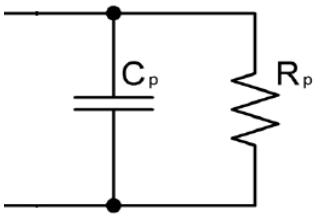
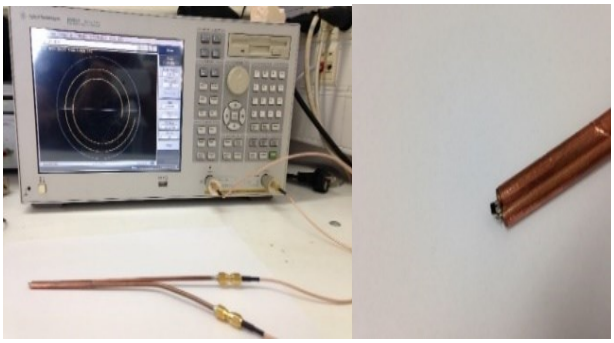


Figure. 5: Tag Chip Linearized R.F. Model

The measurements of the tag antenna and the chip impedances are obtained using vector network analyzer E5062A from Agilent. The tag antenna impedance cannot be directly measured using a 50 Ohms port directly. Therefore, to measure the differential impedance, the tag antenna has been considered as a two-port network and to measure its reflection coefficient, its differential impedance input impedance, Z_{in}, has been calculated represented as:

$$Z_{in} = \frac{2Z_0\,(1-S_{11}S_{22}+S_{12}S_{21}-S_{12}-S_{21})}{(1-S_{11})(1-S_{22})-S_{21}S_{12}} \quad (7)$$

here, Z_0 represents the coaxial cable's characteristic impedance. Figure 6 illustrates the setup to measure the impedance. The differential balun with short-circuited semi-rigid coaxial cables using λ/4 was constructed to measure the impedance. The terminals of constructed balun are soldered to the feed terminals of the designed tag antenna Further, the port extension technique has been applied while measuring differential impedance as detailed in [16].



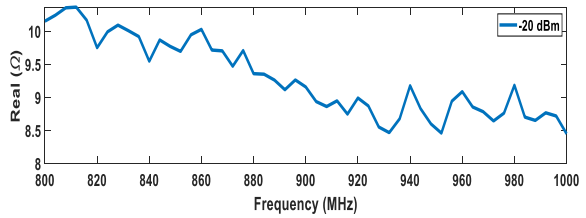
(a) (b)

Figure 6: Equipment for chip and antenna characterization. (a) VNA, (b) Balun

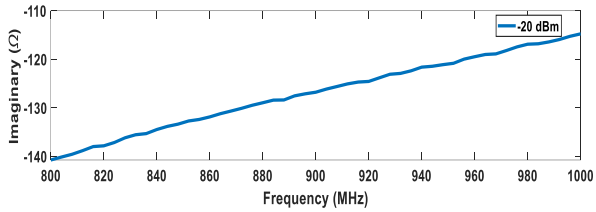
Figure 7 and fig. 8 shows the used chip input impedance measurements versus the frequency. The chip impedance is measured to be 9.7 – j 130 Ω at 866 MHz and 8.7 – j 124 Ω at 915 MHz, respectively. In comparison, the input impedances given in the datasheet are 13.6-j 127 Ω at 866 MHz and 12.2-j 120 Ω at 915 MHz [17]. As observed, the packaged chip input impedance is slightly different and also depends on the input power.

Finally, the reflection coefficient (Γ) is further calculated using measured tag antenna impedance i.e., Zin and chip impedance, i.e., Z_{chip} respectively, and is represented as:

$$\Gamma = \frac{Z_{in}-Z_{chip}^{*}}{Z_{in}+Z_{chip}} \quad (8)$$



(a)



(b)

Figure 7: Chip input impedance for -20 dBm input power

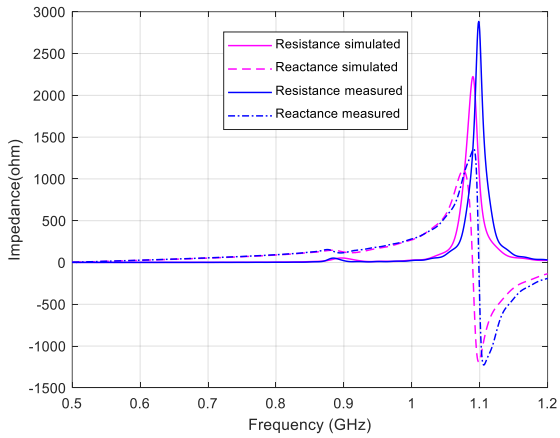


Figure 8: Representation of simulated and measured results of the antenna input impedance

Figure 9 shows a comparison of the simulated and measured reflection coefficient results in the 2 required frequency bands i.e. ETSI & FCC.

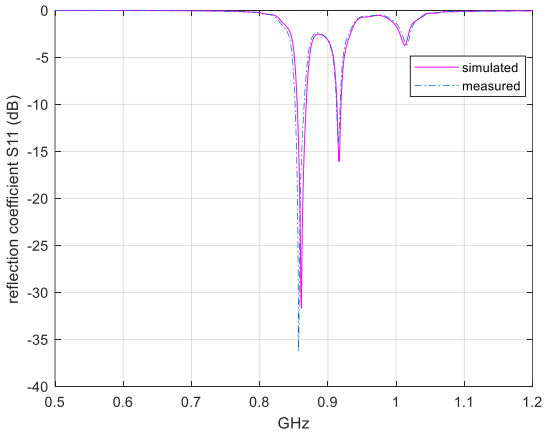


Figure 9: Results of the antenna's reflection coefficient.

It can be seen that a reasonably good value for the reflection coefficient is obtained, which validates the existence of a suitable conjugate impedance matching between the chip' and the designed tag antenna impedance.

7. Read Range Measurements

The reading range or the tag range is defined as the maximum distance from which the tag can be recognized [18], and is the most important parameter for evaluating the RFID tag performance. The equation below shows this reading range following the Friis formula for free space:

$$R = \frac{\lambda}{4\pi}\sqrt{\frac{P_{EIRP}\times G_{tag}\times(1-|\Gamma|^2)}{P_{th}}} \quad (9)$$

where PEIRP is the equivalent isotropic radiated power specified as 3.3 W in the E.U. band and 4 W in the U.S. band.

Further, λ represents the operating free-space wavelength and P_{th} indicates the chip sensitivity at the resonant frequency. The G_{tag} is defined as the antenna tag's gain and Γ is the reflection coefficient defined above in equation (8). The tag's read range can be maximized by achieving a lower-power reflection coefficient defined by $|\Gamma|^2$ and a high-gain antenna G_{tag}.

Also, the ability of the tag to transfer power to the connected chip is represented by $(1-|\Gamma|^2)$ which is also known as the power transmission coefficient of the tag.

Figure 10 describes the devices used for experimental measurement of the reading range in an anechoic chamber.



Figure 10: The Read range measurement setup

The measuring setup consists of an RFID reader, i.e., CAEN RFID, circularly polarized Reader antenna with a realized gain Gr of 8 dBi. The reader is connected to the reader antenna and a host computer. The cable losses and attenuator losses were taken into consideration.

From the measurement results, we deduced that the tag has a reading range of approximately 11 m and 9.7 m at 866 MHz and 915 MHz with an EIRP of 3.3 W, respectively, as illustrated in Fig. 11.

The gain of the tag antenna, as well as the height of the chosen substrate, strongly influences the read range of the tag. a comparison of the read range for two different substrate heights i.e., 3.2 and 1.6 is shown in Figure 12. The read distance was observed to degrade from 11m to 4 m at 866 MHz and 9.7 m to 3.7 m at 915 MHz.

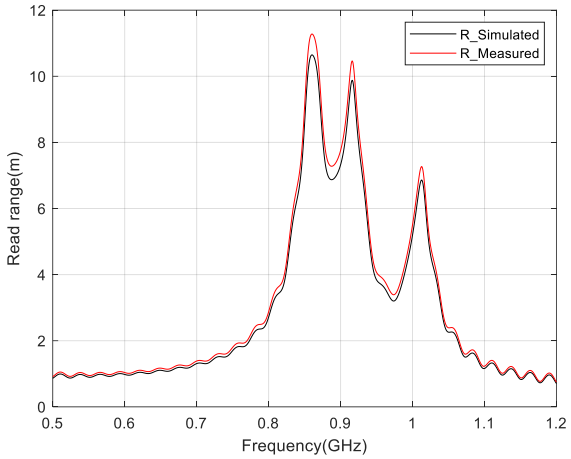


Figure 11: Results of simulated and measured antenna read range

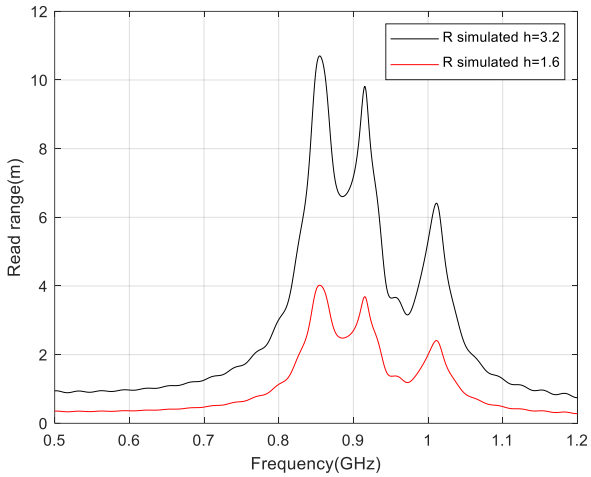


Figure 12: Read-range for two different substrate thickness.

7.1. 3D-Radiation Pattern

As already mentioned, impedance matching between antenna and chip impedances, is a key design requirement. Therefore, its radiation pattern cannot be measured directly using a 50-ohm port. Hence, the performance of the tag antenna is also assessed based on its simulated 3 D (Dimensional) gain and radiation pattern as shown in Figure 13. It has been observed from the figure that the tag antenna exhibits a realized gain of -3 dBi. Also, to minimize the effect of mounted objects, the tag should exhibit a unidirectional radiation pattern. Thus, it is depicted from Fig. 14, that the designed tag antenna exhibits a directional radiation pattern with a directivity of 4.2 dBi. This feature is mainly desirable to design the platform tolerant tag and hence validates its application for metal mountable tracking objects.

A performance comparison of our proposed tag antenna with some previous research studies representing metal-mountable UHF RFID tags is shown in Table 2. From the table, it has been observed that the tags designed in [18-22] are larger as compared to the tag proposed in this work. Moreover, our proposed tag is

characterized by its operability in the two main frequency bands with a better reading range. On the other hand, all the other tags shown in the table cover a single UHF RFID band. Further, the tag designed in [20] is compact, however, still it covers a lower read range in a single UHF RFID band.

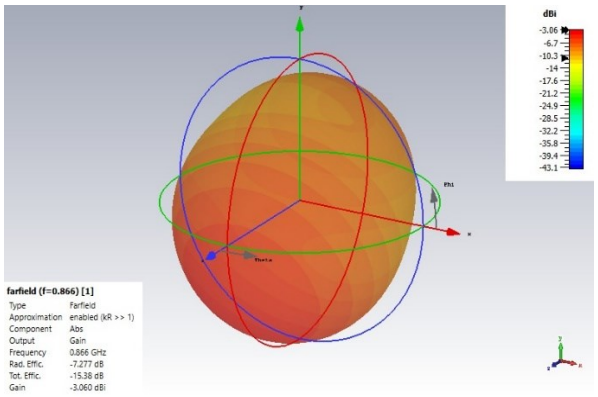


Figure 13: 3 D Gain pattern of the designed tag at 866 MHz

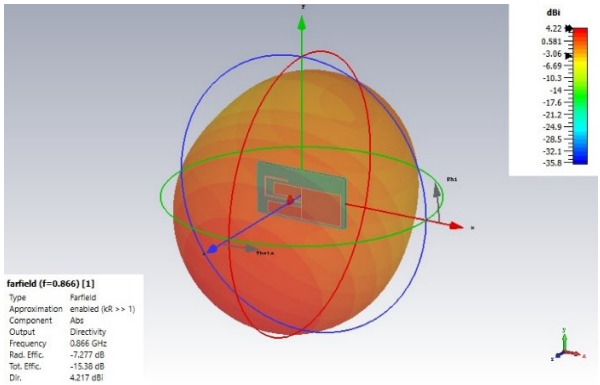


Figure 14: 3 D directivity pattern of the designed tag at 866 MHz

Table 2: A comparison table performances of UHF tags mounted on metallic items

Ref.	Transmitting EIRP Power	Circuit Size (mm)	Backing Plate Size (cm)	Max. Read Range (m)
This work	4 W	85×40×3.2	20×20	11 m (ETSI) 9.7 m (FCC)
[19]	4 W	86×41×3.2	20×20	5.6
[20]	3.3 W	88×60×0.76	15×15	5.4
[21]	4 W	30×30×3	20×20	7.2
[22]	4 W	120×60×0.4	15 × 30	10
[23]	N/A	84×126×1.28	N/A	2.5

Thus, better performances in regard to reading range, size and also covered bands, characterize our proposed RFID tag. Also, it is deduced that the designed tag may be employed for different applications such as health monitoring, logistics, and transport supply chains, and IoT (Internet of Things) network applications, etc.

8. Conclusion

In this research paper, a new planar dual-band compact platform tolerant UHF RFID tag antenna presented in [11] has been simulated, fabricated, and tested for its impedance and read range performance experimentally. To measure the real performance of metallic objects, the tag is mounted on a metal sheet. (20 m x 20 m) while measuring. Our tag antenna is modeled and designed using CST Studio Suite software. The designed tag is specifically designed to be mountable on metallic items by considering a full ground plane. The tag comprises a simple structure and does not include any vias, thus facilitating low-cost, efficient design for mass production. Further, the simulated and measured impedance, reflection coefficient, and read ranges of the tag structure have been presented and are observed to be well in agreement with each other. From the tag antenna reflection coefficient, it is realized that the designed antenna covers both major UHF RFID bands i.e., European (E.U., 865-867 MHz) and North American (the U.S., 902-928 MHz). The designed tag exhibits a read range up to 11 m at 866 MHz and 9 m at 915 MHz.

Hence, the designed tag antenna is observed to perform better in terms of conjugate impedance matching, unidirectional radiation pattern, and high read range in both the required resonant bands. Thus, it is validated that the designed tag is suitable for use in real-time applications such as transport and logistics fields and IoT network applications.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank MESRSFC and CNRST for the financial support.

References

[1] K. Finkenzeller, RFJD handbook, 3rd ed Wiley, New York, 2010
[2] I. Any et al., "Regulatory status for using RFID in the EPC Gen2 (860 to 960 MHz) band of the UHF spectrum," **2**(November), 1–18, 2016.
[3] V. D. Hunt, A. Puglia, and M. Puglia, RFID: A Guide to Radio Frequency Identification. New York: Wiley, 2007.
[4] T. Björninen, L. Sydänheimo, L. Ukkonen, and Y. Rahmat-Samii, "Advances in antenna designs for UHF RFID tags mountable on conductive items," IEEE Antennas Propagation Magazine, **56**(1), 79–103, 2014. doi: 10.1109/MAP.2014.6821761.
[5] D. M. Dobkin, S. M. Weigand, W. J. Communications, and S. Jose, "Environmental Effects on RFID Tag Antennas," 135–138, 2005.
[6] K. Penttilä, M. Keskilammi, L. Sydänheimo, and M. Kivikoski, "Radio frequency technology for automated manufacturing and logistics control. Part 2: RFID antenna utilization in industrial applications," The International Journal of Advanced Manufacturing Technology, **31**(1–2), 116–124, 2006. doi.: 10.1007/s00170-005-0174-y
[7] H. Chen, S. Member, C. Sim, and S. Kuo, "Compact Broadband Dual Coupling-Feed Circularly Polarized RFID Microstrip Tag Antenna Mountable on Metallic Surface," IEEE Transaction Antennas and Propagation, **60**(12), 571–5577, 2012. doi: 10.1109/TAP.2012.2210273.
[8] L. Mo and C. Qin, "Planar UHF RFID Tag Antenna With Open Stub Feed for Metallic Objects," IEEE Transaction Antennas and Propagation, **58**(9),3037–3043, 2010. doi: 10.1109/TAP.2010.2052570
[9] M. Hirvonen, P. Pursula, K. Jaakkola, and K. Laukkanen, "Planar inverted-F antenna for radio frequency identification," ELECTRONIC LETTERS, **40** (14), 2004. doi: 10.1049/el:20045156
[10] H. Chen, S. Member, and Y. Tsao, "Low-Profile PIFA Array Antennas for

UHF Band RFID Tags Mountable on Metallic Objects," IEEE Transaction Antennas and Propagation, **58**(4), 1087–1092, 2010. doi: 10.1109/TAP.2010.2041158

[11] H. Bouazza, A. Lazaro, M. Bouya and A. Hadjoudja, "Dual-band UHF RFID tag for metallic items," 2019 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC), 090-092, 2019, doi: 10.1109/APWC.2019.8870474.

[12] L. Mo and C. Qin, "Tunable compact UHF RFID metal tag based on cpw open stub feed PIFA antenna," International Journal of Antennas and Propagation 2012. doi: 10.1155/2012/167658

[13] L. Reinhold and B. Pavel, RF Circuit Design: Theory and Applications. New York: Pearson Education, 2000. doi: 10.1155/2012/167658

[14] C. A. Balanis, Antenna Theory Analysis and Design, 3rd ed. New Jersey: John Wiley & Sons, 2005.

[15] Datasheet Impinj Monza R6, https://support.impinj.com/hc/en-us/articles/202765328-Monza-R6-Product-Brief-Datasheet

[16] A. Bansal, S. Sharma, Khanna, "platform tolerant dual-band UHF RFID tag antenna with enhanced read range using artificial magnetic conductor structures". Int. J RF Microw Comput Aided Eng. 2019. doi: 10.1002/mmce.22065

[17] H. Bouazza, A. Lazaro, M. Bouya, and A. Hadjoudja, "A Planar Dual-Band UHF RFID Tag for Metallic Items", Radioengineering, **29**(3), 505. 2020, doi: 10.13164/re.2020.0504

[18] A. Bansal, S. Sharma and R. Khanna, "A Spiral Shaped Loop Fed high Read Range Compact Tag Antenna for UHF RFID Applications," 2019 IEEE International Conference on RFID Technology and Applications (RFID-TA), 212-215, 2019, doi: 10.1109/RFID-TA.2019.8892203.

[19] A. P. Sohrab, Y. Huang, M. N. Hussein and P. Carter, "A Hybrid UHF RFID Tag Robust to Host Material," in IEEE Journal of Radio Frequency Identification, **1**(2), 163-169, June 2017, doi: 10.1109/JRFID.2017.2765623.

[20] POLIVKA, M., SVANDA, M. "Stepped impedance coupled patches tag antenna for platform tolerant UHF RFID applications. IEEE Transactions on Antennas and Propagation", **63**(9), 3791–3797. 2015, doi: 10.1109/TAP.2015.2447034

[21] F. Bong, E. Lim and F. Lo, "Flexible Folded-Patch Antenna With Serrated Edges for Metal-Mountable UHF RFID Tag," in IEEE Transactions on Antennas and Propagation, **65**(2), 873-877, 2017, doi: 10.1109/TAP.2016.2633903.

[22] Y. Lin, M. Chang, H. Chen and B. Lai, "Gain Enhancement of Ground Radiation Antenna for RFID Tag Mounted on Metallic Plane," in IEEE Transactions on Antennas and Propagation, **64**(4), 1193-1200, 2016, doi: 10.1109/TAP.2016.2526047.

[23] M. Abu, E. E. Hussin, A. R. Othman, F. M. Johar, N. M. Yatim, and R. F. Munawar, "Design of 0.92 GHz artificial magnetic conductor for metal object detection in RFID tag application with little sensitivity to incidence of angle," Journal of Theoretical and Applied Information Technology., **60**(2), 307–313, Feb. 2014. Id: 73550079

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1072-1078 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Using a safety PLC to Implement the Safety Function

Karol Rástočný, Juraj Ždánsky, Jozef Hrbček*

Department of Control and Information Systems, University of Žilina, Faculty of Electrical Engineering and Information Technology, Žilina, 010 26, Slovak Republic

ARTICLE INFO

Article history:
Received: 06 November, 2020
Accepted: 28 January, 2021
Online: 16 February, 2021

Keywords:
Risk
Safety
Safety Integrity
Safety Function
Safety Plc

ABSTRACT

Nowadays almost every PLC manufacturer offer a so-called safety PLC. It is a specific category of PLC, which in recent years have become a commonly used means of performing safety functions, especially in industrial applications. In this area of specific applications, a maximum of SIL 3 is normally required. However, the guaranteed safety features of the PLC lead to the consideration or discussion, whether they could be used in applications with higher safety requirements. This paper deals with the possibility of using the safety PLC to implement safety functions with SIL 4. The paper presents the long-term experience of the authors in the development of control systems for railway applications with the required level of SIL4.

1. Introduction

This paper is an extension of the work originally presented at the conference [1]. Extension relates to elaboration of the impact of random and systematic failures to the safety of the safety function (SF) realized on the Safety Programmable Logic Controllers (sPLC) in the dual architecture, which is realized by this architecture. The part of realized extension is a specific application example of the realization of SF with the safety integrity level 4 (SIL 4) using two safety PLCs certified to SIL 3.

In practice (in industry, transport, medicine ...) we can quite often encounter that the ongoing process or operation of a machine or equipment can pose a risk to assets (people, environment, property damage, ...), which fall within its remit. It is necessary to deal with the risk analysis in this case. This means identifying hazards, their consequences and calculating (or estimating) the risk. Generally, the risk (R) is given by the sum of the combinations of the occurrence frequency of the i - th hazard (h_i) and its consequences (c_i), it means:

$$R = \sum_{i=1}^{n} h_i \times c_i, \qquad (1)$$

where n is the number of identified hazards.

The appropriate safety measures (technical, organizational) must be applied to reduce this risk to at least a predefined tolerable value (Figure 1), if the calculated (estimated) risk is greater than the predefined tolerable risk ($R>R_T$). The technical safety measures can be passive (covers, fencing, ...) or active. Safety-relevant systems (SRSs) are used as active technical safety measures, which realize the safety functions (SFs).

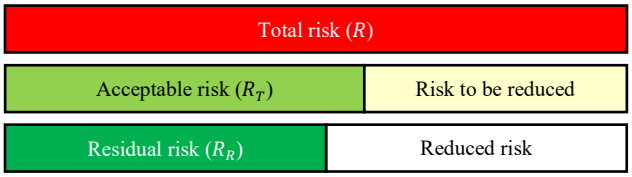


Figure 1: Relations between individual components of risk

A process (equipment, machine) is considered safe if it is valid that

$$SR_R \leq R_T. \qquad (2)$$

The application of safety measures can be focused on:

- only to reduce the frequency of individual hazards;
- only to reduce the consequences of individual hazards;
- to reduce the frequency of occurrence and also to reduce the consequences of individual hazards.

*Corresponding Author: Jozef Hrbček, University of Žilina, +421 041 513 3354, Email: jozef.hrbcek@uniza.sk

This paper deals with the issue of using the technical active measures to reduce the frequency of hazards.

Based on the risk analysis, the safety functions are defined within the specification of safety requirements so that ordinarily one SF covers one hazard (or even more hazards). The aim of SF realization is to achieve risk reduction to the required value (for example, by reducing the frequency of hazards). In order to achieve this objective, the SRS must be characterized by the fact that during fault-free operation it performs the specified SFs and in case of failure the SRS must ensure that the ongoing process either remains in the given operating state (if this state does not endanger the assets within its operation), or has entered a predefined safe state (for example, disconnecting power from the motor and stopping the machine). This is a feature called "fail-safe".

Since safety cannot be understood in absolute but relatively, the possibility of SF failure must also be taken into the account. The degree to which SF is able to achieve the fail-safe property is expressed by SIL. Standard [2] defines four levels of SIL (1 to 4), SIL 4 means the most stringent safety requirements (for continuous operation, the mean frequency of a dangerous safety function failure PFH $10^{-9} \leq$ to $< 10^{-8}$ h^{-1} is required; i.e. an average of one dangerous SF failure is about 11 416 years). The required SIL for SF is determined depending on the size of the risk that SF reduces.

One of the suitable technical means for the implementation of SRS based on processor technology are also sPLCs. These are especially modified PLCs to have a fail-safe feature. These are usually modular systems with which SFs can be implemented not only with the required safety properties (with the required SIL), but also with the required reliability properties (with the required availability).

Commonly available sPLCs are mainly used for the implementation of SFs with a requirement for maximally SIL 3. This is because they are primarily developed for industrial applications where this requirement is usually sufficient (normally the SFs are implemented with desired SIL 2 or SIL 3 in the industry).

However, there are special cases of applications (e.g. in railway transport) where the implementation of SFs with SIL 4 is required, but these sPLCs are relatively expensive and do not always meet the functional requirements of the customer (for example, they do not have suitable I/O modules for wiring special components that are used in a given area of a specific application). Therefore, several manufacturers of the safety-related electronic signaling systems for railway applications are developing their own modular systems (such as generic products), which they use to realize specific applications (for example sPLC type NEXUS from the company PrvíSignalní [3]).

The efforts to increase the efficiency of the development of generic products of electronic signaling systems for railway applications have caused some manufacturers (e.g. [4], [5]) decided to develop a product for the implementation of SFs with SIL 4, which consists of two sPLCs (intended for the realization of SFs with SIL 3). It is a dual structure based on composite fail-safety with fail-safe comparison. A block diagram of such a structure is shown in Figure 2. The Equipment Under Control (EUC) block represents the controlled equipment, resp. monitored equipment (either as a separate element or as a part of the controlled or monitored process).

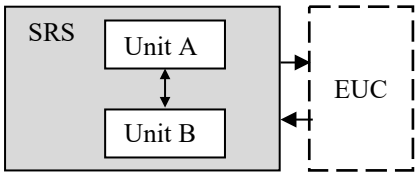


Figure 2: Block diagram of a general dual structure

The technical safety principle of such a solution (Figure 2) is based on mutual exchange and comparison of the data between the unit A (one sPLC) and unit B (second sPLC). The positive result of the comparison is the sign of the correct operation of the system.

Standards [2] and [6] require to prove not only functional safety (correct performance of the function in a fault-free state), but also technical safety of SRS (represents the fail-safe feature). Although functional safety can be in principle proved by the tests, it is impossible to prove technical safety in the same way. Making technical safety prove encounters the problems that are pointed out in this article. Since the safety proves must be performed individually for each SF (SRS can generally be implemented by several SFs, whereas a different SIL may be required for each SF and each SF may be implemented by different parts of the SRS). To clarify the considerations presented in this paper, the authors accepted the simplification that the SRS implements only one SF. Under this assumption is valid, that a dangerous failure of the SRS (due to the occurrence of a failure) is equal to a dangerous failure of the SF.

2. Realization of SF with SIL 4 using sPLCs with SIL 3

SF with SIL 4 must be realized in such a way that the requirements for safety integrity (SI) level 4 are met, which result from [2] resp. [6]. These are the requirements for safety integrity against random failures (RanF-SI) and also for integrity against systematic failures (SysF-SI). In the case of RanF-SI a failure of hardware components is supposed; in the case of SysF-SI it means the software errors, whether embedded software (firmware) or application software, but the hardware failures with a common cause cannot be ruled out (one hardware failure will affect the operation of both units A and unit B).

In principle, the influence of both systematic and random faults to SF failure can be illustrated by the fault tree shown in the Figure 3.

The fault tree in Figure 3 can be described by a logical function:

$$S_DF = S_DRF + S_DSF, \quad (3)$$

whereas:

$$S_DRF = A_RF.B_RF, \quad (4)$$

$$S_DSF = (A_SFA + A_SFE).(B_SFA + B_SFE), \quad (5)$$

where *S_DF* is a dangerous fault of SRS (top event), *S_DRF* is the dangerous random failure of SRS, *S_DSF* is the dangerous systematic failure of SRS, *A_RF* (*B_RF*) is a random fault of unit A (B), *A_SFE* (*B_SFE*) is a systematic failure of the application software of unit A (B), *A_SFE* (*B_SFE*) is a systematic failure of the embedded software or hardware of unit A (B).

The logical function (3) expresses the fact that SRS dangerous failure can be caused by the random or systematic failure.

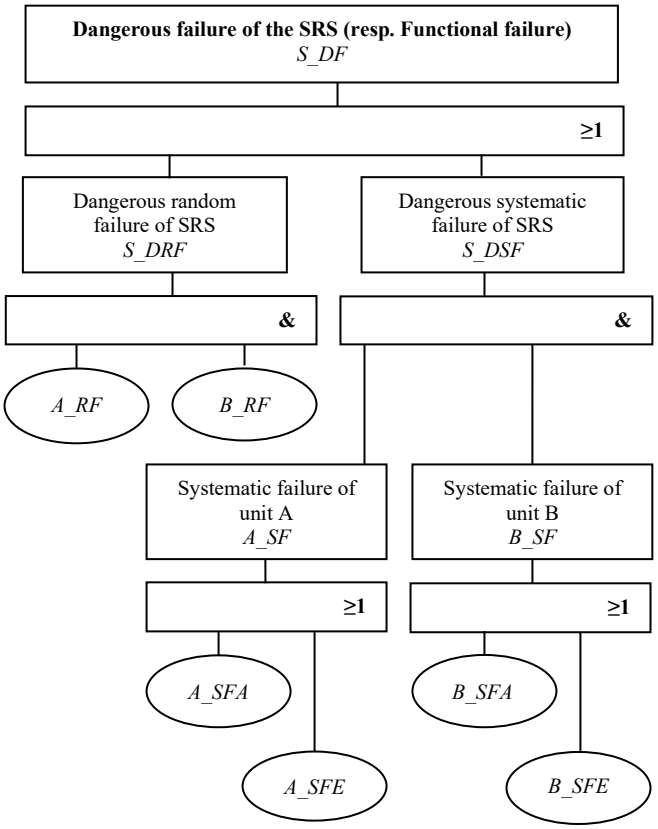


Figure 3: Fault tree for SRS in Figure 2

2.1. The influence of random failures on SF safety integrity

The primary construction elements of PLCs are the electronic components, which are characterized by the occurrence of the random failures. It is generally accepted that the occurrence of random failures of electronic components can be described by the exponential distribution rule. Manufacturers usually declare the failures rates for individual modules (in the case of sPLC, the dangerous failures rate). Based on this information and knowledge of the structure of SRS, the dangerous failures rate of SRS (dangerous failure rate of SF) can be calculated.

If the manufacturer declares the dangerous failure rate for the electronic system or its part (for example, the sPLC module for the implementation of SF with SIL 3), which is designed for industrial applications, then calculating the dangerous failure rate accept to assume that every second failure either alone or in combination with another failure (other failures) is dangerous [7]. Such an assumption may not be in accordance with the requirements for other applications. For example, for railway applications [6], if SIL 4 is required for an SF that is implemented with an electronic SRS, then it is considered that any failure may be potentially dangerous. This fact must be taken into account when calculating the failure rate of SF for a specific application.

If the units A and B are physically independent of each other, then the basic events A_RF and B_RF from the fault tree in Figure 3 are independent too. The probability of the dangerous random failure of the SRS can be expressed by relation 6:

$$SP_{S_DRF}(t) = P_{A_RF}(t).P_{B_RF}(t),$$
$$P_{S_DRF}(t) = \left(1 - e^{-\lambda_{A_RF}.t}\right).\left(1 - e^{-\lambda_{B_RF}.t}\right), \quad (6)$$

where the $P_{S_DRF}(t)$ is the probability of dangerous random failure of SRS, $P_{A_RF}(t)$is the probability of dangerous random failure of unit A, $P_{B_RF}(t)$ is the probability of dangerous random failure of unit B, λ_{A_RF} (λ_{B_RF}) is the dangerous random failures rate of unit A (B).

The dangerous random failure rate of SRS can be calculated using the equation:

$$\lambda_{S_DRF}(t) = \frac{\frac{dP_{S_DRF}(t)}{dt}}{1-P_{S_DRF}(t)}, \quad (7)$$

after substituting (6) into (7) and adjusting, it can be determined that:

$$\lambda_{S_{DRF}}(t) =$$
$$\frac{\lambda_{A_RF}.e^{-\lambda_{A_RF}.t} + \lambda_{B_RF}.e^{-\lambda_{B_RF}.t} - \left(\lambda_{A_{RF}} + \lambda_{B_{RF}}\right).e^{-\left(\lambda_{A_{RF}} + \lambda_{B_{RF}}\right).t}}{e^{-\lambda_{A_RF}.t} + e^{-\lambda_{B_RF}.t} - e^{-\left(\lambda_{A_{RF}} + \lambda_{B_{RF}}\right).t}}. \quad (8)$$

Assuming that $\lambda_{A_RF}.t \ll 1$ and $\lambda_{B_RF}.t \ll 1$, then the SRS dangerous failure rate is:

$$\lambda_{S_DRF}(t) \leq 2.\lambda_{A_RF}.\lambda_{B_RF}.t. \quad (9)$$

The mutual physical independence of unit A and unit B can be achieved by applying appropriate technical measures (e.g.: galvanic separation of the units, separate the power supply of unit A and unit B, ...).

SRS consists not only of sPLC, but also other elements necessary for obtaining information from the monitored (or controlled) process (equipment, machine) and elements for the realization of SF outputs. An example of a frequently used SRS wiring with sensors (S_A, S_B) and contactors (C_A, C_B) represents the Figure 4. Safe disconnection of the EUC from the power supply after affecting the sensors is the purpose of this wiring.

The wiring of sensors and contactors follows the sPLC manufacturer's recommendations and their selection depends on the specific application. In this case, contactors are used which, in addition to the coil (W_A, W_B) and the main contacts (c_{1A}, c_{1B}), also contain control contacts (c_{2A}, c_{2B}). The manufacturer must guarantee the co-operation of the main contact and the mechanically coupled auxiliary contact (it is the characteristic of relay type C according to [8]). The correct function of the

contactor is checked using feedback through application diagnostics. The feedback data is compared with the commands for the contactor in the application software. Test diagnostics in the application software and comparison of feedback data are discussed in more detail, e.g. [9] and [10]. Diagnostics can be performed in distinct forms and can cover a whole system [11], [12] or be specifically oriented on selected system parts [13], [14].

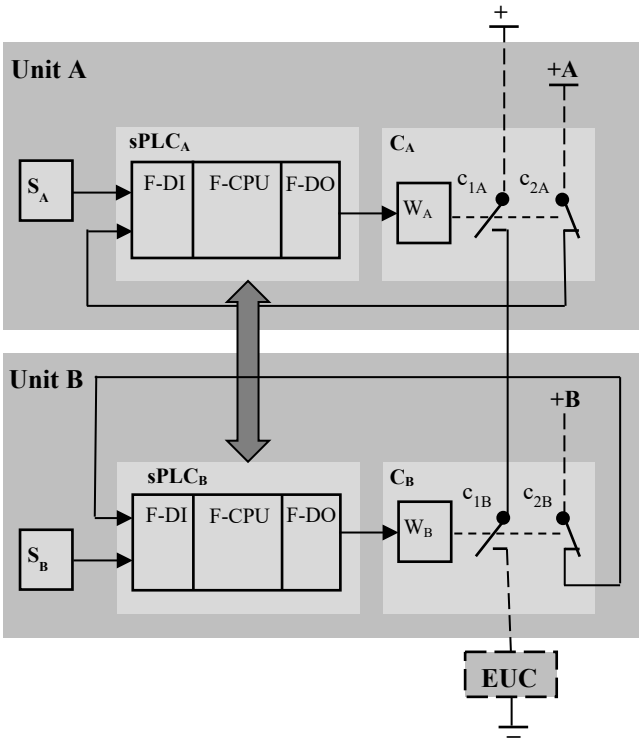


Figure 4: Wiring of the SRS

Occurrence of the dangerous failure due to random failure of SRS in Figure 4 can be described by a fault tree (Figure 5).

If the basic events of the fault tree in Figure 5 are independent (realistic and feasible assumption) then according to (9) it is valid that:

$$\lambda_{S_DRF}(t) \leq 2.\lambda_{SA}.\lambda_{SB}.t + 2.\lambda_{SA}.\lambda_{sPLCB}.t + 2.\lambda_{SA}.\lambda_{CB}.t + 2.\lambda_{sPLCA}.\lambda_{SB}.t + 2.\lambda_{sPLCA}.\lambda_{sPLCB}.t + 2.\lambda_{sPLCA}.\lambda_{CB}.t + 2.\lambda_{CA}.\lambda_{SB}.t + 2.\lambda_{CA}.\lambda_{sPLCB}.t + 2.\lambda_{CA}.\lambda_{CB}.t, \quad (10)$$

whereas:

$$\lambda_{sPLCA} = \lambda_{D_{FDIA}} + \lambda_{D_{FCPUA}} + \lambda_{D_{FDOA}},$$
$$\lambda_{sPLCB} = \lambda_{D_FDIB} + \lambda_{D_FCPUB} + \lambda_{D_FDOB}, \quad (11)$$

where λ_{SA} (λ_{SB}) is the sensor A (B) random failure rate, λ_{sPLCA} (λ_{sPLCB}) is dangerous random failure rate of sPLC$_A$ (sPLC$_B$), λ_{CA} (λ_{CB}) is random failure rate of the contactor A (B), λ_{D_FDIA} (λ_{D_FDIB}) is the module F-DI random failure rate of the unit A (B), λ_{D_FCPUA} (λ_{D_FCPUB}) is the module F-CPU random failure rate of the unit A (B), λ_{D_FDOA} (λ_{D_FDOB}) is the module F-DO random failure rate of the unit A (B).

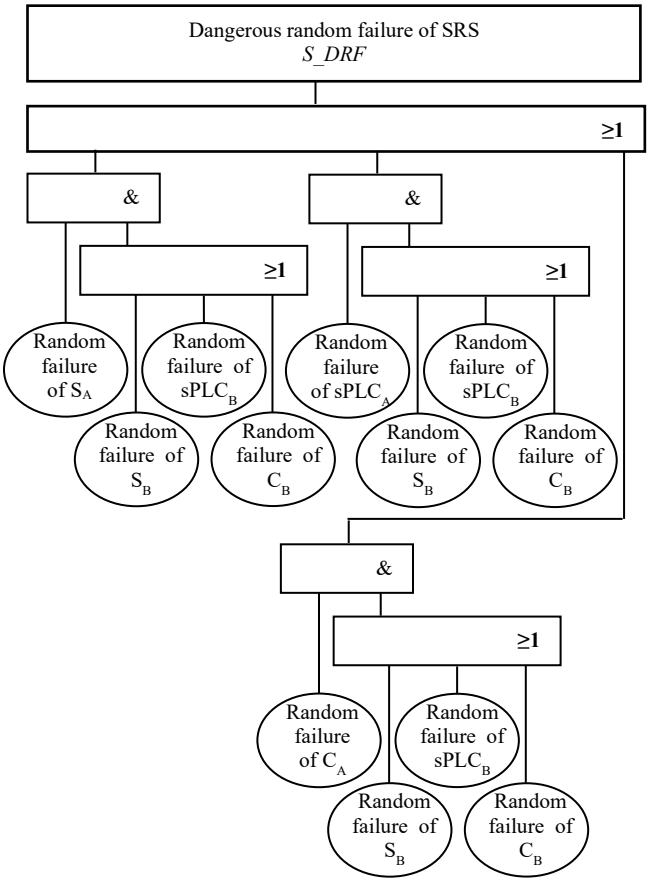


Figure 5: The fault tree of the SRS from Figure 4

If the SF realized by SRS (Figure 4) is to meet the RanF-SI requirement for SIL 4, then it must be true that

$$\lambda_{S_DRF}(t) \leq \lambda_{SIL4}. \quad (12)$$

It follows from (10) and (12) that

$$t_O \leq \frac{\lambda_{SIL4}}{2.X},$$
$$X = (\lambda_{SA}.\lambda_{SB} + \lambda_{SA}.\lambda_{sPLCB} + \lambda_{SA}.\lambda_{CB} + \lambda_{sPLCA}.\lambda_{SB} + \lambda_{sPLCA}.\lambda_{sPLCB} + \lambda_{sPLCA}.\lambda_{CB} + \lambda_{CA}.\lambda_{SB} + \lambda_{CA}.\lambda_{sPLCB} + \lambda_{CA}.\lambda_{CB}), \quad (13)$$

where t_O is the maximum allowed time of failure detection and negation.

2.2. Results of modeling the impact of random failures on the SF safety integrity - a case study

Let the SRS in Figure 4 implements one SF, in the implementation of which all elements participate. The failures rate, or dangerous failures rate of these elements are listed in the Table 1. Figure 6 shows the dependence of the SRS dangerous random failures rate ($\lambda_{S_DRF}(t)$) depending on the time of failure detection and negation (t_O) calculated according to the relation (9) - curve 1 and calculated according to the relation (8) - curve 2. Maximum tolerable dangerous failure rate for SIL 4 is shown by a horizontal line. The intersection of curve 1 or 2 with this line

determines the maximum allowed time of fault detection and negation. For curve 1 it means the time 3830 h and 3845 h for curve 2. The difference between determined times is relatively small and acceptable from a safety point of view because the time determined from the simplified relationship (9) is shorter.

Table 1: The failures rates and dangerous failures rates of realized SF

Element	Failure rate	Dangerous failure rate
S_A	5.10^{-7} h^{-1}	
S_B	5.10^{-7} h^{-1}	
$sPLC_A$		4.10^{-9} h^{-1}
F-DI 16x24VDC		1.10^{-9} h^{-1}
F-CPU 1516F-3PN/DP		1.10^{-9} h^{-1}
F-DQ 8x24VDC/2A PPM		2.10^{-9} h^{-1}
$sPLC_B$		4.10^{-9} h^{-1}
F-DI 16x24VDC		1.10^{-9} h^{-1}
F-CPU 1516F-3PN/DP		1.10^{-9} h^{-1}
F-DQ 8x24VDC/2A PPM		2.10^{-9} h^{-1}
C_A	$6{,}4.10^{-7}$ h^{-1}	
C_B	$6{,}4.10^{-7}$ h^{-1}	

The diagnostic mechanisms available to the SRS must guarantee that all potentially dangerous faults are detectable (diagnostic coverage DC = 100%). Another approach must be used to calculate the dangerous failure rate of SRS if this condition is not fulfilled - for example, a model based on the Markov chain [15] (in this paper, the fault tree analysis (FTA) is used). The issue of using different methods to assess the dangerous failure rate of SRS as well as considering the influence of other factors (besides the random failures rates and DC) on the dangerous failure rate of SRS is discussed in [15]. For the sake of clarity, the fault tree analysis (FTA) method is used in this article.

It can be stated that the realization of SF with SIL 4, with respect to RanF-SI, using a dual structure is realistically achievable.

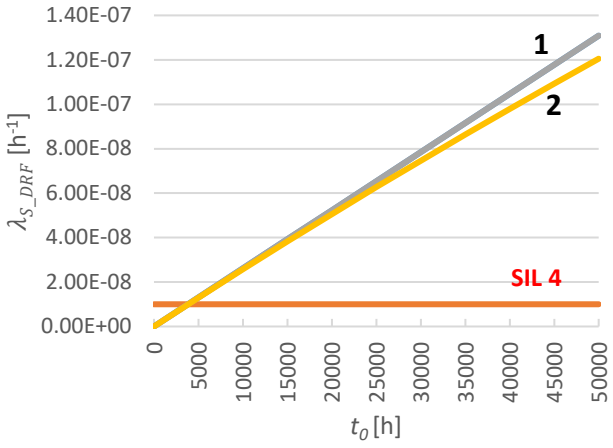


Figure 6: Dependence of the SRS dangerous random failures rate on the failure time detection and negation

2.3. The impact of systematic failures on the safety integrity of SF

While RanF-SI is a quantifiable part of SI, SysF-SI is a non-quantifiable part of SI. Probability calculations cannot be used for its evaluation (assessment), because the rate of the occurrence of the systematic failures is not known (practically, the occurrence of systematic failures is impossible to identify) and the distribution of systematic failures is not known. Achieving the required SysF-SI is based on the effective prevention of failures by applying appropriate and suitable safety measures (depending on the required SIL for SF), which are defined in the relevant standards for the given areas of applications. The evaluation (assessment) of SysF-SI is based on the assessment that the prescribed safety measures are effective and have been applied at a sufficient level [2].

The fault tree in Figure 3 shows the influence of systematic failure on the occurrence of the SRS dangerous failure. This influence is expressed by a logical function (5). In general, it should be assumed that unit A and unit B have only some mutual systematic failures, and thus it cannot be valid that the basic events *A_SF*, *B_SF* are mutually interdependent. Therefore, in this case, it must be evaluated by the general relation that:

$$P_{S_DSF}(t) = P_{A_SF}(t).P_{(B_SF/A_SF)}(t), \text{ resp.}$$
$$P_{S_DSF}(t) = P_{B_SF}(t).P_{(A_SF/B_SF)}(t), \quad (14)$$

where $P_{S_DSF}(t)$ is the probability of dangerous systematic failure of the SRS, $P_{A_SF}(t)$ ($P_{B_SF}(t)$) is the probability of systematic failure of unit A (B), $P_{(B_SF/A_SF)}(t)$is the conditional probability of systematic failure of unit B provided that the systematic failure of unit A has occurred and $P_{(A_SF/B_SF)}(t)$ is the conditional probability of systematic failure of unit A provided that the systematic failure of unit B has occurred.

In general, it is valid that

$$P_{A_SF}(t) \leq P_{(A_SF/B_SF)}(t) \leq 1, \text{ resp.}$$
$$P_{B_SF}(t) \leq P_{(B_SF/A_SF)}(t) \leq 1, \quad (15)$$

where $P_{(A_SF/B_SF)}(t) = 1$, resp. $P_{(B_SF/A_SF)}(t) = 1$, if the units A, B are identical (have the same systematic failures) and $P_{(A_SF/B_SF)}(t) = P_{A_SF}(t)$, resp. $P_{(B_SF/A_SF)}(t) = P_{B_SF}(t)$, if unit A and unit B are mutually independent (units do not have the same systematic failures).

Figure 7 shows in principle the influence of safety measures to prevent of systematic failures in consideration of the degree of interdependence of units A, B and therefore also to the on SI.

If the units A and B are HW and also SW identical (same sPLC and the same application software; Figure 7a), then it can be stated that the units A, B are dependent and the SF, which is realized by these units has the same level of SysF-IS as would have SF if it were realized only by one unit (unit A or unit B). This means that if the sPLC manufacturer states that using this sPLC can be realized SF with max. SIL 3, it must be assumed that the SysF-IS level is 3. The fact is that in such a case it does not make sense to use a dual structure.

Increasing the SysF-IS level can be in principle achieved in the following ways (safety measures):

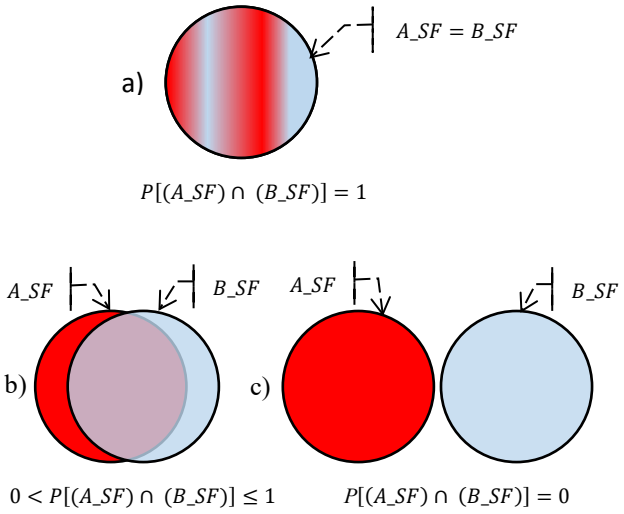


Figure 7: The illustration of the "overlapping" of the sets of systematic failures of unit A and unit B

- software diversification (different software in unit A and in unit B);
- hardware diversification (different hardware in unit A and in unit B, i.e. using of different sPLCs in unit A and in unit B);
- diversification of hardware and software.

Modifying the embedded software is a very effective safety measure to increase the SysF-IS level. However, when using sPLC, such a solution is practically impossible because this software is inaccessible to the user (this is information that the sPLC manufacturer does not provide). Diversification of application software is possible. In addition to functional algorithms related to the realized SFs, the application software can also implement mechanisms that allow the detection of a part of the systematic failures operating in a given unit. These are, for example, the following mechanisms:

- a clear, complete, and verifiable specification of requirements;
- registration of requirements using semi-formal or formal methods;
- compliance with coding standards;
- use of verified function blocks that are verified by "good" practice;
- consistent testing.

The fact is that the number of reduced systematic failures can be expected, but it cannot be demonstrated that the level of remaining systematic failures is acceptable. This is also related to the fact that the set of systematic failures is not known and therefore it is impossible to implement safety measures for their detection and subsequent negation in a targeted manner. In reality, some safety mechanisms implemented in application software may be redundant.

The aim of diversification is not to reduce the number of systematic failures in unit A, resp. unit B, but the goal is to minimize the mutual dependence (Figure 7b). This minimization is achieved by the fact that existing systematic failures do not affect the data manipulated in these units in the same way and at the same time (by acquiring or processing or transforming or storing or transmitting data). In association with diversification, it should be noted that a powerful diagnostic tool in the dual structure presented here is the comparison of data. The larger volume of data that is the subject of the comparison and the frequency of comparison make a greater probability of detecting a failure. Diversification brings the problems of ensuring the compatibility of different sPLCs and with the synchronization requirement to be comparison possible at all. Comparison of output signals only is not considered sufficient.

The ideal situation is shown in Figure 7c, when the interpenetration of systematic failures of unit A and unit B is zero.

3. Conclusion

The use of sPLCs in dual architecture (certified to SIL 3) primarily eliminates random hardware failures. In this way, it is possible to achieve the final SIL of the realized safety functions at level 4.

Fight against the systematic failures can be done best at the embedded software level. Due to the fact that the embedded software is fixed from the user's point of view, the use of sPLC from two different manufacturers and the subsequent application of discrepancy diagnostics may be a suitable solution. Discrepancy diagnostics must be focused on a thorough comparison of internal states, memory states, etc.

Manufacturers of the sPLCs certified to SIL 4 are already starting to appear (e.g. HIMatrix product from the HIMA Company [16]). However, such products are rare and their use is precisely limited to a particular area of application. It can be assumed that in the future, the development of sPLCs will be aimed at increasing their safety and expanding their application possibilities.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work has been supported by the Educational Grant Agency of the Slovak Republic (KEGA) Number 008ŽU-4/2019: Modernization and expansion of educational possibilities in the field of safe controlling of industrial processes using the safety PLC.

References

[1] K. Rástočný, J. Ždánsky, J. Hrbček, "The problems related to realization of safety function with SIL4 Using PLC," in Proceedings of the 30th International Conference on Cybernetics and Informatics, K and I 2020, Institute of Electrical and Electronics Engineers Inc., 2020, doi:10.1109/KI48306.2020.9039878.

[2] EN 61508 - European Standards, Functional Safety of Electrical/ Electronic/Programmable Electronic Safety-Related Systems, 2010.

[3] 1sig - Zabezpečovací systémy pro železnice, tramvaje a metro, První Signální, a.s.

[4] Schneider Electric - Orbus Hardwired Safety Systems, Schneider Electric Global.

[5] Pilz - Automation solution in the railway industry, Pilz GmbH & Co. KG.

[6] EN 50129 - European Standards, Railway Application. Communication, Signalling and Processing Systems. Safety-Related Electronic Systems for Signalling, 2018.

[7] EN 62061 - European Standards, Safety of Machinery – Functional Safety of Safety-Related Electrical, Electronic and Programmable Electronic Control Systems, 2016.
[8] EN 50578 - European Standards, Railways Applications. Direct Current Signalling Relays, 2013.
[9] J. Ždánsky, K. Rástočny, J. Hrbček, "Influence of architecture and diagnostic to the safety integrity of SRECS output part," International Conference on Applied Electronics, **2015**–**Octob**, 297–301, 2015.
[10] K. Rástočný, J. Ždánsky, J. Balák, P. Holečko, "Diagnostics of an output interface of a safety-related system with safety PLC," Electrical Engineering, **99**(4), 1169–1178, 2017, doi:10.1007/s00202-017-0624-1.
[11] A.H. Naghshbandy, H.M. Shanechi, A. Kazemi, I. Pourfar, "Study of fault location effect on the inter-area oscillations in stressed power systems using modal series method," Electrical Engineering, **92**(1), 17–26, 2010, doi:10.1007/s00202-010-0154-6.
[12] W. Velten-Philipp, M.J. Houtermans, "The effect of diagnostic and periodic proof testing on the availability of programmable safety systems," WSEAS Transactions on Systems, **5**(8), 1861–1867, 2006.
[13] J.C. Urresty, J.R. Riba, L. Romeral, J.A. Ortega, "Mixed resistive unbalance and winding inter-turn faults model of permanent magnet synchronous motors," Electrical Engineering, **97**(1), 75–85, 2014, doi:10.1007/s00202-014-0316-z.
[14] M. Moujahed, H. Ben Azza, K. Frifita, M. Jemli, M. Boussak, "Fault detection and fault-tolerant control of power converter fed PMSM," Electrical Engineering, **98**(2), 121–131, 2016, doi:10.1007/s00202-015-0350-5.
[15] K. Rástočný, J. Ždánsky, M. Franeková, I. Zolotová, "Modelling of diagnostics influence on control system safety," Computing and Informatics, **37**(2), 457–475, 2018, doi:10.4149/cai_2018_2_457.
[16] HIMatrix - The Compact Safety Solution, HIMA Paul Hildebrandt GmbH.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1079-1090 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Prioritization of Sustainable Supply Chain Management Practices in an Automotive Elastomer Manufacturer in Thailand

Saruntorn Mongkolchaichana, Busaba Phruksaphanrat*

Thammasat University Research Unit in Industrial Statistics and Operational Research, Industrial Engineering Department, Faculty of Engineering Thammasat University, Pathum-thani, 12121, Thailand

ARTICLE INFO

Article history:
Received: 23 December, 2020
Accepted: 03 February, 2021
Online: 16 February, 2021

Keywords:
Sustainable supply chain management
Multiple attribute decision making
Logarithmic fuzzy preference programming
Prioritization
A case study

ABSTRACT

Nowadays the sustainable awareness trend is increasing. The consumers' attitude has changed, causing companies to pay more attention to management in a sustainable way. Effective sustainable supply chain management (SSCM) can increase social, economic, and environmental benefits. Important factors from literatures were gather and organized to be a framework for SSCM. The proposed framework incorporates the whole supply chain for both internal and external activities, which can be applied to a manufacturer. The case study factory, which is an automotive elastomer producer has planned to adopt SSCM, so it needs to know the main factors for its operations. Logarithmic fuzzy preference programming method (LFPP) was used to rank SSCM criteria. The results of ranking important criteria showed that external factors (government and competition) were the most significant criteria that the factory has determined. Government and competitors are significant drivers that initiate the company to implement SSCM. Regulations and standards were good guidelines to SSCM for the factory. Next, the Triple Bottom Line (TBL) criteria (social, economic, environment) were considered in the overall operations. Not only concerning about cost and profit, but also environmental effect and social responsibility are cooperated. Finally, internal factors (supplier, consumption, and company) were considered with low level of importance. The proposals of actions of the company were also shown as a guideline for a manufacturer.

1. Introduction

This paper is an extension of work of a similar concept presented initially in the ICIEA 2019 conference [1]. The previous work considered the criteria for green supply chain management (GSCM), while the current work extended the scope to sustainable supply chain management (SSCM). In this research, the main criteria for SSCM were collected and organized to be a new framework for an automotive elastomer factory [1–4].

The evolution of process flow management from raw materials to finished goods is known as supply chain management (SCM). SCM plays an important role in any organization's success. However, there are many operations in the supply chain, which cause environmental impact to the social communities. The concept of sustainability has gained increased popularity due to the increase in environmental problems. Organization's management strategies have to improve and adjust to the current world situation for survival and friendly environment by not destroying the environment and not harming social communities. The use of environmentally friendly raw materials in the most effective way is one of the interesting practices [5]. Moreover, the participation of communities by corporate social responsibility (CSR) activities or developing community projects, helps to ensure the longevity of the organization.

SSCM is effective management that considers economic, environmental, and social performances at the same time in the supply chain. The basic aim of GSCM is to eliminate all wastes within the industrial system and limit use of hazardous substances. The aims of SSCM concern not only profit but also environmental and social dimensions, called a Triple Bottom Line (TBL). Companies must consider environmental effects from

*Corresponding Author Busaba Phruksaphanrat, Thailand, +66898962200 & Email: lbusaba@engr.tu.ac.th

processes in the supply chain such as purifying the toxic water or air emitted by the manufacturing process before releasing it to the environment to maintain good health in the community [6]. There are many pieces of research on case studies of SSCM. Some of them consider the specific operation in the supply chain such as supplier selection, reverse logistics, packing, etc [7–9]. However, the overall framework for SSCM for a supply chain needs to be clarified and important factors also need to be highlighted.

SSCM has various factors that affect the operations, which differ in each organization; also, the levels of action within each organization need to be considered differently too. Sari (2017) recommended considering inbound operations, production operations, outbound operations, and reverse logistics as primary criteria for the decision making in GSCM practices [10]. Uygun and Dede (2016) suggested a green design, green purchasing, green transformation, green logistics, and reverse logistics criteria. These activities are mainly concerned with environmental awareness [11]. Wu et al. (2020) and Mastrocinque et al. (2020) considered TBL as primary criteria throughout the SCM [12,13]**.** These criteria need to prioritize and formulate appropriate internal management strategies to achieve more benefit for the stakeholders in the most efficient way.

Multi-criteria decision making (MCDM) method is one of the most commonly used methods for ranking [14]. There are many methods of MCDM such as Analytical Hierarchy Process (AHP), Fuzzy Set Theory, Case-based Reasoning, Data Envelopment Analysis (DEA), Simple Multi-Attribute Rating Technique, Goal Programming, Simple Additive Weighting (SAW), and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) [15,16]. However, in evaluation criteria, the decision-maker may not be able to precisely decide the value for the decision, so the fuzzy set concept was proposed [17]. Then, many fuzzy methods for prioritization were presented [18]; these include the logarithmic least-squares method, geometric mean method, extent analysis method, lambda-max method, fuzzy preference programming method (FPP), linear goal programming method, and logarithmic fuzzy preference programming method (LFPP). Nevertheless, these methods still have some pitfalls to avoid [19–21]. The LFPP method has been chosen in this research to prioritize the criteria for three main reasons. The first reason is realistic computation; there is not an argument about this method from literature reviews. The second reason is that the data's consistency can be calculated, and can be re-evaluated over a short evaluation time. Finally, LFPP can maintain Saaty's AHP assessment rules in all aspects.

In this research, a case study of an automotive elastomer manufacturer was studied. Currently, the case study factory has already applied GSCM, but the management team desire to get a higher level of consideration about sustainability to survive in rapidly changing societies. The new goal of the factory is to implement SSCM in order to focus on social, economic, and environmental aspects. So, the main criteria that influence SSCM implementation should be investigated initially. A new framework of SSCM for the factory is presented. The organization's main objective is to achieve the lowest internal costs (the most profitable) and the least impact on the environment (using environmentally friendly raw materials and the most efficient use) to ensure that the product is green before releasing it to the market and consumers. Criteria were used to construct a new framework composed of SSCM factors for both internal and external factors, it can be used with any factory. All critical factors are integrated into the framework, which can satisfy the qualified policy of the factory. This research aims to prioritize the important criteria that affect the SSCM of the case study factory by use of LFPP, so that the factory can know which criteria should be emphasized and invested firstly. Then, the goals and direction can be clearly set. The proposals of actions of the factory also present in this research.

This research is organized as follows: Section 2 mentions the definition of SSCM and the detail about SSCM activities involved in the operations. Relevant literature was reviewed to determine the important criteria for SSCM of the case study. Section 3, calculation procedure by LFPP method is discussed. Section 4, the results from prioritization of the important criteria that affect the SSCM in this case study are deliberated. Conclusion, discussion, and future research are presented in section 5.

2. Sustainable Supply Chain Management

SSCM is an extension of GSCM, which mainly focuses on environmental, economic, and social dimensions [7]. These three dimensions are related to the Triple Bottom Line (TBL) principles [22], in which the perspectives are focused on the achievement and success of an organization in the economy, environment, and society as shown in Figure 1. They are the key to success of sustainable development, which needs to be balanced [13,22,23].

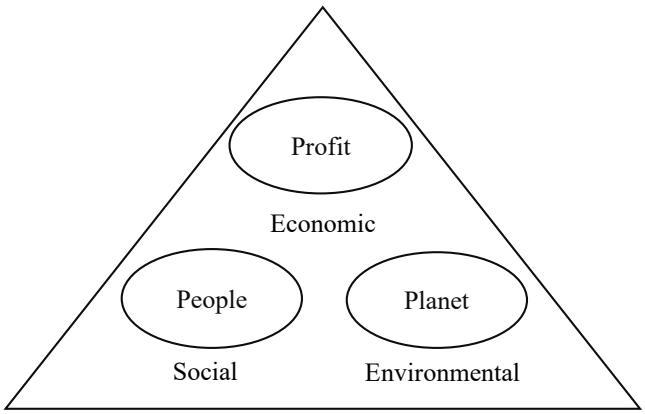


Figure 1: Triple bottom line for SSCM [24]

The objective of the organization previously only focused on profit, which cannot maintain sustainability. Many companies have turned to concern about people, societies, stakeholders, planet, and environmental responsibility. The TBL-based business aim is not only high profit, but also more consideration about the social and environmental problems. TBL shows profit at the top of the triangle. It is an economic dimension that almost all organizations desire.

The concept of people or humans in TBL is to emphasize fair business practices to the communities, employees, and societies by not causing harm to workers and people in the communities, but at the same time, the business has to make a profit for shareholders. Moreover, fair trade must be performed for the people dimension to balance with a social perspective.

Planet or natural dimension means a business should support sustainability and the diversity of the environment. The operations of the company must produce as little waste as possible from resources and must use recycling methods and reduce toxic gas and wastes to secure the environment and biodiversity, so the planet dimension is identical with an environmental perspective.

2.1. Sustainable activities

SSCM has comprehensive activities throughout the workflow processes. SSCM activities must care about social, economic, and environmental dimensions at the same time. These three perspectives incorporated with administrative activities start from planning strategies through to reverse logistics activities. They are described as follows:

- Sustainable strategy and planning: Strategic planning by top management is essential for organizations' initiative to improve and change [25]. If there is no strategic planning or goal setting, there is no direction that makes people understand the management intent [29, 30]. The implementation of SSCM defines the company's goal to sustainability, in which everyone should adhere to the three perspectives: social, economic, and the environment in all operations.
- Sustainable design: The design is not only for recyclable packaging or readily biodegradable products [8,28,29]; SSCM's sustainability design goes beyond that. Green design is a comprehensive design across SCM that is the best starting point for green purchasing, green production, green packaging, and green logistics. In recent years, biodegradable materials from renewable natural resources have received extensive government support.
- Sustainable purchasing: The purchasing managers need to be aware of sustainable raw materials that have to be compared among suppliers. There are many factors to be concerned about, which are cost, delivery, quality, environmentally friendly products, and value. Besides, the company also has the freedom to choose suppliers and avoid bribery from suppliers [7].
- Sustainable production: The production process design has many points to consider [8], such as a suitable location for machinery to reduce energy consumption or less fuel consumption and to get the best performance, type of source of energy, the amount of waste produced by the production [30].
- Sustainable packaging: The essential feature of product packaging is the ability to protect products from contamination damage, and deterioration. The packaging design may include special techniques to make good quality packaging, which is designed according to the 3R theory, which is recyclable (the product can be modified and reuse), reduce (reduce the use of raw materials in the production of packaging or the product itself), and reuse (the product or packaging can be reused). Also, packaging design should be easily disassembled for digestibility and create as low waste material percentage as possible [8,28,30]. Moreover, biodegradable packaging is also a favorite choice for manufacturers.
- Sustainable logistics: The green logistics design process has many points to consider, such as the design of the distribution route for the most efficient transportation, the minimization of the number of trips [31,32], the reduction of transport cost, the volume of the empty return transport called backhauling [33], the selection of transport vehicles to reduce CO_2 emissions that cause environmental impact [34], the selection of third-party logistics service providers [35], the warehouse design for convenient loading and unloading of products without congestion, and the reduction of waiting time and administrative costs [28]. If sustainable logistics can be applied, it can cover three perspectives, which are social perspectives such as management of traffic congestion, economic perspective by reducing the transportation cost and transportation fuel consumption, and environment perspective by reducing the number of transportation trips and distances to reduce the greenhouse gas emission.
- Sustainable consumption: Determine how to consume the raw materials in the right way to minimize the processes that bring environmental pollution. The instructions must be written in words that are easily understood. The company must analyze consumer trends, consumer characteristics [36], customer green preference [37,38], and consumer purchase behavior [39] of the new generation of people who are becoming increasingly influential in the global green and sustainable market [40].
- Sustainable reverse logistics: Due to the increasing environmental impact today, various industries need more control over the amount of waste from the production processes [9]. Controlling waste is a waste management process that needs to delve into great detail to achieve the most efficient reduction in pollution and use resources efficiently [28]. Therefore, setting up a waste center is needed to receive the used products to facilitate consumers and initiate reverse logistics to be more successful [40].

2.2. Sustainable criteria for implementation

Every activity consists of essential components; for example, green purchasing has a vital role in choosing which supplier is suitable to deliver material that satisfies the company's qualification standards. Therefore, a review of the critical criteria that directly affect SSCM activities must be compiled to cover all management practices. In the previous work, the guidelines for GSCM were proposed [1]. There are 6 criteria: government, competitor, social, supplier, customer, and company to prioritize the criteria that are important to GSCM. Further pieces of literature have proposed more criteria for SSCM. These criteria can be incorporated to construct the SSCM framework for defining an action plan as shown in Table 1.

Table 1: The main criteria for SSCM implementation.

Criteria	company	supplier	competition	consumption	govemment	social	economic	environment
[2]	✓	✓	✓	✓	✓	✓	✓	✓
[3]	✓	-	-	-	-	✓	✓	✓
[4]	✓	✓	✓	-	-	✓	✓	✓
[5]	✓	-	✓	✓	✓	✓	✓	✓
[8]	✓	-	-	-	✓	✓	✓	✓
[13]	✓	-	-	-	✓	✓	✓	✓
[26]	✓	-	✓	-	-	-	-	✓
[27]	✓	✓	-	✓	✓	-	✓	-
[29]	✓	✓	-	-	-	-	✓	✓
[41]	✓	✓	-	✓	-	-	-	-
[42]	✓	-	-	-	-	✓	✓	✓
[43]	✓	-	✓	✓	-	-	✓	-
[44]	✓	✓	-	✓	-	-	-	✓
[45]	✓	✓	-	✓	✓	-	-	✓
[46]	✓	-	-	-	-	-	-	✓
[47]	✓	-	-	-	✓	✓	✓	✓
[48]	✓	✓	✓	✓	-	-	-	✓
[49]	✓	-	-	-	-	✓	-	✓
[50]	✓	✓	-	-	-	✓	✓	✓
[51]	✓	✓	✓	-	✓	✓	✓	✓
[52]	✓	✓	-	-	-	-	✓	-
[53]	✓	✓	-	-	-	-	✓	✓
[54]	✓	-	-	✓	✓	-	✓	-

There are 8 main criteria related to SSCM; company, supplier, competition, consumption, government, social, economic, and environment.

2.3. The sustainable supply chain management framework

In the past, the SCM of an organization was not complicated and consisted of a few stakeholders. Most companies focused on developing economic efficiency, such as technical quality development, reducing cost, and improving delivery performance. However, today, business operations are more complicated, leading to a broader range of stakeholders [55]. Supply chain activities are transformed into a complicated network model [56]. The suppliers, manufacturers, distributors, and logistic providers must work together to enhance their competitiveness. Meanwhile, other stakeholders such as consumers, investors, employees, and society pressure the manufacturers to be concerned about social and environmental issues that affect them. Therefore, nowadays organizations have adopted a sustainable development approach to their SCM [6,33], which considers whole parts of the operation throughout the supply chain known as SSCM. SSCM helps the organization to reduce the risk rate and enhances the competitiveness of the business. Promoting good corporate governance throughout the life cycle of products and services can improve environmental performance, social performance, operations performance, and competitiveness of an organization [57].

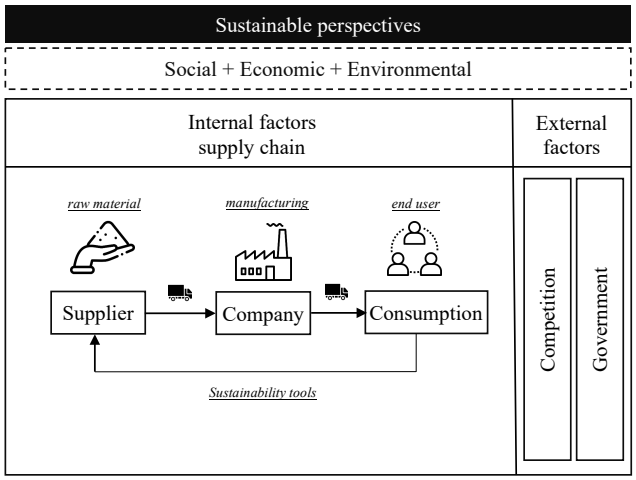


Figure 2: A framework of an SSCM for a manufacturer

A new framework of an SSCM is proposed in this research as shown in Figure 2. It consists of three parts. The first part is external factors, referring to surrounding factors that can indirectly affect the corporate supply chain, which are competition and government. The second part is internal factors, referring to the composition of the main parts of the supply chain, which start from supplier, to company, and end at consumption. Next, the TBL consists of environmental, social, and economic aspects. They are the basis for sustainability. Sustainability tools can be applied to the framework to support the implementation. Whitehead et al. (2020) proposed the tools, which consist of 3 groups; tools for action, tools for analysis and evaluation, and tools for communication [58]. It is also summarized tools and approaches for sustainability, which are cradle to cradle, GSCM, life cycle analysis (LCA), eco-design, reverse logistics, design for environment (DFE), Quality Function Deployment (QFD) for the environment, sustainable value analysis tool (SVAT), sustainability balanced scorecard (SBSC), corporate social responsibility (RSC) and sustainable value stream mapping (Sus-VSM) [59]. These sustainability tools can help an organization's supply chain to achieve successful SSCM. Criteria based on the framework are summarized in Figure 3. They can be described as follows.

Government is the criterion related to enacting laws to enforce and setting standards or regulations that industrial factories must follow. Government criterion is divided into two parts: enforcement of laws or policies and recommendations for

manufacturers from the government. Industries must follow the government laws; if any factory fails to follow the laws, there are penalties. Policies and recommendations from the government might be disregarded if the company is not willing to implement them. It depends on the individual organization's willingness or the opinions of executives within the organization. Many policies require action and enforcement that a business may select by itself or by the government. The role whereby a government can facilitate sustainability will depend on its authority which is different in each country. For example, a government with strong political leaders may generate models which benefit the local community and focus on supporting economic growth. In welfare-state models, the government provides services to firms and nonprofit organizations or takes on social-economic development [60]. Manufacturers determine their profit and level of energy efficiency. The government has sustainable goals that relate to saving energy, seeking profit, and increasing social welfare. Therefore, sustainability is a fundamental issue for both organizations and the government [61].

Competition in the market is one of the critical factors. Competition may be in the form of price, quality, variety of products, responsiveness, or environmental awareness, etc. In the case of SSCM implementation, social trends and substitute goods can affect consumers in making a decision. The organization should be concerned about the market trend to adapt to the competitive market. If the organization does not become active or is slow to adjust to consumption flow, it will lose market share. Competition in each type of market is different. Every organization needs fair competition as an ideal, but in the real situation, it may not be. Therefore, the rapid change of the organization is an advantage and benefit for the organization. The company's choices can affect the decisions of the firm's competitors. The firm invests in sustainability only if it can gain more revenue or reduce cost [62].

Economic sustainability has been defined as gaining more income and stability for society. In the current situation, an economy which does not disturb natural, social, and human societies is needed for sustainability [63]. If the current situation of the economy is getting worse, the investment in SSCM operations will be debated at management board level, and the SSCM operations may be delayed. On the other hand, if the economy is good, a high profit from operations or good cash flow may allow an organization to be able to start SSCM operation quickly. Moreover, the policy formulation from the government is an important factor. For example, if the government requires less tax on the green product, the product price will not be high, it will generate more incentives for consumption and produce more profit. The organization also wants to make more of the products that are best-sellers in the market. All the above result from the first criterion, which is government.

The social factor has direct and indirect effects on SSCM that start with employees on the production line. If the organization has sustainable production management, such as equipped with technology to optimize the use of resources as needed and get the most efficient outputs, minimize energy consumption and air polluting emissions [64], employees do not have the risk of direct exposure to volatile chemicals during production, production accidents will be reduced and the operators of manufacturing plant should have a good quality of life as well. Furthermore, supposing the organization has a policy to apply CSR, it will benefit the surrounding communities. The government is very important in supporting the policy by reducing tax for industrial factories that implement CSR [65, 66], which makes companies willing to undertake CSR projects.

The environment is related to various problems concerning pollutants from water [67], air, and other wastes from factories, which create increasing problems. Reduction of CO_2 emissions, energy consumption [71, 72] and wastes from the processes, products, packaging, and any substances that can increase the world's temperature, are critical issues. Different environmental regulations have started to emerge and be proposed to companies by governments. Companies must comply with environmental legislation existing in each country. Moreover, standards related to environmental consideration are one of the effective tools for the company to apply.

Selecting suppliers that market environmentally friendly raw materials without affecting the surrounding communities can help when implementing SSCM. However, this criterion is difficult to apply in order to make a significant impact on SSCM because organizations cannot enforce or control all suppliers, so choosing a supplier is yet another option in implementing SSCM within the organization. The selection of suppliers requires several factors to be considered [22, 70]. The key principles are the high quality of raw materials, reasonable prices, and on-time delivery.

Consumption can take place only if consumers are more aware of the environment and communities. The eco-label campaign is part of the approachability that can greatly appeal to consumers who are passionate about green consumption. Using its features is an important selling point. Currently, the consumer has a greater preference for healthy and environmentally friendly products [71]. Consumer perspective in sustainability is the driver for a company in the caring economy [72].

The company, including the management team, realize that the operation of SSCM requires a budget for investment [62]. Furthermore, employees within the organization have to adjust to the work with more responsibilities [73]. They must also cooperate to achieve the company strategic goal of SSCM. Moreover, organizations should provide support such as services, knowledge, information, training and facilities for them to achieve efficiency and effectiveness as soon as possible.

Criteria

Company
- Information
- Top management support and motivation
- Process /System operation
- Strategy and goal
- Incentives and rewards
- Reputation loss
- Business characteristics
- Organizational culture
- Innovation

Supplier
- Raw material use (material toxicity and chemicals)
- Trade groups
- Certification

Consumption
- Green image
- Green product
- Consumers' characteristics
- Reverse logistics
- Feedback

Competition
- Market segment
- Product pricing
- Competitive advantage
- Fair competition

Government
- Laws / Regulation / Standard
- Government support
- Disposal green policies
- Environmental policies of government
- Transparency
- Pressure

Social
- Communities
- Corporate social responsibility

Economic
- Cost and benefit
- Tax on green product
- Production cost

Environment
- CO_2 emission
- Risk management
- Pollution prevention
- Energy reduction
- Waste reduction

3. Logarithmic fuzzy preference programming method

Wang and Chin (2011) have proposed a logarithmic fuzzy preference programming method (LFPP) for finding the weight to prioritize the fuzzy number by comparing pairs of the dataset from the decision matrix [74]**.** The basis of this method was developed from the FPP method of Mikhailov (2004) by using nonlinear functions computation to find the weight with the following [75]:

$$ln\ \tilde{a}_{ij} \approx (ln\ l_{ij}, ln\ m_{ij}, ln\ u_{ij}),\ i, j = 1,..., n. \quad (1)$$

where $\tilde{a}_{ij}$ is an approximate triangular fuzzy number.

The membership function is shown below.

$$\mu_{ij}\left(ln\left(\frac{w_i}{w_j}\right)\right) = \begin{cases} \dfrac{ln\left(\frac{w_i}{w_j}\right) - ln\ l_{ij}}{ln\ m_{ij} - ln\ l_{ij}}, ln\left(\frac{w_i}{w_j}\right) \le ln\ m_{ij} \\ \dfrac{ln\ u_{ij} - ln\left(\frac{w_i}{w_j}\right)}{ln\ u_{ij} - ln\ m_{ij}}, ln\left(\frac{w_i}{w_j}\right) \ge ln\ m_{ij} \end{cases} \quad (2)$$

$\mu_{ij}\,(ln\ (w_i/w_j))$ is a membership of function $ln\ (w_i/w_j)$ belonging to the approximate triangular fuzzy judgment $ln\ \tilde{a}_{ij}$. A crisp priority vector to maximize the minimum membership degree $\lambda = \min\{\mu_{ij}(\ln(w_i/w_j)) \mid i = 1,\ldots, n-1; j = i+1,.., n\}$. The subsequent model can be constructed as

Maximize λ

$$\text{Subject to} \begin{cases} \mu_{ij}\left(\ln\left(\frac{w_i}{w_j}\right)\right) \ge \lambda, i = 1, \ldots, n-1; j = i+1, \ldots, n, \\ w_i \ge 0, i = 1, \ldots, n, \end{cases} \quad (3)$$

or as

Maximize 1 - λ

$$\text{Subject to} \begin{cases} \ln w_i - \ln w_j - \lambda \ln(m_{ij}/l_{ij}) \ge \ln l_{ij}, \\ -\ln w_i + \ln w_j - \lambda \ln(u_{ij}/m_{ij}) \ge -\ln u_{ij}, \\ i = 1, \ldots, n-1; j = i+1, \ldots, n, \\ w_i \ge 0, i = 1, .., n. \end{cases} \quad (4)$$

To avoid λ from taking a negative value, Wang and Chin (2011) proposed nonnegative deviation variables δ_{ij} and η_{ij} for $i = 1,\ldots,n-1$ and $j = i+1,\ldots, n$ such that they meet the following inequalities:

$\ln w_i - \ln w_j - \lambda\ln(m_{ij}/l_{ij}) + \delta_{ij} \geq \ln l_{ij}$, $i = 1,\ldots, n-1; j = i+1,\ldots, n$,

$-\ln w_i + \ln w_j - \lambda\ln(u_{ij}/m_{ij}) + \eta_{ij} \geq -\ln u_{ij}$, $i = 1,\ldots, n-1; j = i+1,\ldots, n$.

Then, the model to maximize the minimum membership degree becomes

MODEL: LFPP

$$Minimize J = (1-\lambda)^2 + M \cdot \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\delta_{ij}^2 + \eta_{ij}^2\right)$$

$$Subject\ to \begin{cases} x_i - x_j - \lambda ln\left(m_{ij}/l_{ij}\right) + \delta_{ij} \geq lnl_{ij}, \\ i = 1,\ldots,n-1; j = i+1,\ldots,n, \\ -x_i + x_j - \lambda ln\left(u_{ij}/m_{ij}\right) + \eta_{ij} \geq -ln\, u_{ij}, \\ i = 1,\ldots,n-1; j = i+1,\ldots,n, \\ \lambda, x_i \geq 0, i = 1,\ldots,n, \\ \delta_{ij}, \eta_{ij} \geq 0, i = 1,\ldots,n-1; \\ j = i+1,\ldots,n, \end{cases} \quad (5)$$

Let $x_i = ln\, w_i$ for $i = 1,\ldots, n-1$,

M is a large constant value, x_i^* is an optimal solution, λ is a membership degree, δ_{ij}, η_{ij} are nonnegative deviation variables for $i = 1,\ldots, n$-1 and $j = i+1,\ldots, n$ and

$l_{ij} = 1/u_{ji}$, $m_{ij} = 1/m_{ji}$, $u_{ij} = 1/l_{ji}$ *and* $0 < l_{ij} \leq m_{ij} \leq u_{ij}$ for all $i,j = 1,\ldots,n; j \neq i$.

Normalize the value of x_i^* and sorting fuzzy pairwise comparison matrices by (6)

$$w_i^* = exp(x_i^*)/\ \Sigma^n{}_{j=1}\, exp(x_i^*),\ i = 1,\ldots, n. \quad (6)$$

where $exp(\ )$ is the function of exponential for which the calculation is $exp(x_i^*) = e^{x_i^*}$ for $i = 1,\ldots, n$.

w_i^* is the weight of each criterion from $i = 1,\ldots, n$.

The following equation is used to check the consistency of data.

$$l_{ij} \leq w_i/w_j \leq u_{ij},\ i = 1,\ldots, n-1, j = i+1,\ldots, n. \quad (7)$$

4. Prioritizing of SSCM criteria for the case study

The SCM process of an automotive elastomer manufacturer has a flow chart as shown in Figure 2. The supply chain flow process begins with the purchasing process, where sources of the best raw materials for the company such as a polymer, fillers, softeners, and other additives are procured. A supplier who has the lowest cost, which high-quality products and delivery on time was selected. After receiving raw materials, the production is done according to production planning. After the products are finished and pass the quality inspection, they are transferred to the packing process. Then, finished goods are delivered to customers. There are environmental concerns in all of the above activities, which need to be reviewed through three aspects: social, economic, and environmental. These processes are internal factors of the SCM. Additionally, the process would not be complete if it did not take the external factors into account. The external factors are competitors and the government. The competition between the companies is considered in terms of price, quality, strategies, and market share. Also, the company needs to follow the laws and regulations of the government. However, the government also provides support to the company to reach international standards. The three sustainable perspectives are integrated with both internal and external factors. The company can control all processes by setting the procedures for each aspect by using sustainability tools that can help an organization's supply chain to achieve successful SSCM.

The automotive elastomer manufacturer has decided to apply SSCM in the company, so the importance of each criterion needs to be clarified. Then, everyone in the company can move in the same direction emphasizing the critical criteria. Eight criteria in the previous section were considered and ranked. A fuzzy decision matrix was constructed and evaluated by five experts who are key men in the factory: the procurement supervisor, planning supervisor, quality control supervisor, production engineer, and research and development engineer who has worked in this company for more than 10 years. The decision matrix is shown in Table 2. Model in equation (5) for this fuzzy pairwise comparison matrix can be written as

$$Minimize\ J = (1-\lambda)^2 + M \cdot \sum_{i=1}^{7}\sum_{j=i+1}^{8}\left(\delta_{ij}^2 + \eta_{ij}^2\right)$$

$$Subject\ to \begin{cases} x_1 - x_2 - \lambda\ln\left(\frac{0.35}{0.26}\right) + \delta_{12} \geq \ln(0.26), \\ -x_1 + x_2 - \lambda\ln\left(\frac{0.57}{0.35}\right) + \eta_{12} \geq -\ln(0.57), \\ x_1 - x_3 - \lambda\ln\left(\frac{0.34}{0.24}\right) + \delta_{13} \geq \ln(0.24), \\ -x_1 + x_3 - \lambda\ln\left(\frac{0.46}{0.34}\right) + \eta_{13} \geq -\ln(0.46), \\ x_1 - x_4 - \lambda\ln\left(\frac{1.52}{0.80}\right) + \delta_{14} \geq \ln(0.80), \\ -x_1 + x_4 - \lambda\ln\left(\frac{2.41}{1.52}\right) + \eta_{14} \geq -\ln(2.41), \\ \vdots \\ x_7 - x_8 - \lambda\ln\left(\frac{2.09}{0.99}\right) + \delta_{78} \geq \ln(0.99), \\ -x_7 + x_8 - \lambda\ln\left(\frac{3.19}{2.09}\right) + \eta_{78} \geq -\ln(3.19), \\ \lambda, x_1, x_2, x_3, \ldots, x_8 \geq 0, \\ \delta_{12}, \delta_{13}, \delta_{14}, \ldots, \delta_{78} \geq 0, \\ \eta_{12}, \eta_{13}, \eta_{14}, \ldots, \eta_{78} \geq 0. \end{cases} \quad (8)$$

Taking a sufficiently large number for M, say M=1000, to solve this model with Lingo 17.0 as shown in Figure 4, the optimal solution can be obtained as

Table 2: Fuzzy comparison matrix of the aggregated weights of the criteria

Criteria		C_1 Company			C_2 Supplier			C_3 Competition			C_4 Consumption		
C_1	Company	1.00	1.00	1.00	0.26	0.35	0.57	0.24	0.34	0.46	0.80	1.52	2.41
C_2	Supplier	1.74	2.83	3.87	1.00	1.00	1.00	0.29	0.37	0.51	1.25	2.00	3.06
C_3	Competition	2.19	2.93	4.11	1.97	2.69	3.47	1.00	1.00	1.00	1.32	2.05	3.10
C_4	Consumption	0.42	0.66	1.25	0.33	0.50	0.80	0.32	0.49	0.76	1.00	1.00	1.00
C_5	Government	1.55	2.64	3.68	2.76	3.82	4.85	0.49	0.80	1.35	2.30	3.32	4.34
C_6	Social	1.64	2.09	3.13	1.07	1.52	2.22	0.40	0.64	0.94	1.00	1.52	2.46
C_7	Economic	1.43	2.00	5.00	1.05	1.28	2.78	0.45	0.81	1.00	1.01	1.56	2.56
C_8	Environment	1.59	2.94	3.33	1.27	1.61	2.50	0.45	0.66	0.99	1.28	1.79	4.35
Criteria		C_5 Government			C_6 Social			C_7 Economic			C_8 Environment		
C_1	Company	0.27	0.38	0.64	0.32	0.48	0.61	0.20	0.50	0.70	0.30	0.34	0.63
C_2	Supplier	0.21	0.26	0.36	0.45	0.66	0.93	0.36	0.78	0.95	0.40	0.62	0.79
C_3	Competition	0.74	1.25	2.05	1.06	1.55	2.49	1.00	1.23	2.21	1.01	1.51	2.22
C_4	Consumption	0.23	0.30	0.44	0.41	0.66	1.00	0.39	0.64	0.99	0.23	0.56	0.78
C_5	Government	1.00	1.00	1.00	1.52	2.27	3.37	1.49	2.07	3.07	1.09	1.98	2.98
C_6	Social	0.30	0.44	0.66	1.00	1.00	1.00	0.98	0.99	1.00	1.01	1.12	1.32
C_7	Economic	0.33	0.48	0.67	1.00	1.01	1.02	1.00	1.00	1.00	0.99	2.09	3.19
C_8	Environment	0.34	0.51	0.92	0.76	0.89	0.99	0.31	0.48	1.01	1.00	1.00	1.00

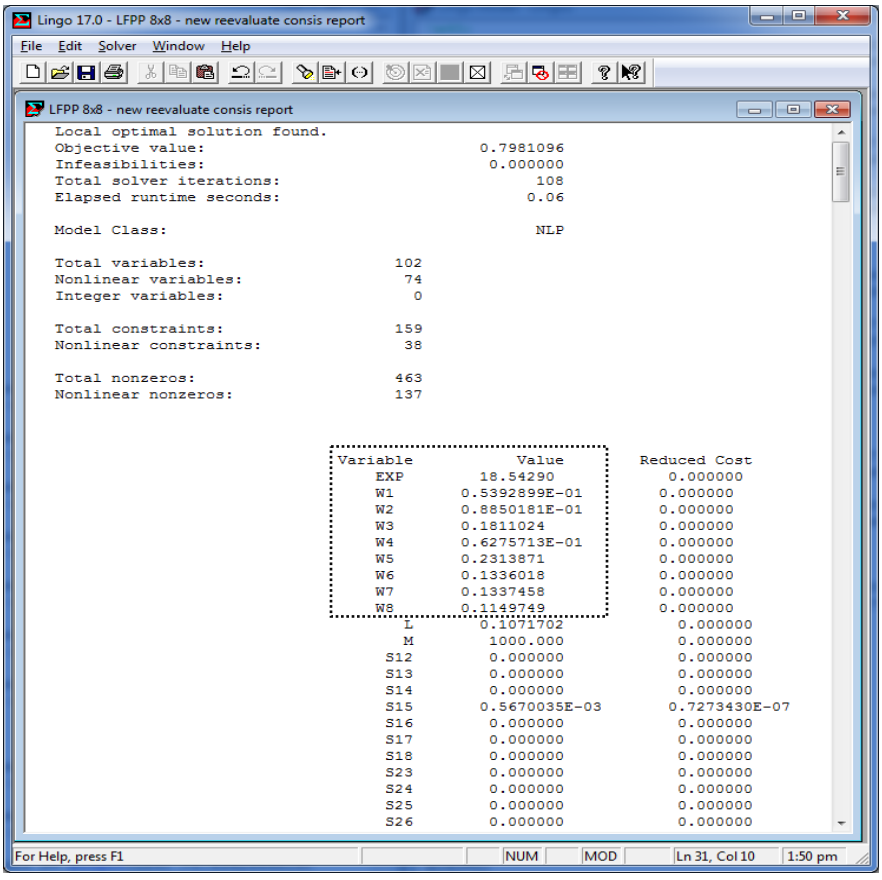


Figure 4: The result obtained from Lingo 17.0.

$$x_1^* = 0, x_2^* = 0.4656, x_3^* = 1.1459, \dots, x_8^* = 0.5329,$$
$$\delta_{12}^*, \delta_{13}^*, \delta_{14}^*, \dots, \delta_{78}^* = 0,$$
$$\eta_{12}^* = 0.0889, \eta_{13}^*, \eta_{14}^*, \dots, \eta_{78}^* = 0,$$

based on which, normalization of LFPP priorities as

$$w_1^* = \frac{EXP(x_1^*)}{\sum_{i=1}^{8} EXP(x_i^*)} = 0.0620,$$
$$w_2^* = \frac{EXP(x_2^*)}{\sum_{i=1}^{8} EXP(x_i^*)} = 0.0988,$$
$$w_3^* = \frac{EXP(x_3^*)}{\sum_{i=1}^{8} EXP(x_i^*)} = 0.1951,$$
$$\dots,$$
$$w_8^* = \frac{EXP(x_8^*)}{\sum_{i=1}^{8} EXP(x_i^*)} = 0.1057.$$

According to (7), the results of the consistent test found that there were four pairs outside the specified limits: a_{12}, a_{15}, a_{23}, and a_{25}, so the decision-makers needed to reevaluate the pairwise comparison judgment as shown in Table 3.

After reevaluating, weights were found and the importance could be sorted as shown in Table 4 and Figure 5. Lingo 17.0 with Intel (R) Core (TM) i5-3570 CPU @ 3.40GHz RAM 8.00 GB 64-bit was used to calculate weights of each criteria. The total calculation time for weights criteria was 0.06 seconds.

The prioritizing results clearly show that external factors, which are government and competition, are the main criteria for applying SSCM in the factory. Then, the economic, social, and environment criteria, which are TBL factors for sustainability, are considered. Their weights are similar which means that the emphasis of the factory on TBL is balanced. Finally, supplier, consumption, and company, which are internal factors have lower weightings.

Table 3: Reevaluate fuzzy comparison matrix

Criteria		C_1 Company			C_2 Supplier			C_3 Competition			C_4 Consumption		
C_1	Company	1.00	1.00	1.00	0.26	0.35	0.65	0.24	0.34	0.46	0.80	1.52	2.41
C_2	Supplier	1.54	2.83	3.87	1.00	1.00	1.00	0.29	0.37	0.52	1.25	2.00	3.06
C_3	Competition	2.19	2.93	4.11	1.92	2.69	3.47	1.00	1.00	1.00	1.32	2.05	3.10
C_4	Consumption	0.42	0.66	1.25	0.33	0.50	0.80	0.32	0.49	0.76	1.00	1.00	1.00
C_5	Government	1.55	2.64	4.55	2.50	3.82	4.85	0.49	0.80	1.35	2.30	3.32	4.34
C_6	Social	1.64	2.09	3.13	1.07	1.52	2.22	0.40	0.64	0.94	1.00	1.52	2.46
C_7	Economic	1.43	2.00	5.00	1.05	1.28	2.78	0.45	0.81	1.00	1.01	1.56	2.56
C_8	Environment	1.59	2.94	3.33	1.27	1.61	2.50	0.45	0.66	0.99	1.28	1.79	4.35
Criteria		C_5 Government			C_6 Social			C_7 Economic			C_8 Environment		
C_1	Company	0.22	0.38	0.64	0.32	0.48	0.61	0.20	0.50	0.70	0.30	0.34	0.63
C_2	Supplier	0.21	0.26	0.40	0.45	0.66	0.93	0.36	0.78	0.95	0.40	0.62	0.79
C_3	Competition	0.74	1.25	2.05	1.06	1.55	2.49	1.00	1.23	2.21	1.01	1.51	2.22
C_4	Consumption	0.23	0.30	0.44	0.41	0.66	1.00	0.39	0.64	0.99	0.23	0.56	0.78
C_5	Government	1.00	1.00	1.00	1.52	2.27	3.37	1.49	2.07	3.07	1.09	1.98	2.98
C_6	Social	0.30	0.44	0.66	1.00	1.00	1.00	0.98	0.99	1.00	1.01	1.12	1.32
C_7	Economic	0.33	0.48	0.67	1.00	1.01	1.02	1.00	1.00	1.00	0.99	2.09	3.19
C_8	Environment	0.34	0.51	0.92	0.76	0.89	0.99	0.31	0.48	1.01	1.00	1.00	1.00

Table 4: Comparison result for reevaluate

Criteria		Weight
C_5	Government	0.2314
C_3	Competition	0.1811
C_7	Economic	0.1337
C_6	Social	0.1336
C_8	Environment	0.1150
C_2	Supplier	0.0885
C_4	Consumption	0.0628
C_1	Company	0.0539

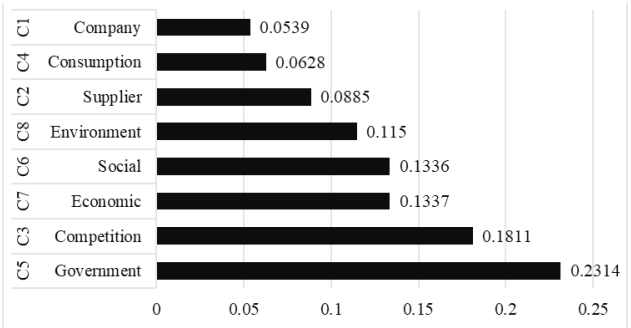


Figure 5: SSCM criteria weights

External pressures from the government and competitors' impact on SSCM for the automotive elastomer manufacturer in Thailand. There are some regulations and laws that control environmental problems that are caused by small particles and hazardous chemicals. The factory has followed the standards of the Ministry of Industry and certified ISO 14001:2015 for many years. The environmental issue is one of the company's strategies. It has an operation plan to protect the environment such as use fewer chemical substances, reduce wastes and pollution, etc. These two factors are very important in leading the factory to the implementation of SSCM.

In the competitive market with rapid change, the factory needs to adjust itself to satisfy the target market, which is the global market. Most of the management team in this factory are foreigners who mainly focus on the global trend. Sustainable awareness for the manufacturer is currently in practice.

The economic issue is important for all profit organizations to survive. However, it should be balanced with social and environment issues. Every project about sustainable management in the factory has to be evaluated by feasibility analysis before implementing it.

The company implemented GSCM before desiring to switch to SSCM, so environmental concerns have existed in almost all of the supply chain. It has been certified to ISO14001:2015 for many years and still keeps track to reduce environmental problems of the factory.

The factory has a policy to select high-quality suppliers. These selected suppliers can help the company to improve the firm's competitiveness across its supply chain. Reasonable price, high quality of raw materials, and fast response are the main factors of the selection. The existing suppliers rarely consider sustainability.

Most customers of the factory are foreign customers, who have less concern about the sustainability of the products. The company products are small parts for automotive manufacturers. Most of the automotive manufacturers are considering clean technology involving combustion in the engine, which is not related to the case study company's products.

The company has the least weight of importance for SSCM. It operates for maximizing profits but the regulations, standards, laws, and stakeholders have forced the company to be concerned not only with profit but also the environment and social issues.

5. Conclusions

This research proposes a framework for SSCM that combines the whole supply chain. It is applied to the case study factory, which is an automotive elastomer producer. Criteria are collected and factors (government and competition), TBL factors (environment, rearranged into the framework, which is suitable for a manufacturer. Three main groups of criteria, which are external social, and economic), and internal factors (supplier, company, and customer) are reviewed. These factors are prioritized by the LFPP method, which has advantages over other methods. Government and competition are the most important criteria for the case study factory that force it to implement SSCM. Laws, regulations, and standards are carefully determined and set for the company to achieve. They are set has the highest priority for the company to achieve. Market and competitors are studied and the operation plan related to these criteria is defined. Environment, social, and economic criteria are the next factors that the factory is concerned about. They are considered to be at the same level of importance, which is balanced according to the concept of SSCM. Profit is still the goal that the factory wants to achieve, but there is also an environmental concern and the certified environmental standard to be satisfied. Moreover, the factory has plans for corporate social responsibility activities with the communities in its action plan. Existing SCM involves suppliers, the company, and customers. They are considered as the less important criteria. The proposals of actions of the case study factory was shown as an example for the other factories.

For future research, the level of SSCM performance in each practice tool should be evaluated to know the efficient tools for the implementation of SSCM.

Acknowledgment

This research was support financially by Ph.D. Scholarship and research unit funding from Thammasat University, Thailand.

References

[1] S. Mongkolchaichana, B. Phruksaphanrat, "A Comparison of Prioritization Methods for Green Supply Chain Management Criteria Evaluation," 2019 IEEE 6th International Conference on Industrial Engineering and Applications, ICIEA 2019, (2015), 313–317, 2019, doi:10.1109/IEA.2019.8715191.

[2] S. Seuring, M. Müller, "From a literature review to a conceptual framework for sustainable supply chain management," Journal of Cleaner Production, **16**(15), 1699–1710, 2008, doi:10.1016/j.jclepro.2008.04.020.

[3] C.R. Carter, D.S. Rogers, "A framework of sustainable supply chain management: Moving toward new theory," International Journal of Physical Distribution and Logistics Management, **38**(5), 360–387, 2008, doi:10.1108/09600030810882816.

[4] M. Pagell, Z. Wu, "Building a more complete theory of sustainable supply chain management using case studies of 10 exemplars," Journal of Supply Chain Management, **45**(2), 37–56, 2009, doi:10.1111/j.1745-493X.2009.03162.x.

[5] M. Pakseresht, B. Shirazi, I. Mahdavi, N. Mahdavi-Amiri, "Toward sustainable optimization with stackelberg game between green product family and downstream supply chain," Sustainable Production and Consumption, **23**, 198–211, 2020, doi:10.1016/j.spc.2020.04.009.

[6] S.A.R. Khan, Z. Yu, H. Golpira, A. Sharif, A. Mardani, A state-of-the-art review and meta-analysis on sustainable supply chain management: Future research directions, Journal of Cleaner Production, **278**, 2021, doi:10.1016/j.jclepro.2020.123357.

[7] J. Li, H. Fang, W. Song, "Sustainable supplier selection based on SSCM practices: A rough cloud TOPSIS approach," Journal of Cleaner Production, **222**, 606–621, 2019, doi:10.1016/j.jclepro.2019.03.070.

[8] L. Meherishi, S.A. Narayana, K.S. Ranjani, "Sustainable packaging for supply chain management in the circular economy: A review," Journal of Cleaner Production, **237**, 117582, 2019, doi:10.1016/j.jclepro.2019.07.057.

[9] J. Jemai, B. Do Chung, B. Sarkar, "Environmental effect for a complex green supply-chain management to control waste: A sustainable approach," Journal of Cleaner Production, **277**, 122919, 2020, doi:10.1016/j.jclepro.2020.122919.

[10] K. Sari, "A novel multi-criteria decision framework for evaluating green supply chain management practices," Computers and Industrial Engineering, **105**, 338–347, 2017, doi:10.1016/j.cie.2017.01.016.

[11] Ö. Uygun, A. Dede, "Performance evaluation of green supply chain management using integrated fuzzy multi-criteria decision making techniques," Computers and Industrial Engineering, **102**, 502–511, 2016, doi:10.1016/j.cie.2016.02.020.

[12] C. Wu, Y. Zhang, H. Pun, C. Lin, "Construction of partner selection criteria in sustainable supply chains: A systematic optimization model," Expert Systems with Applications, **158**, 2020, doi:10.1016/j.eswa.2020.113643.

[13] E. Mastrocinque, F.J. Ramírez, A. Honrubia-Escribano, D.T. Pham, "An AHP-based multi-criteria model for sustainable supply chain development

in the renewable energy sector," Expert Systems with Applications, **150**, 2020, doi:10.1016/j.eswa.2020.113321.
[14] X. Yu, S. Zhang, X. Liao, X. Qi, "ELECTRE methods in prioritized MCDM environment," Information Sciences, **424**, 301–316, 2018, doi:10.1016/j.ins.2017.09.061.
[15] M. Alvandi, S. Fazli, L. Yazdani, M. Aghaee, "An Integrated MCDM Method in Ranking BSC Perspectives and key Performance Indicators (KPIs) ," Management Science Letters, **2**(3), 995–1004, 2012, doi:10.5267/j.msl.2012.01.024.
[16] M. Velasquez, P. Hester, "An analysis of multi-criteria decision making methods," International Journal of Operations Research, **10**(2), 56–66, 2013.
[17] C. Kahraman, B. Öztayşi, I. Uçal Sari, E. Turanoğlu, "Fuzzy analytic hierarchy process with interval type-2 fuzzy sets," Knowledge-Based Systems, **59**, 48–57, 2014, doi:10.1016/j.knosys.2014.02.001.
[18] R.R. Kumar, S. Mishra, C. Kumar, "Prioritizing the solution of cloud service selection using integrated MCDM methods under Fuzzy environment," Journal of Supercomputing, **73**(11), 4652–4682, 2017, doi:10.1007/s11227-017-2039-1.
[19] S. Kubler, J. Robert, W. Derigent, A. Voisin, Y. Le Traon, "A state-of the-art survey & testbed of fuzzy AHP (FAHP) applications," Expert Systems with Applications, **65**, 398–422, 2016, doi:10.1016/j.eswa.2016.08.064.
[20] Y.M. Wang, Y. Luo, Z. Hua, "On the extent analysis method for fuzzy AHP and its applications," European Journal of Operational Research, **186**(2), 735–747, 2008, doi:10.1016/j.ejor.2007.01.050.
[21] K. Zhü, "Fuzzy analytic hierarchy process: Fallacy of the popular methods," European Journal of Operational Research, **236**(1), 209–217, 2014, doi:10.1016/j.ejor.2013.10.034.
[22] K. Rashidi, A. Noorizadeh, D. Kannan, K. Cullinane, Applying the triple bottom line in sustainable supplier selection: A meta-review of the state-of-the-art, Journal of Cleaner Production, **269**, 2020, doi:10.1016/j.jclepro.2020.122001.
[23] C.S. Goh, H.Y. Chong, L. Jack, A.F. Mohd Faris, Revisiting triple bottom line within the context of sustainable construction: A systematic review, Journal of Cleaner Production, **252**, 2020, doi:10.1016/j.jclepro.2019.119884.
[24] M.C. Arslan, H. Kisacik, "The Corporate Sustainability Solution: Triple Bottom Line," The Journal of Accounting and Finance, (July), 18–34, 2017.
[25] H.L. Lam, W.P.Q. Ng, R.T.L. Ng, E.H. Ng, M.K.A. Aziz, D.K.S. Ng, "Green strategy for sustainable waste-to-energy supply chain," Energy, **57**, 4–16, 2013, doi:10.1016/j.energy.2013.01.032.
[26] J. Dai, F.L. Montabon, D.E. Cantor, "Reprint of 'Linking rival and stakeholder pressure to green supply management: Mediating role of top management support,'" Transportation Research Part E: Logistics and Transportation Review, **74**, 124–138, 2015, doi:10.1016/j.tre.2014.12.003.
[27] S.M. Lo, S. Zhang, Z. Wang, X. Zhao, "The impact of relationship quality and supplier development on green supply chain integration: A mediation and moderation analysis," Journal of Cleaner Production, **202**, 524–535, 2018, doi:10.1016/j.jclepro.2018.08.175.
[28] Stella Despoudi, Green supply chain, Elsevier Inc., 2020, doi:10.1016/b978-0-12-816449-5.00002-3.
[29] P. Du, X. Yang, L. Xu, Y. Tan, H. Li, "Green design strategies of competing manufacturers in a sustainable supply chain," Journal of Cleaner Production, **265**, 121853, 2020, doi:10.1016/j.jclepro.2020.121853.
[30] M. Kharaji Manouchehrabadi, S. Yaghoubi, J. Tajik, "Optimal scenarios for solar cell supply chain considering degradation in powerhouses," Renewable Energy, **145**, 1104–1125, 2020, doi:10.1016/j.renene.2019.06.096.
[31] L. Ren, S. Zhou, X. Ou, "Life-cycle energy consumption and greenhouse-gas emissions of hydrogen supply chains for fuel-cell vehicles in China," Energy, **209**, 118482, 2020, doi:10.1016/j.energy.2020.118482.
[32] J. hua Zhao, D. lin Zeng, L. ping Che, T. wei Zhou, J. yi Hu, "Research on the profit change of new energy vehicle closed-loop supply chain members based on government subsidies," Environmental Technology and Innovation, **19**, 100937, 2020, doi:10.1016/j.eti.2020.100937.
[33] H. Allaoui, Y. Guo, J. Sarkis, "Decision support for collaboration planning in sustainable supply chains," Journal of Cleaner Production, **229**, 761–774, 2019, doi:10.1016/j.jclepro.2019.04.367.
[34] Y. Cao, Y. Zhao, L. Wen, Y. Li, H. li, S. Wang, Y. Liu, Q. Shi, J. Weng, "System dynamics simulation for CO2 emission mitigation in green electric-coal supply chain," Journal of Cleaner Production, **232**, 759–773, 2019, doi:10.1016/j.jclepro.2019.06.029.
[35] H. Baligil, S.S. Kara, P. Alcan, B. Özkan, E. Gözde Alar, "A distribution network optimization problem for third party logistics service providers," Expert Systems with Applications, **38**(10), 12730–12738, 2011, doi:10.1016/j.eswa.2011.04.061.
[36] S. Wei, T. Ang, V.E. Jancenelle, "Willingness to pay more for green products: The interplay of consumer characteristics and customer participation," Journal of Retailing and Consumer Services, **45**(June), 230–238, 2018, doi:10.1016/j.jretconser.2018.08.015.
[37] Y. Wang, G. Hou, "A duopoly game with heterogeneous green supply chains in optimal price and market stability with consumer green preference," Journal of Cleaner Production, **255**, 120161, 2020, doi:10.1016/j.jclepro.2020.120161.
[38] X. Zhang, H.M.A.U. Yousaf, "Green supply chain coordination considering government intervention, green investment, and customer green preferences in the petroleum industry," Journal of Cleaner Production, **246**, 118984, 2020, doi:10.1016/j.jclepro.2019.118984.
[39] V. Carfora, C. Cavallo, D. Caso, T. Del Giudice, B. De Devitiis, R. Viscecchia, G. Nardone, G. Cicia, "Explaining consumer purchase behavior for organic milk: Including trust and green self-identity within the theory of planned behavior," Food Quality and Preference, **76**(March), 1–9, 2019, doi:10.1016/j.foodqual.2019.03.006.
[40] D. Chen, J. Ignatius, D. Sun, S. Zhan, C. Zhou, M. Marra, M. Demirbag, "Reverse logistics pricing strategy for a green supply chain: A view of customers' environmental awareness," International Journal of Production Economics, **217**(May 2018), 197–210, 2019, doi:10.1016/j.ijpe.2018.08.031.
[41] H. Gholizadeh, H. Fazlollahtabar, "Robust optimization and modified genetic algorithm for a closed loop green supply chain under uncertainty: Case study in melting industry," Computers and Industrial Engineering, **147**(January 2019), 106653, 2020, doi:10.1016/j.cie.2020.106653.
[42] M. Ramirez-Peña, A.J. Sánchez Sotano, V. Pérez-Fernandez, F.J. Abad, M. Batista, "Achieving a sustainable shipbuilding supply chain under I4.0 perspective," Journal of Cleaner Production, **244**, 2020, doi:10.1016/j.jclepro.2019.118789.
[43] M.A.N. Agi, X. Yan, "Greening products in a supply chain under market segmentation and different channel power structures," International Journal of Production Economics, **223**(October 2019), 107523, 2020, doi:10.1016/j.ijpe.2019.107523.
[44] A.Y. Uemura Reche, O. Canciglieri Junior, C.C.A. Estorilio, M. Rudek, "Integrated product development process and green supply chain management: Contributions, limitations and applications," Journal of Cleaner Production, **249**, 119429, 2020, doi:10.1016/j.jclepro.2019.119429.
[45] T. Hadi, S.K. Chaharsooghi, M. Sheikhmohammady, A. Hafezalkotob, "Pricing strategy for a green supply chain with hybrid production modes under government intervention," Journal of Cleaner Production, **268**(December 2015), 121945, 2020, doi:10.1016/j.jclepro.2020.121945.
[46] P. Gautam, A. Kishore, A. Khanna, C.K. Jaggi, "Strategic defect management for a sustainable green supply chain," Journal of Cleaner Production, **233**, 226–241, 2019, doi:10.1016/j.jclepro.2019.06.005.
[47] A.C. Ng, Z. Rezaee, "Business sustainability factors and stock price informativeness," Journal of Corporate Finance, **64**(June), 101688, 2020, doi:10.1016/j.jcorpfin.2020.101688.
[48] S. Shoukohyar, M.R. Seddigh, "Uncovering the dark and bright sides of implementing collaborative forecasting throughout sustainable supply chains: An exploratory approach," Technological Forecasting and Social Change, **158**(April), 120059, 2020, doi:10.1016/j.techfore.2020.120059.
[49] C.R. Carter, P.L. Easton, "Sustainable supply chain management: Evolution and future directions," International Journal of Physical Distribution and Logistics Management, **41**(1), 46–62, 2011, doi:10.1108/09600031111101420.
[50] S. Seuring, M. Müller, "Core issues in sustainable supply chain management - A Delphi study," Business Strategy and the Environment, **17**(8), 455–466, 2008, doi:10.1002/bse.607.
[51] E. Koberg, A. Longoni, "A systematic review of sustainable supply chain management in global supply chains," Journal of Cleaner Production, **207**, 1084–1098, 2019, doi:10.1016/j.jclepro.2018.10.033.
[52] G. Yadav, S. Luthra, S.K. Jakhar, S.K. Mangla, D.P. Rai, "A framework to overcome sustainable supply chain challenges through solution measures of industry 4.0 and circular economy: An automotive case," Journal of Cleaner Production, **254**, 120112, 2020, doi:10.1016/j.jclepro.2020.120112.
[53] M. Abdel-Basset, R. Mohamed, "A novel plithogenic TOPSIS- CRITIC model for sustainable supply chain risk management," Journal of Cleaner Production, **247**, 119586, 2020, doi:10.1016/j.jclepro.2019.119586.
[54] S. Yadav, S.P. Singh, "Blockchain critical success factors for sustainable supply chain," Resources, Conservation and Recycling, **152**(May 2019), 104505, 2020, doi:10.1016/j.resconrec.2019.104505.

[55] J.R. Brown, R.P. Dant, C.A. Ingene, P.J. Kaufmann, "Supply chain management and the evolution of the 'Big Middle,'" Journal of Retailing, **81**(2 SPEC. ISS.), 97–105, 2005, doi:10.1016/j.jretai.2005.03.002.

[56] M.S. Shaharudin, Y. Fernando, C.J. Chiappetta Jabbour, R. Sroufe, M.F.A. Jasmi, "Past, present, and future low carbon supply chain management: A content review using social network analysis," Journal of Cleaner Production, **218**, 629–643, 2019, doi:10.1016/j.jclepro.2019.02.016.

[57] D. Das, "The impact of Sustainable Supply Chain Management practices on firm performance: Lessons from Indian organizations," Journal of Cleaner Production, **203**, 179–196, 2018, doi:10.1016/j.jclepro.2018.08.250.

[58] J. Whitehead, C.J. MacLeod, H. Campbell, "Improving the adoption of agricultural sustainability tools: A comparative analysis," Ecological Indicators, **111**(June 2019), 106034, 2020, doi:10.1016/j.ecolind.2019.106034.

[59] C.A.L. Vanegas, G.A. Cordeiro, C.P. de Paula, R.E.C. Ordoñez, R. Anholon, "Analysis of the utilization of tools and sustainability approaches in the product development process in Brazilian industry," Sustainable Production and Consumption, **16**, 249–262, 2018, doi:10.1016/j.spc.2018.08.006.

[60] E. Dawkins, K. André, K. Axelsson, L. Benoist, Å.G. Swartling, Å. Persson, "Advancing sustainable consumption at the local government level: A literature review," Journal of Cleaner Production, **231**, 1450–1462, 2019, doi:10.1016/j.jclepro.2019.05.176.

[61] M. Rasti-Barzoki, I. Moon, "A game theoretic approach for car pricing and its energy efficiency level versus governmental sustainability goals by considering rebound effect: A case study of South Korea," Applied Energy, **271**(May), 115196, 2020, doi:10.1016/j.apenergy.2020.115196.

[62] J. Veldman, G. Gaalman, "On the design of managerial incentives for sustainability investments in the presence of competitors," Journal of Cleaner Production, **258**, 120925, 2020, doi:10.1016/j.jclepro.2020.120925.

[63] M. Mofidi Chelan, A. Alijanpour, H. Barani, J. Motamedi, H. Azadi, S. Van Passel, "Economic sustainability assessment in semi-steppe rangelands," Science of the Total Environment, **637**–**638**, 112–119, 2018, doi:10.1016/j.scitotenv.2018.04.428.

[64] Y. Zhao, Y. Tan, S. Feng, "Does reducing air pollution improve the progress of sustainable development in China?," Journal of Cleaner Production, **272**(September 2013), 122759, 2020, doi:10.1016/j.jclepro.2020.122759.

[65] B. Xia, A. Olanipekun, Q. Chen, L. Xie, Y. Liu, "Conceptualising the state of the art of corporate social responsibility (CSR) in the construction industry and its nexus to sustainable development," Journal of Cleaner Production, **195**, 340–353, 2018, doi:10.1016/j.jclepro.2018.05.157.

[66] E. Staniškienė, Ž. Stankevičiūtė, "Social sustainability measurement framework: The case of employee perspective in a CSR-committed organisation," Journal of Cleaner Production, **188**, 708–719, 2018, doi:10.1016/j.jclepro.2018.03.269.

[67] H. Ding, H. Huang, O. Tang, "Sustainable supply chain collaboration with outsourcing pollutant-reduction service in power industry," Journal of Cleaner Production, **186**, 215–228, 2018, doi:10.1016/j.jclepro.2018.03.039.

[68] S.A. Neves, A.C. Marques, J.A. Fuinhas, "Is energy consumption in the transport sector hampering both economic growth and the reduction of CO2 emissions? A disaggregated energy consumption analysis," Transport Policy, **59**(May), 64–70, 2017, doi:10.1016/j.tranpol.2017.07.004.

[69] M. Olfati, M. Bahiraei, F. Veysi, "A novel modification on preheating process of natural gas in pressure reduction stations to improve energy consumption, exergy destruction and CO2 emission: Preheating based on real demand," Energy, **173**, 598–609, 2019, doi:10.1016/j.energy.2019.02.090.

[70] S. Hendiani, H. Liao, R. Ren, B. Lev, "A likelihood-based multi-criteria sustainable supplier selection approach with complex preference information," Information Sciences, **536**, 135–155, 2020, doi:10.1016/j.ins.2020.05.065.

[71] J.M. Polimeni, R.I. Iorgulescu, A. Mihnea, "Understanding consumer motivations for buying sustainable agricultural products at Romanian farmers markets," Journal of Cleaner Production, **184**, 586–597, 2018, doi:10.1016/j.jclepro.2018.02.241.

[72] B. Hartl, T. Sabitzer, E. Hofmann, E. Penz, "'Sustainability is a nice bonus' the role of sustainability in carsharing from a consumer perspective," Journal of Cleaner Production, **202**, 88–100, 2018, doi:10.1016/j.jclepro.2018.08.138.

[73] H. Dixon-Fowler, A. O'Leary-Kelly, J. Johnson, M. Waite, "Sustainability and ideology-infused psychological contracts: An organizational- and employee-level perspective," Human Resource Management Review, **30**(3), 100690, 2020, doi:10.1016/j.hrmr.2019.100690.

[74] Y.M. Wang, K.S. Chin, "Fuzzy analytic hierarchy process: A logarithmic fuzzy preference programming methodology," International Journal of Approximate Reasoning, **52**(4), 541–553, 2011, doi:10.1016/j.ijar.2010.12.004.

[75] L. Mikhailov, "A fuzzy approach to deriving priorities from interval pairwise comparison judgements," European Journal of Operational Research, **159**(3), 687–704, 2004, doi:10.1016/S0377-2217(03)00432-6.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 1084-1090 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Supporting the Management of Predictive Analytics Projects in a Decision-Making Center using Process Mining

Marlene Ofelia Sanchez-Escobar*, Julieta Noguez, Jose Martin Molina-Espinosa, Rafael Lozano-Espinosa

School of Engineering and Science, Tecnologico de Monterrey, Mexico City, 14380, Mexico

ARTICLE INFO

Article history:
Received: 11 January, 2021
Accepted: 12 April, 2021
Online: 28 April, 2021

Keywords:
Process design
Process modeling
Project Management
Data Mining
Best Practices
Decision making

ABSTRACT

A Decision-Making Centers (DMCs) Environment facilitates stakeholders' decision-making processes using predictive models and diverse what-if scenarios. An essential element of this environment is the management of Decision Support Components (e.g., models or systems) that need to be created with mature methodologies and good delivery time. However, there has been a gap in the understanding of project management best practices in DMC environments and in the application of methodologies to ease project execution. In the following paper, we address that gap by analyzing six predictive analytics projects executed in a Mexican DMC using Process Mining techniques. We perform process discovery using a detailed activity event log, which has not been possible in previous studies. Additionally, we perform a compliance evaluation versus the de facto methodology to identify the current process alignment gaps, and finally, we analyze the social networks present in the process execution. The research reveals that (1) process mining models are helpful to address management issues of PA/DM projects (2) PA/DM projects require alignment to mature methodologies to improve process performance and avoid execution problems (3) PA/DM project execution should be revised at the activity level to identify issues and to propose specific strategies. This study's findings can help project managers to perform process analyses and to make informed decisions in PA/DM projects. The following paper is an extension of the article "Applying Process Mining to Support Management of Predictive Analytics/Data Mining Projects in a Decision-Making Center" presented in the 2019 International Conference on Systems and Informatics (ICSAI 2019).

1. Introduction

Decision-Making Centers (DMCs) are immersive virtual environments used to understand complex problems, simplify decision-making, and visualize the results of predictive and scenario-based models [1]. These environments depend on the creation process of tools like: Predictive Analytics/Data Mining (PA/DM) models to operate [2]. Nevertheless, the authors have demonstrated in previous studies that DMC processes focus on high-level tasks and exclude detailed and standard PA/DM activities [2]. The absence of commonality in PA/DM project execution, generates issues, since (1) models are built using empiric methodologies and (2) managers cannot follow up specific technical activities since they are different in every project.

In this research, we propose three approaches to overcome the mentioned issues and help managers and modelers make informed decisions about PA/DM projects. In the first, we apply process mining techniques to a set of PA/DM processes to discover the timing, flow, frequency, and performance of activities from diverse perspectives (e.g., process, organizational, and case). Second, we compare a real PA/DM project execution with an accepted PA/DM methodology, to identify how aligned are the real processes to the formal methodology (i.e., CRISP-DM) and what gaps need to be closed to achieve compliance. Third, we perform complementary human resources analyses to visualize the relationship between resources and communication channels during process execution.

We expect that managers in DMCs use the models presented in this study to evaluate their processes and to consider the implementation of specific management strategies.

*Corresponding author: Marlene Ofelia Sanchez-Escobar, Tel.: +521 5544995733, Email address: A00704709@itesm.mx

Finally, the organization of the paper is as follows: Section 2 and 3 presents the background and the literature review respectively. Section 4 describes the experimental design. Section 5 provides final results and discussion, and section 6 describes conclusions and future work.

2. Background

2.1. Process Mining Techniques and Project Management applications

The Process Mining (PM) technique is a reverse engineering approach where process models are generated using event logs [2]. In [3], the author classifies the following PM techniques that we use for our analysis: discover, conformance, and enhancement.

The process discovery technique aims to mine process models using discovery algorithms, so the process helps managers answer specific questions [4]. Examples of discovery algorithms include alpha algorithm, heuristic miner, fuzzy miner, genetic miner, region miner, and integer linear programming (ILP). Differently, the process conformance technique aims to measure the process quality through metrics like: fitness, precision, generalization, and simplicity [4]. In this category, conformance checking is used to compare the expected model and the reality obtained from event logs. Likewise, it is possible to identify processes, commonalities, similarities, and deviations [3]. Finally, the process enhancement technique aims to extend the model with relevant information [4]. For instance, statistical metrics based on timestamps (e.g., throughput time, working time, and waiting time) or the use of replay analysis to visualize process execution. In this research, we use Disco and ProM 6 tools to implement the process mining techniques previously explained.

Finally, in [5], the authors explain that the project management field requires the process mining discipline to identify optimal workflows within project life cycles. In this regard, we consider that the following managerial issues identified in PA/DM projects can be analyzed using process mining techniques: establishing realistic goals, the creation of good teams, gaining knowledge of data, lack of infrastructure, poor project communication methodology, lack of risks management and change management [6], [7].

2.2. CRISP-DM Framework

The Cross-Industry Standard Process for Data Mining (i.e., CRISP-DM) is the most accepted methodology in the field for executing data mining projects [8]-[10]. In the framework, the project life cycle includes the following key phases [11]: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. The methodology provides a universal process with generic tasks that can be executed for all data mining projects. The CRISP-DM reference model can be regarded in Figure 1, and the generic tasks to be performed in each phase are listed in Table 1. As can be noticed, the generic tasks are defined at the activity level, which facilitates its integration to high-level DMC process. In this study, we examine the data understanding, data preparation, modeling, and evaluation phases of CRISP-DM methodology, given information constraints.

2.3. PA/DM processes with CRISP-DM

We discussed in [2] the importance of integrating PA/DM processes with DMC processes, and we make an integration effort. However, PA/DM processes at the level of activity have not been studied separately. In this study, we focus just on PA/DM processes, as a part of DMC processes, since there is no work in the literature that performs such analysis using process mining techniques. Finally, we assume that the CRISP-DM methodology matches DMC's PA/DM processes.

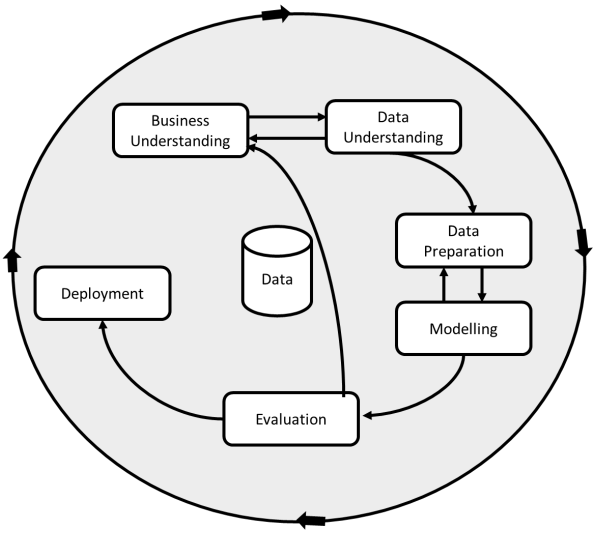


Figure 1: CRISP-DM phases and activity flow. Adapted from (Chapman, 2000)

Table 1: CRISP-DM phases and generic tasks

Phase	Generic Task
Business understanding	Determine business objectives Assess situation Determine Data Mining goals Produce Project Plan
Data Understanding	Collect Initial Data Describe Data Explore Data Verify Data Quality
Data Preparation	Select Data Clean Data Construct Data Integrate Data Format Data
Modeling	Select Modeling Technique Generate Test Design Build Model Assess Model
Evaluation	Evaluate Results Review Process Determine Next Steps
Deployment	Plan Deployment Plan Monitoring and Maintenance Produce Final Report Review Project

3. Literature review

Limited research has focused on Project Management using Process Mining techniques. In the literature, studies concentrate on the analysis of project management processes using data mining [12]-[14] and predictive analytics [15]-[17] techniques, but not process mining techniques. Likewise, we identify that most analyses are focused on software life cycles [12]-[18], but not on DA/PM processes. For instance, in [18], the authors use conformance checking techniques to reveal aspects of processes and identify deviations in software project execution. Additionally, the author presented an application to optimize the software development life cycle of projects using process mining [5]. However, in both cases, research is not focused on PA/DM projects or DMC environments.

Finally, the authors developed a previous study to analyze PA/DM processes in a DMC [2]. Nevertheless, the paper is oriented to analyze project phases and not project activities. Thus, that limits the possibility to discover low-level issues and applying target strategies. Besides, the study uses only in its majority enhancement techniques and a limited number of discovery algorithms. In the present study, we address those limitations.

4. Experimental Design

The DMC located at Tecnologico de Monterrey, Mexico City campus also referred as "Decision Laboratory", is a room with seven big format and high-definition screens that offers a space to make consensual decisions and to present solution proposals to a group of decision-makers [19]. This last with the goal of selecting the best possible solution. A picture of Tecnologico de Monterrey's DMC can be regarded in Figure 2.



Figure 2: Tecnologico de Monterrey's DMC

In the Mexican DMC, managers execute PA/DM projects to create models that support decision-making. The team organization is defined according to the knowledge of resources and their affinity to develop specific models. The project manager role is performed by one resource, and one or more product owners define the business requirements. After project execution, the modelers report that no formal methodology is applied to create the models.

Even though project managers try to deliver models with quality and on time, modelers and the supervisor report the following issues during project execution: (1) Wrong selection of modeling technique and (2) lack of standardization of data glossary among models.

With this in mind, we examine the possible causes behind the reported issues and perform a complete analysis of the PA/DM process execution in the Mexican center.

4.1. Question to be answered

For this experiment, we aim to answer the following questions about the execution of PA/DM projects in a DMC.

1) RQ1:What do the dependency, frequency, and performance statistics of the process model reveal?

2) RQ2: How compliant is the discovered model vs. the CRISP-DM reference model?

3) RQ3: How is the interaction among resources during process execution?

4) RQ4: What are the possible causes for the reported issues?

4.2. Information Gathering

We obtained qualitative and quantitative information from six real PA projects executed at Tecnologico de Monterrey DMC by interviewing modelers and managers. The format utilized for quantitative data gathering is available in Appendix A.

During this phase, five modelers and one manager were interviewed. The requested data include information from four CRISP-DM process stages (i.e., data understanding, data preparation, modeling, and evaluation), since we do not have access to data from the business understanding and deployment phases

Finally, the following data was obtained from stakeholders: start and finish dates of activities, the average number of hours per activity per day, and the number of resources involved in each activity. At the end of the interview, we request impressions about execution processes to identify specific issues.

4.3. Event Log Generation

We create 4945 records with timestamps based on the provided information. The corresponding records per project can be regarded in Table 2. In this phase, no assumptions were considered since the modelers provide specific times and dates for each activity.

Table 2: Event Log record generation by the project.

Project	Records generated
P1	504
P2	1262
P3	1496
P4	463
P5	599
P6	621

4.4. Event Log Analysis

We use Disco and ProM 6 applications to perform process mining. The first is a commercial tool that provides accurate process models [20]. While the second supports other types of functionalities like Petri nets and Social Networks [21]. Table 3 shows the modules used in each application.

Table 3: Models used in PM applications

Application	Models used
Disco	(1) Map (2) Statistic 3) Filter by case
ProM	(1) Social Network Miner

We use the Map view from the disco application to visualize the flow of activities, dependencies, frequencies, and performance. Likewise, the statistics view is used to identify the process event distribution, the activities, and the frequency of resources. Finally, we use the Filtering functionality to analyze the process model by specific cases. From ProM application, we use Social Network Miner to identify relationships among resources.

5. Results and Discussion

During process analysis, we document the global statistics shown in Table 4. As can be noted, the number of events represents the total records in the event log, and the cases correspond to the number of processes. The activities represent 16 generic tasks of the following phases of the CRISP-DM methodology: data understanding, data preparation, modeling, and evaluation. Finally, we examine projects that were executed in the next time range (April 25th, 2016 to June 30th, 2019). The statistics reveal that, on average, the project duration is 30 weeks.

Table 4: Process global statistics

Metric	Value
Events	4945
Cases	6
Activities	16
Median Case Duration	27 weeks
Mean Case Duration	30 weeks
Start	4/25/2016 9:00:00
End	06/30/2019 18:00:00

On the other hand, the statistics per activity showed in Table 5 reveal the most, the average, and the least executed activities in the project.

Table 5: Statistics per activity

Activity	Relative Frequency
Build model	19.38%
Evaluate results	10.11%
Select modeling technique	8.15%
Verify data quality	7.28%
Construct data	6.23%
Assess model	6.15%
Explore data	6.09%
Integrate data	5.46%
Collect initial data	5.42%
Review process	5.04%
Determine next steps	4.99%
Generate test design	4.77%
Select data	3.94%
Describe data	3.28%
Clean data	2.93%
Format data	.79%

In the following subsections, we respond to the defined research questions.

RQ1: What do dependency, frequency, and performance statistics of the process model reveal?

Figure 3 shows the process map of the event log. As can be noticed, there are four thick arrows in the diagram that represents significant dependence among activities. For instance, the most substantial and unique bidirectional dependency is present between the review process and the determination of Next steps activities. Likewise, a significant reliance is visible between the process review and next steps activities, which means that those tasks execution order is the same in cases majority. Besides, managers should pay attention to the iteration that involves all data manipulation activities (i.e., collect, explore, verify, select, clean, construct, and integrate data) with the modeling selection technique. The evidence support that the team is having trouble with gaining knowledge of the data, which is a common problem in these kinds of models. We can assume that the lack of consistent execution of the description and selection of data could be the cause of the described problem.

Lastly, the diagram shows a dependency between the modeling technique selection and the initial data collection, which should be revised. Strangely, a change in data impacts the modeling technique and also the model construction. We recommend the inclusion of roles with expertise in modeling techniques to break that dependency.

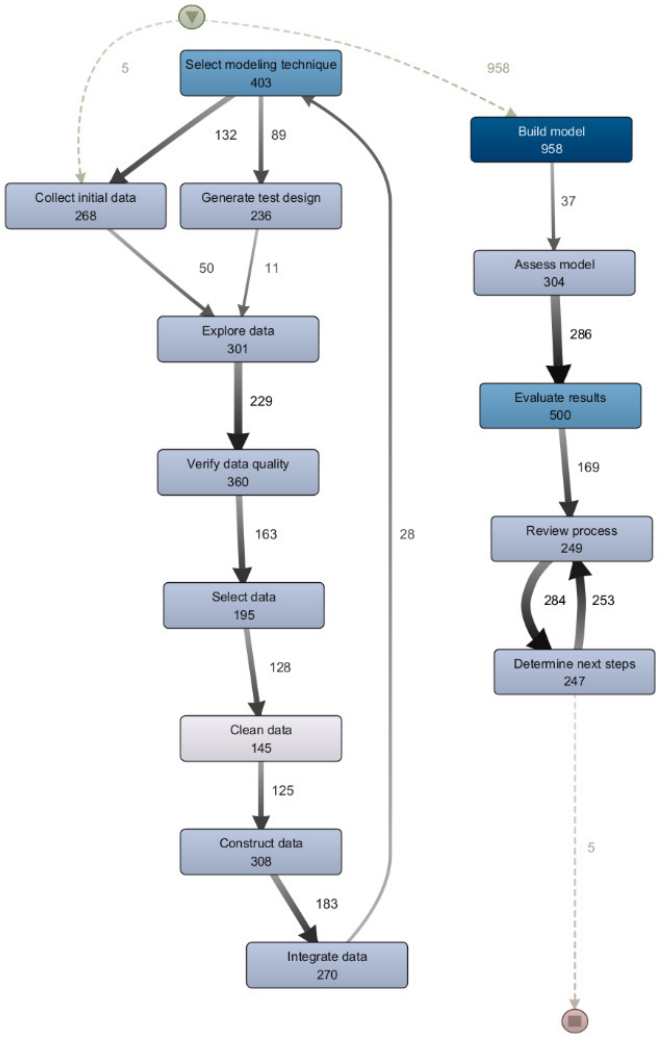


Figure 3: Process map by absolute frequency

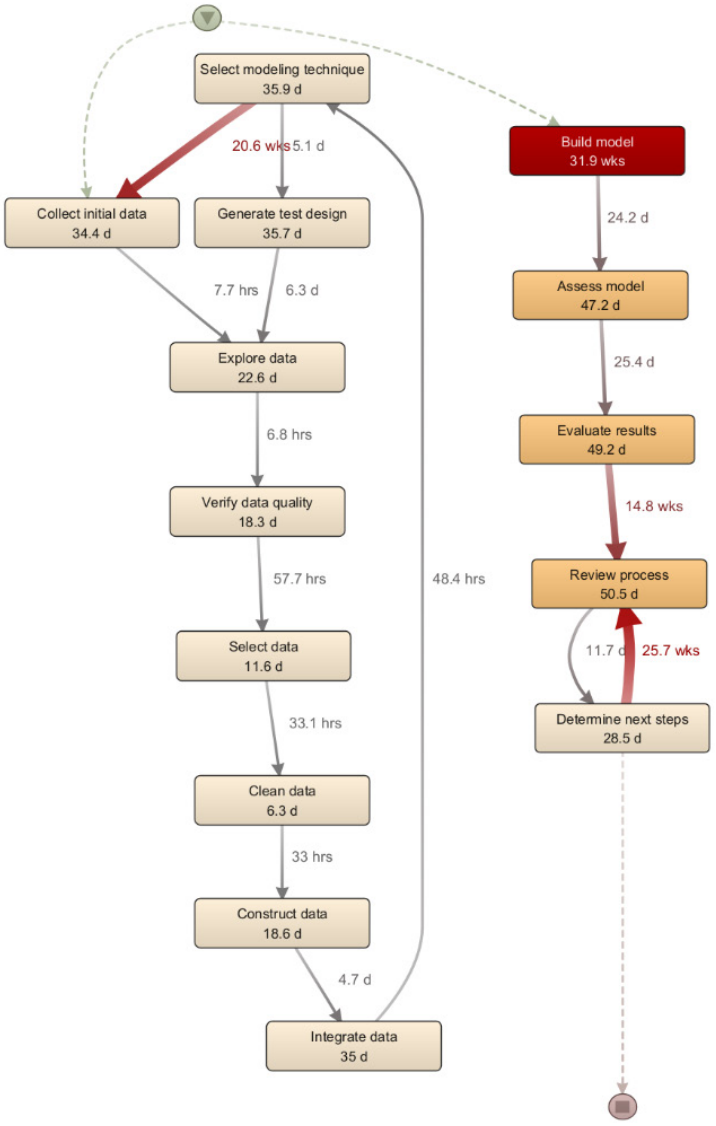


Figure 4: Process map by total duration

On the other hand, the absolute frequency of activities is represented with color. A high-frequency task is painted with intense blue, while one with low frequency is depicted with light blue. For this process model, the activity with the most significant frequency is the model's construction, followed by evaluating results and selecting the modeling technique. As we have mentioned, the model construction is affected by previous executed or non-executed activities. We can assume that previous activities improvement has a positive impact on the construction activity. In this case, we recommend using lean prototypes to facilitate the technique selection and diminish the time devoted to the model construction.

Finally, the performance of the model can be regarded in Figure 4. The model shows the total task duration and delays between activities. The model´s construction is the most significant task with 31.9 wks. Likewise, a delay of 20.6 weeks is present between the select modeling technique and collect initial data activities. In this case, managers should focus on diminishing the time between those two activities, by involving more resources or/and experts to the project.

A second delay is exposed between the process review and the definition of the next steps; however, this case should be analyzed separately since all resources execute these activities simultaneously, and that variable could affect the metric and not represent the real delay.

RQ2: How compliant is the discovered model vs the CRISP-DM reference model?

Figure 5 shows the process map by case frequency that is useful to analyze compliance. For instance, we are examining six cases, and theoretically, all activities should be present in all cases; however, in reality, this is not the case. Specifically, for this DMC, we discovered that the activities with a lower presence in the execution are the data formatting and data description. This last represents an issue in subsequent tasks since modelers report that the lack of data description has delayed the model's development and integration processes. Likewise, data cleaning and construction activities have problems with compliance in one of the cases, so DMC managers need to review these deviations.

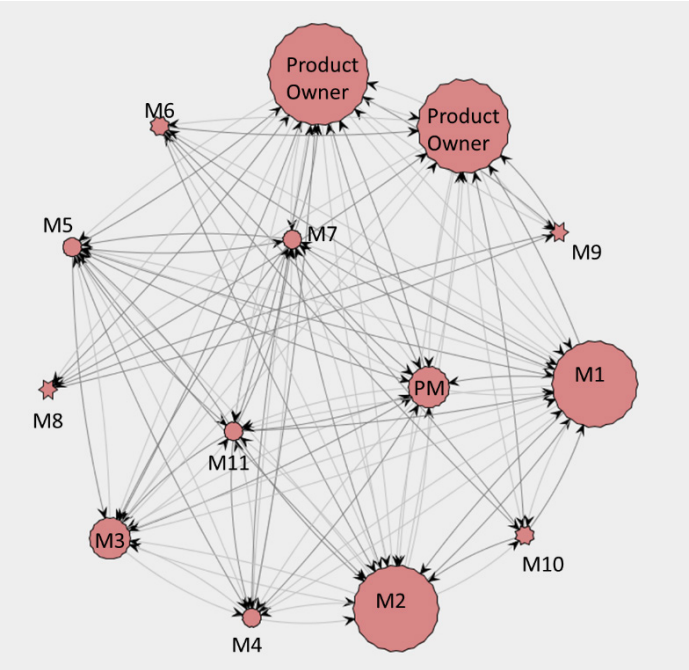


Figure 6: Social Networks of the process execution

Finally, the flow of activities has some compliance issues. In contrast to CRISP-DM, the actual execution is iterative in the data understanding and preparation phases. With this information, managers can create initiatives to align the model to CRISP-DM by stages and increase its performance.

RQ3: How is the interaction among resources during process execution?

To answer this question, we use the Social Network mining capability of ProM. As can be regarded in Figure 6, the two product owners are critical intermediaries in the project execution. In this case, the PM role is key in the process; however, the manager seems distant from individual modelers. On the other hand, M1 and M2 modelers seem to be more connected to the group. This last can have two explanations: (1) The existence of a functional dependency among parties (e.g., infrastructure,

software, etc.) or (2) the modeler has participated in several projects which allow him to collaborate with more people. Finally, managers must pay attention to isolated modelers M8 and M9 and understand why they are separated from the group.

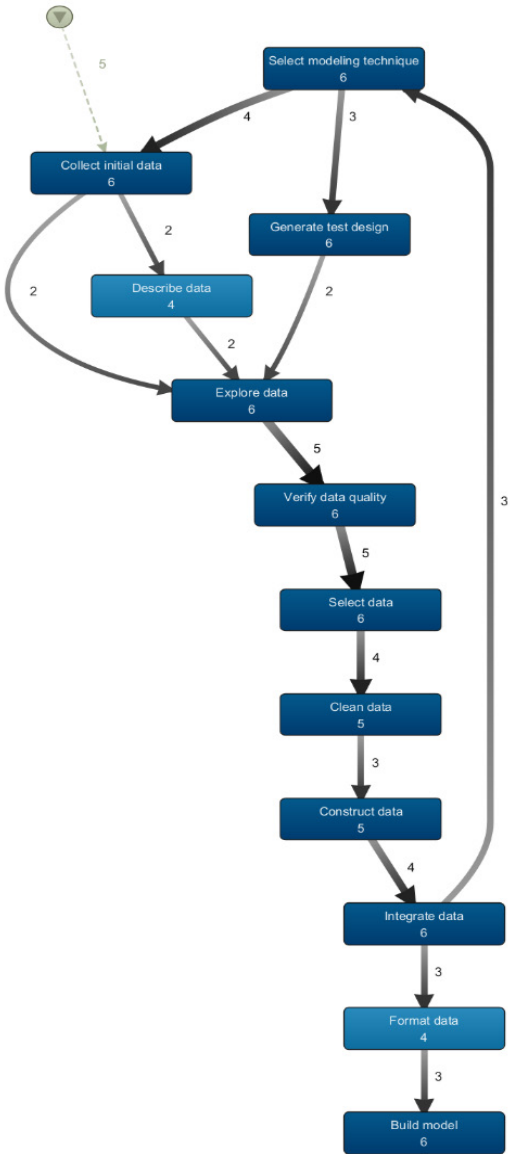


Figure 5: Process map by case frequency

RQ4: What are the possible causes for the reported issues?

We analyze the possible causes of the reported issues using previous process maps.

Wrong selection of modeling technique. It seems that the modeling technique selection is an exploratory process that takes too much time to define. As we have recommended before, inclusion of an expert and the implementation of prototypes can address this problem. Since the modeling technique can be evaluated with a prototype and with the expert support.

Lack of standardization of data glossary among models: This problem is caused by the lack of execution of the data documentation activity. We believe that an alignment to CRISP-DM methodology can solve this problem.

Finally, it is relevant to mention the limitations of the present study, which can be addressed in future research.

First, the presented model represents PA/DM execution processes of DMCs exclusively, so other PA/DM processes outside this environment are not represented in the research. Second, we don't include the business understanding and deployment phases in the modeling given data restrictions. So we represent part of the PA/DM process in this study. Lastly, the absence of previous research limits the possibility to compare and contrast our model with others and evaluate its completeness.

6. Conclusions and Future Work

In this research, we reveal the value of process mining as a tool to support project management of PA/DM projects in DMCs. Likewise, we expose the need to implement mature PA/DM processes in DMCs that facilitate (1) project management and (2) process improvement.

In this study, we create a process model to identify project execution issues, gaps in compliance, and the interaction of resources. We perform interviews to obtain detailed data from the PA process execution from modelers and managers. An event log at the level of activities was created considering CRISP-DM generic tasks, timestamps, and resources. Disco application was used to apply process discovery and process enhancement techniques. ProM application was used to perform Social Network mining. The results of the study reveal that: (1) Process mining models are helpful to analyze and address common management issues of PA/DM projects (2) PA/DM projects require alignment to mature methodologies to improve process performance and avoid execution problems (3) PA/DM project execution should be revised at the activity level to identify issues and to propose specific strategies (4) PA/DM projects should be analyzed from different perspectives to obtain valuable information for the management team.

Although it has been proved that Process Mining techniques are useful tools to support the management of PA/DM projects, there is work that needs to be addressed in the future. For instance, we need to use ProM tool to obtain compliance metrics of process models. Likewise, we need to use additional social network algorithms to analyze other organizational relationships.

Conflict of Interest

The authors of this paper certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Acknowledgment

This work was supported by CONACYT Project 266632, identified as "Binational Laboratory on Smart Sustainable Energy Management and Technology Training." Additionally, the

authors would like to thank the following experts for the provided information about PA project execution: Jose Hugo Carmona, Luis Alberto Lira, Ricardo Granados, Aban Moreno, Eduardo Meza, and Paolo Paez.

References

[1] R. Edsall, K. Larson, "Decision making in a virtual environment: effectiveness of a semi-immersive 'Decision Theater' in understanding and assessing human-environment interactions," 2006.

[2] M.O.S. Escobar, R.L. Espinosa, J.M.M. Espinosa, J.J.N. Monroy, G. V Solar, "Applying Process Mining to Support Management of Predictive Analytics/Data Mining Projects in a Decision Making Center," in 2019 6th International Conference on Systems and Informatics (ICSAI), 1527–1533, 2019, doi:10.1109/ICSAI48974.2019.9010135.

[3] W. van der Aalst, "Process Mining: Overview and Opportunities," ACM Trans. Manage. Inf. Syst., **3**(2), 2012, doi:10.1145/2229156.2229157.

[4] C. dos S. Garcia, A. Meincheim, E.R. Faria Junior, M.R. Dallagassa, D.M.V. Sato, D.R. Carvalho, E.A.P. Santos, E.E. Scalabrin, "Process mining techniques and applications – A systematic mapping study," Expert Systems with Applications, **133**, 260–295, 2019, doi: 10.1016/j.eswa.2019.05.003.

[5] J. Joe, T. Emmatty, Y. Ballal, S. Kulkarni, "Process mining for project management," in 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), 41–46, 2016, doi:10.1109/SAPIENCE.2016.7684142.

[6] G. Cordeiro, F. Deschamps, E. Pinheiro de Lima, Managing a Big Data/Analytics project: a systematic literature review, 2017.

[7] G. Kabanda, "An Evaluation of Big Data Analytics Projects and the Project Predictive Analytics Approach," Oriental Journal of Computer Science and Technology, **12**, 132–146, 2020.

[8] R. Nisbet, G. Miner, K. Yale, The Data Mining and Predictive Analytic Process, Elsevier: 39–54, 2018, doi:10.1016/b978-0-12-416632-5.00003-7.

[9] V. Kotu, B. Deshpande, Data Science Process, 19–37, 2019, doi:10.1016/B978-0-12-814761-0.00002-2.

[10] A. Azevedo, M. Santos, KDD, semma and CRISP-DM: A parallel overview, 2008.

[11] R. Wirth, J. Hipp, "CRISP-DM: Towards a standard process model for data mining," Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 2000.

[12] R. Nayak, T. Qui, "Data Mining Application in a Software Project Management Process," in Proceedings 3rd Australasian Data Mining Conference (AusDM04), University of Technology Sydney: 99–109, 2004.

[13] P. Pospieszny, "Application of Data Mining Techniques in Project Management – an Overview," Collegium of Economic Analysis Annals, 199–220, 2017.

[14] J. Balsera, V. Montequín, F. Ortega-Fernández, C. Alba, Data Mining Applied to the Improvement of Project Management, 2012, doi:10.5772/48734.

[15] G. Schuh, M. Riesener, C. Dölle, "Implementation and assessment of a predictive analytics model for development project management," 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 696–700, 2017.

[16] G. Guillaume-Joseph, J. Wasek, "Improving software project outcomes through predictive analytics: Part 2," IEEE Engineering Management Review, **43**, 39–49, 2016, doi:10.1109/EMR.2015.2469471.

[17] E.S. Awolumate, Using Predictive Analytics to Deliver an Improved IT Project Cost Performance Model, 2020.

[18] E. Kouzari, L. Sotiriadis, I. Stamelos, "Process mining for process conformance checking in an oss project: An empirical research," in IFIP International Conference on Open Source Systems, Springer: 79–89, 2018.

[19] C.N. Corella, J. Noguez, E.J.M. Molina, E. Sotkoeva, R. Salla, "Use of a decision-making laboratory to support student's visual analysis for the solution of a transportation problem in Mexico City," in 2019 IEEE Frontiers in Education Conference (FIE), 1–5, 2019, doi:10.1109/FIE43999.2019.9028550.

[20] C.W. Günther, A. Rozinat, "Disco: discover your processes," in: Lohmann, N. and Moser, S., eds., in Proceedings of the Demonstration Track of the 10th International Conference on Business Process Management (BPM 2012), CEUR-WS.org: 40–44, 2012.

[21] B. Dongen, H. Verbeek, A. Weijters, W. Aalst, The ProM Framework: A New Era in Process Mining Tool Support, 2005, doi:10.1007/11494744_25.

Appendix A. Data gathering format

Project Name					
Start date		**End date**			
Phase: Data Understanding					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Collect initial data					
Describe data					
Explore data					
Verify data quality					
Phase: Data Preparation					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Select data					
Clean data					
Construct data					
Integrate data					
Format data					
Phase: Modeling					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Select modeling technique					
Generate test design					
Build model					
Assess model					
Phase: Evaluation					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Evaluate results					
Review Process					
Determine Next Steps					

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1091-1099 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design of Platform to Support Workflow Continuity in Multi-Device Applications

Oscar Chacón-Vázquez[1], Luis G. Montané-Jiménez[1,*], Carlos Alberto Ochoa-Rivera[1], Betania Hernández-Ocaña[2]

[1]*FEI, Universidad Veracruzana, Xalapa, 91020, México*

[2]*Universidad Juárez Autónoma de Tabasco, Villahermosa, 86040, México*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 31 January, 2021
Online: 16 February, 2021

Keywords:
Continuity
Platform
Multi-device

ABSTRACT

Nowadays, the Internet has become an indispensable tool for the realization and continuity of activity at a different time, place, and technological context (e.g., mobile, pc, tablet), so that interaction techniques through the use of multi-device support have become of great interest. From this perspective, continuity in interactions is an essential concept in the face of changes in context environments where an interaction develops. There are works related to the continuity and support of multi-device environments through software platforms that are useful to improve continuity support; however, reducing steps to resume an activity on a different device is an aspect that needs to be studied in greater detail. In support of the above, this paper presents an exploratory study that shows that continuity is a useful feature for users; however, there are still aspects that need to be studied. Therefore, in this paper, we propose a platform to implement continuity in a workflow that reduces the steps necessary to continue and resume activity in a different device context and a case study which serves as a method to evaluate the platform proposal and the models on which it is based.

1. Introduction

The Internet has become one of the main tools for carrying out different activities in various sectors, for example, work, entertainment and communication [1]. The evolution of the Internet has had effects in many areas of information and communication technologies such as computers [2] that make use of it, and the paradigms and models of software development [3].

The evolution of computers towards small devices with considerable computing capacity and intelligence, have brought about changes in the design and interaction paradigms seeking a better integration in our environments [4]. The constant evolution of computers towards small devices with considerable computing capacity and intelligence, have brought about changes in the design and interaction paradigms seeking a better integration in our environments [5]. In general, these developments use tools that make it easier to carry out software projects.

Interaction techniques through the use of platforms with support for multi-device environments have become one of the fields of greatest interest thanks to the great adoption of the internet and the evolution of information and communication technologies, taking advantage of its advantages [6]. The tools that have emerged derived from these techniques are broad and useful, focusing on different areas of interaction such as the use of single sign-on systems, the use of multiple devices to interact with the same system at the same time, and the ability to continue certain activities with a long duration across different contexts [2].

In multi-device interactions, an activity can be carried out through different devices and in other contexts; this is much more complex to keep a user-focused on an activity and recover the activity's status was left. For example, a student who writes a document on a computer at his school may interrupt his work and continue it later from a different computer at home. This activity implies having to transfer the information from one computer to another and recovering the last state (e.g., page of a document where the student was, available tools, and information on the clipboard).

Continuity has become an essential concept in recent years for commercial information systems. We can find these systems such as Netflix, Amazon Video, and Spotify, which allow through a login and constant synchronization between devices to maintain a state between them and allowing the user to resume their activity (playback of some content multimedia) in a different context. Similarly occurs with Apple operating systems, which, in the case of system-specific applications, allow resuming an activity between devices, omitting several additional steps to do so.

Continuity has been one response to the interruptions brought by changes in context, reducing the number of steps necessary to

*Corresponding Author: Luis G. Montané-Jiménez, Email: lmontane@uv.mx

continue an activity; however, its implementation is complex due to multiple aspects [4], [6]. Therefore, we propose a conceptual platform that helps implement continuity in a software project, and it helps reduce the number of steps necessary to continue an activity when changing context. This proposed platform allows programmers and designers to focus their efforts completely on said design and reducing the workload to implement continuity. It also reduces the users' cognitive load in applications complex by quickly migrating an activity between different devices, reducing interruptions in the workflow.

This paper is structured as follows. Section 2 reviews the terms related to continuity in the workflow, and Section 3 presents a comparison of the work related to continuity and workflow. Section 4 presents a study of continuity in commercial applications. Section 5 presents a conceptual model of continuity in software projects. Section 6 presents a case study where the platform was implemented for testing. Finally, Section 7 presents the conclusions and future work.

2. Continuity in workflow

2.1. Workflow

The term workflow is a concept applicable to all areas where the human being develops. Specifically, information systems refer to a series of tasks that must be followed and completed so that a user can obtain a result (e.g., complete an activity). In [7], the authors refer to the workflow simply as flow. The workflow tends to be confused with the activity itself. This activity is the aim pursued by the user. For example, when someone to read a book, the workflow is necessary to complete it: i) read chapter 1, ii) read chapter 2, among others. The workflow has the characteristics shown in Table 1.

2.2. Interaction between devices

Adopting multiple devices in people's daily lives creates an environment where information is easily shared between them [8]. Currently, users are not limited to owning only a computer and a mobile phone, systems that in most cases are isolated from each other. In these environments, a user can have multiple computers that are used in different places and with other devices such as tablets or Smart TVs that have tools and applications that have the ability to interact and communicate with each other to offer the user a user interface that is seen as a whole, regardless of the equipment or device where users perform an activity.

Interaction between devices is an emerging branch of Human-Computer Interaction (IHC) focused on the way in which users interact between the different devices to which they have access, as well as the way in which applications take advantage of these forms interaction to offer much more dynamic user experiences [9]. There are two types of interaction between devices, i) interaction with multiple devices and ii) interaction across multiple devices [6]. The i) interaction with multiple devices is based on the ability of each of the devices to operate a system or application at the same time [10], allowing to create an environment where each device takes a specific role within the application and thereby better distributing the options and capabilities of the system. It focuses especially on collaborative environments where each person can take a specific role within the application, adapting the interface to each role and thereby avoiding overloading the interface with unnecessary elements. On the other hand, ii) interaction through multiple devices focuses especially on monitoring activities that were started on a certain device, and due their duration, it may need to be interrupted and continued on a different device [7], facilitating the change of context for the user using concepts such as continuity and workflow to provide better user experiences.

Table 1: Workflow characteristics

Characteristic	Description
Divisible in tasks	The activities can be divided into more specific and shorter tasks, the completion of these tasks has the consequence of approaching the end of the activity.
Duration	It is an important factor in the definition of an activity and its flow, it defines the time that is invested in carrying out the activity and generally greatly affects the methods and moments of interaction.
Focus	It refers to the number of people involved in the activity; it affects the way in which the activities are carried out. It is possible to catalog the activity flows in two: one user and multi-user.

2.3. Continuity

The continuity term is defined as the ability to perform an activity on devices used sequentially and without presenting major interruptions in their workflow (context changes) [6]. It is a capacity and characteristic that we can find present in many commercial systems and applications and that take advantage of the benefits of the Internet and the capabilities of the devices with which we interact to reduce the steps necessary to continue an activity between them.

The experiences that implement continuity have 6 main characteristics that can be seen as objectives to be achieved [6]: i) Privacy, provide the user with the assurance that the data is safeguarded. ii) Appropriation, operate inherently in conjunction with the user and their work context. iii) Personalization, provide users with the possibility to choose which and when the continuity functions are activated. iv) Awareness, ability to "infer" the next steps that the user will perform. v) Inclusion, broad compatibility with current and future devices of the user, increasing interaction possibilities. Finally, vi) Troubleshooting, offering the ability to solve problems arising from the use of continuity to the user.

In a specific context in which a user interacts with a system that implements continuity and requires interrupting their work and continuing it in a different context, it is possible to distinguish two types of continuity [7]: monoactivity, and sequential activities. In monoactivity, continuity refers to activities such as: reading a book, watching a movie, or writing a text that may require a long time to complete. The advent of technology and the great expansion of mobile devices have contributed significantly to users changing how they performed these actions [2]. These activities can be carried out in short periods, where the context of use and the time available have an essential effect on the

interaction. Continuity in one-time activities focuses on this type of interaction, trying to minimize the necessary steps so that a user who left an activity pending completion can resume it when he deems it appropriate.

Continuity in sequential activities refers to when the user faces activities whose time and necessary effort allow them to be divided into the performance of sub-activities [7]. The continuity in these types of activities does not depend on being able to carry out the same activity through different contexts, but rather on maintaining a state between all the sub-activities, allowing the user to minimize the mental processes necessary to continue working. The design of these interactions is not based on its duration but rather on the place where a sub-activity occupies the general activity. In general, these types of interaction focus on offering the user information related to the last time they interacted with the system rather than minimizing the steps necessary to continue with the activities.

3. Continuity related work analysis

Multi-device environments and the accelerated evolution of information and communication technologies have led to the development of a great variety of works with multi-device and continuity implications; each of them has its characteristics and qualities. Among the works carried out, the most noteworthy include frameworks, development tools, and evaluations of applications that implement continuity.

3.1. Frameworks

The frameworks are a recurring theme in computer science because they allow us to establish future work bases and propose different approaches to solve a problem. In the multi-device and continuity field, there are several works where proposals for theoretical frameworks are presented. Recent works conceptually propose a framework for co-located devices in the same environment distributing the interface of an application through the users' roles when using the interface simultaneously [10].

In [11] propose a conceptual framework where for each device a role must be established and used in some way within the adaptation process, continuity is included at the conceptual level. More recent works introduce the concept of continuity as in "Design a framework to support the development of smart cross-device applications" [9], where a framework is proposed which provides guidelines and tools to distribute interfaces through nearby devices. Continuity is introduced as a state maintenance that improves the user experience.

A recurring theme in framework development is maintaining state through different methods. In [12,13] present frameworks that allow data to be automatically migrated between applications. However, they do not have current tools that maintain the state completely. On the other hand, [14] propose a tool to migrate data through interaction with Kinect to migrate data during a work session with several devices easily.

3.2. Development tools

Application development is a complex task that uses many tools to facilitate the process. In the field of continuity and multi-device interaction, work has also been carried out based on tools' development.

In [15] present a studio-type tool focused on the design of multi-device interfaces, and focused on its interactivity. The tool focuses on designing web-based interfaces that can be distributed and function in a coordinated manner between various devices; it also allows defining contexts for the devices to adapt according to them.

Several tools focus on data migration; among them, we can find tools that allow data to be migrated between "neighboring" applications of a mobile device [16]. In [2] propose a tool that allows for migrating information searches between different devices while maintaining history. On the other hand, there are works such as present in [17–19] that allow for migrating the state through different triggers such as biometric sensors and JavaScript methods. Finally, in [20] present a tool that allows data to be migrated through NFC tags, mainly focused on web searches.

Other authors have focused on the development of tools that support the internal procedures of the applications. In [5], the author propose a middleware type tool to implement a single sign-on between devices; continuity is a crucial part of the tool since, through this sign-on, the state can be transferred between devices.

3.3. Evaluations to applications

Continuity is a tool that has been implemented in both software projects and projects for the commercial field, verifying the effectiveness of applications that implement continuity. A multimedia application prototype to continue multimedia consumption through different devices was developed in [21]. This prototyping's main objective is to demonstrate the interaction techniques' effectiveness between the most used devices. Through the study and the prototype, it was possible to demonstrate that it developed software with a user-centered approach is necessary. However, there are still processes that can be implemented to improve interaction.

On the other hand, a conceptual model is presented in [22] where long-term searches on the web are supported across different devices; their approach is not just based on allowing long-term searches across different devices. It also provides intelligence by offering suggestions based on search patterns. The study carried out on the model showed significant differences in user interaction, depending on the context and type of device used.

In [6], the author evaluated one of the most representative applications of continuity in a commercial context: *Apple Continuity*. The study carried out contrasts the characteristics of Apple's service against user comments; the results indicate that continuity is accepted and desired by users in general. However, adapting to these new interaction methods can become complex, so the focus on user-based development is important to avoid such complexity. Further, [23] present an evaluation of automatic code generation to implement continuity with a strong emphasis on commercial applications.

Finally, there are work that present the applications of a method developed by authors to evaluate the migratory aspect of multi-device application such as presented by [24], [25].

3.4. Continuity related work comparative

By reviewing the work related to continuity and multi-device applications, Table 2 is presented, which shows that several

elements come into play when continuity is implemented in devices and systems.

The concept of multi-devices is applied in many ways throughout the different jobs, allowing the creation of different forms of interaction in the systems. It is also proven that it is a key point to give way to the continuity of activities in applications.

Table 2: Comparative of papers related to continuity

Paper	Year	Type	Emphasis	Multi-device	Continuity	
					Explicit	Implicit
Mobile to TV migration	2008	Tool	Backend	Yes	Yes	Yes
Remote AppBus	2009	Tool	Backend	Yes	Yes	No
Flexible framework	2009	Framework	Backend	Yes	No	Yes
JavaScript state	2011	Tool	Frontend	No	No	Yes
Multimedia prototype	2011	Evaluation	Frontend	Yes	Yes	Yes
Deepshot	2011	Framework	Backend	Yes	Yes	Yes
Single session start	2012	Tool	Backend	Yes	No	Yes
Virtual Browser	2012	Tool	Backend	No	Yes	No
Taskshadow-W	2013	Tool	Backend	No	Yes	Yes
Communication evaluation	2013	Evaluation	Backend	Yes	Yes	Yes
Interface IDE	2014	Tool	Frontend	No	No	No
XDKinect	2014	Framework	Backend	Yes	Yes	Yes
Search model for continuity	2015	Evaluation	Backend	Yes	Yes	No
Code generator	2015	Evaluation	Backend	Yes	Yes	No
Apple Continuity	2016	Evaluation	Frontend	Yes	Yes	Depends
CustomConsole	2016	Framework	Backend	Yes	No	No
Moving context kit	2017	Framework	Backend	Yes	Yes	No
Yanux framework	2018	Framework	Frontend	Yes	No	Yes
CORMORANT	2019	Tool	Backend	Yes	Yes	Yes
Model 4C's	2019	Framework	Backend	Yes	No	Yes

In the same way, works have been created that support the implementation of a part of continuity in applications. The main problem found in them is the poor compatibility with applications due to their focus on a specific aspect makes it difficult for continuity to be implemented by developers except in the aspects for which these tools were developed. It is also noteworthy that, in recent years, jobs such as [9–11] have focused on the creation of frameworks both at a conceptual and technical level for the creation of tools with interactions in continuity, showing a huge advantage over applications because they allow developers to expand the range of end uses that may have an application.

Based on the information presented in Table 2 and the reviewed works, we can see that, although explicit continuity and implicit continuity (called by many authors as maintenance of a state) are characteristics that do not always go hand in hand. The results, especially from those studies focused on evaluations of continuity implementations, suggest maintaining a state in applications that implement continuity is a very desirable feature.

With the analysis of the related works, it can be affirmed that one of the ways to implement continuity in the applications is through a framework or technological platform that provides developers with the necessary tools to create ecosystems with continuity characteristics. Although forays into these types of work have previously been made, there are aspects that remain to be covered, such as compatibility with modern platforms and with future platforms and the generalization of activities.

One of the most desirable features in continuity is maintaining an application state, which, in the long run, reduces the steps to resume an activity. This aspect is a topic little explored by researchers.

4. Exploratory study

In support of the points identified in the review of works, an exploratory study was designed to identify points of improvement in the user experience in multi-device applications, with a specific focus on applications that implement one or more aspects of continuity. The methods and materials used for this study are described below.

4.1. Experiment design

For this study, an exploratory experiment was designed and executed using the Google Forms tool; a questionnaire was constructed with a total of 20 items or questions that encompassed four different categories. Users were chosen randomly, regardless of their age, gender, or degree of academic studies. A total of 30 people were chosen for the application of the study.

The categories considered for the exploratory study are the following: i) general user data for classification and future consideration during data analysis; ii) continuity analysis in multi-device applications focused on music streaming; iii) continuity analysis in multi-device video streaming applications; and iv) continuity analysis in multi-device applications for document editing.

4.2. Evaluation tools

For the study, a questionnaire with 20 items was designed, the first four dedicated to categorical information from the user; the rest are described in Table 3. For its execution, the Google Docs tool was used; specifically, the Google Forms application allowed the questionnaire to be distributed quickly and effectively to the respondents. Another reason for its choice is that it provides a graphical analysis of the results and avoids the results' repetition by monitoring the respondents with their email.

Table 3: Questions applied in the exploratory study

ID	Question	Detail
5	How often do you use these devices to play music?	Question on Likert 5 points scale

6	It would be useful if the application automatically changes places to play your music on the most suitable device	Question on Likert 5 points scale
7	Would you like the application to be displayed as you interrupted it on another device when changing devices?	Question on Likert 5 points scale
8	What would you expect your player app to do when you want to switch devices?	Multiple option question
9	How often do you use these devices to play video?	Question on Likert 5 points scale
10	Of the following characteristics of an application for video playback, which of them are most important to you	Multiple option question
11	Is it useful to you that a video player application allows you to continue a movie / series right where you left off?	Question on Likert 5 points scale
12	You would like a video application to allow you to continue with a movie / series on another device only by turning it on and opening the application	Question on Likert 5 points scale
13	When it comes to playback across multiple devices, what would you improve on a video app?	Open question
14	How often do you use these devices to work on a document?	Question on Likert 5 points scale
15	Have you ever used the "pick up where you left off" feature in Microsoft Office	Multiple option question
16	How useful is this feature for you when switching from one workplace to another or between different devices?	Question on Likert 5 points scale
17	How much would you like Microsoft Office to allow you to continue a document without having to indicate it using the "continue where you left off" function?	Question on Likert 5 points scale
18	Would it be useful for you if, in addition to taking you to the section where you worked in the document, also the settings, clipboard and tools kept their last state?	Question on Likert 5 points scale
19	Would you like that when you open the Microsoft application (Office, Excel, PowerPoint) it would automatically open the last document you had open?	Question on Likert 5 points scale
20	When it comes to writing documents, what would you expect from an application when switching between contexts or devices?	Open question

4.3. Process

The participants' selection was carried out randomly, a short introduction to the subject of the study was given, and the instructions for filling out the tool; their participation was completely voluntary.

The tool was sent by electronic means to the users chosen for the test who had the approximate time of 1 day to answer them without affecting their working hours or daily activities. Finally, after a week, the data obtained through the platform was collected for subsequent analysis.

4.4. Results

Regarding music playback applications, it is notable that one of the characteristics most sought after by users is the "intelligence" of the application to play on the most suitable device available.

The results translate into that it would be instrumental for most users if they are playing music, and this automatically chooses the most appropriate device to play on. For example, a user who plays content on their cell phone, and when they get to Your home could automatically transfer this playback to your TV or computer, without the need for intermediate steps.

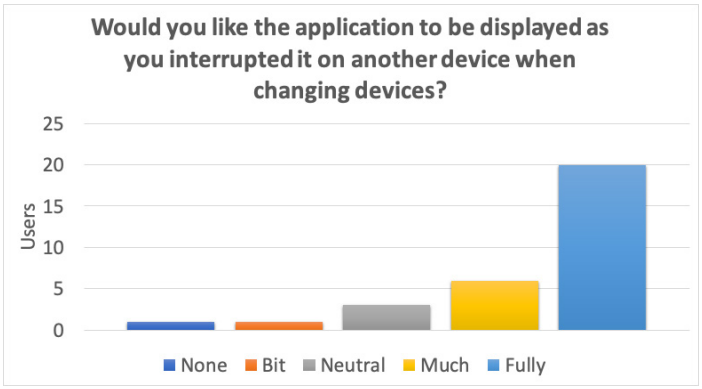


Figure 1: Opportunities of continuity in music streaming apps

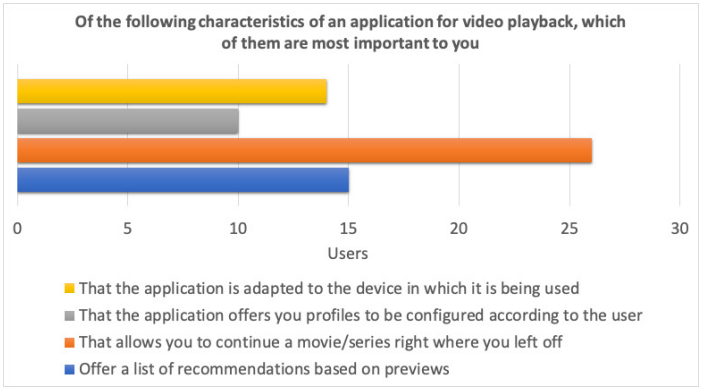


Figure 2: Continuity in video streaming apps

In addition, an aspect that music applications lack and that users expect from applications is the maintenance of a state, that is, that the application remembers where it was stopped the last time it was used, as can be seen in the second graph of Figure 1.

Regarding video applications, continuity has a greater presence in the application processes and turns out to be one of the characteristics that users like the most, as shown in Figure 2. In the same way as playback applications, music users are interested in the possibility of maintaining the state of the content reproduction when exchanging between devices. However, the approach applied in implicit continuity is not as relevant for users as seen in the second and third graphs of the survey conducted. Finally, for productivity applications, an analysis of the continuity implementation was carried out for an office suite widely used in the world and whose continuity implementation is explicit. Each time the user starts the application and opens a document you have been working on recently, you are presented with a button to go directly to the place where you last edited the document.

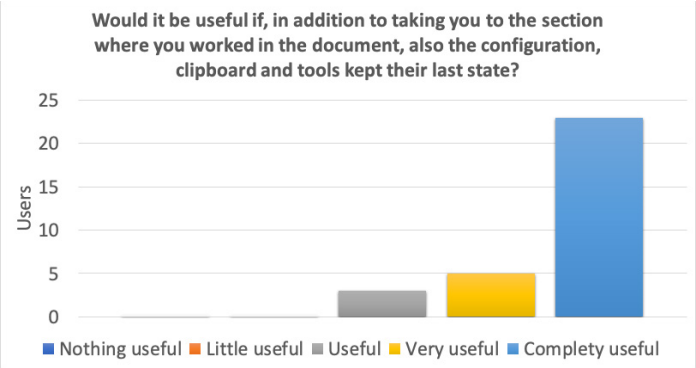


Figure 3: Continuity in document edition apps

In this type of application, two important aspects can be distinguished with regard to continuity. An important aspect mentioned by users is that the application directly remembers the status of the application with the documents they are working on, without the need for this function to be explicitly activated by the users. On the other hand, users would also find it useful to retrieve the configuration of tools for editing documents such as clipboards, temporary settings of the application, and additional tools such as a calculator or a browser; the results of the latter can see in Figure 3. With what was previously described, some users differ in the case that the application implements implicit continuity. It could be noted in the review of the comments made by the users; not all agree that the last document worked should be opened.

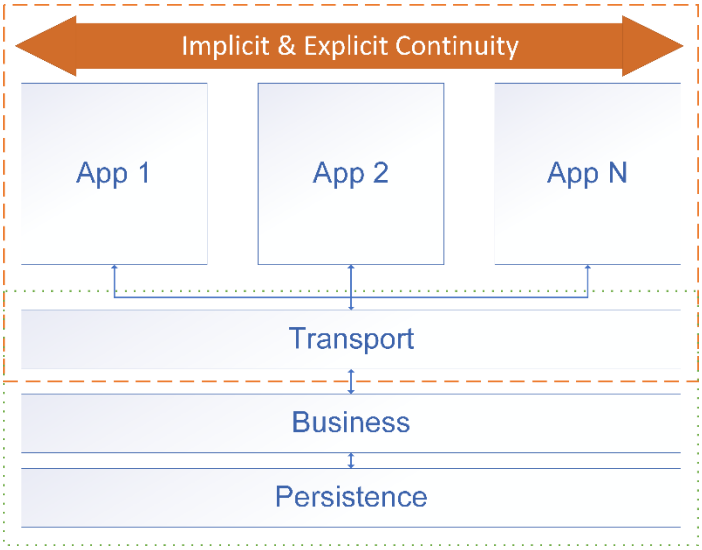


Figure 4: General model of the platform for implementation of continuity

5. Conceptual model for continuity implementation in software projects

Based on the models proposed by [5,9–11,15] and in the results of the continuity study carried out, a technological platform proposal was made at a conceptual level for the implementation of continuity in the activity flows shown in Figure 4 and detailed in its components in Figure 5. This model proposes an architecture for a platform where applications compatible with different devices can be easily implemented, and one of its goals is to allow its users to continue their task in a simple way when faced with a context change. The objective and the structure of the model for the platform allows the reduction of steps necessary to continue an activity between several devices, that is, to facilitate the implementation of implicit continuity in several applications.

In Figure 4, the general behavior of the platform can be observed. Through a set of layers, the necessary support is provided to create multi-device environments, taking advantage of the possibilities of explicit continuity and, above all, of implicit continuity, regardless of the type and number of applications being developed.

The persistence layer focuses on the storage of information corresponding to the application domain and the context where it operates (for example, the type of device used and the environment where it is done), typically composed of three elements, i) the application's usage status data, that is, configuration and usage information, this data contains information about how the last device was used ii) the device status data that considers device-specific details such as type of last device or the app version used which helps to adapt the interface of next used device and iii) the domain data referring to the application's own information (documents, photos, among others) which is used to present the data that were working, is based on unstructured language for the storage of information. This layer also contains the basic operational engine, which provides essential functions (create, update, delete, and read) over the data storage in order to operate with the basic application information.

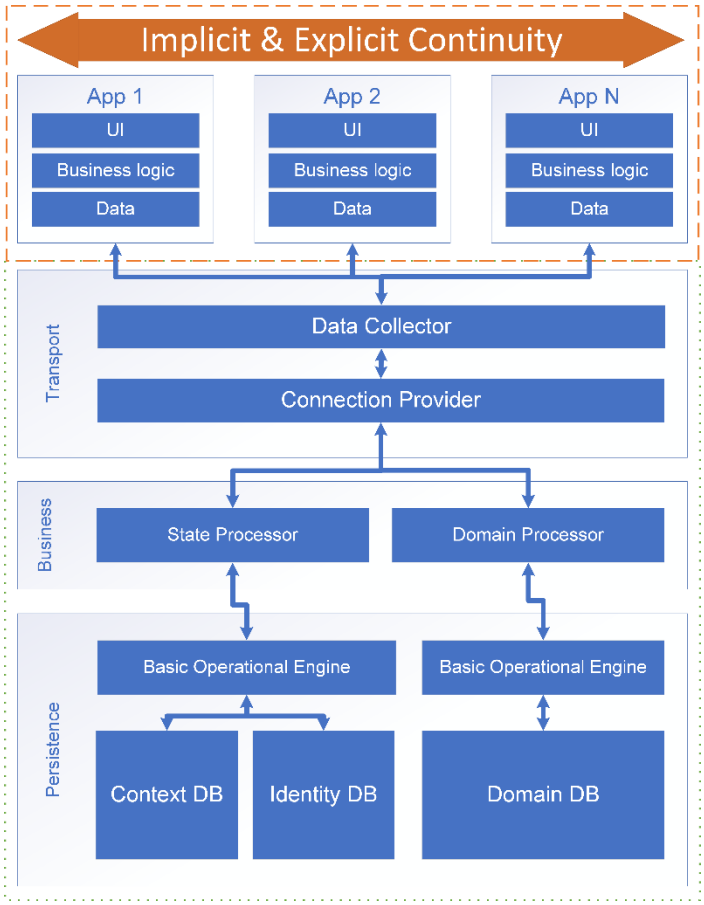


Figure 5: Components of the technological platform for the implementation of continuity in software projects

Table 4: Description of the conceptual model of a technological platform for the implementation of continuity in workflow

Characteristic	Description
Context DB	Storage component that allows to save the information of the context of use of the devices, it is a long-term data storage software, it contains elements that function as tools for the selection of device sequence and the configuration of interfaces
Identity DB	Storage component that stores information corresponding to users and that is used as a starting point to create state representations, this database must contain data from the user that help to identify the information that is stored in the context and domain database, its purpose is to maintain control over the information that the platform operates
Domain DB	Storage component focused on the operational information of the applications, it contains representations of the information that the user enters or obtains from the application.
Basic Operational Engine	It is a logical component that provides access to the basic functions on the storage components, the proposed implementation design defines an engine for domain data and another for identification and context information in order to reduce workloads over other components.
State Processor	Logical component that allows composing and decomposing the state representations and the information obtained from the operational engines and sending that information to the next layer depending on whether it is retrieved or entered into the platform, its purpose is to offer an interpreter for both directions of the platform.
Domain Processor	Logical component that provides functionality on domain information and that must establish the data transformation bases for the changes that are necessary between the jumps of multiple applications. This element works as a translator when it is necessary. It must also be responsible for communicating the domain data to the upper and lower layers in order to reduce the information contained by state representations
Connection provider	High-level component that functions as a data input to the lower layers, receives the information from the data collector and is responsible for passing it to one of the two processors of the business layer. This component is proposed to function as an API in order to be able to be integrated into software projects without having to modify the base structure of the project
Data Collector	Its a high-level component that is proposed to be responsible for direct communication with the final applications, it must be in charge of obtaining the necessary information to create the representation of the state that will be sent to the connection provider and also of presenting the information when a representation is retrieved from the platform towards a final device, It is proposed that this element be built as libraries for the specific programming languages of each final application
App 1,2… N	These are the final applications where the platform can be implemented, which can go from 1 to N, each one can have its own logic processes, information storage media and graphic interface.

The persistence layer communicates with the business layer, which comprises two processing services (state and domain logic). These processing elements are responsible for making the necessary changes to the data; being the most important elements since a large part of the adaptation depends on them. Finally, in the highest part, the transport layer is composed of two elements (data collector and connection provider). These elements are those that are exposed to the developers and are integrated into the final applications, built in the form of APIs, extracting the necessary information from the applications and sending them to the processing models so that the rest of the platform works. It is also are responsible for making the necessary changes to the application interfaces. It is through these two elements that applications reduce the steps necessary to resume activities. These elements have to be able to scale to serve as many final applications as a program has been developed.

In order to support the understanding of this platform, a diagram was developed with the components of the model and which is presented in Figure 5, as well as the description of these components which is found in Table 4.

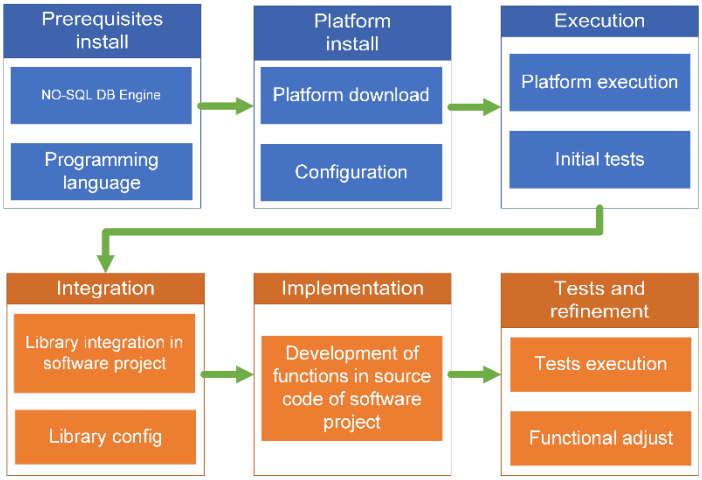


Figure 6: Platform implementation roadmap

In Figure 6, the platform implementation is shown; in the upper part the work stages are shown directly on the platform which are: i) pre-requisite installation in which the execution environment of the platform through the configuration of tools for database management and the programming language used by the platform, ii) installation of the platform that consists of downloading the platform code and its configuration with the database management tool, iii) Execution which consists of the start-up and rapid tests of the correct operation of the platform.

At the bottom of the implementation sheet, the steps after the installation of the platform are shown, which apply directly to the software project; these steps are: iv) Integration where the libraries that are added to the software projects allow communication to the platform and its configuration, v) Implementation where the views and functions of the software are developed, and it is indicated where the platform libraries will work and vi) Testing and refinement where the tests should be carried out on each final application and adjusting functions in case these changes are necessary.

6. Case Study

The platform is an intermediate software that communicates with the different components of other applications; due to the nature of this type of software. It is impossible to carry out functional and use tests directly on its construction, so the software is implemented in a Case study where both the function and the user experience when using continuity could be evaluated.

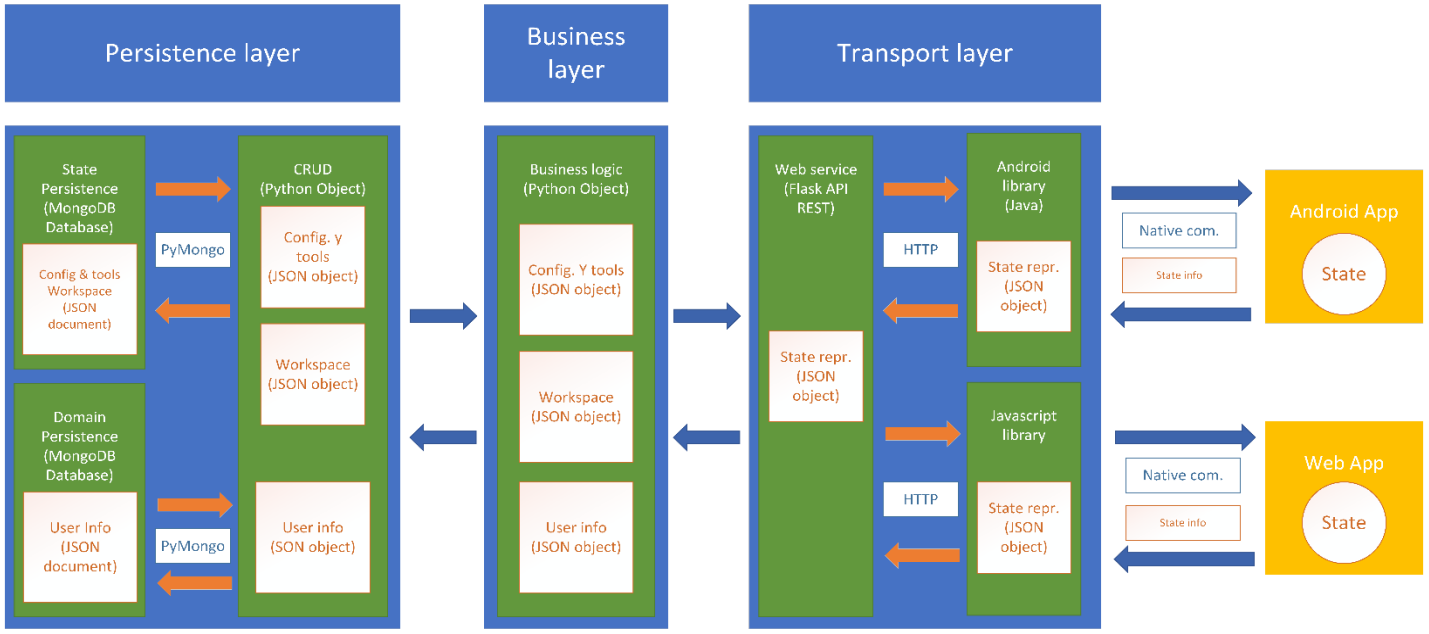


Figure 7: Implementation architecture for study case

The selected case study is an application that allows you to create exams according to a catalog of questions previously evaluated and approved. The process of editing an exam can be a complex task for those who take it, and because of this, in many cases, it is necessary to move it between one device and another.

The case study has three elements: a direct administration system that allows the question catalog to be generated, edited, and debugged. A web application is attached to the administration system to generate the exams from any browser. Finally, an Android application that also allows generating the exams.

The structure of the selected application makes it perfect to test the proposed solution because it's possible to change between two different types of apps and two different contexts, also, the selected applications can be executed in many modern mobile devices and even in any device that can accessing internet through a web browser.

On the other hand, the limitations of this case study are that it does not allow the analysis of an interaction that is too deep, such as using a text editor, in addition to the fact that it is an application in its final stages of development, so an error that is not related to the use of continuity and biases the study in a certain way.

Two libraries were built that allowed to integrate the platform in the case study, the first in JavaScript for the web sections and the second in Java-Android for the android section of the case study. Figure 7 shows in detail the structure and elements that make up the platform in its implementation for the case study.

The persistence layer is implemented through two databases that represent the storage components, one for the domain and one for the state. These connect a first layer of python that offers basic functions on databases (basic functional engines). At the second level there is a business layer that implements methods and processes to arm and disarm the state representations and finally the transport layer is responsible for implementing a RESTful-type web service and the two libraries mentioned above, while the REST service receives and sends the information when the libraries so require, the libraries control directly and in a practically automatic way the events in the applications of the case study that save and retrieve the information.

7. Conclusions and future work

The development of computer systems has changed over time, adapting to the different tools that arise daily. The notable increase in the use of mobile devices is one of the main reasons that have led to the emergence of new interaction and development paradigms.

One of the main topics that researchers and the market related to systems development have focused on is multi-device interaction. Continuity arises as part of the multiple ways in which a multi-device interaction occurs. The possibilities that continuity offers to improve the different ways in which a user interacts with a system are vast, yet there is still a long way to go.

One of the problems that researchers currently focus on solving through the use of continuity is the minimization of interruptions in the flow of an activity, something that could be seen through the review of works related to this research and the exploratory study carried out is a feature that users are beginning to expect from the systems.

The design of a platform to implement this type of interaction is complex due to the components that are required for correct operation, as well as the very logical structure that is required to maintain the state of the applications between the changes made by the user.

The results obtained once the implementation was carried out in the case study were favorable, the platform facilitated the continuity implementation in the case study and also allowed to reduce the necessary steps to resume an activity, which facilitated context changes for users.

In contrast to other works that were reviewed during the writing of this document, we can see that a platform that provides full support for continuity has a deeper impact than applications such

as [2,4,5,12–14], whose emphasis is strongly focused on the login and the migration of user data over the migration of operation data.

On the other hand, a better performance can also be noted than the proposals made by authors such as [3,9,11,16,18], where migration Information is a present feature, but small details such as the configuration or the last view requested are not considered, forcing the user to have to remember this when resuming an activity. It is still necessary to make improvements with respect to the platform's ability to adapt based on usage preferences and configurations that affect continuity, as well as to carry out tests that guarantee its effectiveness, with continuity being a useful tool in different contexts of use. However, it still requires a long and dedicated development process.

Finally, as part of the future work, it is expected to carry out a series of tests through the case study in order to validate the functionality and design of the Platform.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors thank the reviewers of this article for their helpful comments. The corresponding author also thanks to CONACYT-México for the grant No. 715387 used for his postgraduate studies.

References

[1] L. Tan, N. Wang, "Future Internet: The Internet of Things," in ICACTE 2010 - 2010 3rd International Conference on Advanced Computer Theory and Engineering, Proceedings, 2010, doi:10.1109/ICACTE.2010.5579543.

[2] B. Cheng, "Virtual browser for enabling multi-device web applications," Proceedings of the Workshop on Multi-Device App Middleware - Multi-Device '12, 1–6, 2012, doi:10.1145/2405172.2405175.

[3] C. Ye, J. Wei, H. Zhong, T. Huang, "Middleware Support for Internetware : A Service Perspective," Systems Research, 2010, doi:10.1002/ardp.201300158.

[4] H. Kim, E.A. Lee, "Authentication and Authorization for the Internet of Things," IT Professional, **19**(5), 27–33, 2017, doi:10.1039/b904090k.

[5] P.A. Cabarcos, F. Almenares, R. Sánchez Guerrero, A. Marin, D. Díaz-Sánchez, "Multi-device Single Sign-on for cloud service continuity," Digest of Technical Papers - IEEE International Conference on Consumer Electronics, 644–645, 2012, doi:10.1109/ICCE.2012.6162011.

[6] D. Raptis, J. Kjeldskov, M.B. Skov, "Continuity in Multi- - Device Interaction: An Online Study," Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16, 2016, doi:10.1145/2971485.2971533.

[7] M. Levin, Designing Multi-Device Experiences, First, O'Reilly Media, Inc, 2014.

[8] J. Schöning, H. Reiterer, C. Holz, J. Vermeulen, N. Marquardt, C. Klokmose, S. Houben, "Opportunities and challenges for cross-device interactions in the wild," Interactions, **24**(5), 58–63, 2017, doi:10.1145/3121348.

[9] A. Pedro, C. Nuno, N. Rui, "Designing a framework to support the development of smart cross-device applications," Mum18 Rw, 367–374, 2018.

[10] A. Marzo, "CustomConsole : A Framework for Supporting Cross-device Videogames," 1952–1958, 2016.

[11] K. O'Leary, T. Dong, J.K. Haines, M. Gilbert, E.F. Churchill, J. Nichols, "The Moving Context Kit: Designing for Context Shifts in Multi-Device Experiences," Proceedings of the ACM Conference on Designing Interactive Systems (DIS 2017), 309–320, 2017, doi:cshg.

[12] M. Barisch, J. Kögel, S. Meier, "A flexible framework for complete session mobility and its implementation," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **5733 LNCS**, 188–198, 2009, doi:10.1007/978-3-642-03700-9_20.

[13] T.H. Chang, Y. Li, "Deep shot: A framework for migrating tasks across devices using mobile phone cameras," in Conference on Human Factors in Computing Systems - Proceedings, 2011, doi:10.1145/1978942.1979257.

[14] M. Nebeling, E. Teunissen, M. Husmann, M.C. Norrie, "XDKinect: Development framework for cross-device interaction using kinect," in EICS 2014 - Proceedings of the 2014 ACM SIGCHI Symposium on Engineering Interactive Computing Systems, 2014, doi:10.1145/2607023.2607024.

[15] M. Nebeling, T. Mintsi, M. Husmann, M. Norrie, "Interactive development of cross-device user interfaces," Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14, 2793–2802, 2014, doi:10.1145/2556288.2556980.

[16] N. Narasimhan, C. Janssen, M. Pearce, "Remote AppBus - Enabling seamless access to short term memory on mobile devices," 2009 6th IEEE Consumer Communications and Networking Conference, CCNC 2009, 1–5, 2009, doi:10.1109/CCNC.2009.4784773.

[17] D. Hintze, J. Kepler, "CORMORANT : Ubiquitous Risk-Aware Multi-Modal Biometric Authentication across Mobile Devices," **3**(3), 2019.

[18] F. Bellucci, G. Ghiani, F. Paternò, C. Santoro, "Engineering JavaScript state persistence of web applications migrating across multiple devices," in Proceedings of the 2011 SIGCHI Symposium on Engineering Interactive Computing Systems, EICS 2011, 2011, doi:10.1145/1996461.1996502.

[19] F. Paternò, C. Santoro, A. Scorcia, "User interface migration between mobile devices and digital TV," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **5247 LNCS**, 287–292, 2008, doi:10.1007/978-3-540-85992-5-28.

[20] S. Li, G. Pan, Y. Xu, Z. Ye, L. Chen, "Taskshadow-W: NFC-triggered migration of web browsing across personal devices," in UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing, 2013, doi:10.1145/2494091.2494119.

[21] R. Bernhaupt, M. Abdellatif, T. Mirlacher, "Cross-device continuous media consumption," 217, 2011, doi:10.1145/1941007.1941046.

[22] S. Han, Y. Zhen, H. Daqing, "Understanding and supporting Cross-Device Web Search For Exploratory Task with Mobile Touch Interactions," ACM Transactions on Information Systems, **33**(4), 2015.

[23] E. Umuhoza, H. Ed-Douibi, M. Brambilla, J. Cabot, A. Bongio, "Automatic code generation for cross-platform, multi-device mobile apps: Some reflections from an industrial experience," in MobileDeLi 2015 - Proceedings of the 3rd International Workshop on Mobile Development Lifecycle, 2015, doi:10.1145/2846661.2846666.

[24] R. De A. Maués, S.D. Junqueira Barbosa, "Cross-communicability: Evaluating the meta-communication of cross-platform applications," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **8119 LNCS**(PART 3), 241–258, 2013, doi:10.1007/978-3-642-40477-1_15.

[25] T. Dong, E.F. Churchill, J. Nichols, "Understanding the Challenges of Designing and Developing Multi-Device Experiences," Designing Interactive Systems, 62–72, 2016, doi:10.1145/2901790.2901851.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1103-1108 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Evaluation of the Physico-Chemical Properties of Soil and Apple Leaves (Malus Domestica) in Beni Mellal-Khenifra Region, Morocco

Berrid Nabyl[1], Lougraimzi Hanane[2], El-Khabbazi Houda[3], Abidli Zakaria[3], Hamidi Otman[1], Keltoum Rahali[1], El Mahjoub Aouane[1,*]

[1]*Laboratory of Biotechnology, Quality and Environment, Faculty of Science, Ibn Tofail, University, Kenitra, Morocco*

[2]*Laboratory of Nutrition, Health and Environment, Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*

[3]*Laboratory of Genetic and Biometric, Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*

ARTICLE INFO

Article history:
Received: 01 September, 2020
Accepted: 26 November, 2020
Online: 14 December, 2020

Keywords:
Physico-chemical properties
Soil
Apple leaves
Mineral elements
Organic matter
Morocco

ABSTRACT

The apple tree (Malus domestica) is an agricultural species of great importance in Morocco. Currently, it occupies about a quarter of the surface of fruit-bearing rosaceae. Apple tree production has evolved rapidly, stimulated by a buoyant market, a varietal range which is tending to diversify and a dynamic profession. The present study aims to determine the physico-chemical properties of soil and apple leaves in Beni Mellal-Khenifra region, Morocco. Data collection was consisted to sample the soils at different depths, three Pearson correlation coefficients were determined: i) Correlation between organic matter and soil mineral elements, ii) Correlation between clay content and soil mineral elements and correlation between CaCO3 content and soil mineral elements. The analytical results of soils were compared to reference values. The results show that there is no correlation between these mineral elements and the rate of CaCO3 and clay fraction. Analysis of the main components showed that the apple leaf contains more sodium, copper and zinc and less magnesium, phosphorus and potassium. The coefficient of variation (CV) of manganese content expressed as a percentage is significantly higher than that of the other elements, while the potassium content is the lowest, in particular for magnesium and sodium the coefficient of variation is zero. In apple orchards in the Middle Atlas, there are great variations in terms of yield levels, fertilizer inputs, soil richness in nutrients and their concentration at leaf level. These variations are due to differences in cultivation practices, in particular fertilization, because the relationship between apple yield and nutrient content at soil and leaf level is significant.

1. Introduction

The agricultural sector plays a key role in economic activities in Morocco. Likewise, it represents a strong support to the development of the country because it constitutes a higher percentage of 20% of the gross domestic product and employs about 40% of the active population [1, 2].

In Morocco, the apple tree (*Malus domestica*) is the species, which occupies the largest surface in terms of fruit rosaceae. Currently, orchards planted with apples total 32,000 ha, or about a quarter of the surface area of fruit-bearing rosaceae. Annual production is estimated at 560-600,000 tonnes, based on an average yield of around 20 T/ha [3]. However, Intensive agriculture and the search for new fertile land put pressure on the ecosystem, which in turn leads to reduced soil fertility [4, 5]. As a result, yields are relatively low and land productivity decreases. There are several types of soils in Morocco which have developed on different geological substrates and in various climatic, topographical and biotic contexts. These soils must be protected as a natural resource because it contains substances essential to the maintenance of ecosystems and human populations, such as essential minerals. this study is carried out in the Beni Mellal-Khenifra region which is characterized by a variety of soils and by the predominance of fertile types rich in organic matter which allows an important agricultural activity, in particular for apple trees [6]. So far, interpretation of soil and leaf analyzes for

*Corresponding Author, El Mahjoub Aouane, University Campus BP 133 Kenitra, Morocco. Email: aouane_mahjoub@yahoo.fr

Moroccan apple orchards has been limited. For this reason, the main objective of our study is to determine the physicochemical properties of soil analysis and apple leaves in the Beni Mellal-Khenifra region in Morocco. More specifically, i) Determining the correlation coefficients between organic matter and soil mineral elements, ii) Correlation between clay content and soil mineral elements and correlation between CaCO3 content and soil mineral elements.

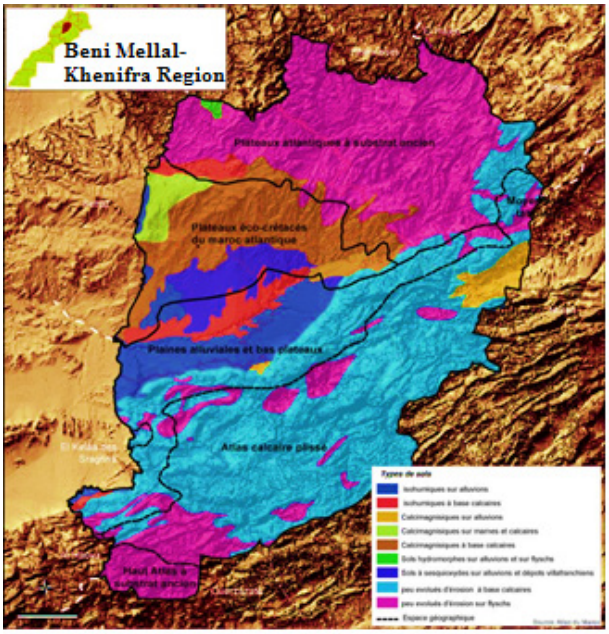


Figure 1: Geographical location and soil types in Beni Mellal-Khenifra region, study site [7].

2. Materials and methods

2.1. Presentation of the study area

The Beni Mellal-Khenifra region is a major contributor to national agriculture. From field crops to local products, its environmental diversity allows it to have a very varied agriculture. Between mountains, foothills and plains, the diversity of climates and soils offers the Beni Mellal-Khenifra region many agricultural opportunities. In addition to its environmental and natural wealth, this region has the advantage of having significant water resources unlike other regions of the Kingdom. The study area is part of the perimeter of the Beni Mellal-Khenifra region, which is made up of four large distinct geographic areas, which maintain relations of exchange, complementarity and forms of solidarity, particularly in the field of agriculture and extensive farming [7]. It has a total surface area of 2.8 million hectares, of which 960,000 hectares of useful agricultural area (205,000 hectares under irrigation and 700,000 hectares of rangelands). Moreover, the region has adopted a specific agricultural plan designed to remove the constraints that hamper agricultural development. It is a specific region, inserted between the Atlas Mountains and the Atlantic coastal provinces, with a much-contrasted physiognomy and a marked character. However, these contrasts, however clear-cut they may appear, present undeniable advantages for the implementation of mutually beneficial solidarity, the rich natural heritage (hydraulic and forest) which is highly diversified, the agricultural potential of the plain which is being reconverted (Beni Amir and Beni Moussa), the opportunities offered by mining activity, the major structuring projects (connecting infrastructures and agro-industrial and commercial platforms) and the hinge position in the national system [7].

2.2. Soil types in the area

The soils in the region are diverse due to the nature of the source rocks. They are shallow, shallow, and low in humus. Their water reserves are sometimes limited. Moreover, the dominance of holm oak, a very plastic species from an ecological point of view, confirms this situation when the ecological atmosphere is difficult. The substrates are colonized by thermo-xerophilic (*Aleppo pine*, thuja and red juniper) and xerophilic (*Thuriferous juniper*) species (Figure 1). Even at altitude, the soils generally remain shallow to shallow, humus and humid for a fairly long period of the year, allowing the establishment of friendly forest species (zene oak, cedar and holm oak) [7].

2.3. Soil sample collection

The study covered 19 observation sites (10 samples by site). Soil samples were taken at different depths (0-60 cm) using an auger on plots of apples. These samples were air dried overnight or longer and sieved through a 2 mm sieve to obtain a fine soil. The results of the soil analyzes were compared to limit values and reference standards. We also measured the electrical conductivity and pH of the main elements using the same techniques as below:

Table 1: Analysis technique.

Analysis	Methods
Granulometry	Sieving (2 mm to 45 µm sieve)
pH 1/ 2.5	pH to put
CE 1/5 25 °C	Granulometry
Cation exchange capacity CEC and exchangeable shares	Cationic exchange capacity CEC and exchangeable shares
Na+ and K+ Cations	Flame spectrophotometer
Ca++ and Mg++ Cations	Volumetric titration
Chlorides Cl-	MOHR method (silver nitrate precipitation in the presence of chromates)
Sulphate SO4--	Precipitation at BaCl2

2.4. Sampling of plant material

The foliar analysis of the apple tree (or foliar diagnosis of the apple tree), is the best way to determine the quantity and the type of fertilizer to give to the fruit trees. Thirty samples of healthy and homogeneous trees were chosen (70 to 90 days after full flowering (stage F2)) and the samples were representative of the plot. We chose the woody branches representative of the vigor and the average growth of the tree (eliminate the greedy ones), the leaf in the middle of the branch with its stem.

In total 200 sheets. The collected leaves are placed in new perforated plastic bags and awaiting transport to the analysis

laboratory; they are stored cool or cold at 4-5 ° C so that they retain all their characteristics. The analysis is carried out the day after harvest at the latest. It focuses in particular on the nitrogen content expressed as a percentage of dry matter (% DM). The results of soil analyzes were compared to values thresholds and reference standards specific to apple production (Table 2).

Table 2: Reference thresholds [8].

	N (% MS)	P (% MS)	K (% MS)	Ca (% MS)	Mg (% MS)
Apple	2.2-2.4	0.22	1.13 à 1.75	1.25	0.22 à 0.3
Prunes	2.4 -3	0.14-0.25	1.6 - 3	1.5 – 3	0.3 - 0.8

2.5. Statistical methodology

Statistical analysis was performed using SPSS software. Descriptive statistics were used and analyzes of variance and comparison of means were performed at a probability level of $p<0.05$.

3. Results

3.1. Physico-chemical characteristics of the soil studied for each site

The results of the particle size analysis of the studied soils are presented in table 3. On the samples analyzed, the texture tests did not reveal significant clay contents except for site 2, which showed values barely greater than 45%. The clay content of the studied soils varies between 26% and 46.57%. Under these conditions, the deep migration of nutrients is slow and leaching losses are low. According to the Texture Triangle (FAO), which is used to determine the textural class of soils, the study area mainly has loamy clay soils. Considering also that apple trees prefer healthy and deep loamy-clayey-loamy soils with a permeable subsoil, it can be assumed that the studied soils are favorable for this apple crop.

Table 3: Physical constituents of the soils studied.

Observation sites	Depth	Granulometry (%)		
		Clay	Silt	Sand
1	20-50	35	30	35
2	0-20	35	30	35
3	0-20	45	25	30
4	20-50	45	30	25
5	30-60	26	18.77	55.23
6	0-30	40.73	25.98	33.29
7	30-60	39.91	26.1	33.99
8	0-30	26.92	47.52	25.56
9	30-60	27.13	46.69	26.18
10	0-30	31.97	39.2	28.83
11	30-60	32.1	38.99	28.91
12	0-30	44.81	43.24	11.95
13	30-60	45.02	43.45	11.53
14	0-30	46.49	42.54	10.97
15	30-60	46.96	41.99	11.05
16	0-30	40.46	35.71	23.83
17	30-60	41	35.5	23.5
18	40-60	42.7	44.05	12.95
19	0-40	41	43.5	15.5

The pH is an important parameter of soil dynamics, the degree of acidity or basicity plays a very essential role in the absorption of nutrients by the plant. The average pH of the soil samples studied is eight, which shows that the majority of sites are basic with an alkaline tendency (Table 4).

Table 4: General Statistics of different physico-chemical characteristics of the samples studies.

Characteristics	Min	Max	Median	Average		Standard deviation
				Statistique	Erreur std	Statistique
MO	0.48	3	2.00	1.71	.187	0.772
pH	8	9	8.00	8.06	.063	0.250
CaCO3	6	41	10.00	12.11	1.838	8.013
Clay	26	47	41.00	38.58	1.596	6.955
Silt	19	48	37.50	36.44	2.033	8.624
Sand	12	55	26.00	26.11	2.567	10.889
Fe mg/kg	1	8	6.50	5.83	.452	1.917
Bm g/kg	0	0	0.00	0.00	0.000	0.000
Zn mg/kg	0	6	3.00	2.72	0.523	2.218
Cu mg/kg	0	1	1.00	0.65	0.119	0.493
P2O5 mg/kg	2	47	5.00	10.40	2.953	11.438
K2O meq/100g	0	551	340.00	301.94	38.147	157.283
CaO meq/100g	12	6850	1388.00	3217.27	689.230	2669.376
MgO meq/100g	0	380	320.00	269.53	36.272	140.479

The soils studied exhibited an organic matter content in the range 0.48-3%. Analysis of the elements of Iron, zinc and copper revealed average contents of 6.5 mg/kg, 3 mg/kg and 1 mg/kg respectively. For all the soils studied and sampled at different depths, three Pearson correlation coefficients were determined: one between the different soil types and CaCO3 content, the second between the different soil types and the percentage of organic matter, and the third between the different soil types and the clay content. The coefficients of determination obtained show

that variation of soil types and their sampling at different depths explains 12% of the variation of clay content, 11% of CaCO3 and 8% of the material content. In terms of physico-chemical characteristics, descriptive statistics are the set of methods and techniques that allow us to present, describe and summarize, numerous and varied data; they allow us to know the characteristics of our sample.

For all variables, the standard deviation is more or less low, so the dispersion of the measurements around the mean is narrow. The results are therefore consistent, except for the last three variables, namely the K2O, CaO and MgO content, where the standard deviation is very high, which shows that there is a wide dispersion of the results for these contents (Table 5,6 and 7).

Table 5: Correlation between organic matter and soil mineral elements

		Fe mg/kg	B mg/kg	Zn mg/kg	Cu mg/kg	P2O5 mg/kg	K2O meq/100g	CaO meq/100g	MgO meq/100g
MO %	Correlation Pearson	0.402	.a	-0.073	0.239	-0.018	0.573*	0.433	0.285
	Sig.(bilatérale)	0.123		0.788	0.391	0.951	0.026	0.094	0.285
	N	16	5	16	15	14	15	16	16

***The correlation is significant at the 0.01 level (bilateral).*
**The correlation is significant at the 0.05 level (bilateral).*
at. Calculation impossible because at least one variable is a constant

Table 6: Correlation between clay content and soil mineral elements

		Fe mg/kg	B mg/kg	Zn mg/kg	Cu mg/kg	P2O5 mg/kg	K2O meq/100g	CaO meq/100g	MgO meq/100g
Arg %	Correlation Pearson	-0.142	.a	-0.108	-0.135	-0.054	-0.329	0.369	-0.312
	Sig.(bilateral)	0.575		0.671	0.605	0.850	0.197	0.132	0.207
	N	18	6	18	17	15	17	18	18

***The correlation is significant at the 0.01 level (bilateral).*
**The correlation is significant at the 0.05 level (bilateral).*
at. Calculation impossible because at least one variable is a constant.

Table 7: Correlation between CaCO3 content and soil mineral elements

		Fe mg/kg	B mg/kg	Zn mg/kg	Cu mg/kg	P2O5 mg/kg	K2O meq/100g	CaO meq/100g	MgO meq/100g
CaCO3	Correlation Pearson	0.224	.a	0.377	0.375	-0.201	-0.010	0.171	0.321
	Sig.(bilateral)	0.372		0.123	0.138	0.473	0.970	0.499	0.194
	N	18	6	18	17	15	17	18	18

***The correlation is significant at the 0.01 level (bilateral).*
**The correlation is significant at the 0.05 level (bilateral).*
at. Calculation impossible because at least one variable is a constant.

In order to study the relationship between the mineral element content and the CaCO3, MO and clay fraction, we calculated the Pearson correlation coefficient (Figure 2). The results show that there is no correlation between these mineral elements and the rate of CaCO3 and clay fraction. The results for the organic matter content show that the latter explains 57% ($p<0.05$) of the variation in K2O content.

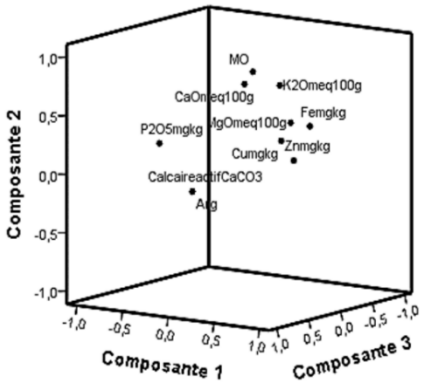


Figure 2: Principal component analysis of the physico-chemical properties of the soils sampled

The ANOVA test application, to compare the average characteristics of 19 sites (without distinction of depths) is presented in the table above. It shows that in the different sites studied, the difference in mean is insignificant for the parameters pH, organic matter, CaCO3 content, clay content, B, CaO; it is also insignificant for the silt and sand variables, so the null hypothesis H0 is accepted as long as there is no difference in the means. For the variables Fe, Zn, Cu, P2O5, K2O, MgO the difference of the means is significant and, in this case, the null hypothesis H0 is rejected.

3.2. Physico-chemical properties of apple leaf soil

The differences in the average composition of the three varieties are shown in the table 8. The apple leaf contains more Na, Cu and zinc and less Mg, P and K. The coefficient of variation (CV) of the manganese content expressed as a percentage is much higher than that of the other elements, and the coefficient of variation (CV) of the potassium content is the lowest, especially for magnesium and sodium the coefficient of variation is zero (Figure 3).

Table 8: Global apple leaf nutrient content statistics.

	Average	Standard deviation	CV	Min	Max
Magnesium, Mg (%)	0.1900	0	0	0.19	0.19
Phosphorus, P (%)	0.2200	0.01732	7.873	0.20	0.23
Potassium, K (%)	0.553333	0.0152753	2.761	0.5400	0.5700
Calcium, Ca (%)	1.4333	0.25580	17.846	1.19	1.70
Manganese, Mn (ppm)	24.033	6.5577	27.286	20.0	31.6
Zinc, Zn (ppm	32.733	6.0343461	18.435	25.8000	36.8000
Copper, Cu (ppm)	129.33	12.662	9.790	118	143
Boron, B (ppm)	26.333	2.5482	9.677	23.4	28.0
Sodium, Na (%)	249.00	0.000	0	249	249

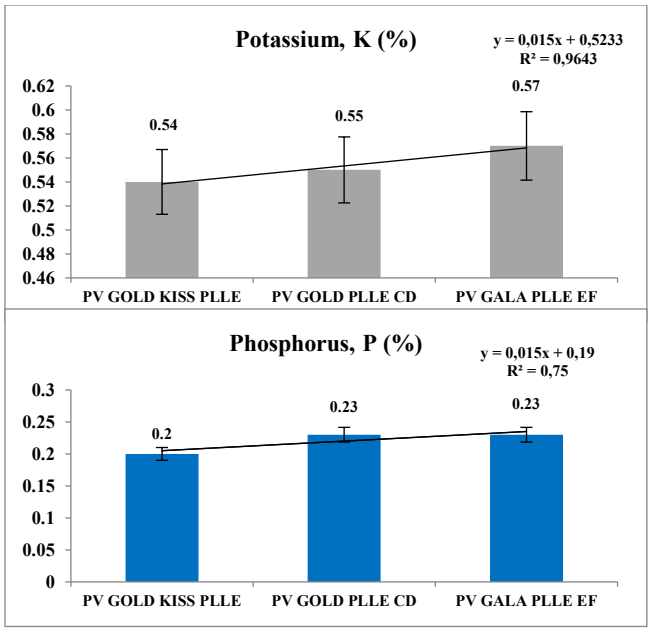


Figure 3: Phosphorus and potassium content in different apple varieties.

The differences in the average composition of the four varieties are shown in the table 9 and figure 4. Apple leaf contains more Na, Cu and zinc and less Mg, P and K. The coefficient of variation (CV) of the copper content expressed as a percentage is significantly higher than that of other elements, and the coefficient of variation (CV) of boron content is the lowest.

Table 9: Global statistics on nutrient levels in apple fruit twigs.

	Average	Standard Deviation	CV	Min	Max
Azote, N (%)	0.42000	0.151217	36.004	0.2100	0.560
Magnesium, Mg (%)	0.612500	0.4090130	66.778	0.2500	1.100
Phosphorus, P (%)	0.060000	0.0529150	88.192	0.0100	0.130
Potassium, K (%)	2.4800	0.84573	34.102	1.31	3.26
Calcium, Ca (%)	1.8000	1.07316	59.620	0.60	3.10
Manganese, Mn (ppm)	4.5100	1.98741	44.067	2.10	6.30
Zinc, Zn (ppm	11.9675	7.93202	66.280	2.60	19.70
Copper, Cu (ppm)	2.2225	2.05596	92.506	1.02	5.30
Boron, B (ppm)	4.9475	0.94026	19.005	4.12	6.30
Sodium, Na (%)	0.2100	0.07659	36.473	0.15	0.31
Chlorine, Cl (%)	0.1525	0.0377492	24.754	0.1100	0.200

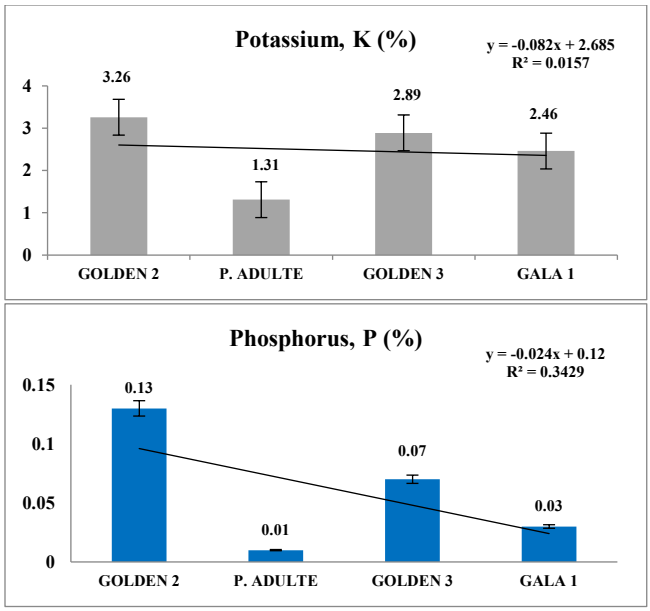


Figure 4: Phosphorus and potassium content in different varieties of apple fruit.

4. Discussion

The apple tree represents in Morocco an agricultural species of great importance. Growing apples requires a long period of vegetative rest to meet its cold needs; it also requires very specific conditions with regard to the various production factors, in particular for the soil / irrigation water combination. The main areas of apple plantations in Morocco are; Meknes, Midelt, Khenifra, Haouz Marrakech, Fes, Ouarzazate. Our study was carried in Beni Mellal-Khenifra region. The results of the granulometric analysis of the studied soils showed that the samples have a silty clay texture, which is often very favourable for apple tree cultivation. A good yield could therefore be obtained if the mineral elements provide the essential and important quantities for the good development of the crop. According to the literature review, the physicochemical properties of the soil play a major role in the yield of apples, in particular the pH [10,11]. It should be noted that apple trees generally prefer slightly acidic soils with a pH between 6.5 and 6.8 [12]. However, the soil samples analyzed recorded rather average alkaline pH values of up to 8.85. Despite the measured pH, qualified as alkaline, the yields obtained on the farm are very satisfactory. Indeed, even if we often talk about an optimal pH for a particular plant, in fact the pH does not meet a physiological need of the plant. It is the action of pH on the physical, chemical and biological properties of the soil that creates a more favourable environment for mineral nutrition and plant growth. In fact, plants are most often able to cope with a pH difference of up to one unit. Most plants grow very well in soil with a pH of 6.5 [12]. Doucet in 2006 states that agricultural species that grow at pH 6 most often prefer a pH of 6.5 and that all plants that grow in neutral soil with a pH of around 7 can tolerate a pH of 6.5 [13]. Regarding active limestone, apple tree fears soils with a high content of active limestone. This is not the case for our study soil with variable proportions which can sometimes be very high or even twice the tolerated standards. This can easily lead to iron deficiency. Furthermore, the more abundant the total CaCO3 rate in the soil, the greater the risk of blocking trace elements and phosphorus, hence the need to take these constraints into account

when developing fertilization programs. For a good interpretation of the rate of organic matter, it is necessary to take into account the quantity, but also the quality of this organic matter, the ventilation, humidity and temperature conditions. According to the guide reasoned fertilization in fruit arboriculture, presented by the Chamber of Agriculture of Tarn-et-Garonn, the value of the desirable organic matter content according to a clay content> 30 % must be between 2.5 and 3%. However, the organic matter content in the present study is generally low, with contents varying between 0.48% and 3 %. Nevertheless, it is therefore necessary to add MO to the soil. As long as the OM content is high, the production potential of the soil is great. In addition, it will improve the physical properties (structure, porosity, water and air circulation, water retention capacity) and chemical (cation exchange capacity, reversible anionic bond, release of trace elements. by decomposition in the soil) [15]. Compared to the interpretation standards given by the INRA (National Institute of Agronomic Research) for Contribution to the establishment of standards of fertilization of apple trees in the Middle Atlas. According to Dr. Rachid Razouk (researcher, CRRA Meknes), the optimal levels of nutrients found in clay-loam soils (the case of the study) are classified as low for phosphorus, in the “sufficient” class for potassium and in the "high" class of Ca and Mg [16]. However, for K2O, the values found correspond in fact to the optimum intakes for a high apple yield. Internationally, experts agree that 1.5% is the theoretical critical limit, below which fertility declines rapidly. It is often desirable to aim to maintain a minimum rate of 2.5% and even 3.5 to 4% in heavy soils [13]. A similar study was carried out to assess the quality of agricultural soils in the northeastern area of the plain Tadla (Morocco).The results indicated that 1.12% of the study area has poor suitability for sustainable intensive agriculture due to their unsuitable texture and low salinity, while about 77% of cultivated soils are adapted to agricultural production [17]. According to Barakat et al, 2017 the diagnostic of the quality of agricultural land in Beni-Moussa, Tadla plain (Morocco), showed that about 1.69% (1324.38 ha) of studied area are classified in the clas ‘excellent suitable’, 17.14% (13449.81 ha) in ‘good suitable’, 75.70% (59393.30 ha) in ‘medium suitable’ and 5.47% (4293.11 ha) in ‘poor suitable’ for sustainable intensive agriculture. Reading the results of the leaf analysis and their comparison with international standards for the nutrient content of apple leaves shows that for the elements phosphorus, calcium and magnesium, their concentration levels in the leaves are too satisfactory. On the other hand, the nutrient potassium has levels in the leaves below the desired limits compared to the required reference standards. A slight potassium deficiency is probably due to excessive nitrogen fertilization, as nitrogen-induced vegetative growth generates high potassium requirements. In general, leaf levels of N, P, K, Ca, and Mg are significantly correlated with their levels in the soil. However, foliar analysis is an effective tool for assessing the nutritional status of trees and for readjusting fertilizer requirements taking into account factors that can affect nutrient availability and root uptake.

5. Conclusion

The soils studied and sampled at different depths (0-60 cm) have a clay-loam texture. They are relatively poor in organic matter, the content of organic matter and major elements are not sufficient to optimize the optimum yield of the apple tree culture. The study shows a correlation between apple yield and nutrient content in the soil and leaves. The nutrient deficiencies observed can be corrected by the application of fertilization and adequate maintenance. However, further studies are needed to fully understand other variables such as the nature of rootstocks, the varieties cultivated and the method of irrigation which strongly influences apple tree production and yield. A prospect may be distant; indeed, it would be very interesting to generalize our paradigm to other sectors related to apple growing. As well as tackling research on phytosanitary techniques in order to move from the conventional vision to the integrated vision. Further research could also focus on other crops.

References

[1] R. Moussadek Impacts of conservation agriculture on the properties and productivity of vertisols in Central Morocco. Afrika focus. **25** (2) 147-151, 2012.

[2] N. Akesbi. Evolution and perspectives of agriculture Moroccan. In Report 50 years of human development and perspectives for 2025. Casablanca. Maghrebi. **85**-198, 2006.

[3] Agricultural credit in Morocco: Accessed on November 09, 2020. Available at: https://www.fellah-trade.com/fr/ plant-sector / technical-sheets / apple tree.

[4] P. Milleville, G. Serpantie. Agrarian dynamics and the question of the intensification of farming in the Sahelian and savanna zones of Africa. Accounts of the French Academy of Agriculture **80**, 149-161, 1994.

[5] B. Diwediga, K. Hounkpe, K. Wala, K. Batawila, T. Tatoni, Akpagana, K. Off-season agriculture on the banks of the oti and its tributaries. African Crop Sci. J. **20**, 613-624, 2012.

[6] R. Moussadek. Impacts of conservation agriculture on the properties and productivity of vertisols in Central Morocco. Afrika focus. **25**(2), 147-151, 2012.

[7] Regional Directorate of Agriculture of Beni Mellal-Khenifra, agricultural monograph of the Beni Mellal-Khenifra region. 2016. Accessed on March 26, 2020.

[8] P. Martin –Prevel, J. Gagnard, P. Gautier. Plant analysis in the control of the diet of temperate and tropical plants. 131, 1984.

[9] A. Finigue. Water resources in the direction of the Beni Mellal Atlas: management issues and local development. 2018. University of Angers Doctoral School DEGEST Angers Thesis. Accessed on March 26, 2020. Available on https://tel.archives-ouvertes.fr/tel-01708289/document.

[10] Ministry of Agriculture and Maritime Fisheries. Technology transfer in agriculture. 2008. Consulted on March 26, 2020. Available at https://www.agrimaroc.net/bulletins/btta_165.pdf.

[11] K. Aggelopoulou, D. Wulfsohn, S. Fountas, G. Nanos, T. Gemtos, S. Blackmore. Spatial variability of yield and quality in an apple orchard. Precision Agriculture **11**(5), 538–556, 2010.

[12] K. Aggelopoulou, S. Fountas, D. Pateras, G. Nanos, Gemtos, T. Soil spatial variability and site-specific fertilization maps in an apple orchard. Precision Agriculture **12**(1), 118–12, 2011.

[13] K. Pontikis, Applied Pomology. Stamoulis Publications, Athens, Greece, 74.

[14] R. Doucet. Climate and agricultural soils. ed. Berger, Eastman, Quebec. xv, 443, 2003.

[15] Tarn and Garonne. Chamber of Agriculture. Guide To Reasoned Fertilization In Fruit Arboriculture. 2019. Accessed March 27, 2020. Available at: [https://mag.inrameknes.info/? P = 1400.

[16] R. Razouk Contribution to the establishment of apple tree fertilization standards in the Middle Atlas. 2018. Accessed March 27, 2020. Available at: https://mag.inrameknes.info/?p=1400.

[17] W. Ennaji, A. Barakat, M. El Baghdadi, H. Oumenskou, M. Aadraoui, L. A. Karroum, & A. Hilali. GIS-based multi-criteria land suitability analysis for sustainable agriculture in the northeast area of Tadla plain (Morocco). Journal of Earth System Science, **127**(6), 79, 2018.

[18] A. Barakat, W. Ennaji, A. El Jazouli, R. Amediaz, & F. Touhami. Multivariate analysis and GIS-based soil suitability diagnosis for sustainable intensive agriculture in Beni-Moussa irrigated subperimeter (Tadla plain, Morocco). Modeling Earth Systems and Environment, **3**(1), 3, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1106-1111 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Investigation of the LoRa Transceiver in Conditions of Multipath Propagation of Radio Signals

Dmytro Kucherov[1,*], Andrei Berezkin[2], Volodymyr Nakonechnyi[3], Olha Sushchenko[4], Ihor Ogirko[5], Olha Ogirko[6], Ruslan Skrynkovskyy[7]

[1]*Department of Computerized Control Systems, National Aviation University, Kyiv, 03058, Ukraine*

[2]*Pukhov Institute for Modelling in Energy Engineering, National Academy of Science of Ukraine, Kyiv, 03164, Ukraine*

[3]*Department of Cyber Security and Information Protection, Taras Shevchenko National University, Kyiv, 01033, Ukraine*

[4]*Department of Aerospace Control Systems, Institute, National Aviation University, Kyiv, 03058, Ukraine*

[5]*Department of Mathematic, Kazimierz Pułaski University of Technology and Humanities in Radom, Radom, 26600, Poland*

[6]*Department of Information Technologies, Lviv State University of Internal Affairs, Lviv, 79020, Ukraine*

[7]*Department of Business Economy and Information Technology, Lviv University of Business and Law, Lviv, 79021, Ukraine*

ARTICLE INFO

Article history:
Received: 08 November, 2020
Accepted: 09 February, 2021
Online: 16 February, 2021

Keywords:
Radio communication
Underground propagation
LoRa module

ABSTRACT

The article presents some results of the research of the LoRa module. These modules can be the basis of possible IoT technologies are implementing, providing enough good range of receiving and transmitting messages. The SX1276 transceiver has been testing to determine the signal loss in the propagation channel. These experiments took in a highly-populated Kyiv district and one of the passageways of a Podbryantsevsky salt mine near the Solar town. The measured parameters are the maximum radio communication range in the mine, the signal-to-noise ratio, the number of bit errors and losses of the signal transmission. The data of the study we plan to use for the engineering of the radio-messaging networks based on LoRa radio modules.

1. Introduction

Multipath propagation is obtaining as a result of re-reflections of the signal source antenna beam from ground objects. Namely from buildings, a large number of rooms and corridors inside them, hills, vegetation, mine tunnels, etc., which create except the one direct some indirect directions of arrival of the beam. The receiving signal is an additive mixture of signals with different amplitudes, phases, and arrival directions. Multipath propagation of radio signals results in signal fading and frequency shift of the receiving signal.

Multipath propagation decreases the quality of the data reception. In the same circumstances, radar reception is resolved by introducing a small frequency additive into the low power oscillator of the frequency converter then receiving signal frequency deviation is synchronous with the transmitting signal frequency. The next way is the creation of additional reception channels. Another situation arises in technique for communicating, where noise has a strong effect, and the signal-to-noise ratio (SNR) is less than unity. A natural solution pointed problem is the complication of the receiving set by introducing additional receive channels and complex signal processing. However, this approach is not acceptable for the Internet of Things (IoT) technology. The peculiarity of the application of this technology in mines is the small size of antennas and equipment, also low power consumption.

Optimal filtering is one way to expand the SNR at the receiver output. However, in processing technology, this approach is not always suitable. The signal spectrum expands due to interference, and by increasing the acceptance threshold, and the signal may be lost. In this case, for improving the signal quality at the input of the matched filter, the noise protection is offered. Unluckily, well-known methods of suppression of narrow-band interference do not apply as well.

*Corresponding Author: Dmytro Kucherov, Email: d_kucherov@ukr.net

Due to its ability to maintain communications and lighting simultaneously, some authors consider the visible light communications (VLC) system like potential opportunity to implement effective communications in mines, which is basing on the phenomenon of light propagation. In this case, the quality of communication devices is highly dependent on the available channel propagation models, including long-term fading, which was critical for the project and assessing the operation of VLC systems. But, the basic features of the VLCs have not yet been enough explored. Also, dust and moisture block the propagation of light signals.

This paper is an extended version of our work [1] presented at the 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T) in which the research purpose was to study the features of LoRa (Long Range) signal processing in multipath propagation. The remaining part of the paper is organized as follows: Section 2 presents the works that provide an overview of the subject area; Section 3 devotes the general formulation of the problem; Section 4 briefly describes the features of the radio module when used in free space. Section 5 presents the results of module tests in multipath conditions, and conclusions from the studies are presented in Section 6.

2. Paper Review

Due to frequent disasters that occur in mining sites caused by landslides and rockslides of underground mine workings, the creation of a reliable system for communication in places of underground mining has recently become important.

In this case, the traditional solution is the use of radio communication systems. In the document [2], prepared in response to the MINER Act, an analysis of existing communication systems in different frequency ranges has been presented, and the basic requirements for existing radio communication facilities are formulating. Their functions are not only messaging but also tracking the actions of miners. Thus, radio communication is recognizing as the principal method of communication. In this respect, it is also worth noting that the work to establish the requirements for frequency ranges and radio communication range based on the study of the average attenuation characteristics of radio signals in mines of various types of minerals by this time was partially done (for example, see [3, 4]).

Reviews of communication facilities in underground mines have also presented in papers [5, 6], where were manifest the problems of wireless communication under the ground, associated with attenuation and loss of communication that is greater than when they have on the surface. The attenuation parameter for a frequency of 900 MHz has been refined in [7] and was concluding the possibility of using mobile communications in coal mines.

Recently, as an alternative to conventional wired and radio transmitting systems, researchers [8-11] are considering the possibility of using visible light as a means of communication for mine communication. So, in the article [8], based on this approach, a method is proposed for modeling the visible light propagation channel for two types of underground utilities: directly at the bottom hole and on the roadway, and the results of studies of the number of indicators spread signal in the visible light channel have presented. The main achievement of the authors of [9] is assessment loss as a function of distance.

In [10], the use of angular diversity receivers has been proposed that was focus on reducing inter-cell interference. It also considers new mathematical models of light distribution: pyramidal and hemi-dodecahedral. Numerical values for data rate, signal energy, and signal-to-noise-plus-interference ratio have also been presented, which can be useful for evaluating new solutions.

It is worth noting, the use of optics for communication has studied earlier. One of the earliest sources for time division multiple access (TDMA) broadcast communications dates back to 2005 [11]. Described here are a malfunction resistant system start and restart of the system with the dynamic self-configuration capability and a controlled internal diagnostic system.

The authors of [12, 13] propose the technology of detecting personnel and equipment in tunnels of underground manufactories by data transmission through light-emitting diodes and a photodetector; their approach also uses the trilateration method. So, in [12], an algorithm for determining the location of targeting objects is proposed. It is built based on data from several VLC systems located on the cover of an underground tunnel, as well as VLC mobile devices placed in the interests of ensuring the safety of miners, and the same problem is solved in three-dimensional space by authors [13].

To detect underground facilities, it is also possible to use ground penetration radar (GPR). So, in [14], it is proposed to divide the transmission paths into two types of signals: blocked and unblocked. Since the volume of blockaded signals is less than the estimated signal space, their routes are using in signal processing. A similar approach was proposed by the authors [15] based on the ULF (Ultra Low Frequency) radio channel for a mine specializing in the extraction of non-ferrous metals. In [16], GPR was used to assess soil moisture in the interests of agriculture. However, the range of propagation of electromagnetic waves underground is not long, does not exceed several meters; their use is advisable only in the interests of agriculture, which has been established by numerous studies and is confirming in [17 - 19]. The article [20] discusses the problems of robot navigation for coal mines.

This research continues the work presented in [21, 22].

3. Problem Statement

A multipath reception for a receiver based on LoRa technology is considering. In an open space, the module manufacturer guarantees the service of remote users at a distance of up to 20 km for a long life of the time, which makes the technology comparable with common Wi-Fi and LTE (Long-Term Evolution) [1].

The set of modules forms a system for picking reports from users and transmitting them to the server for further treatment and, maybe, answering queries. The messages exchanged between users; the service is providing by the open air at frequencies below GHz.

A message is a data package, which includes a preamble, a header and a payload, its size is dependent on the spreading factor

(S_f). Length of the package varied from 51 to 256 bits. The information transfer rate is between 22 and 27 kbps. The selection of this parameter depends upon the range, bandwidth and expansion factor. The signal that is emitted is a chirp signal, which is now assumed as the standard in the area of short-range wireless communication systems (IEEE 802.15.4a).

A distinguishing characteristic of signal receiving in multipath propagation is a significant deterioration in the signal quality, up to its complete loss. Examples of single radiation without interference and the total effect of four chirp signals without interference shown in Fig. 1, 2.

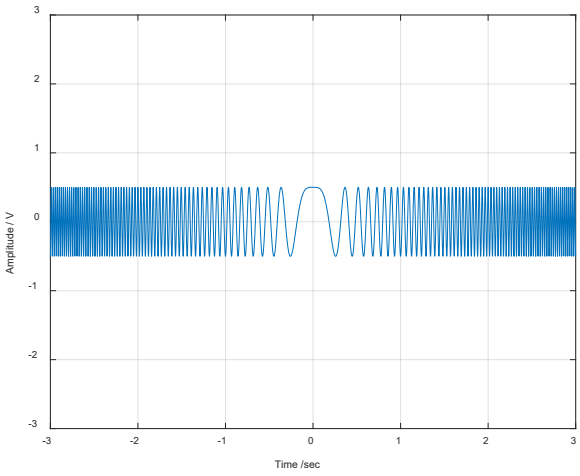


Figure 1: The one chirp signal

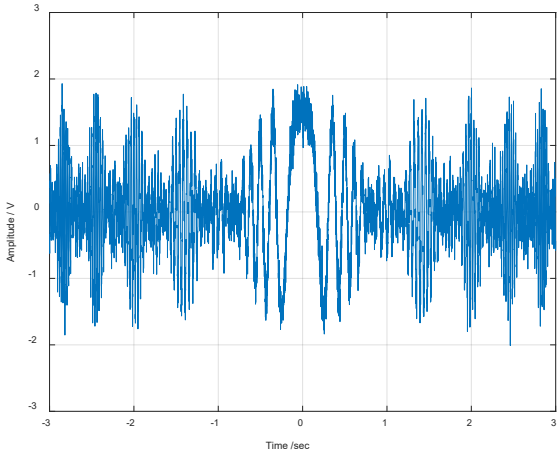


Figure 2: The four chirp signal with different phase shifts

A traditional study in the course of multipath reception is to clarify losses in the transmitter-receiver section and guarantee the receiving range [23]

$$R = \sqrt{\frac{P_t F(\theta) G_t A_r e^{-2\alpha d}}{4\pi P_r}} \quad (1)$$

where R is the radio communication range; P_t is the transmitter signal power; $F(\theta)$ is the antenna directivity in the θ direction, $F(0) = 1$; G_t is the transmitting antenna gain; A_r is the effective receiver area; P_r is the signal power at the receiving set input; α is the attenuation of the signal along the path of the beam propagation, taking into account the free space and propagation environment; d is the distance between two antennas.

Channel losses determine the budgeting of the wireless communication channel, which is calculated by the formula [23]:

$$P_r = P_t + G_t - L_s - L_{ch} - M \quad (2)$$

where G_t is the gain determined by the directivity of the antennas, L_s is the system's loss (for instance, in short antennas used in most remote devices), etc.; L_{ch} is the loss in the signal spillover channel, M is the attenuation signal is computed by empirical data. All values in the equation are expressing in dB.

If we know value R radio communication ranges in free space and the range in the loss conditions D, the value of losses in dB can be specified using the formula

$$L = 40 \lg\left(\frac{D}{R}\right) \quad (3)$$

SNR and bit error (BER) rates are auxiliary indicators of reception quality. The SNR is determining as follows

$$SNR = P_{signal} / P_{noise} = (a_{signal} / a_{noise})^2 \quad (4)$$

The purpose of the research is to assess the effectiveness of using LoRa modules in a confined space, where it is possible to observe multipath propagation.

4. Preliminary remarks

A typical LoRaWAN radio network integrates end devices into architecture type a star. Messages are underlying double encrypted to ensure security. Devices that have small power consumption are ensuring by using a short time for transferring in a comparatively long time interval when the terminal device is in a "sleep" regime. Information is transferring both asynchronous and synchronous modes. The information from the end node of the network is spreading through the gateway to the network server and then deliver to the application server. LoRa modules have better receiver sensitivity because it uses chirp signals, error correction, high transmit noise immunity and insensibility to frequency changes of resonators.

In experiments, we used a module based on the SX1276 microcircuit (Fig. 3), in which messages are transmitting at a frequency of 868 MHz. Other parameters of the transceiver based on the SX1276 microcircuit [24] are presented in Table 1.

Table 1: Some parameters of SX1276

Option	Value
Δf, kHz	20.8
S_f	8
BR, bps	1562
P_r, dBm	– 128

The channel budget, calculated by the formula (2), for a transmitter with a power of 100 mW in the case of ideal transmit-receive dipoles, for which the gain was chosen to be 2.1 dB and the minimum loss and signal attenuation turns out to be 130 dB.

The maximum range of radio communication in line-of-sight conditions for a lifting height of the receiving and transmitting antennas of 1 m under experiment conditions and normal refraction does not exceed 25 km.

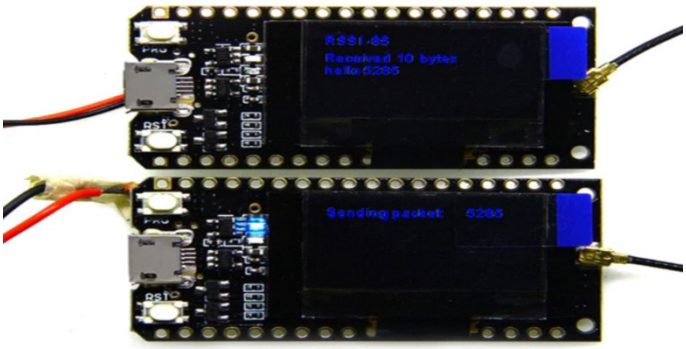


Figure 3: The module SX1276

The force of the received signal as a function of the range without interference [25, formula (5.15)] has shown in Fig. 4 with a blue line; the threshold sensitivity of the studied module is a black straight line. Crossing the lines gives the radio communication range corresponding to the real sensitivity of the receiving module, which turns out to be 4.8 km.

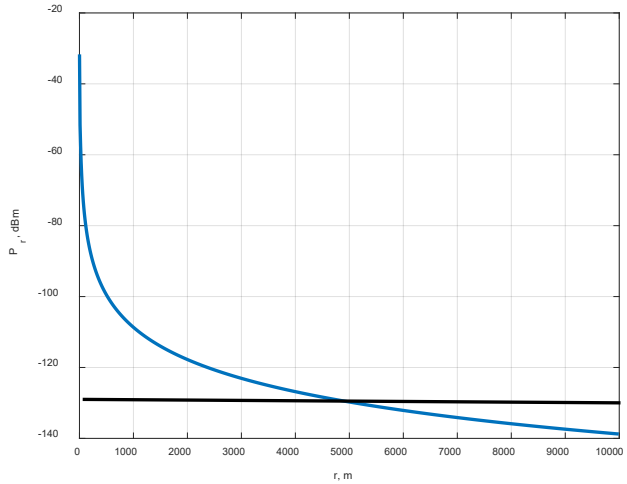


Figure 4: Estimation of the maximum radio communication range for a given sensitivity

5. Research results

The LoRa chipset study consisted of a sequence of surface experiments in a highly-populated city spot, where mobile communications and radio internet interfere with the receiving channel, and underground, in a mine where spurious radiation is absent. Ground tests based on a point-to-point scheme and a point-gateway with treatment on a remote server and underground tests were according to a point-to-point layout.

5.1. Interference radio

In the beginning, the transmitter installs on the last floor of a hard-wall building in Kyiv. A few Wi-Fi networks were deploying inside this building, on the roof of the building were also installed antennas of three operators' mobile communication. The indicator of the receiving device showed the message quality. At the experiment, the LoRa radio module was deposited from the 11th floor of the building to the first floor and to the downstairs located two floors below the first floor for creating multipath reception. The measurement framework reveals in Fig. 5. In Fig. 5, the abbreviations MS are used for the marking of the antenna of the stations of mobile operators, Tr is the transmitting module, Rc is the receiving device.

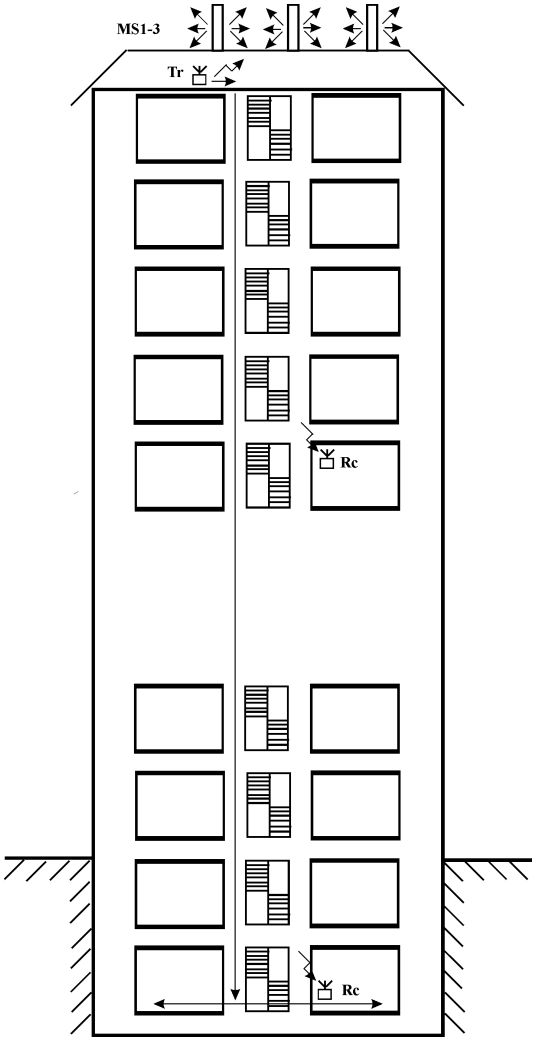


Figure 5: Layout of devices in the experiment

Rooftop antennas of mobile communications and Wi-Fi routers causing interference. These signals create a noise background, which chose as the "white". Crosstalk is causing by multiple reflections of beams inside the building from reinforced concrete structures. Using signal and noise level were fixed on every building floor, and the legibility of the text message was controlled as well. Additional documenting of the measurements has been carried out by a panoramic receiver, which chose as a reference. A initial report of the interference condition presents in Table 2.

The functioning of the transceiver in the "point-to-gateway" regime is fundamentally no different from the "point-to-point." The main difference is data processing, which carries out on a cloud server, and therefore the information has been slightly delayed (tens of milliseconds) compared to the point-to-point regime. In measurements, an auxiliary radio communication

channel was used by a mobile phone. The results of measurements of the BER and SNR for the LoRa radio module shows in Fig. 6, 7.

Table 2: Interference Condition

Type of sources	*Frequency, MHz*	*Power, W*	*Effect*
CDMA	800-900	2-60	strong
LTE	900-2400	2-60	middle
Wi-Fi	2400	0,1	weak
Other	400-5000	<0,01	weak

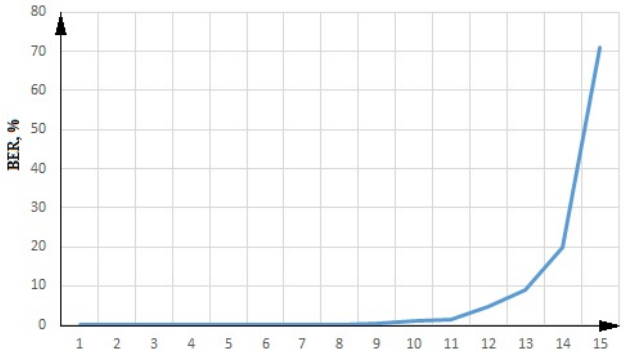


Figure 6: BER as a function of the number of the floor

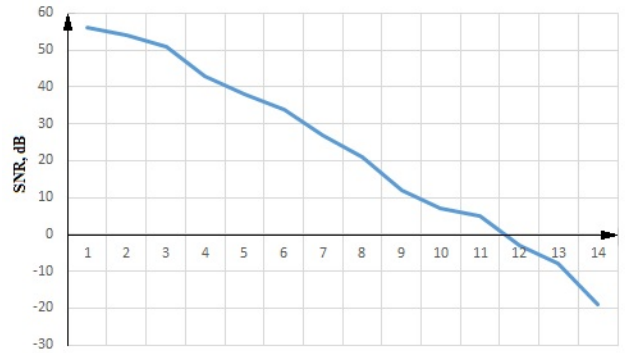


Figure 7: SNR as a function of the number of the floor.

By the experiment results, it has established that text messages receive on all floors and within 2 m from the staircase in the basement; the loss connection was at a distance of 8-10 m from the stairs.

5.2. Using the LoRa module in a mine

This subsection contains data on the results of measurements of the quality of message reception by the LoRa radio module at the conveyor slope of the Podbryantsevsky salt mine, located in the vicinity of the Soledar town.

The measurements were fulfilled in tunnels of the mine, in which are two parallel corridors connected by passages, Fig. 8. The experiment was carried out by step-by-step removal of the receiving set from a stationary transmitter with a fixed distance. The motion has been done from the initial point, corresponding to the transmitter location, clockwise around the salt layer 23 m wide. The transmitter has installed 100 m from the comer.

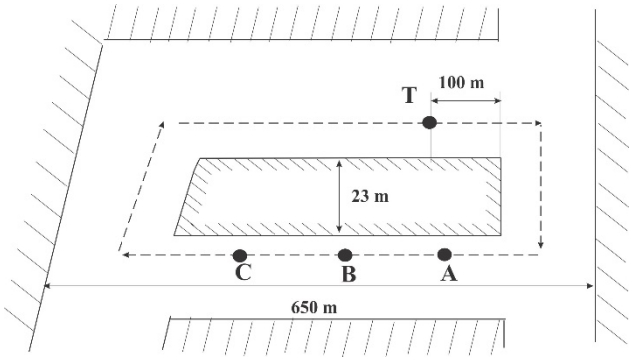


Figure 8: Section of the Podbryantsevsky formation mine, where measurements have been done

Point A in the diagram corresponds to a growth in the signal level around the comer, point B matches the signal maximum, and point C to a total signal loss. Also, the information in the mine was transmitted by voice; it was also repeated by the Motorola radio station functioning in the decimeter range and its output signal power of 1 W.

The quality of communication in LoRa assesses by the number of packets received and by voice; the last way was estimated according to the degree of intelligibility of radio phrases. The results of the measurements present in Table 3, where value *D* is the distance from the LoRa receiving set to the transmitter antenna (T); Received Signal Strength Indicator (RSSI) serves as an indicator of the signal level at the receiving set input; Packets is the number of transmitted packets; Quality of Receiving is a category parameter of reception. The last parameter was the estimated sign "+" in case of satisfying reception, and "-" otherwise.

Table 3: Report of Receiving Quality

D, m	*RSSI, dBm*	*Packets*	*Quality of Receiving*
100	-94	4	+
115	-96	4	+
129	-107	4	+
162 (A)	-98	4	+
187	-97	4	+
211	-96	4	+
230 (B)	-93	4	+
243	-101	4	+
263	-113	4	+
283	-124	-	±
303 (C)	-128	-	–

Based on the measurement data, a graph of the loss function *L*(*D*) is built, where *D* is the range from the signal source to the

measurement point, Fig. 9. The resonant character of the curve in Fig. 9 is a consequence of the superposition of beams resulting from multipath propagation. Signal power losses in the section of indirect signal propagation with a length of 150-230 m have an average value of 15 dB.

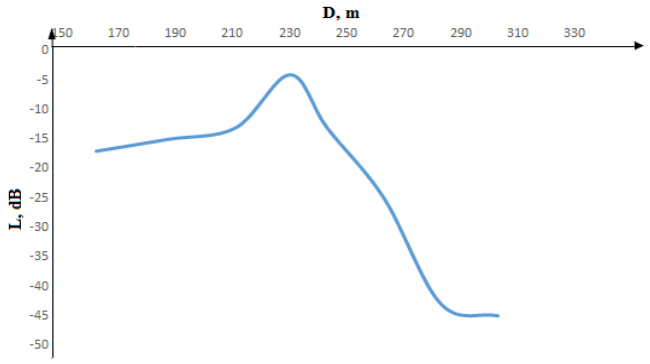


Figure 9: Graph of the function $L(D)$.

Analysis of various forms of communication in mine made it possible to establish that the range of voice communication, with due regard for phrasal awareness, no more than 30 m; the distance of radio communication no more than 100 m; the transmission range by LoRa reached 150 m.

6. Conclusions

The paper presents a study of the LoRa module on two models of multipath propagation with external interference in a city building and without them in a salt mine. In the course of the studies carried out in a high-rise building, it was found that for the LoRa transmitter with a decrease in the SNR by 70 dB, the BER deteriorates by only 15%. The transmission range of messages in the mine with indirect beam propagation (Rayleigh channel) exceeded the distance of conventional radio communications by 30%, voice communications by five times. The average loss at the frequency under consideration is 15 dB for a Rayleigh radio propagation channel, provided the messages are intelligible. These results are planning to be the basis for the choice of the LoRaWAN radio network topology as an element of mine communication.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors thank both the authorities of the Taras Shevchenko National University, and National Aviation University the especially leadership of the Faculty of Cybersecurity, Computer and Software Engineering for their support during the preparation of this paper. The authors also thank the anonymous reviewers, whose comments significantly improved the content of the paper.

References

[1] D. Kucherov; A. Berezkin; O. Sushchenko, "Features of Signal Processing by Means of LoRa Technology," in 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Kyiv, Ukraine, 19-24, 2019. https://doi.org/10.1109/PICST47496.2019.9061410

[2] "Basic Tutorial on Wireless Communication and Electronic Tracking: Technology Overview," https://www.cdc.gov/niosh/mining/content/emergencymanagementandresponse/commtracking/commtrackingtutorial1.html

[3] T.S. Cory, Propagation of EM Signals In Underground Mines, Rockwell International Commercial Telecommunications Group, 1977. https://www.cdc.gov/niosh/mining/works/coversheet331.html

[4] T.S. Cory, R.J. Mahany, Mining Publication: Propagation of EM Signals in Underground Metal/Non-Metal Mines, U.S. Bureau of Mines, 1981. https://www.cdc.gov/niosh/mining/works/coversheet828.html

[5] J.N. Murphy, H.E. Parkinson, "Underground Mine Communications," P IEEE, **66** (1), 26–50, 1978. https://www.cdc.gov/NIOSH/mining/UserFiles/works/pdfs/umc.pdf

[6] S. Yarkan, S. Guzelgoz, H. Arslan, R.R. Murphy, "Underground Mine Communications: A Survey," IEEE COMMUN SURV TUT, **11**(3), 125–142, 2009. https://doi.org/10.1109/SURV.2009.090309

[7] Y.P. Zhang; G.X. Zheng; J.H. Sheng, "Radio propagation at 900 MHz in underground coal mines," IEEE T ANTENN PROPAG, **49**(5), 757–762, 2001. https://doi.org/10.1109/8.929630

[8] J. Wang, A. Al-Kinani, W. Zhang, and C.-X. Wang, "A New VLC Channel Model for Underground Mining Environments," in 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 2017. https://doi.org/10.1109/IWCMC.2017.7986613

[9] J. Wang, A. Al-Kinani, J. Sun, W. Zhang, and C.-X. Wang, "A Path Loss Channel Model for Visible Light Communications in Underground Mines," in 2017 IEEE/CIC International Conference on Communications in China (ICCC), Qingdao, China, 2017. https://doi.org/10.1109/ICCChina.2017.8330479

[10] P.P. Játiva, M.R. Cañizares, C.A. Azurdia-Meza, D. Zabala-Blanco, A.D. Firoozabadi, F. Seguel, S. Montejo-Sánchez and I. Soto, "Interference Mitigation for Visible Light Communications in Underground Mines Using Angle Diversity Receivers," Sensors, **20** (2), 367, 1-29, 2020. https://doi.org/10.3390/s20020367

[11] T.-P. Wilfredo, Malekpour, Mahyar R., Miner, Paul S., "ROBUS-2: A Fault-Tolerant Broadcast Communication System," in NASA technical report, 2005, p. 188, https://ntrs.nasa.gov/citations/20050158766

[12] D. Iturralde, F. Seguel, I. Soto, C. Azurdia, S. Khan, "A new VLC System for Localization in Underground Mining Tunnels," IEEE LAT AM T, **15**(4), 581-587, 2017. https://doi.org/10.1109/TLA.2017.7896341

[13] A.D. Firoozabadi, C. Azurdia-Meza, I. Soto, F. Seguel, N. Krommenacker, D. Iturralde, P. Charpentier and D. Zabala-Blanco, "A Novel Frequency Domain Visible Light Communication (VLC) Three-Dimensional Trilateration System for Localization in Underground Mining," APPL SCI, **9**(7), 1488; 1-15, 2019. https://doi.org/10.3390/app9071488

[14] S. Suherman, A.H. Rambe, E. Wijayanto and N. Mubarakah, "Received signal consideration for the through the earth radiobased underground object detection," J PHYS CONF SER, **1566**, 1-7, 2020. https://doi.org/10.1088/1742-6596/1566/1/012003

[15] D.S. Kudinov, O.A. Maikov, E.A. Kokhonkova and V.S. Potylitsyn, "Development of a mathematical model and analysis of the receiver for emergency alerts in a mine," IOP CONF SER-MAT SCI, **862**, 1–7, 2020. https://doi.org/10.1088/1757-899X/862/3/032111

[16] A. Salam, A. Ahmad, "Underground Soil Sensing Using Subsurface Radio Wave Propagation," Purdue e-Pubs, 255–260, 2019. https://docs.lib.purdue.edu/cit_articles/19

[17] I.F. Akyildiz, E.P. Stuntebeck, "Wireless Underground Sensor Networks: Research Challenges," Ad Hoc Networks, **4**, 669–686, 2006, https://doi.org/10.1016/j.adhoc.2006.04.003

[18] I.F. Akyildiz, Z. Sun, M.C. Vuran, "Signal Propagation Techniques for Wireless Underground Communication Networks," PHYS COMMUN-AMST, **2**(3), 167–183, 2009. https://doi.org/10.1016/j.phycom.2009.03.004

[19] N. Saeed, M.-S. Alouini, T.Y. Al-Naffouri, "Towards the Internet of Underground Things: A Systematic Survey," IEEE COMMUN SURV TUT, **21**(4), 2019, 3443–3466. https://doi.org/10.1109/COMST.2019.2934365

[20] X. Chen, R. Liu, S. Zhang, H. Zeng, X. Yang, H. Deng, "Development of millimeter wave radar imaging and SLAM in underground coal mine environment," J CHINA COAL SOC, **45**(6), 2182-2192, 2020. https://doi.org/10.13225/j.cnki.jccs.ZN20.0316

[21] D. Kucherov, A. Berezkin, L. Onikienko, V. Nakonechnyi, "Processing Signals in The Receiving Channel for the Lora System," Data-centric business and applications. ICT Systems-Theory, Radio-Electronics, Information Technologies and Cybersecurity, **5**, 423–445. https://doi.org/10.1007/978-3-030-43070-2_19

[22] D. Kucherov, A. Berezkin, "Identification approach to determining of radio signal frequency," in Int. Conf. Antenna Theory and Techniques (ICATT), Kyiv, Ukraine, 2017. https://doi.org/10.1109/ICATT.2017.7972668

[23] B. Sklar, Digital Communications Fundamentals and Applications, Second Ed, Prentice Hall PTR, USA, 2002.

[24] SEMTECH SX1276. https://pdf1.alldatasheet.com/datasheet-pdf/view/800239/SEMTECH/SX1276.html

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1109-1124 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Multi-Criteria Decision Analysis Coupled with GIS and Remote Sensing Techniques for Delineating Suitable Artificial Aquifer Recharge Sites in Tafilalet Plain (Morocco)

Aicha Ousrhire[*,1], Hassane Oulidi Jarar[2], Abdessamad Ghafiri[1]

[1]*Laboratory of geomatics applied to geosciences and the environment, Department of geology, Hassan II university, Casablanca, 20660, Morocco*

[2]*SGEO research team, Department of mathematics, computer science and geomatics, Hassania school of public works, Casablanca, 20230 ,Morocco*

ARTICLE INFO

Article history:
Received: 12 October, 2020
Accepted: 17 November, 2020
Online: 14 December, 2020

Keywords:
AHP
Artificial aquifer recharge
Aquifer depletion
GIS
Remote sensing
Tafilalet plain
Morocco

ABSTRACT

Despite that groundwater is an important and vital water resource, it is not well managed; depletion of aquifers around the world due to overexploitation is of serious concern especially in arid regions where the situation is much more alarming. Tafilalet plain in Morocco which belongs to this type of environment is certainly no exception and is viewing its groundwater disappearing. Artificial aquifer recharge (AAR) is found to be appropriate to such an urgent issue. Thus, the objective of this paper is the exploration of suitable sites to process Artificial aquifer recharge in Tafilalet plain by the joint use of remote sensing (RS), geographic information systems (GIS), and the analytic hierarchy process (AHP) method. For doing so, eight parameters were considered as groundwater influencing parameters such as slope, soil, geology, land Cover/land Use, depth to the water table, aquifer transmissivity, electrical conductivity and drainage density. The laters were integrated and processed in a GIS, their thematic layers were created, and their relative weights were generated by The AHP method based on their significance in recharging the aquifer. Afterward, thematic layers were reclassified and assigned their weights, so the GIS overlay tool was used for inferring artificial groundwater recharge potential regions in the study area. 47 % were identified as suitable while only 12 % were identified as unsuitable. Such studies facilitate groundwater management for stakeholders and water managers because important decisions may be taken in a record time which will preserve water resources and prevent them from being in an alarming situation.

1. Introduction

Groundwater is a valuable natural resource and constitutes a significant freshwater reservoir as it accounts for 30 % of the earth's freshwater while surface water accounts for just 0.3 % [1, 2]. Besides, groundwater offers a multitude of benefits in different forms. In fact, groundwater meets the need for drinking water to relatively half of the world's population [3, 4] and around 38 % of the world's irrigated lands are supplied by groundwater [5], not to mention its contribution to the industry sector and its paramount role for the environment. Nevertheless, there is growing evidence that groundwater is under threat of numerous difficulties such as climate change impacts, contamination, pollution in addition to depletion [6–10]. Actually, many aquifers throughout the world suffer from groundwater depletion [11–17] and this situation is much more serious in arid regions since groundwater constitutes their main source of water supply. Mismanagement and namely over-exploitation are considered as its major cause alongside many other factors. In fact, rapid population growth, agricultural and industrial activities that require a considerable availability of water together with climatic conditions notably of arid environments consisting of high evapotranspiration rates, low precipitation, and low vegetation cover are placing greater pressure on limited aquifer resources which has led to their intensive extraction in many parts of the world [9, 12, 17–25]. Moreover, If this abstraction is greater than the natural groundwater recharge, it is referred to as an overdraft; so, groundwater levels will drop sharply over time, triggering depletion [8, 9, 14, 17, 25–30], which can make the situation complicated as the replenishment of the aquifers

*Corresponding Author: Aicha Ousrhire, Email: aicha.ousrhire@gmail.com

to the original levels takes a lot of time [3] and it is more complicated in cases of fossil or compacting aquifers because of their irreversible and non-renewable character [14, 31]. In such a way, countless negative impacts and many political, social, environmental, and economic issues, such as conflicts, land subsidence, degradation of groundwater dependent ecosystems, deterioration of water quality, aquifers drying and obviously water scarcity will take place particularly in countries which are groundwater dependent [10, 11, 14, 23–25, 27, 28, 32–39].

The above situation has been a major point of concern of water researchers, water managers, and water decision-makers, seeing that in many parts of the world a lot of measures have been taken to properly manage their groundwater resources and to prevent the worsening of the situation. The implementation of monitoring networks to control groundwater levels, the establishment of groundwater laws, and the setting up of exploitation rules for regulating its use, in addition to the organization of awareness-raising companies to urge on the importance of water and also education on water conservation are a good illustration of the undertaken measures [9,37,38,40,41], but even so, the decline of groundwater resources levels has been noticed and is still ongoing in many places around the globe. This is especially due to the violations of the established rules and the intensive mining of underground water notably in case of droughts. Added to that, the natural recharge of aquifers by streams, precipitation, rivers, pond seepage, reservoirs, irrigation canals, and lakes has become overwhelmed by the intensity of groundwater withdrawals [25] and has become difficult to occur especially in urbanized areas [42], because the presence of impervious surfaces such as roads, sidewalks, driveways and parking lots has led to many modifications to natural processes like impeding the infiltration of water. Thereby, the natural recharge of aquifers has proved to be insufficient to fill the quantity of water coming out of the aquifers in many parts of the world. Given these trends, water researchers, water managers, and water decision-makers have been prompted to think about solutions and innovative scientific analyses in order to bring better and sound management strategies that may address this groundwater issue and guarantee its sustainable use. If decreasing discharge and regulating abstraction has not brought important results, then this depletion can only be alleviated by increasing and boosting recharge.

Thus, artificial or managed aquifer recharge is a promising solution to this problem. Actually, artificial aquifer recharge (AAR) or managed aquifer recharge (MAR) is defined as an intentional human intervention arranged to enhance groundwater resources through processes implemented to heighten and improve the replenishment of depleted aquifers or depleting aquifers [22, 43–47]. The (AAR) has many gains and benefits, the major ones are the restoration of groundwater balance and the revival of the overexploited aquifers [11, 19, 25, 39, 48, 49], the other ones are enhancement and raising the water table which will cut down its operating costs, improving groundwater quality and prevention of seawater intrusion in the case of coastal aquifers [28, 46, 49, 50], alleviating floods adverse impacts given that during strong storms aquifers can be managed to play the role of a water storage reservoir which will mitigate the magnitude of flooding [22, 45, 51–53], besides (AAR) can limit the loss of land covered with surface water storage structures and limit the loss of water by evaporation in dams, ponds, and other surface reservoirs by storing it in aquifers [12, 39]. In order to artificially recharge aquifers, treated wastewater, floodwater, or water pumped from another aquifer can be used and redirected across the land surface. So, a variety of ways and systems have been developed such as infiltration basins, canals, ponds, ditches, sprinkler systems, rills, and irrigation furrows where the surface water is put on and subsequently infiltrates into the soil and moves downward to recharge aquifers or simply artificial recharge can be achieved by injecting water directly into the aquifer through injection wells especially in case of confined aquifers and each of these ways is implemented according to specific conditions. Applying the (AAR) needs some criteria to be met notably, the presence of highly permeable geological formations allowing percolation of recharge water, endowed with high capacity of water storage (unsaturated zone thickness should be deep enough to increase the piezometric level) and have high transmissivity. Moreover, the presence of a water source ensuring volumes of water needed for the recharge such as dams, stormwater runoff, wastewater treatment station, and desalination station besides, the water intended for artificial groundwater recharge as well as those of receiving aquifer must be of good quality and meet a certain number of standards and also good knowledge of the climatic characteristics of the region is essential in (AAR) projects (temperature, evaporation, rainfall, frequency of floods, periods of drought, snowmelt, etc.). Consequently, performing the (AAR) requires a suitable place whose selection depends on solid scientific analysis. Actually, selecting optimal sites for (AAR) involves integrating several parameters proved to have an influence on the infiltration of water to aquifers. To do this, a combination of geospatial technologies and multi-criteria decision analysis (MCDA) is used in this study.

Tafilalet plain was chosen as the study area in the present study as it is one of the Moroccan regions calling for urgent intervention and requiring important measures to be taken at all levels, including the environmental one; it suffers from enormous environmental difficulties mainly scarcity and poor water quality, due essentially to long periods of drought that has affected the region in recent years, combined to severe nature of the climate which makes groundwater resources the main source of water supply in low-flow periods. In 1967, the Hassan Addakhil dam was constructed in the north of Errachidia city and located at a distance of about 80 km from the Tafilalet plain, it was built on wadi Ziz in order to regulate floodwaters and, simultaneously, it serves a variety of goals, ranging from irrigation to water supply for domestic and industrial uses and remarkably it helps to establish a certain balance of groundwater of Tafilalet aquifer both quantitatively and qualitatively, but in front of considerable population growth and an agrarian-based economy resulting in increasing the number of wells notably prived ones, the exploitation of groundwater resources is proceeding at a very fast pace which has lead to their overdraft, so the implementation of artificial aquifer recharge in this region is essential to recover its overexploited groundwater.

Thus, the prime focus of the current paper is to delineate the artificial groundwater recharge potential regions in Tafilalet plain (Morocco), using the analytic hierarchy process (AHP) method which is a forme of Multiple-criteria decision analysis (MCDA) and geospatial technologies represented here by remote sensing (RS) and geographic information systems (GIS).

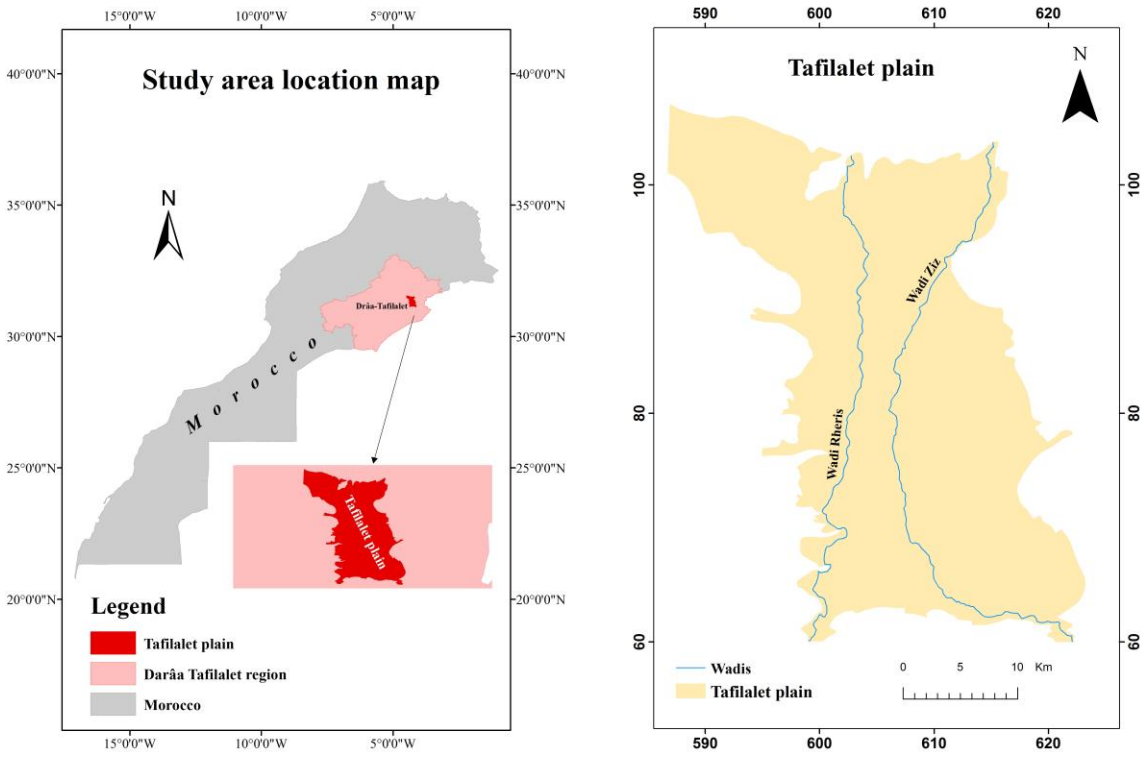


Figure 1: Study area location map

After this introduction, the following parts are organized in five sections, the first one is an overview of the study area, the second one is a presentation of the materials and definition of the methods used, the third one is a description of the major results, the fourth one is a discussion about the outcomes and the last one is a conclusion of the main ideas of this study.

2. Study Area

Tafilalet plain belongs to the pre-Saharan region located in the extreme southeast of the kingdom of Morocco, particularly in the Drâa-Tafilalet region, at an altitude of 700 to 800 m, it lies in between the coordinates of projection system Lambert zone Morocco (Zone I) X = 580000 m to X = 630000 m and Y = 50000 m to Y = 120000 m (Figure 1), covers an area of about 700 km², and the majority of its population are rural. Besides, Tafilalet plain is part of the quaternary basin of Tafilalet; it is an erosional depression arising out of the secondary and tertiary coverage clearing during the quaternary era, the bottom of this depression was, afterward, covered by the alluvium of wadis Ziz and Rheris. In fact, it constitutes the zone of gathering and accumulation of alluvial deposits, conveyed by surges of the Ziz and Rheris wadis and their tributaries. Moreover, Tafilalet plain extends over three different geological formations, it is enclosed by the High Atlas in the north, it crosses the pre-Africain furrow and is surrounded at the south by the old massif of the Anti-Atlas domain as well as overlying an impermeable Palaeozoic and pre-Cambrian substratum.

Concerning the climate, Tafilalet plain is dominated by an arid to semi-arid climate, typified by a great rainfall scarcity as the average annual precipitation barely reaches 70 mm all over the area, this is mainly due to the presence of Atlas reliefs in the north which constitute a barrier to oceanic influences. The climate is also characterized by a great irregularity of the temperature, rising to very high values in summer (about 40 degree C ° especially in July) and dropping to very low values in winter (surroundings 0 degree C ° especially in January), this great difference in temperature brings out two seasons: a hot season from May to October and a cold season from November to April so, in the Tafilalet plain, The dry periods last almost all the year with extremely hot summers and highly cold winters which is mainly due to its openness on the Sahara. In addition, The mean annual evapotranspiration is close to 500 mm which is higher than the mean annual rainfall, this means that Tafilalet plain suffers from water deficit all over the year. Adding to that, the dominant winds are commonly known as Chergui, they are characterized by their hotness and dryness and they blow from the southeast to the northeast.

Regarding the water resources, Tafilalet plain is, on the one hand, underlain by confined and unconfined aquifers, the north of the region encloses the infra Cenomanian aquifer which is a totally confined aquifer, whereas the entire area is almost occupied by an unconfined, shallow and very extensive aquifer named as "Tafilalet aquifer" which is the aquifer targeted in this study. In fact, this aquifer is not very thick and the water below the ground surface is 2 m to more than 30 m in some places, its shape matches the slightly convex shape of the plain and its slope is very weak. Consequently, this makes the water of this aquifer sensitive to hydrological and atmospheric conditions such as pollution and evapotranspiration. The spatial distribution of Tafilalet aquifer transmissivities indicates dispersed values ranging from 10^{-2} m²/s

to 10^{-4} m²/s and groundwater flows mainly from the north-east to south-west, south, and south-east characterized by an average hydraulic gradient that declines from upstream to downstream [54]. From a lithological point of view, the base of the aquifer is constituted by conglomerates and lacustrine limestones including gravelly levels and at the top of the aquifer, there are sands, gravel, and pebbles, covered with unevenly distributed powerful silts. Moreover, the main sources of groundwater recharge of the Tafilalet aquifer are essentially from the High Atlas mountains surrounding the plain in the north, from surface runoff including rivers, mainly Ziz and Rheris wadis, as well as irrigation water which infiltrates to the aquifer, and without forgetting precipitation and dams contribution. Conversely, the Tafilalet aquifer discharges mostly by evapotranspiration, by extraction from wells, and by Khettarats and springs. On the other hand, Tafilalet plain is crossed by many ephemeral wadis, the major ones are wadi Ziz in the center, wadi Rheris in the west. Besides, the region holds some springs resulting from the emergence of groundwater resources to the surface, such as Aïn Dolla-Attrouss and Aïn Chorfa.

Furthermore, the heart of the Tafilalet plain is made up of a very large palm grove, which occupies most of its surface so, a great part of the plain is devoted to agriculture, especially irrigated agriculture, making it the main economic activity in this region and is the main source of income for a large population. Because of climate aridity, the majority of crops are water-demanding which implies a strong mobilization of water notably during the months of June and July and due to the scarcity of surface waters in the region, growers are turning to groundwater exploitation to fill this water gap and meet their water needs. This has made the groundwater balance in aquifers of Tafilalet plain quickly in deficit.

3. Materials and Methods

The overall process followed in this study of the artificial aquifer recharge site selection has been schematically presented in (Figure 2).

3.1. Influencing factors Selection reason

The prediction of suitable sites for artificial aquifer recharge depends on a number of parameters that have an influence in either a direct or indirect way on the process of aquifer recharge. The selection of these factors was based firstly on a deep review of studies of artificial recharge carried out in other areas in the world, secondly on a complete understanding of the characteristics and specificity of the study area "Tafilalet Plain" and thirdly on the point of view of experts. Thus, eight factors were revealed to have control over the recharge of the Tafilalet aquifer, including slope, soil, geology, land cover/ land use, depth to the water table, aquifer transmissivity, electrical conductivity and drainage density as mentioned in (Table 1).

- Slope: The slope is one of the essential factors influencing the artificial aquifer recharge because the slope directly affects the penetration of water into the soil. Actually, the flatter the ground, the more time the water has to sink in, unlike the steep slopes where the water runs over the ground rather than entering it [25, 55–57]. However, the penetration of water is not guaranteed even if the slope is suitable because it depends on the soil's nature.
- Soil: Soil is a factor of significant importance in aquifer recharge because soil's nature determines the percolation ability of the soil as well as the water infiltration quantity and eventually governs the recharging capacity of the bedrock [58]. Moreover, soil acts as a purificator of water [29] that is why it should contain high rates of loam and sand to avoid groundwater contamination and to ensure a high infiltration rate.
- Geology: Taking into consideration geology in artificial aquifer recharge site selection is of major importance because geological formations which must be selected should enjoy high permeability so that water can flow seamlessly downward towards the groundwater besides, they should have adequate transmissivity and high water storage capacity.
- Depth to the water table: Depth to the water table or unsaturated zone thickness is the distance from the ground surface to the groundwater, it is another major factor to consider in artificial groundwater recharge site selection since, the deeper the water table, the greater the amount of groundwater storage, and the more efficient the artificial aquifer recharge will be Moreover, the deeper the water table, the more purified the water will be, because it will pass through a very large layer of soil, which is called soil aquifer treatment (SAT) [19, 29]. Acceptable depth is about 3 m consequently, sites, where the depth is less than 3 m, should not be considered [59].
- Transmissivity: transmissivity is an important hydrodynamic parameter of an aquifer which makes its consideration paramount in the selection of suitable sites for artificial groundwater recharge. Actually, transmissivity determines the rate of water flowing per unit width of the fully saturated thickness of an aquifer under a unit hydraulic gradient [25, 60–62] so, to avoid the excessive formation of groundwater mounds that interfere with the infiltration process, aquifers should be unconfined and sufficiently transmissive to allow the lateral flow of infiltrated water well away from the recharge zone [49].
- Drainage density: Drainage density is defined as the total length of streams of all orders per unit area and provides an indication of the watershed draining state. Drainage density was heeded as another factor of artificial aquifer recharge site selection because it indirectly signals the suitability of a site. In fact, high drainage density allows less penetration and more surface runoff and areas with great infiltration capacity are endowed with low drainage density [25, 39, 51, 58, 63–68] hence appropriate to artificial aquifer recharge.
- Electrical Conductivity: Groundwater quality is a paramount parameter that should be investigated in the artificial aquifer recharge projects because it is a determinant factor of the future usage of that water. Usually, the water intended for artificial aquifer recharge is of good quality but when it penetrates across a geological formation and mingles with different dissolved materials in it, its quality may get altered. Electrical conductivity (EC) is considered in this study because its variations have similar trends to those of total dissolved solids (TDS) as shown in the following formula :

$$EC = TDS / 0.65 \quad (1)$$

In fact, the higher the electrical conductivity, the poorer the water quality and vice versa [69]. Thus, the quality of groundwater will be assessed through the EC in the artificial aquifer recharge site selection.

- Land Cover / Land Use (LCLU) : (LCLU) is one of the important aquifer recharge influencing factors, it affects both the infiltration capacity and the water restraining of the land. Actually, areas with an important density of vegetation cover are appropriate for undertaking artificial aquifer recharge projects because the denser the vegetation cover, the more the runoff is slowed down and the more water infiltration takes place and also the less evapotranspiration and soil surface erosion will occur. In addition, organisms present in the soil and plant roots form openings in the subsoil, allowing water to infiltrate easily, unlike in areas without vegetation and in built-up areas, where the soil is mostly impervious, which favors runoff and prevents infiltration.

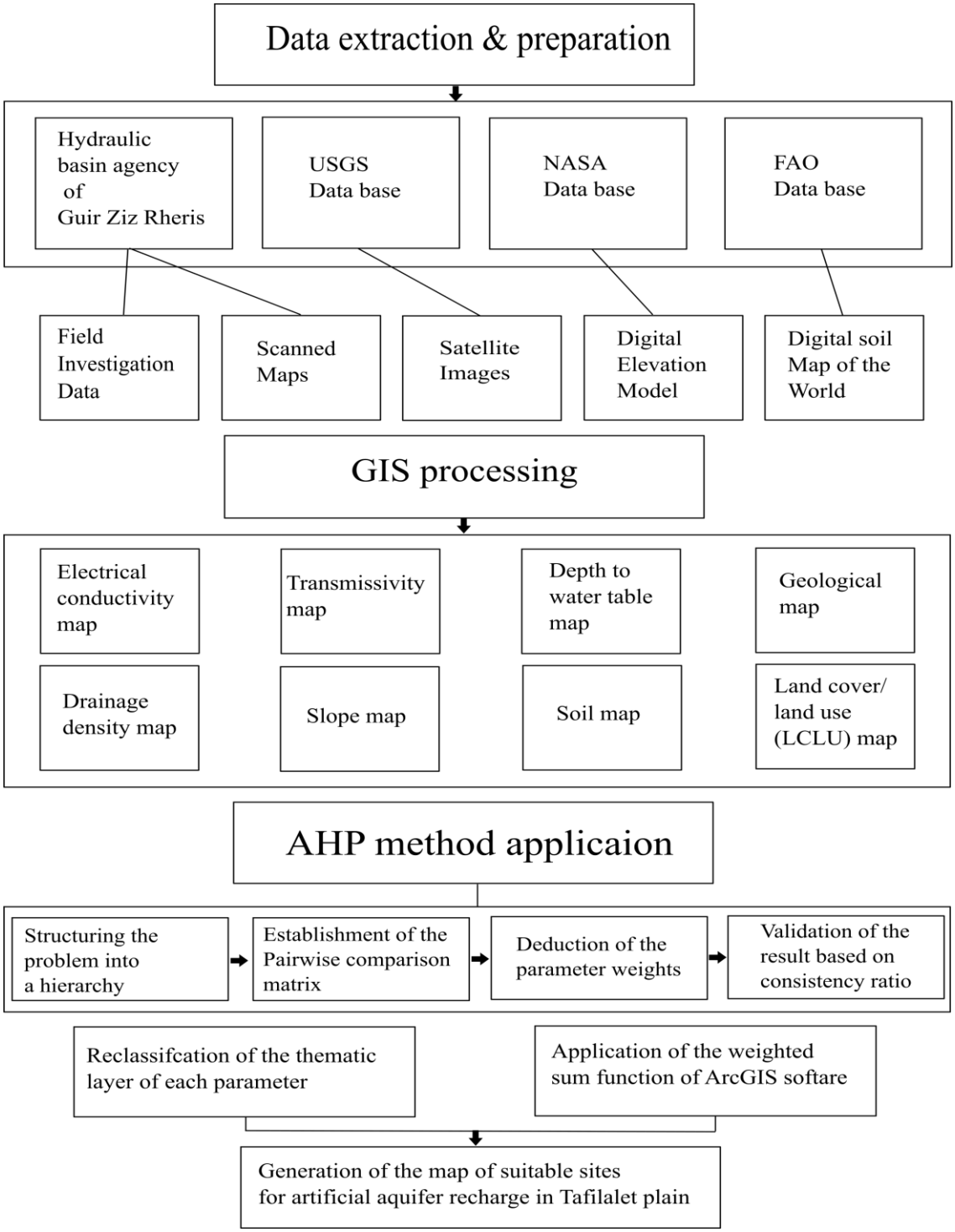


Figure 2: Methodology workflow

Table 1: Data Extraction and Preparation

Data	Origin format	Data source	Acquisition mode	Output
Slope	ASTER DEM at 30 m	National Aeronautics and space administration (NASA) website	Extraction	Slope map of Tafilalet plain
Soil	Digital soil map of the world	Food and Agriculture Organization of the United Nations (FAO) website	Extraction	Soil map of Tafilalet plain
Geological data	geological map of Tafilalet-Taouz (1/200000 scale)	Hydraulic basin agency of Guir Ziz Rheris	Digitalization	Geological map of Tafilalet plain
land Cover/land Use (LCLU)	satellite image (Landsat 8)	United states geological survey (USGS) institute website	Classification	LCLU map of Tafilalet plain
Depth to the water table	Excel sheet of field investigation data	hydraulic basin agency of Guir Ziz Rheris	Interpolation	Depth to the water table map of Tafilalet aquifer
Transmissivity	Excel sheet of field investigation data	hydraulic basin agency of Guir Ziz Rheris	Interpolation	Transmissivity map of Tafilalet aquifer
Electrical conductivity	Excel sheet of field investigation data	hydraulic basin agency of Guir Ziz Rheris	Interpolation	Electrical conductivity map of Tafilalet aquifer
Drainage density	ASTER DEM at 30 m	National Aeronautics and space administration (NASA) website	Extraction of the hydrographic network then application of Line Density	Drainage density map of Tafilalet plain

3.2. Data extraction, preparation, and GIS processing:

Given the interdisciplinary nature of groundwater, using remote sensing (RS) and geographic information systems (GIS) techniques are required because RS is a technology that provides important and information-rich geospatial data, especially for unreachable and very large areas. Moreover, GIS is commonly known for its usefulness to handle, analyze, integrate, and process various thematic geospatial data.

Therefore, in order to get the geographic data related to the above-mentioned factors, various data sources were used: FAO database, NASA database, USGS database, and other data were given by the hydraulic basin agency of Guir-Ziz-Rheris as detailed in Table 1. Furthermore, in order to process and to analyze those data, the geographic data processing functions of the ArcGIS software were used. In fact, as indicated in Table 1, slope and hydrographic network were extracted from ASTER DEM at 30 m, then the drainage density was formed by applying the Line Density function to the hydrographic network, the soil was extracted from the digital soil map of the world taken from the FAO database, geology was digitized from the geological map of Tafilalet-Taouz (1/200000 scale), electrical conductivity, transmissivity and depth to the water table were taken from field surveys carried out with the hydraulic basin agency of Guir-Ziz-Rheris and land cover/land use (LCLU) was classified from satellite image (Landsat 8) drawn from USGS database. Thus, a geospatial database was constructed and filled with the aforementioned data. Afterward, using ArcGIS software a thematic layer of each factor was created (Figure 3) and each thematic layer has been converted to raster format and reclassified into five classes pertaining to its importance for artificial groundwater recharge (Figure 4), each class is translated to a category and each category has assigned a score as it is detailed in Table 2. Actually, the five categories adopted are very good, good, moderate, poor, and very poor in accordance with a specific evaluation scale (Table 3).

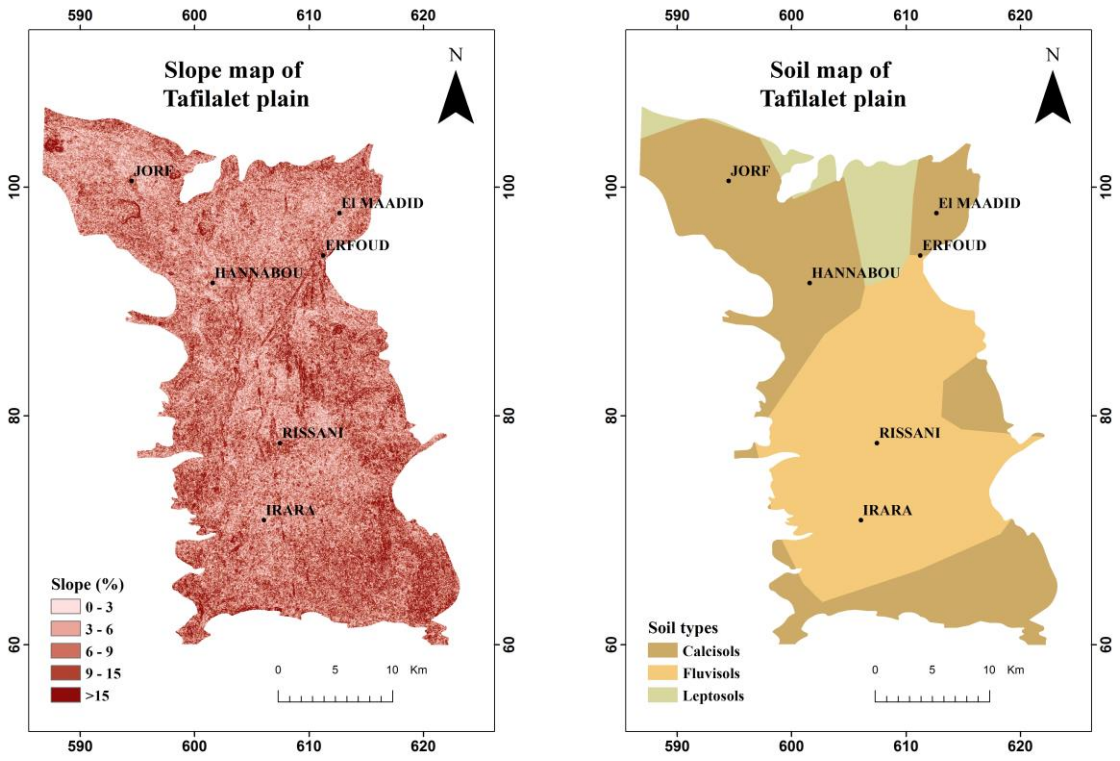
3.3. AHP method application

After reclassification of all the thematic layers, appropriate weights were generated by the analytic hierarchy process method (AHP).

Indeed, AHP is a multi-criteria decision analysis method and a mathematically-based technique [70], the mathematical underpinnings and the philosophy of which were developed in the seventies of the last century by the mathematician Thomas L. Saaty [71], AHP involves breaking down a convoluted multi-criteria decision problem into a hierarchy also called decision modeling in order to organize and better understand the problem and analyze the decision to be made.

a) Slope map

b) Soil map



c) Geological map

d) LCLU map

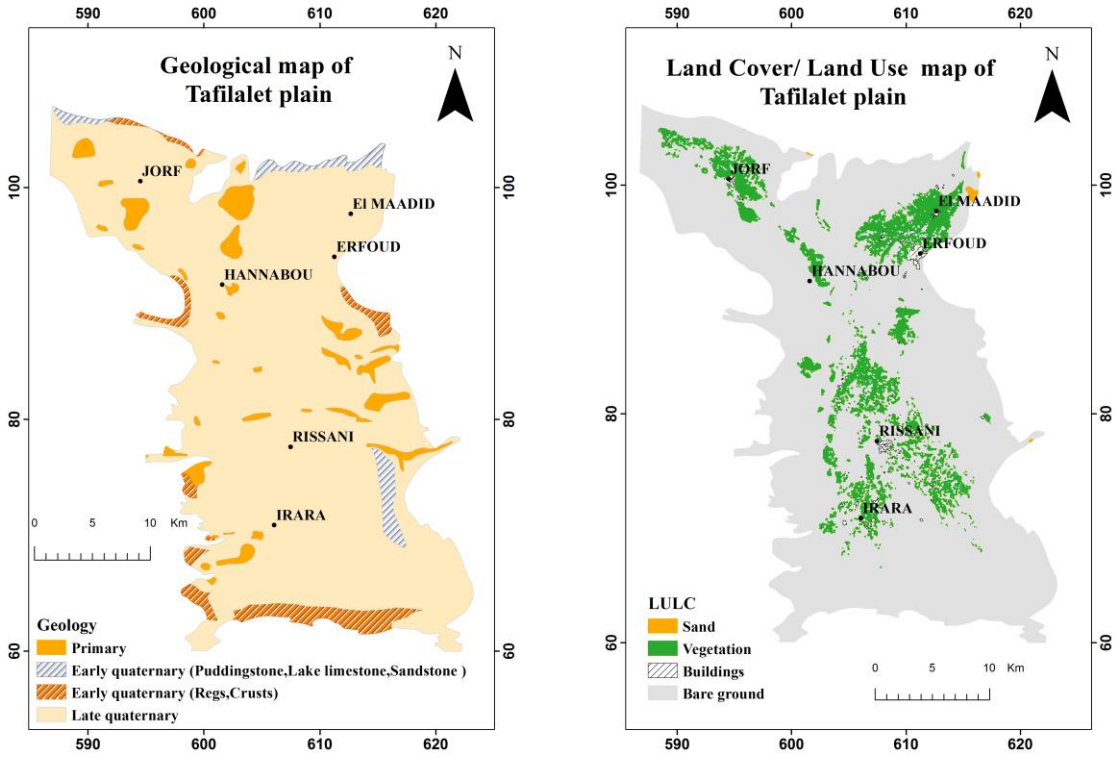
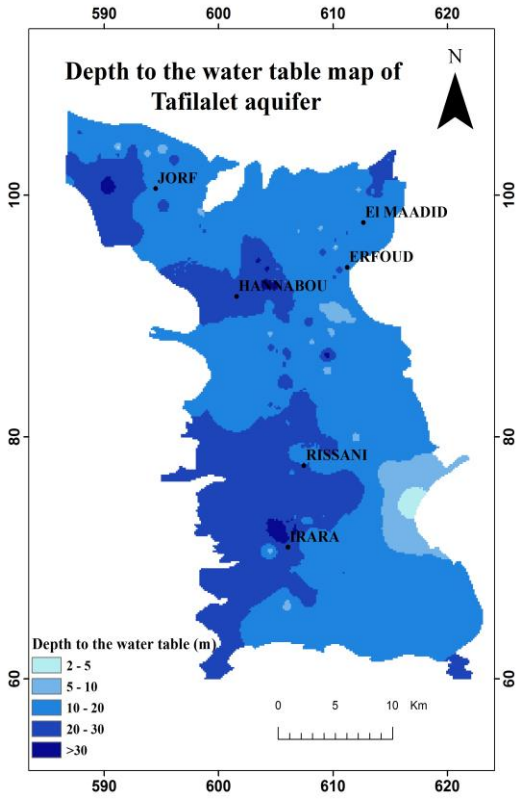
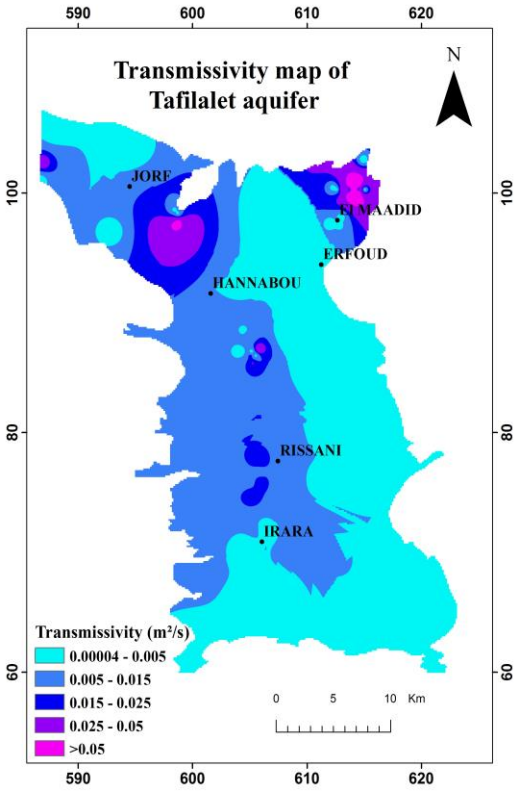


Figure 3: Thematic maps

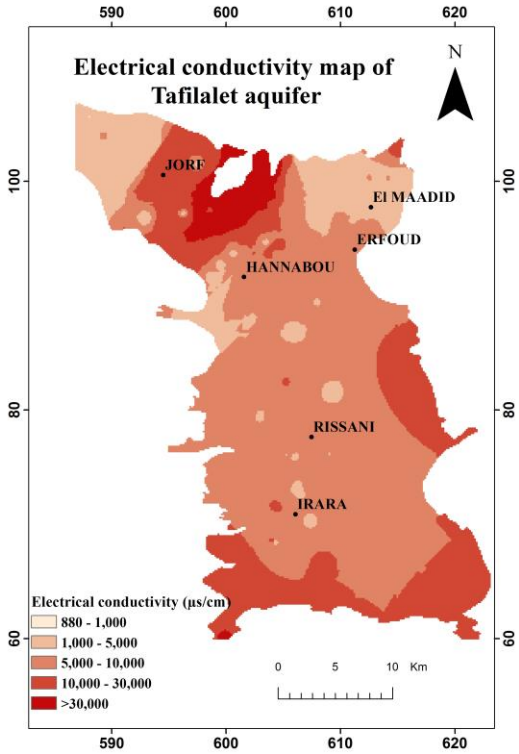
e) Depth to the water table map



f) Transmissivity map



g) Electrical Conductivity map



h) Drainage Density map

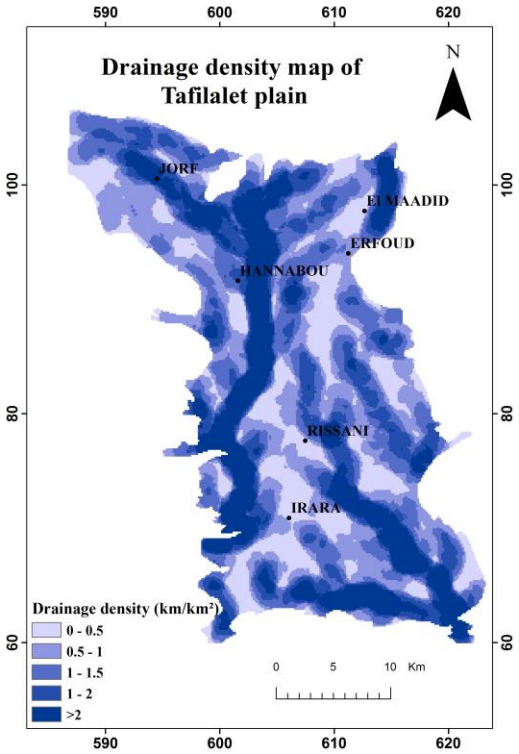
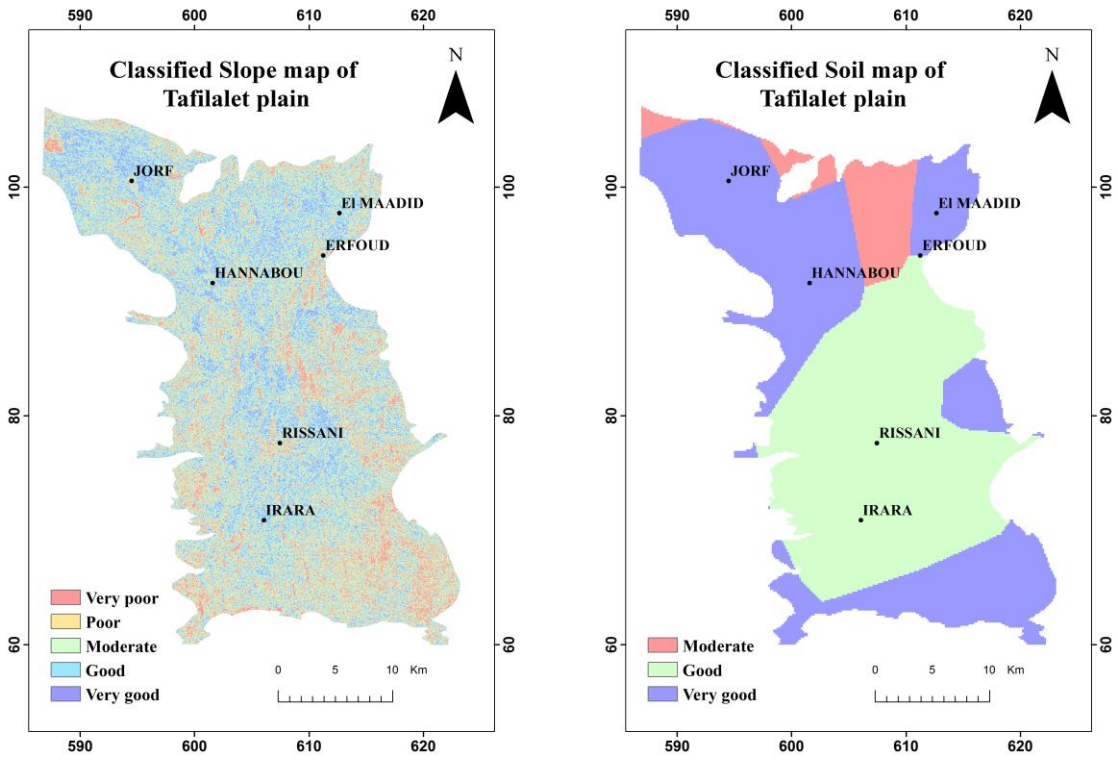


Figure 3: (Continued)

a) Classified Slope map

b) Classified Soil map



c) Classified Geological map

d) Classified LCLU map

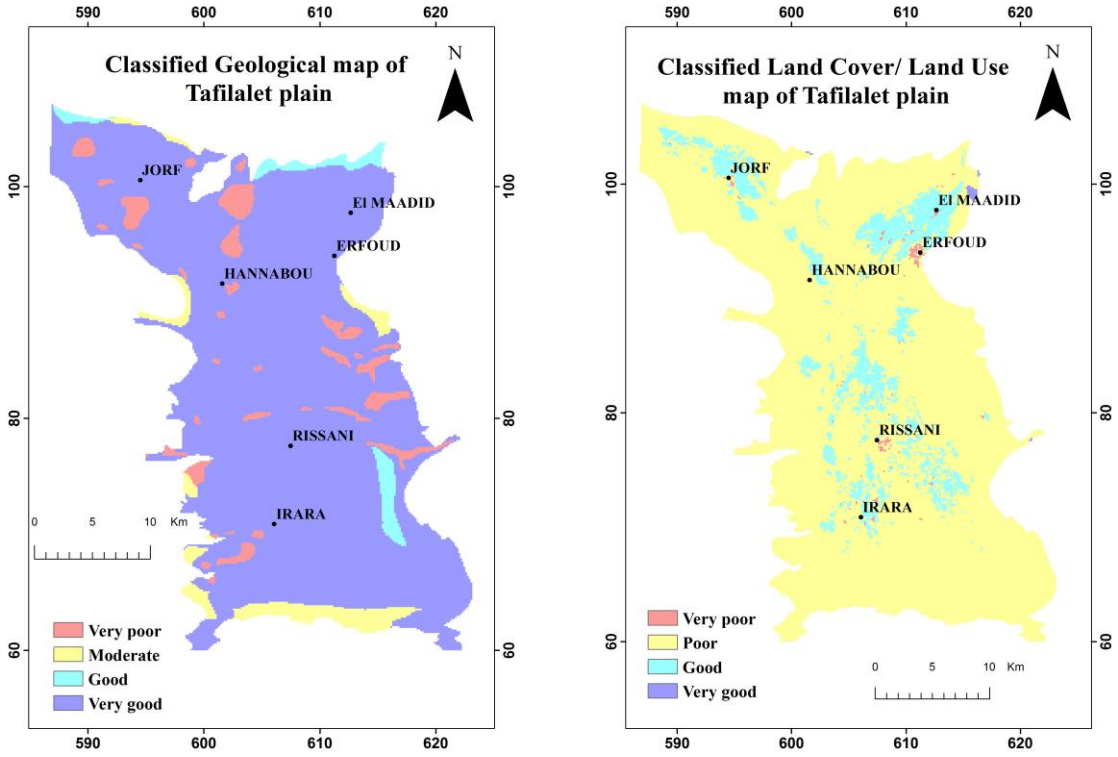
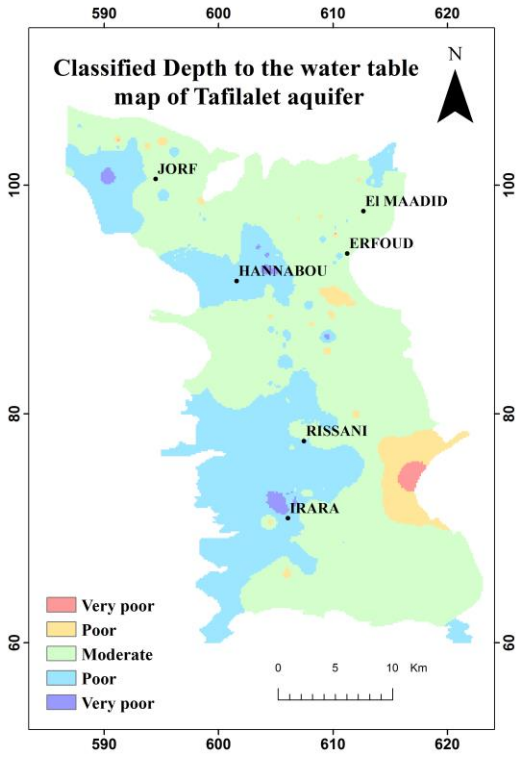
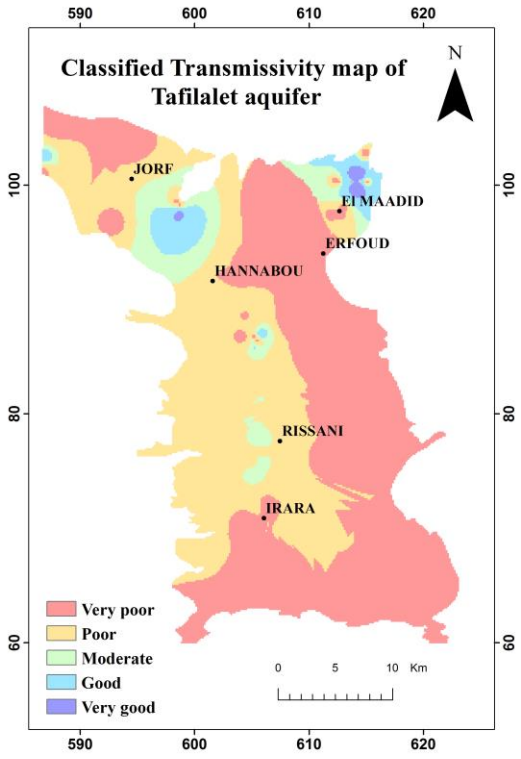


Figure 4: Classified thematic maps

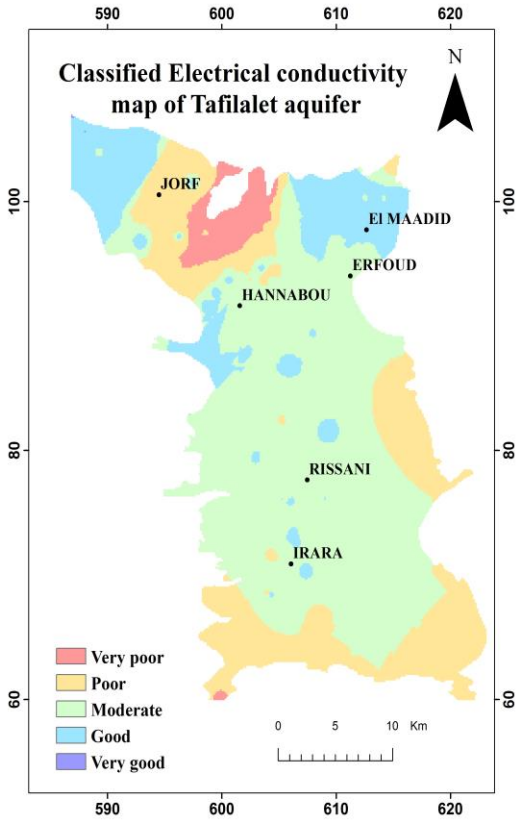
e) Classified Depth to the water table map



f) Classified Transmissivity map



g) Classified Electrical Conductivity map



h) Classified Drainage Density map

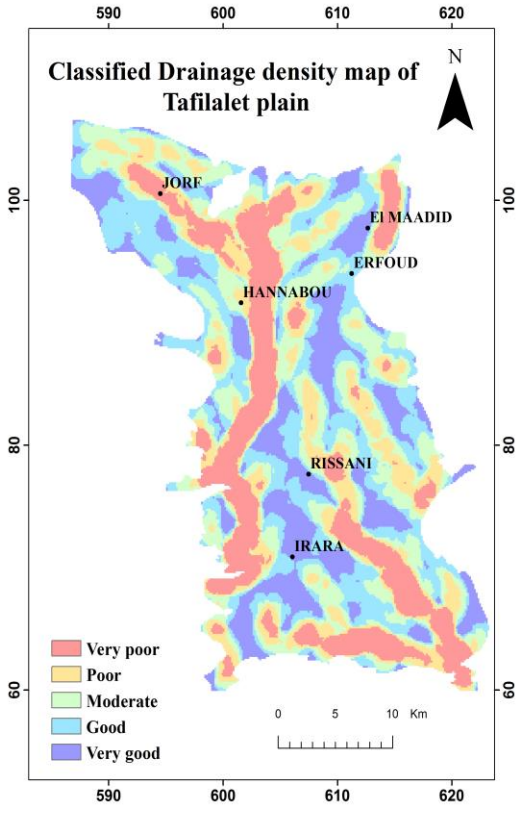


Figure 4: (Continued)

Table 2: Classes of each parameter and assigned categories

Parameter	Class	Category	Score
Slope (%)	0 - 3	Very good	9
	3 - 6	Good	8
	6 - 9	Moderate	5
	9 - 15	Poor	2
	> 15	Very poor	1
Soil	Calcisols	Very good	9
	Fluvisols	Good	8
	Leptosol	Moderate	5
Geology	Primary	Very poor	1
	Early quaternary (Regs, crusts)	Moderate	5
	Early quaternary (Puddingstone, Lake Limestone, Sandstone)	Good	8
	Late quaternary	Very good	9
LCLU	Sand	Very good	9
	Vegetation	Good	8
	Buildings	Very poor	1
	Bare ground	Poor	2
Depth to the water table (m)	2 - 5	Very poor	1
	5 - 10	Poor	2
	10 - 20	Moderate	5
	20 - 30	Good	8
	> 30	Very good	9
Transmissivity (m^2/s)	0.00004 - 0.005	Very poor	1
	0.005 - 0.015	Poor	2
	0.015 - 0.025	Moderate	5
	0.025 - 0.05	Good	8
	> 0.05	Very good	9
Electrical conductivity (µs/cm)	880 - 1000	Very good	9
	1000 - 5000	Good	8
	5000 - 10000	Moderate	5
	10000 - 30000	Poor	2
	> 30000	Very poor	1
Drainage density (km/km^2)	0 - 1	Very good	9
	1 - 2	Good	8
	2 - 3	Moderate	5
	3 - 4	Poor	2
	> 4	Very poor	1

After structuring the problem in such a way, the second step is to evaluate, compare and prioritize in a pairwise comparison matrix, based on decision-makers' judgments each criterion against the others pertaining to their importance in the problem decision making. To represent the magnitude of judgments, Saaty's fundamental scale should be used (Table 4), in doing so, relative weights of each criterion is deduced using Saaty's eigenvector technique and the generated weights must be validated because judgments are subjective and can typically be made with errors, so Saaty defined an index for measuring the degree of inconsistency which is the consistency ratio C. R. and it should be computed.

Table 3: Adopted evaluation scale for reclassification

Category	Very poor	Poor	Moderate	Good	Very good
Score	1	2	5	8	9

Actually, C. R. is a way of measuring the number of "errors" created when making the judgments, the rule-of-thumb is that if the C.R. is below 0.1, the errors are relatively small, thus the final estimate is acceptable [70,72–76].

The above-explained concept of AHP could be explained mathematically as follows:

Suppose $c_1, c_2, \ldots, c_n$ are n criteria influencing a decision with respect to a given problem, the pairwise comparisons matrix $B=[b_{ij}]$ of that problem is an $n \times n$ matrix and it may be represented as follows :

$$\begin{array}{c} \\ c_1 \\ \vdots \\ c_n \end{array} \begin{array}{c} \begin{array}{ccc} c_1 & \cdots & c_n \end{array} \\ \begin{bmatrix} 1 & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ 1/b_{1n} & \cdots & 1 \end{bmatrix} \end{array}$$

comparing the n criteria in pairs should comply with the following conditions:

$b_{ij}>0$ and $b_{ij}=\frac{1}{b_{ji}}$ *for all i and j* which means that B should be a positive and a reciprocal matrix. Besides, the values that can be attributed to b_{ij} *for all i and j* must be on the scale of intensities defined by Saaty (Table 4).

Table 4: The fundamental Scale

Intensity of importance	Definition	Explanation
1	Equal importance	Two criteria contribute equally to the objective
3	Moderate importance	one criterion is slightly favored over another
5	Strong importance	one criterion is strongly favored over another
7	Very strong importance	one criterion is highly favored over another
9	Extreme importance	one criterion is extremely favored over another
2,4,6,8	Intermediate values	when compromise is needed
Reciprocals	if criterion i has one of the above values assigned to it when compared with criterion j, then j has the reciprocal value when compared with i.	

Deriving the weights $w_1, w_2, \ldots, w_n$ relative to the n criteria $c_1, c_2, \ldots, c_n$ can be done by Saaty's eigenvector technique when judgments are inconsistent. Saaty's eigenvector technique stipulate that weights can be computed as the principal right eigenvector of the matrix B:

$$BW = \lambda_{max} W \tag{2}$$

where λ_{max} is the maximum eigenvalue of the matrix B, or

$$w_i = \frac{\sum_{j=1}^{n} b_{ij} w_j}{\lambda_{max}} \quad for\ all\ i = 1, 2, \ldots, n \tag{3}$$

After generating the weights, the consistency of the matrix should be measured by C.R. as explained above but first consistency index C.I. should be computed by the following formula:

$$C.I = \frac{\lambda_{max} - n}{n - 1} \tag{4}$$

because C.R. is the ratio of C.I. to R.I. :

$$C.R = \frac{C.I}{R.I} \tag{5}$$

with:

R.I. is the random index and is defined for each size of the matrix which is illustrated in Table 5.

C.R. must be less than 0.1 to consider the result as acceptable, larger values require the decision-maker to revise judgments in order to reduce the inconsistencies.

Therefore, for the current study, application of the AHP gives firstly the hierarchy structure illustrated in Figure 5, secondly, the pairwise matrix (Table 6) was developed by comparing pairs of factors with the help of expert point of views and using the fundamental scale of Saaty (Table 4) to translate judgments into values. Comparisons were done with regard to the relative influence of each factor on aquifer recharge and on implementing the (AAR) technique in Tafilalet plain. R.I. used is 1.41 because the dimension of the pairwise matrix of the current study is 8×8 (Table 6) so, after computation, the obtained C.R. is equal to 0.041 which is less than 0.1 and according to the AHP method this result is satisfactory and the generated weights are acceptable.

Based on the generated weights, geology is the highest eigenvector while drainage density is the lowest (Table 6). In fact, geology, soil type, depth to the water table, and transmissivity are the factors that were given the highest weights. This was expected because, according to the analysis that was made during pairwise comparisons, the first thing to look out for in a project for artificial recharge of an aquifer suffering from overexploitation is the geological formations, which must be sufficiently permeable and transmissive, the soil, which must have a high infiltration capacity and the groundwater which must be deep enough to provide sufficient space for water storage to allow the water table to regenerate its level.

3.4. AAR index generation

After generating the relative weights for each factor by the AHP method, the reclassified thematic layers of the factors were multiplied by their weights and finally added up using the weighted sum function of ArcGIS software.

Thus, the potential sites for the artificial aquifer recharge map was produced by calculating the AAR index according to the following formula:

$$I_{AAR} = \sum_{i=1}^{n} F_i w_i \tag{6}$$

where:

-I_{AAR} is the AAR index

-n is the number of factors influencing (AAR) which is eight in the current study

-F_i stands for factor

-w_i stands for weight

Table 5: Random inconsistency index

n	1	2	3	4	5	6	7	8	9	10
R.I.	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

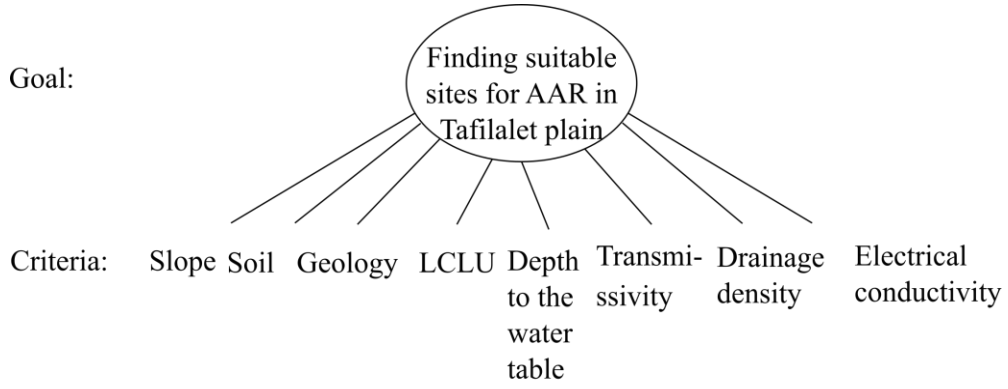


Figure 5: Hierarchy structure of the current study

Table 6: Pairwise matrix developed for the current study

Criteria	Slope	Soil	Geology	LCLU	Depth to the water table	Transmissivity	Electrical conductivity	Drainage density	Generated Weights
Slope	1	1/2	1/5	1/3	1/2	1/2	1	5	0.07
Soil	2	1	1/2	2	1	1	3	6	0.15
Geology	5	2	1	4	2	3	2	6	0.28
LCLU	3	1/2	1/4	1	1/2	1/2	2	5	0.10
Depth to the water table	2	1	1/2	2	1	1	3	7	0.16
Transmissivity	2	1	1/3	2	1	1	3	6	0.15
Electrical conductivity	1	1/3	1/2	1/2	1/3	1/3	1	5	0.07
Drainage density	1/5	1/6	1/6	1/5	1/7	1/6	1/5	1	0.02

4. Result

The map of potential sites for artificial aquifer recharge which is illustrated in Figure 6 was produced as explained above and then reclassified according to five categories which are: very poorly suitable, poorly suitable, moderately suitable, suitable, and very suitable. It indicates that the most suitable areas for the artificial groundwater recharge of Tafilalet aquifer are located notably in the northeast and in the west of the plain. Statistically, About 47 % of the study area is suitable for artificial aquifer recharge and about 40 % are moderately suitable. Thus, those results can be used to formulate an efficient groundwater management plan for sustainable utilization of these limited resources because identified areas will be prioritized in the implementation of artificial aquifer recharge projects.

Based on the produced map of potential sites for artificial recharge of Tafilalet aquifer (Figure 6), most suitable zones are underlain by high permeable geological formations that mostly belong to the late quaternary and covered with soils with high infiltration capacity most often constituted by alluvium and basic aeolian reliefs rich in meteoric materials and located in areas where the water table is very deep more than 20 m deep in addition the slope of the plain is generally not too steep and the above-mentioned description of the Tafilalet aquifer revealed that it is an alluvial and unconfined aquifer so, the artificial aquifer recharge technique most conducive with regard to the aforementioned conditions is infiltration basins nevertheless, the climatic characteristics of the region known by very high evapotranspiration rates will constitute a big challenge because more water will be lost by evaporation so less water will infiltrate unless the artificial aquifer recharge is conducted just during wet seasons when evapotranspiration rate decreases. Thus, several other factors must be taken into account when deciding whether a specific (AAR) system is suitable to be implemented, even if the appropriate locations are identified.

5. Discussion

The present-day situation of Tafilalet aquifer should urge water managers to work hard to avoid its worsening so, great efforts in risk mitigation need to be greatly accelerated and they should be centered on a range of scientific and engineering-based solutions. The current study presents one of those solutions which is the (AAR) and how to identify, in a scientific way, the most suitable locations for its implementation. In fact, A GIS-based multi-criteria analysis was performed, using eight thematic maps in order to produce an appropriate map for the identification of sites for (AAR) in Tafilalet plain. The resulting map shows that the most suitable areas are identified in the northeast and in the west of the region while inappropriate locations are limited to small portions in the northern, southern, and eastern parts of the study area. This result is extremely important because it will help taking appropriate decisions for mitigating the overexploitation problem of the Tafilalet aquifer that is why the scientific approach should be taken into consideration in such important projects because it gives nearly accurate results and it brings time savings.

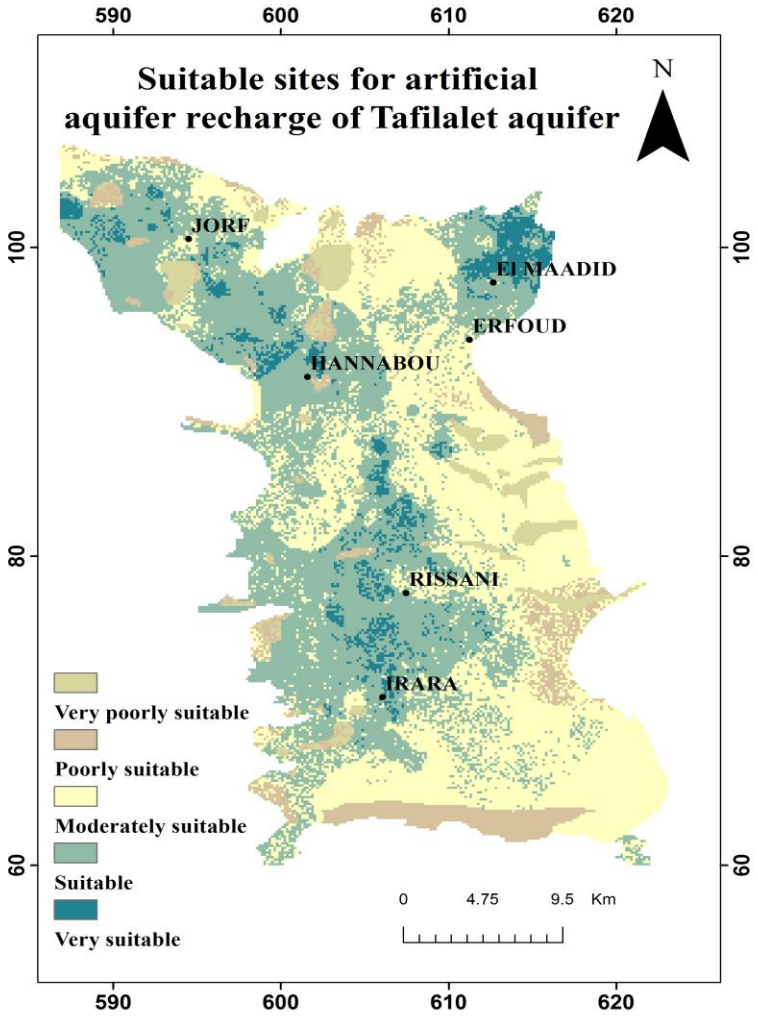


Figure 6: Potential artificial aquifer recharge sites map

The (AAR) will certainly sort out the problem of depletion of groundwaters in the Tafilalet aquifer as it was proved in some other aquifers in Morocco and in other countries. In fact, the (AAR) has already initiated and applied to some aquifers suffering from the same problem in Morocco such as the Charf El Akab aquifer in the region of Tangier, the Haouz Mejjate aquifer in the region of Marrakech, the Souss aquifer in the region of Agadir and the Berrechid aquifer located in the south of Casablanca, those aquifers have already started to reap its benefits. Even if the (AAR) practices in Morocco are still very much in their infancy, the technology is beginning to be recognized at the national level as having a value in tackling groundwater depletion issues.

Some attempts have been made in Tafilalet plain by the hydraulic basin agency of Guir-Ziz-Rheris in this regard and artificial recharging techniques have already begun to emerge on the plain, such as the construction of dikes or sills in the bed of the wadi Ziz in order to slow down the speed of the water and increase the wetted areas along the bed of the wadi, however, if the working field is extended to other locations that must be well chosen, these locations must be identified by means of scientific decision-making like GIS and multi-criteria analysis taking into consideration several criteria such as geological, pedological, hydrogeological, and many other factors to make sure that they are suitable for artificial recharge, in doing so, other artificial recharge techniques can be performed and can better and quickly restore the groundwater level situation of Tafilalet aquifer to its normal stage.

6. Conclusion

The aridity of the climate together with the excessive need for water is putting pressure on groundwater resources of Tafilalet plain so, the water level of the Tafilalet aquifer is declining to a threateningly low level which may undermine its productivity and causing its depletion, problems which are more likely to have irreversible adverse effects may occur like imperiling groundwater-dependent ecosystems, human activities as well as economic development of the region. Thus, the core objective of this study is to investigate suitable places for artificially recharge the Tafilalet aquifer to restore its levels to normal and fill the gap caused by overexploitation. For this purpose, thematic layers of factors that control groundwater recharge like depth to the water table, drainage density, electrical conductivity, slope, aquifer transmissivity, geology, soil, LCLU were prepared mapped, and reclassified in a GIS (ArcGIS software), after that, AHP method

was applied to compute weights of each factor and the map of appropriate sites of artificial groundwater recharge was generated by the application of weighted sum tool of ArcGIS software. The result has revealed that around 47% of the study area has been identified as a high potential zone for groundwater recharge which is a very important outcoming. However, identifying suitable places is essential but not sufficient for implementing (AAR) because as mentioned above even if infiltration basins are estimated as the appropriate (AAR) system to implement in Tafilalet plain, there are other factors to consider like climatic conditions of the region, need for periodic maintenance of the basin in order to sustain infiltration rates over the long term because organic matter build-up, silting and/or clogging can make the surface of the basin impervious and the capacity of infiltration will decrease over time in addition to the existence of a good quality water source nearby that will ensure the supply of water to be put into the basin and many other conditions that if not met will render the project impractical. Lastly, the present study has shown an example of the benefits of coupling multi-criteria analysis with GIS and remote sensing techniques in taking relevant decisions to better manage and protect groundwater resources.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors wish to record their deep sense of gratitude and profound thanks to all the employees of hydraulic basin agency of Guir-Ziz-Rheris for their warm welcome and their support to bring this work to fruition.

References

[1] I.A. Shiklomanov, "World water resources: a new appraisal and assessment for the 21st century; 1998," 40, 1998.

[2] Gleick Peter H, World fresh water resources, 1993.

[3] T. Oki, Shinjiro. Kanae, "Global Hydrological Cycles and World Water Resources," Science, **313**(5790), 1068–1072, 2006, doi:10.1126/science.1128845.

[4] S. Tushaar, Burke J, Villholth K, Angelica, M, Custodio, E, Daibes, F, Hoogesteger, J, Giordano, Mark, Girman, J, van der Gun, J, Kendy, E, Kijne, J, Llamas, R, Masiyandima, Mutsa, Margat, J, Marin, L, Peck, J, Rozelle, S, Sharma, Bharat R, Vincent, L, Wang, J, Groundwater: a global assessment of scale and significance, 395-423., 2007.

[5] S. Siebert, V. Henrich, K. Frenken, J. Burke, Update of the digital global map of irrigation areas to version 5., Institute of Crop Science and Resource Conservation, University of Bonn, Germany; Land and Water Division, Food and Agriculture Organization of the United Nations, Rome, Italy, 2013, doi:10.13140/2.1.2660.6728.

[6] T. Havril, Á. Tóth, J.W. Molson, A. Galsa, J. Mádl-Szőnyi, "Impacts of predicted climate change on groundwater flow systems: Can wetlands disappear due to recharge reduction?," Journal of Hydrology, **563**, 1169–1180, 2018, doi:10.1016/j.jhydrol.2017.09.020.

[7] E. Sener, A. Davraz, "Assessment of groundwater vulnerability based on a modified DRASTIC model, GIS and an analytic hierarchy process (AHP) method: the case of Egirdir Lake basin (Isparta, Turkey)," Hydrogeology Journal, **21**(3), 701–714, 2013, doi:10.1007/s10040-012-0947-y.

[8] M. Giordano, "Global Groundwater? Issues and Solutions," Annual Review of Environment and Resources, **34**(1), 153–178, 2009, doi:10.1146/annurev.environ.030308.100251.

[9] A.S. Qureshi, P.G. McCornick, A. Sarwar, B.R. Sharma, "Challenges and Prospects of Sustainable Groundwater Management in the Indus Basin, Pakistan," Water Resources Management, **24**(8), 1551–1569, 2010, doi:10.1007/s11269-009-9513-3.

[10] A. Jódar-Abellán, J.A. Albaladejo-García, D. Prats-Rico, "ARTIFICIAL GROUNDWATER RECHARGE. REVIEW OF THE CURRENT KNOWLEDGE OF THE TECHNIQUE," 12, 2017.

[11] S. Kaliraj, N. Chandrasekar, N.S. Magesh, "Identification of potential groundwater recharge zones in Vaigai upper basin, Tamil Nadu, using GIS-based analytical hierarchical process (AHP) technique," Arabian Journal of Geosciences, **7**(4), 1385–1401, 2014, doi:10.1007/s12517-013-0849-x.

[12] P.K. Naik, M. Mojica, F. Ahmed, S. Al-Mannai, "Storm water injection in Bahrain: pilot studies," Arabian Journal of Geosciences, **10**(20), 2017, doi:10.1007/s12517-017-3232-5.

[13] A. Rawluk, A. Curtis, E. Sharp, B.F.J. Kelly, A.J. Jakeman, A. Ross, M. Arshad, R. Brodie, C.A. Pollino, D. Sinclair, B. Croke, M.E. Qureshi, "Managed aquifer recharge in farming landscapes using large floods: an opportunity to improve outcomes for the Murray-Darling Basin?," Australasian Journal of Environmental Management, **20**(1), 34–48, 2013, doi:10.1080/14486563.2012.724785.

[14] L.F. Konikow, E. Kendy, "Groundwater depletion: A global problem," Hydrogeology Journal, **13**(1), 317–320, 2005, doi:10.1007/s10040-004-0411-8.

[15] M. Giordano, K.G. Villholth, eds., The agricultural groundwater revolution: opportunities and threats to development, CABI, Wallingford, 2007.

[16] D. Molden, al, eds., Water for food, water for life: a comprehensive assessment of water management in agriculture, Earthscan, 2007.

[17] M.I. Al-Hamdi, Competition for scarce groundwater in the Sana'a Plain, Yemen : a study on the incentive systems for urban and agricultural water use, CRC Press / Balkema, Netherlands, 2000.

[18] Yoshihide Wada, Ludovicus P. H. van Beek, Cheryl M. van Kempen, Josef W. T. M. Reckman, Slavek Vasak, Marc F. P. Bierkens, "Global depletion of groundwater resources," GEOPHYSICAL RESEARCH LETTERS, 37, 2010.

[19] F. Pedrero, A. Albuquerque, H. Marecos do Monte, V. Cavaleiro, J.J. Alarcón, "Application of GIS-based multi-criteria analysis for site selection of aquifer recharge with reclaimed water," Resources, Conservation and Recycling, **56**(1), 105–116, 2011, doi:10.1016/j.resconrec.2011.08.003.

[20] A.S. Qureshi, P.G. McCornick, M. Qadir, Z. Aslam, "Managing salinity and waterlogging in the Indus Basin of Pakistan," Agricultural Water Management, **95**(1), 1–10, 2008, doi:10.1016/j.agwat.2007.09.014.

[21] N. Iqbal, F. Hossain, H. Lee, G. Akhter, "Satellite Gravimetric Estimation of Groundwater Storage Variations Over Indus Basin in Pakistan," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **9**(8), 3524–3534, 2016, doi:10.1109/JSTARS.2016.2574378.

[22] P. Pavelic, K. Srisuk, P. Saraphirom, S. Nadee, K. Pholkern, S. Chusanathas, S. Munyou, T. Tangsutthinon, T. Intarasut, V. Smakhtin, "Balancing-out floods and droughts: Opportunities to utilize floodwater harvesting and groundwater storage for agricultural development in Thailand," Journal of Hydrology, 55–64, 2012, doi:10.1016/j.jhydrol.2012.08.007.

[23] W. Martín-Rosales, J. Gisbert, A. Pulido-Bosch, A. Vallejos, A. Fernández-Cortés, "Estimating groundwater recharge induced by engineering systems in a semiarid area (southeastern Spain)," Environmental Geology, **52**(5), 985–995, 2007, doi:10.1007/s00254-006-0541-5.

[24] Y.B. Katpatal, A.M. Pophare, B.R. Lamsoge, "A groundwater flow model for overexploited basaltic aquifer and Bazada formation in India," Environmental Earth Sciences, **72**(11), 4413–4425, 2014, doi:10.1007/s12665-014-3342-2.

[25] A. Singh, S.N. Panda, K.S. Kumar, C.S. Sharma, "Artificial Groundwater Recharge Zones Mapping Using Remote Sensing and GIS: A Case Study in Indian Punjab," Environmental Management, **52**(1), 61–71, 2013, doi:10.1007/s00267-013-0101-1.

[26] Y.S. Prasad, B.V. Rao, "Groundwater depletion and groundwater balance studies of Kandivalasa River Sub Basin, Vizianagaram District, Andhra Pradesh, India," Groundwater for Sustainable Development, **6**, 71–78, 2018, doi:10.1016/j.gsd.2017.11.003.

[27] C.J. Nevill, "Managing Cumulative Impacts: Groundwater Reform in the Murray-Darling Basin, Australia," Water Resources Management, **23**(13), 2605–2631, 2009, doi:10.1007/s11269-009-9399-0.

[28] L.J. Frookh, L. Froukh, "Artificial Aquifer Recharge with Treated Wastewater in North Gaza," Riyadh,Saudi Arabia: 9, 2019.

[29] R. Spandre, Artificial Groundwater Recharge, EOLSS: 20, 2009.

[30] D. Das, Drainage and lineament analysis towards artificial recharge of groundwater, Springer Berlin Heidelberg, Berlin, Heidelberg: 37–44, 2011, doi:10.1007/978-3-642-24076-8_5.

[31] K.G. Villholth, E. Lopez-gunn, J. Van Der Gun, A. Garrido, K. Conti, Advances in Groundwater Governance, CRC Press, 2019.

[32] M.R. Llamas, P. Martínez-Santos, "Intensive Groundwater Use: Silent Revolution and Potential Source of Social Conflicts," Journal of Water Resources Planning and Management, **131**(5), 337–341, 2005, doi:10.1061/(ASCE)0733-9496(2005)131:5(337).

[33] S.B. Gingerich, C.I. Voss, A.G. Johnson, "Seawater-flooding events and impact on freshwater lenses of low-lying islands: Controlling factors, basic management and mitigation," Journal of Hydrology, **551**, 676–688, 2017, doi:10.1016/j.jhydrol.2017.03.001.
[34] W.-C. Hung, C. Hwang, J.-C. Liou, Y.-S. Lin, H.-L. Yang, "Modeling aquifer-system compaction and predicting land subsidence in central Taiwan," Engineering Geology, 147–148, 78–90, 2012, doi:10.1016/j.enggeo.2012.07.018.
[35] T.J. Burbey, "The Influence of Geologic Structures on Deformation due to Ground Water Withdrawal," Ground Water, **46**(2), 202–211, 2008, doi:10.1111/j.1745-6584.2007.00395.x.
[36] Q. Feng, G. Liu, L. Meng, E. Fu, H. Zhang, K. Zhang, "Land subsidence induced by groundwater extraction and building damage level assessment — a case study of Datun, China," Journal of China University of Mining and Technology, **18**(4), 556–560, 2008, doi:10.1016/S1006-1266(08)60293-X.
[37] M.J. Abdulrazzak, P.D. Oikonomou, N.S. Grigg, "Transboundary Groundwater Cooperation among Countries of the Arabian Peninsula," Journal of Water Resources Planning and Management, **146**(1), 05019023, 2020, doi:10.1061/(ASCE)WR.1943-5452.0001140.
[38] R.L. Nelson, "Assessing local planning to control groundwater depletion: California as a microcosm of global issues: LOCAL PLANNING TO CONTROL GROUNDWATER DEPLETION," Water Resources Research, **48**(1), 2012, doi:10.1029/2011WR010927.
[39] R.A. Freeze, J.A. Cherry, Groundwater, Prentice-Hall, Englewood Cliffs, N.J, 1979.
[40] E.A. Varouchakis, P.G. Theodoridou, G.P. Karatzas, "Decision-Making Tool for Groundwater Level Spatial Distribution and Risk Assessment Using Geostatistics in R," Journal of Hazardous, Toxic, and Radioactive Waste, **24**(1), 04019031, 2020, doi:10.1061/(ASCE)HZ.2153-5515.0000464.
[41] C.K. Singh, Y.B. Katpatal, "A GIS Based Design of Groundwater Level Monitoring Network Using Multi-Criteria Analysis and Geostatistical Method," Water Resources Management, **31**(13), 4149–4163, 2017, doi:10.1007/s11269-017-1737-z.
[42] W.B. Cutter, "Valuing Groundwater Recharge in an Urban Context," Land Economics, **83**(2), 234–252, 2007.
[43] P. Dillon, "Future management of aquifer recharge," Hydrogeology Journal, **13**(1), 313–316, 2005, doi:10.1007/s10040-004-0413-6.
[44] P. Dillon, P. Stuyfzand, T. Grischek, M. Lluria, R.D.G. Pyne, R.C. Jain, J. Bear, J. Schwarz, W. Wang, E. Fernandez, C. Stefan, M. Pettenati, J. van der Gun, C. Sprenger, G. Massmann, B.R. Scanlon, J. Xanke, P. Jokela, Y. Zheng, R. Rossetto, M. Shamrukh, P. Pavelic, E. Murray, A. Ross, J.P. Bonilla Valverde, A. Palma Nava, N. Ansems, K. Posavec, K. Ha, et al., "Sixty years of global progress in managed aquifer recharge," Hydrogeology Journal, **27**(1), 1–30, 2019, doi:10.1007/s10040-018-1841-z.
[45] S. Rahimi, M. Shadman Roodposhti, R. Ali Abbaspour, "Using combined AHP–genetic algorithm in artificial groundwater recharge site selection of Gareh Bygone Plain, Iran," Environmental Earth Sciences, **72**(6), 1979–1992, 2014, doi:10.1007/s12665-014-3109-9.
[46] A.K. Bhattacharya, "ARTIFICIAL GROUND WATER RECHARGE WITH A SPECIAL REFERENCE TO INDIA," 8, 2010.
[47] T. Asano, Artificial Recharge of Groundwater, Butterworth Publishers, Boston, 1985.
[48] A. Mahdavi, S.H. Tabatabaei, M. Nouri, R. Mahdavi, "Identification of groundwater artificial recharge sites using Fuzzy logic: A case study of Shahrekord plain, Iran," 2012.
[49] H. Bouwer, "Artificial recharge of groundwater: hydrogeology and engineering," Hydrogeology Journal, **10**(1), 121–142, 2002, doi:10.1007/s10040-001-0182-4.
[50] E. Widad, "Etude de la zone favorable pour la recharge artificielle de la nappe phréatique Mnasra," 55, 2014.
[51] H. Nasiri, A.D. Boloorani, H.A.F. Sabokbar, H.R. Jafari, M. Hamzeh, Y. Rafii, "Determining the most suitable areas for artificial groundwater recharge via an integrated PROMETHEE II-AHP method in GIS environment (case study: Garabaygan Basin, Iran)," Environmental Monitoring and Assessment, **185**(1), 707–718, 2013, doi:10.1007/s10661-012-2586-0.
[52] H.A. Faraji Sabokbar, H. Nasiri, M. Hamze, S. Talebi, Y. Rafiei, "Identification of suitable areas for artificial groundwater recharge using integrated ANP and pair wise comparison methods in GIS environment, (case study: Garbaygan Plain of Fasa)," Geography and Environmental Planning, **44**(4), 2012.
[53] J. Ghayoumian, B. Ghermezcheshme, S. Feiznia, A.A. Noroozi, "Integrating GIS and DSS for identification of suitable areas for artificial recharge, case study Meimeh Basin, Isfahan, Iran," Environmental Geology, **47**(4), 493–500, 2005, doi:10.1007/s00254-004-1169-y.
[54] I. Bouaamlat, A. Larabi, M. Faouzi, "Assessment of Water Resources in the Tafilalet Oasis System by a Mathematical Model (South East of Morocco)," 1, 2012.
[55] I.P. Senanayake, D.M.D.O.K. Dissanayake, B.B. Mayadunna, W.L. Weerasekera, "An approach to delineate groundwater recharge potential sites in Ambalantota, Sri Lanka using GIS techniques," Geoscience Frontiers, **7**(1), 115–124, 2016, doi:10.1016/j.gsf.2015.03.002.
[56] J. Ghayoumian, M. Mohseni Saravi, S. Feiznia, B. Nouri, A. Malekian, "Application of GIS techniques to determine areas most suitable for artificial groundwater recharge in a coastal aquifer in southern Iran," Journal of Asian Earth Sciences, **30**(2), 364–374, 2007, doi:10.1016/j.jseaes.2006.11.002.
[57] A. Chowdhury, M.K. Jha, V.M. Chowdary, "Delineation of groundwater recharge zones and identification of artificial recharge sites in West Medinipur district, West Bengal, using RS, GIS and MCDM techniques," Environmental Earth Sciences, **59**(6), 1209–1222, 2010, doi:10.1007/s12665-009-0110-9.
[58] C. william Carlston, Drainage density and streamflow, Professional Paper, 1963.
[59] Guide on artificial recharge to groundwater, central ground water board ministry of water resources, new delhi, 2000.
[60] Z. Şen, Basic Porous Medium Concepts, Elsevier: 43–97, 2015, doi:10.1016/B978-0-12-800075-5.00002-9.
[61] J. Bear, Hydraulics of groundwater, McGraw-Hill, Inc., New York, 1979.
[62] G.P. Kruseman, N.A. de Ridder, Analysis and evaluation of pumping test data, 2. ed. (compl. rev.), repr, International Institute for Land Reclamation and Improvement, Wageningen, 1994.
[63] W. Luo, D.T. Pederson, "Hydraulic conductivity of the High Plains Aquifer re-evaluated using surface drainage patterns: HYDRAULIC CONDCTIVITY RE-EVALUATED," Geophysical Research Letters, **39**(2), n/a-n/a, 2012, doi:10.1029/2011GL050200.
[64] J.J. De Vries, "The groundwater outcrop-erosion model; evolution of the stream network in The Netherlands," Journal of Hydrology, **29**(1–2), 43–50, 1976, doi:10.1016/0022-1694(76)90004-4.
[65] R.E. Horton, "EROSIONAL DEVELOPMENT OF STREAMS AND THEIR DRAINAGE BASINS; HYDROPHYSICAL APPROACH TO QUANTITATIVE MORPHOLOGY," Geological Society of America Bulletin, **56**(3), 275, 1945, doi:10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2.
[66] J. Krishnamurthy, A. Mani, V. Jayaraman, M. Manivel, "Groundwater resources development in hard rock terrain - an approach using remote sensing and GIS techniques," International Journal of Applied Earth Observation and Geoinformation, **2**(3–4), 204–215, 2000, doi:10.1016/S0303-2434(00)85015-1.
[67] B. Rambert, Recherche sur la signification hydrogéologique de la densité du drainage: Application à la détermination et à la cartographie de l'écoulemment souterrain., Bureau de recherche géologique et minières, France, 1973.
[68] T. Dunne, Chapter 1. Hydrology mechanics, and geomorphic implications of erosion by subsurface flow, Geological Society of America: 1–28, 1990, doi:10.1130/SPE252-p1.
[69] J.A. Sandoval, C.L. Tiburan, "Identification of potential artificial groundwater recharge sites in Mount Makiling Forest Reserve, Philippines using GIS and Analytical Hierarchy Process," Applied Geography, **105**, 73–85, 2019, doi:10.1016/j.apgeog.2019.01.010.
[70] B.L. Golden, E.A. Wasil, P.T. Harker, eds., The Analytic Hierarchy Process: Applications and Studies, Springer Berlin Heidelberg, Berlin, Heidelberg, 1989, doi:10.1007/978-3-642-50244-6.
[71] T.L. Saaty, "A scaling method for priorities in hierarchical structures," Journal of Mathematical Psychology, **15**(3), 234–281, 1977, doi:10.1016/0022-2496(77)90033-5.
[72] T.L. Saaty, "How to make a decision: The Analytic Hierarchy Process," European Journal of Operational Research, 48, 1990.
[73] R.W. Saaty, "The analytic hierarchy process—what it is and how it is used," Mathematical Modelling, **9**(3–5), 161–176, 1987, doi:10.1016/0270-0255(87)90473-8.
[74] T.L. Saaty, "Decision making with the analytic hierarchy process," International Journal of Services Sciences, **1**(1), 83, 2008, doi:10.1504/IJSSCI.2008.017590.
[75] E. Mu, M. Pereyra-Rojas, Understanding the Analytic Hierarchy Process, Springer International Publishing, Cham: 7–22, 2017, doi:10.1007/978-3-319-33861-3_2.
[76] T.L. Saaty, What is the Analytic Hierarchy Process?, Springer Berlin Heidelberg, Berlin, Heidelberg: 109–121, 1988, doi:10.1007/978-3-642-83555-1_5.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1112-1119 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Challenges and New Paradigms in Conservation of Heritage-based Villages in Rural India-A case of Pragpur and Garli Villages in Himachal Pradesh

Preeti Nair[1,*], Devendra Pratap Singh[1], Navneet Munoth[2]

[1]*School of Architecture and Planning, Amity University, Noida, 201301, India*

[2]*Maulana Azad National Institute of Technology,* Department of Architecture and Planning*, Bhopal, 462003, India*

ARTICLE INFO

Article history:
Received: 14 November, 2020
Accepted: 23 January, 2021
Online: 16 February, 2021

Keywords:
Rural Heritage
Settlement Morphology
Community Perception
Community Awareness

ABSTRACT

The research paper aims to focus on the issues and challenges in developing a sustainable model of an ideal heritage village project by using descriptive and empirical investigation methods. To capture the perception and understanding of the concept of sustainability of a Heritage Village, a mixed-methods approach was conducted by the researcher where document where reviewed, observations were done, structured interviews and a questionnaire survey was conducted involving resident's in the heritage villages of Pragpur and Garli in Himachal Pradesh, India. Through this research, the objective was to catalogues the resident's outlook and understand their belongingness towards their rural settlement. The analysis conducted, was also to understand their attachment to the heritage fabric which would act as a catalyst for their sustainable development. Due to the diversification in terms of architecture, social and cultural aspects, it was important to analyze the resident's perception towards the built heritage as it may vary to be more or less important to different people, community groups, or generations.

1. Introduction

The term "Heritage" proposes a wider based concept that includes both tangible and intangible heritage as well as the natural ecosystem of the settlement. It includes built environment, historic places and monuments, settlement pattern, settlement morphology, heritage collections, past and present cultural practices including living experiences of the residents of these villages. It can act as a catalyst for the overall sustainable growth and change in these rural settlements [1]. The collective memory of each locality or community or a particular heritage is unique. The preservation, conservation, documentation, presentation of the built heritage, and cultural diversity of any particular place are important challenges.

Manmade structures which possess heritage character and significance should become a part of the built heritage of any settlement and the society [2]. Due to the rapid modernization of the built environment, there is a need for focusing attention to preserve the cultural and architectural heritage, before it disappears completely. Buildings and settlements of historic or architectural value in all urban and rural areas are disappearing very fast due to neglect and decay or reconstruction in newer materials which results in changing the specific characters of these village cores, leading to monotonous and repetitive new modern constructions.

The built heritage is the prominent aspect to transform the current situation of rural areas because it has the potential to diversify people's perception and acts as a common denominator for the development of these settlements. Heritage can act as a catalyst for the future development and sustainability of a village [3]. It indicates that by enhancing the heritage character of a village, a sustainable tool can be devised for the residents who live in these settlements to make them part of the rural development process. Additionally, the tool is likely to give rural areas a positive, vibrant image and in this way, a trend can be developed for bringing the migrated youth back into these rural areas. It also indicates that cultural meanings and construction assets such as vernacular construction techniques can be mobilized for rural and regional development, and cultural identities can become important for uplifting local economies.

*Preeti Nair, Research Scholar, Amity School of Architecture and Planning, Amity University, Noida, Uttar Pradesh, India, +91 -9711433201, preetinair30@gmail.com

https://dx.doi.org/10.25046/aj0601124

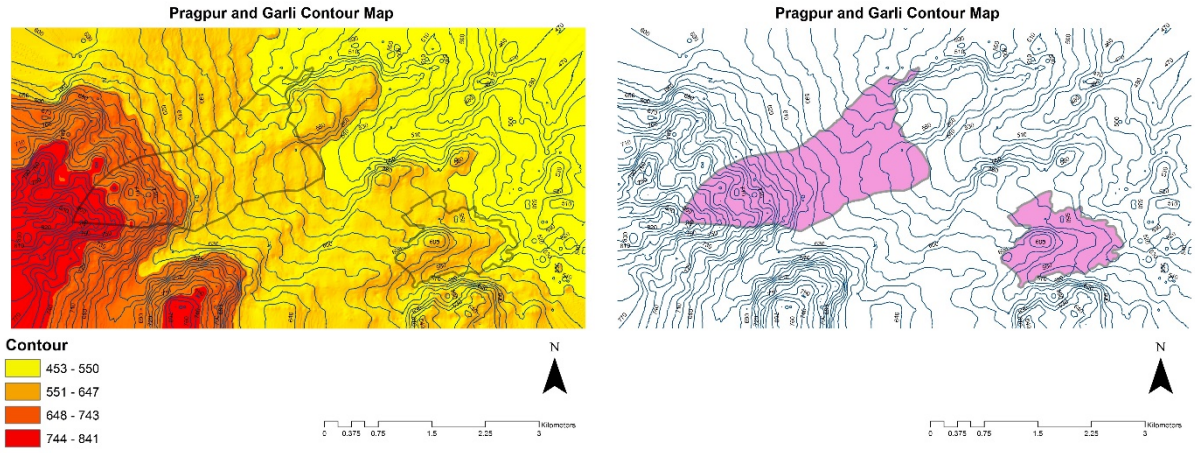


Figure1: Contour Map of Pragpur and Garli geo-referenced on GIS

2. Methodology

In the following study, a mixed method of analysis is conducted to investigate the status quo of these villages and also to understand the perception of the residents. In this mixed method of analysis, a three-stage layered analysis is conducted where the first stage is to conduct a spatial analysis using GIS, and then a descriptive and empirical analysis is conducted upon the responses received from the entire sample area. Questionnaire Surveys and interview-based surveys were used as tools to obtain variables. The relevance and public reaction to the concept of being a "Heritage village "models and theories regarding the sustainable built environment were questioned primarily. A total of 100 responses were gathered from these two villages and Turkey Post-Hoc test was conducted on the responses gathered using a five-point Likert scale. The questionnaire was developed on seven major parameters as Economic Viability, Infrastructure, Heritage, Social and cultural Character, Governance/ Policy framework, Environment, Settlement Pattern.

3. Study Area

The villages of Pragpur and Garli in Himachal Pradesh are situated at two thousand feet which are approx. 550 Meters above sea level, it comes under the region of district Kangra, Himachal Pradesh. The heritage-based settlement of Pragpur and Garli with a pristine view ideally suited to explore the distinct characteristics of Kangra valley. This area has several water channels that mix into river Beas with its pleasant and comfortable climate, better connectivity, rich flora, and fauna (Figure 1). The Dhauladhar mountain range raised behind this picture of pastoral perfection and needy frames it with tall peaks that are draped with snow for the better part of the year. Recognized as Heritage Village by a notification dated 9th December 1997 the state Government has classified Pragpur and Garli as a "Heritage Zone".

Heritage villages of Pragpur and Garli also fall under the ecozone, ideal for the development of tourism. These villages are also promoted by the administration as an example for community development in the hospitality sector. The core area of Pragpur is notified as "Heritage Villages" by the state government. This makes it one of the villages in India with such a distinctive title. Understanding the unique architectural style and settlement pattern the Government has declared the creation of SADA (Special Area Development Authority) for improving the infrastructure in the area in the field of tourism[4][5].

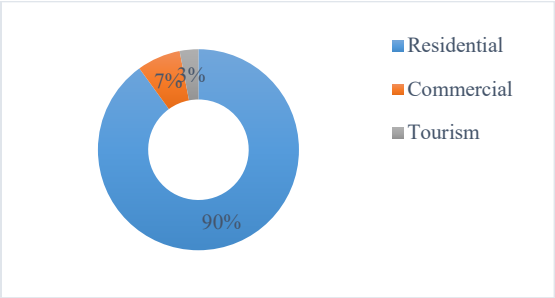


Figure 2: Land use percentage for the village of Pragpur and Garli

4. Architectural features of the built form

Pragpur and Garli can be divided into three major sectors based on the activities, such as residential, commercial, and tourism-based buildings (Figure 2). The Residential area is the zone where all the heritage-based residences are built[6]. The commercial area is where all the small-scale activities such as retail activity, weavers, shoemakers, etc. are housed. Tourism based infrastructure is where the residential houses have been converted into small scale hotel. Some of the major features that can be observed within the settlement are,15ft. that main street is around 15 feet in width which also acts as the main spine for connectivity from the urban centres.7-10 ft wide secondary streets are used for core connectivity towards the interior of the settlement having around 6010 feet width. Low height structures can be observed along the main street usually used for retail purposes.

The layout of the streets is in a linear pattern and the secondary roads radiate from the main street along the natural contour of the settlement. The main axial road of the village consists of built with variable land uses. As per the land use characteristics observed it was seen that the ground floors are used for commercial and retail. The lower floors are designed such as to create a recessed structure and the above floors act as cantilevers. A row of the columnar corridor is constructed supporting the projected floors of the upper storey. A colonnaded facade is recessed which provides easy entry and forms a semi-

transparent visual screen that creates a unified image of the structures. These plinth areas or semi-open areas act as recreational spots for the shop owners. The upper cantilever storey is used for residential purposes. Most of the residential and commercial structures are not more than two stories. A series of residences having vernacular architectural characteristics are built along the connecting streets of the village which diverge from the main spine. Many of these houses have a central courtyard. The common material used for the construction of these heritage houses are mud plaster, dry and seasoned bamboos with a layer of bitumen for roofing purposes. In the case of repetitive floors, the floors are finished with mud mixed with reed or bhusa to create a thick paste-like consistency. Wood is primarily used for doors and windows. The balustrades on all floors are constructed using seasoned timber. The doors, windows, and even the railings have magnificent ornamentation and carvings designed on them. The door frames are also adorned with carved patterns such as local flowers and flowering buds, animal frescos, and geometrical patterns. Window and door arches are constructed to support the lintel beams. A mixed variety of arches are observed such as segmental arches, flat arches, gothic spires, etc. The chajjas over the openings are also intricately carved. The streets are designed using cobbled stone in interesting patterns. The verandas are also constructed using stones in various patterns like linear, rectangular and circular patterns. The houses are arranged in a row formation around an open to sky courtyard. On the upper floors, the windows are reduced gradually in size. Half-timbered construction is mainly observed in these villages for the construction of residential buildings.

The half-timber construction method is, in which the external and internal walls are constructed using the timber-framed structure and the space between them is filled with materials such as mud plaster, brick, or daub and wattle. In these hilly regions the half-timber construction is usually made of squared oak timbers joined by rafters, tenons, mortises, and wooden pegs; the building's structural skeleton is strengthened using braces in the corners (Table 1).

To understand the evolution of the rural settlement and to catalogue its existence in the rural areas, the help of GIS (Geographic Information System) is taken. First, we created a base map of the spatial database of the Garli village. Then specific attributes such as the timeline of the construction, the year of the last restorations or reconstructions over time, etc. were attached to each vector. The main objective was to identify the residential buildings which were constructed through vernacular style and which residential buildings were renovated with their material study used in construction. In this way, it was possible to generate several maps that emphasize the old village system and the houses which have been abandoned by its residents (Figure 3 and 4). This attribute and the database are updated in real-time, and in this way the government can decide their restoration policies in a more precise way, comparing with the traditional methods. The updated data regarding the preservation and conservation status of residential houses represent an important factor in monitoring, protection and documentation of local heritage and, at the same time, offer the possibility to involve the local community in heritage management for sustainable development.

Table 1: Visual Survey of Pragpur and Garli depicting architectural features

Parameters	Images	Architectural Features
Settlement pattern	 	1. Streets paved with dressed cobbled stone.
Buildings	 	2. Gothic spires 3. Gables with decorative and ornate finials 4. Wooden pillars and heavy wooden beams 5. Roofs are sloping, made of local slate on wooden framework 6. Gabbled and cross gabbled
Materials	 	7. Mud plastered 8. Slate roofed houses 9. Bamboo 10. Reed or bhusa 11. Cobbled stone

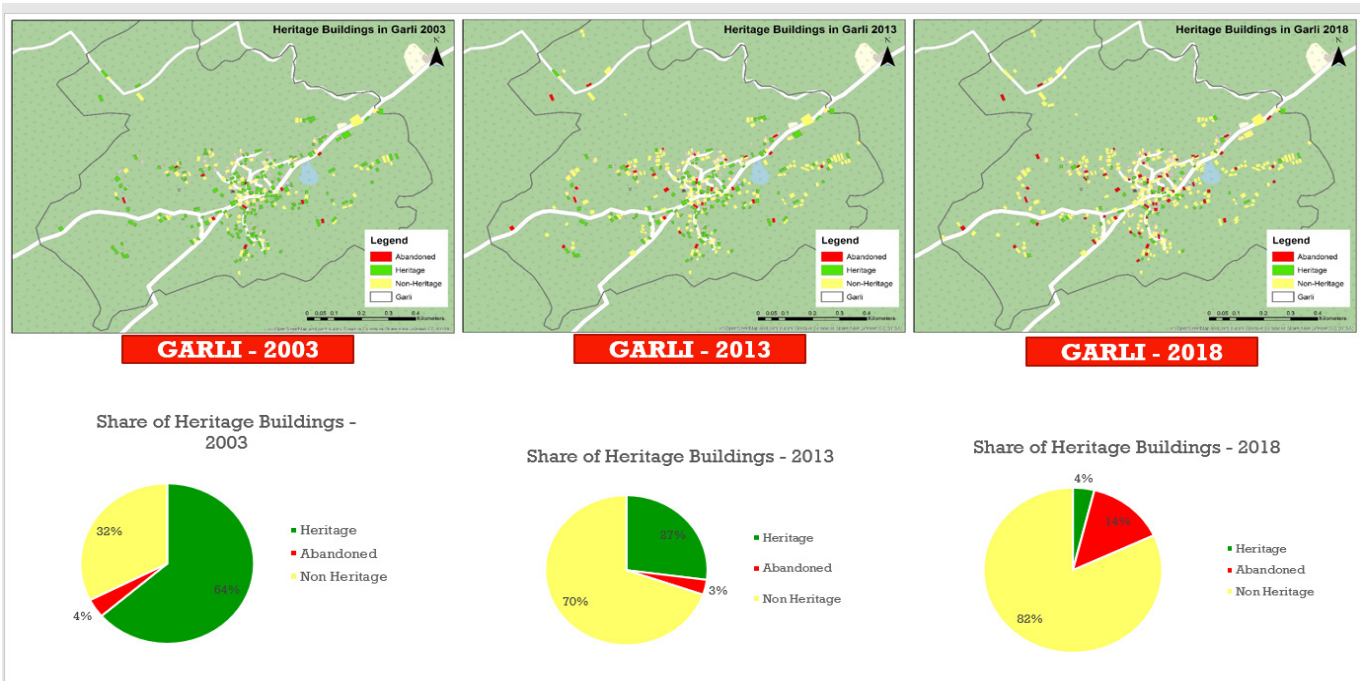


Figure 3- Settlement mapping of built heritage for year 2003,2013,2018 of Garli

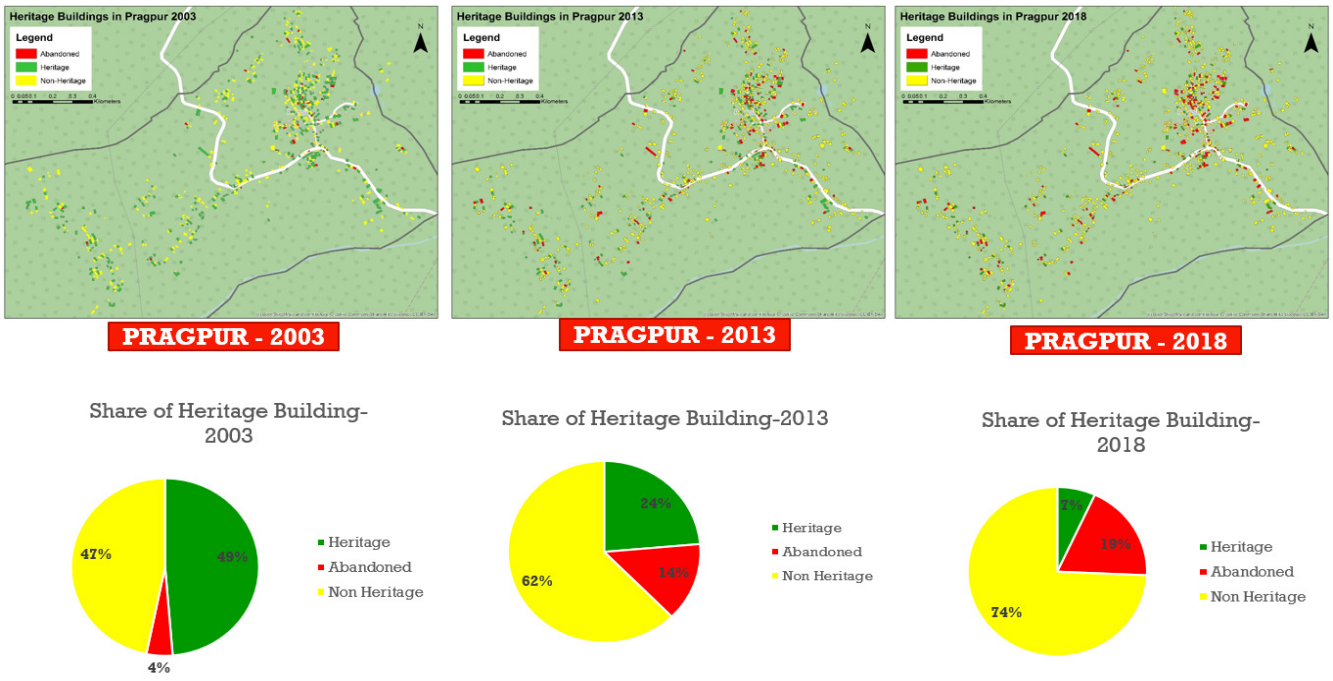


Figure 4: Settlement mapping of built heritage for year 2003,2013,2018 of Pragpur

5. Data Collection

To attain a holistic view, 100 interviews (50 in Pragpur and 50 in Garli) were carried out with groups of specialists (including academicians and conservationist), cultural reference groups (including community heads or chieftains, heritage managers, cultural groups, the private hospitality sector, and local NGOs). It is also important to understand that these traditional villages are located near the core and buffer zones of their urban district[7]. The composition of the sample of respondents was taken according to people who have a direct impact due to the development of these villages (Table 2).

The respondents were asked varied questions regarding the status of a "Heritage Village". Also, it reflects the understanding of the respondents as to whether after getting the said status whether the employment opportunities have increased[8]. This study also aims to enrich the literature with an empirical investigation on issues and challenges in developing a sustainable model of heritage village projects[9]. Thus, the objectives are to identify the most common and fundamental issues and challenges that currently exist in these villages and to determine if these issues can be resolved with certain interventions[10].

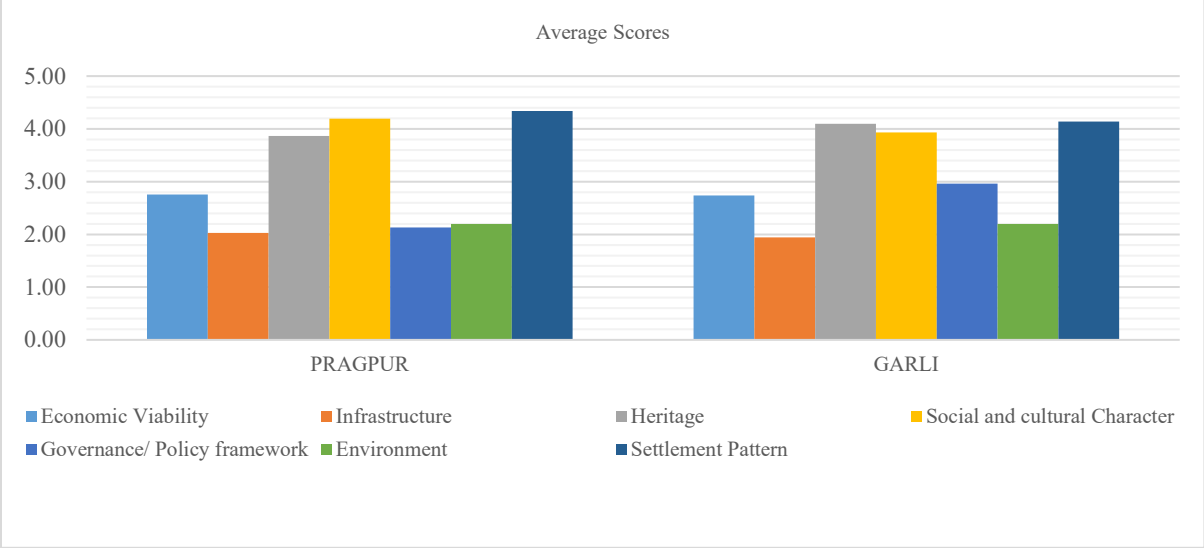


Figure 5: Descriptive analysis for the responses

Table 2: Sample size for the questionnaire survey in Pragpur and Garli

	Pragpur	Garli
Total Residential Buildings	354	118
Heritage Building	93	21
New Construction	261	97
Abandoned Residences	67	17
Sample Size for Survey	50	50

5.1. Statistical analysis

Statistical analysis is performed by using descriptive and inferential statistics. Pearson correlation was applied to see the relationship between the variables. One-way Anova followed by Tukeys HSD post-hoc test for multiple comparisons was applied to see the mean difference between the villages.

Correlation in the broadest sense is a measure of an association between variables. In correlated data, the change in the magnitude of 1 variable is associated with a change in the magnitude of another variable, either in the same (positive correlation) or in the opposite (negative correlation) direction. The Pearson correlation coefficient is typically used for jointly normally distributed data. Then to understand the study One-way Anova was used.

- To compute large and varied data sets, which included data collected from two different villages and having varied responses. It was also done to verify the variability of the data acquired.
- To test which method yields the most suitable result in the sample of data collected. The sample consisted of stakeholders comprising of the residents, government officials, and tourists in these two villages.

If data are analyzed using ANOVA, and a significant F value obtained, a more detailed analysis of the differences between the treatment means will be required. In this case, the P-value is less than 0.05 considered as significant at 95% confidence level.

The statistical software SPSS version 24.0 was used in the analysis. The questions in the questionnaire invited them to indicate the level of agreement of each factor in relation to the seven main parameters based on a five-point scaling, i.e., Strongly Agree = 5, Agree = 4, neutral = 3, Disagree = 2, and Strongly disagree = 1.

6. Result and Discussion

The Economic Viability parameter, the focus was to correlate the economic fabric [11]of these villages as a community and whether this can act as a prominent factor that governs the sustainable development of a rural area[12]. The Infrastructure parameter focuses on the infrastructure needed for developing a sustainable village model which may include roads, tourism infrastructure, connectivity, etc. The parameter also addresses the negative impact of creating an ideal village model and that may hamper the heritage fabric of the villages[13]. Understanding the relationship and assessing the loss and degradation of the core's-built heritage is the focus behind the Heritage parameter. This is to analyse, investigate, and discuss, the concerns, obstacles, and issues of sustainability to this rural heritage conservation. In the Social and cultural Character parameter the focus is to examine the influence of socio-cultural factors on the formation of architectural spaces in Heritage based villages by applying a descriptive-analytical method and using literature study.

The results of this study indicate that the structural system of traditional houses is based on a hierarchical system and is intricate for the development of the settlement. The Governance/Policy Framework parameter attempts to focus on the holistic system or dynamic model to analyze the interaction of the social–economic–heritage sectors to examine policies to achieve the goal of development and conservation. The Literature study shows that conservation alone cannot lead to its own goal. Economic development strategies or policies could be a better choice if the development is carefully planned and effectively controlled to avoid over-consuming resources.

Table 3: Statistical analysis

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	p-value
						Lower Bound	Upper Bound			
Economic Viability	PRAGPUR	50	19.28	3.33	0.47	18.33	20.23	14.00	32.00	<0.001
	GARLI	50	19.18	2.97	0.42	18.33	20.03	12.00	25.00	
Infrastructure	PRAGPUR	50	16.20	2.74	0.39	15.42	16.98	11.00	23.00	<0.001
	GARLI	50	15.54	2.21	0.31	14.91	16.17	11.00	20.00	
Heritage	PRAGPUR	50	23.20	3.16	0.45	22.30	24.10	15.00	29.00	<0.001
	GARLI	50	24.58	3.14	0.44	23.69	25.47	15.00	29.00	
Social and cultural Character	PRAGPUR	50	16.78	1.64	0.23	16.31	17.25	13.00	20.00	<0.001
	GARLI	50	15.74	2.42	0.34	15.05	16.43	10.00	20.00	
Governance/ Policy framework	PRAGPUR	50	4.26	1.69	0.24	3.78	4.74	2.00	7.00	<0.001
	GARLI	50	5.92	2.15	0.30	5.31	6.53	2.00	9.00	
Environment	PRAGPUR	50	2.20	1.03	0.15	1.91	2.49	1.00	4.00	0.017
	GARLI	50	2.20	0.97	0.14	1.92	2.48	1.00	4.00	
Settlement Pattern	PRAGPUR	50	4.34	0.80	0.11	4.11	4.57	3.00	5.00	0.012
	GARLI	50	4.14	0.93	0.13	3.88	4.40	2.00	5.00	

Understanding the influence of human behavior on the landscape is the prime objective of the Environment parameter. Also, to analyze the reason for village landscape transition which come from road construction, the transformation of building materials, higher living level, the inflection of family structure and social relationship, different culture and self-identity [14]. To investigate or to identify the challenges and hurdles towards the growth or deterioration of the settlement of heritage villages, it was gathered that through cataloguing the settlement pattern the evolution of the community and its historical background can be documented with its various changes in all its bearings.

To understand the aim of the Economic Viability parameter was to determine their economic status of the respondents and whether they have benefitted from the status of "heritage village" [15]. The respondents of the villages of Garli and Pragpur had major concerns while developing their residential buildings. The families owning these houses lacked the economic resources and finances to maintain these houses [16]. While enquiring about the infrastructure development the residents were not content and strongly disagrees with the increase in any land value in villages of Garli and Pragpur in Himachal where the circle rate for 100 sq. meters are around 886 Rs and 1751 Rs (District Kangra, Government of India website) respectively from last one decade.

It indicated the stagnation in terms of development. In these villages, it was also observed that mostly the migrant population is residing in the heritage houses were on a rental basis [17]. They are mostly not aware of the heritage status and do not have the economical means to maintain these residential buildings. The infrastructure needed to sustain and develop a heritage village was not visible in these Villages.

Community Participation has been observed as a major factor for the sustainable development of a settlement [18]. In villages like Garli, some residential patrons have tried to restore their residential buildings and convert them into commercial and recreational spaces for the tourist, parallel developing the village and generating employment for the locals. The basic amenities such as fire protection and police protection also lack in these villages. The villagers were governed under Panchayat System and they themselves resolve the petty issues related to the governance of the villages [19].

The villages lacked the basic infrastructure to support a tourist influx. It was determined by the study that, many people outside of these villages are not aware of the heritage status. The upkeep the vernacular architectural style of the residences, not many skilled labours are available. The construction and maintenance in these villages are done according to the new construction style using materials such as RCC and Bricks. In Pragpur and Garli, timber construction and wood such as teak and sheesham were used for the construction of residential houses.

7. Findings

The major factors highlighted in this study were as detailed below:

- The first major criteria were to identify the heritage-based settlements through the extensive literature study. A multidisciplinary approach was initiated where along with the architectural characters certain parameters were combined such as economy, social conditions, settlement morphology, etc.
- It was also observed that different policy development schemes or frameworks for the inclusion of rural heritage preservation measures is not highlighted or mentioned. There is no incorporation of heritage-based rural development in the physical planning processes as part of a holistic approach towards sustainable development [20]. This research wants to focus on and promote the inclusion of preservation of heritage-based residential buildings representing the local architectural style.

- Finally the study wants to highlight the importance of heritage in rural development. Through the descriptive and empirical study of the data, it was found that for the development of these types of settlements the role of the local residents and community plays an important part [21]. There is a need for introducing public awareness as a tool to generate employment in these rural areas.

The data acquired through the questionnaire survey along with the interview survey gave many insights to identify the hurdles and challenges while implementation of heritage-based development policy. A multidisciplinary approach as mentioned earlier can be used where collaboration with tourism, hospitality industry all related sectors can be initiated [22]. This was considered important to resolve any issues which may cause the undermining preservation of the built resources. It is also stressed that preservation of cultural and natural heritage should be perceived together and they must not be promoted with activities that can cause hindrances towards sustainability. To make the residents conscious about the values of rural heritage certain programs can be initiates such as studies on vernacular architecture, promotion of knowledge, and training about traditional techniques and materials.

It was also observed during the survey that the majority of participants were of the suggestion that a balance should be maintained between the retained heritage-built fabric and its economic viability. The entire exercise was conducted to understand the importance of the residential buildings and should actually be protected. Secondly, it was to address its importance for the future generations of these villages. The second factor also ponders upon the different methods of adaptability and reuse of these heritage buildings. The process also helped to understand the hurdles along with the decision-making towards the protection and conservation of these buildings. Many respondents have also agreed that the policymakers and government officials should ensure that if the value of a residential building is under heritage significance then relevant heritage policies and bye-laws should be formulated.

8. Conclusion

Through the extensive literature study regarding the preservation of this built heritage, the researcher had identified core parameters that are vital for sustainable development. The qualitative study and empirical study evaluated the study and the observation done to understand the ground reality. The major parameters were categorized into several sub-parts depending on the focus, such as historical value, economic viability, infrastructure, socio-cultural characteristics, and the overall aesthetic value. The research aimed to understand the perception of the local community and analyse their approach and consensus to further develop a holistic policy framework. The detailed analysis led to the determination that the infrastructure lacks the capacity to promote tourism or hospitality industry to create a sustainable model to promote the heritage of that area. In very few cases, the architectural value (vernacular architectural style) is also considered but all are linked to the history of a particular settlement along with its cultural values.

It is evident from the analysis from the sample groups and interviews that diversity plays an important role when discussing the importance of their heritage settlement. It varies across the communities, geographical location. It also varied because of their economic status, socio- cultural parameters as seen in these villages of Pragpur and Garli. As per the investigation, it was also highlighted that administration plays an important role as policymakers and implementors. It is important for the local government to create awareness towards the importance of these heritage-based residential buildings and that they may gain value in the future.

9. References

[1] S.H. Zolfani, E.K. Zavadskas, "Sustainable development of rural areas' building structures based on local climate," Procedia Engineering, **57**, 1295–1301, 2013, doi:10.1016/j.proeng.2013.04.163.

[2] E. Erdogan, S.D.B. Erkiş, "Sille settlement in the context of sustainable historical fabric and façade analysis of its traditional houses," Archnet-IJAR, **8**(3), 117–135, 2014, doi:10.26687/archnet-ijar.v8i3.214.

[3] I. Syahrul, M. Radzuan, N. Fukami, Y. Ahmad, "Cultural heritage, incentives system and the sustainable community: lessons from Ogimachi Village, Japan.," Cultural Heritage, Incentives System and the Sustainable Community: Lessons from Ogimachi Village, Japan., **10**(1), 130–146, 2014.

[4] R.P.B. Singh, M. Fukunaga, "The {World} {Heritage} {Villages} of {Shirakawa}-{Gō} and {Gokayama}, {Japan}: {Continuing} {Culture} and {Meeting} {Modernity}," Heritagescape and Cultural Landscapes. Planet Earth & Cultural Understanding Series, Pub, **6**(Special Lecture 3), 129–150, 2010.

[5] S. Kim, "World Heritage Site designation impacts on a historic village: A case study on residents' perceptions of Hahoe Village (Korea)," Sustainability (Switzerland), **8**(3), 2016, doi:10.3390/su8030258.

[6] C. Holtorf, "Embracing change: how cultural resilience is increased through cultural heritage," World Archaeology, **50**(4), 639–650, 2018, doi:10.1080/00438243.2018.1510340.

[7] N. Mitsche, F. Vogt, D. Knox, I. Cooper, P. Lombardi, D. Ciaffi, "Intangibles: Enhancing access to cities' cultural heritage through interpretation," International Journal of Culture, Tourism, and Hospitality Research, **7**(1), 68–77, 2013, doi:10.1108/17506181311301381.

[8] N.A. Haddad, L.A. Fakhoury, "Towards developing a sustainable heritage tourism and conservation action plan for irbid's historic core," Archnet-IJAR, **10**(3), 36–59, 2016, doi:10.26687/archnet-ijar.v10i3.1035.

[9] S. Prompayuk, P. Chairattananon, "Preservation of Cultural Heritage Community: Cases of Thailand and Developed Countries," Procedia - Social and Behavioral Sciences, **234**, 239–243, 2016, doi:10.1016/j.sbspro.2016.10.239.

[10] R. Askarizad, "Influence of Socio-Cultural Factors on the Formation of Architectural Spaces (Case Study : Historical Residential Houses in Iran) Influence of Socio-Cultural Factors on the Formation of Architectural Spaces," **1**(May), 44–54, 2019.

[11] O.S. Asfour, A.T. Hathat, "Strategies of Sustainable Rural Development in the Gaza Strip : Wadi Gaza Town as a Case Study STRATEGIES OF SUSTAINABLE RURAL DEVELOPMENT IN THE GAZA STRIP : WADI GAZA TOWN AS A CASE STUDY," (August), 2016.

[12] G.A.. Tanguay, E. Berthold, J. Rajaonson, "A Comprehensive Strategy to Identify Indicators of Sustainable Heritage Conservation," Les Cahiers Du CRTP (Workin Paper), (September), 2014.

[13] S. Roy, "Regional Disparities in Growth and Human Development in India," (September), 2012.

[14] I.S. Mat Radzuan, S. Inho, Y. Ahmad, "A rethink of the incentives programme in the conservation of south korea's historic villages," Journal of Cultural Heritage Management and Sustainable Development, **5**(2), 176–201, 2015, doi:10.1108/JCHMSD-02-2014-0006.

[15] M. A. S., "Efficient Rural Development Strategies for the Improvement of Indian Economy," International Journal of Pure & Applied Bioscience, **5**(6), 566–570, 2017, doi:10.18782/2320-7051.5462.

[16] N.A. Haddad, F.Y. Jalboosh, L.A. Fakhoury, R. Ghrayib, "Urban and rural umayyad house architecture in Jordan: A comprehensive typological analysis at Al-Hallabat," Archnet-IJAR, **10**(2), 87–112, 2016, doi:10.26687/archnet-ijar.v10i2.835.

[17] I.K. Suarthana, N. Madiun, S.O. Yuniarsa, "Exploring The Community Participation, Tourism Village, And Social-Economic To Environment Impact (Case Study: Pentingsari Village, Yogyakarta)," International Journal of Business and Management Invention, **4**(9), 85–90, 2015.

[18] A.J. Imperiale, F. Vanclay, "Using Social Impact Assessment to Strengthen Community Resilience in Sustainable Rural Development in Mountain Areas," Mountain Research and Development, **36**(4), 431–442, 2016, doi:10.1659/MRD-JOURNAL-D-16-00027.1.

[19] I.S. Mat Radzuan, Y. Ahmad, "Synthesising an effective incentives system in safeguarding the heritage village of melaka and george town," Planning Malaysia, (5), 157–168, 2016, doi:10.21837/pmjournal.v14.i5.200.

[20] X. Honggang, "Exploring Sustainable Policies for Xidi, the World Heritage Village ," Proceedings of the 28th International Conference of the System Dynamics Society, 1–16, 2010.

[21] J. Hosagrahar, J. Soule, L.F. Girard, A. Potts, "International Council on Monuments and Sites (ICOMOS) Cultural Heritage, the UN Sustainable Development Goals, and the New Urban Agenda," 2016.

[22] L. Yan, C. Liang, H. Zhiwen, "Conservation of Historical and Cultural Towns and Villages in China," Aussie-Sino Studies, **3**(4), 85–97, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1120-1127 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Electronically Tunable Triple-Input Single-Output Voltage-Mode Biquadratic Filter Implemented with Single Integrated Circuit Package

Natchanai Roongmuanpha, Taweepol Suesut, Worapong Tangsrirat*

School of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok 10520, Thailand

ARTICLE INFO

Article history:
Received: 30 September, 2020
Accepted: 31 January, 2021
Online: 25 February, 2021

Keywords:
Commercially available IC
LT1228
Voltage mode circuit
Biquadratic filter

ABSTRACT

This article proposes a compact and simple design of electronically adjustable voltage-mode biquadratic filter using fundamental active cell implemented on a single integrated circuit (IC) package as LT1228. The proposed circuit having triple inputs and single output (TISO) employs namely one resistor and two capacitors as the passive components. All the five possible biquadratic filtering responses, namely low-pass (LP), band-pass (BP), high-pass (HP), band-stop (BS) and all-pass (AP), are realized by the appropriate selection of the relevant input signals. The pole angular frequency and the quality factor of the proposed TISO filter are electronically tunable through the bias current of the IC chip LT1228. Non-ideal effects and sensitivity performance are carried out. The theoretical results are satisfactorily validated by both PSPICE simulation results and experimental measurements using commercially available LT1228.

1. Introduction

Over the decade, analog filters always play a role in many important analog signal processing applications, i.e. communication systems, measurement and instrumentation systems, etc. Nowadays, the realization of an active analog filter using versatile active building blocks has been focused by many researchers due to many advantage features, such as simple circuitry, high linearity, and wide dynamic range. In the literature, many modern active electronic elements have been utilized in analog active filter design, such as current conveyor (CC) [1–7], differential difference current conveyor (DDCC) [8–12], differential voltage current conveyor (DVCC) [13–16], fully differential second-generation current conveyor (FDCCII) [17], current differencing buffered amplifier (CDBA) [18–20], current feedback operational amplifier (CFOA) [21–27], current follower transconductance amplifier (CFTA) [28–29], operational transconductance amplifier (OTA) [30–34], voltage differencing buffered amplifier (VDBA) [35–36], voltage differencing inverting buffered amplifier (VDIBA) [37], fully balanced voltage differencing buffered amplifier (FB-VDBA) [38], voltage differencing transconductance amplifier (VDTA) [39–41], and voltage differencing gain amplifier (VDGA) [42–44]. However, so many of them require at least two or more active elements for their realizations [1–5, 7, 8, 10, 11, 13, 18–27, 30–39, 41, 43, 44]. Moreover, the voltage-mode filters presented in [1–17, 19–29, 38, 42] need a large number of passive resistors, while the articles in [1, 4, 16, 22] also contain three passive capacitors. It is also to be emphasized that the realizations of [1–27] suffer from the lack of electronic tuning capability of their important parameters. Even though some similar works were developed by based on various active building blocks in either bipolar junction transistor or (BJT) or complementary metal oxide semiconductor (CMOS) technologies, they are not commercially available chips and reachable in general. Besides, the performances of the research developments in [1–4, 7–20, 22, 24, 25, 28–37, 39–44] have been demonstrated through only simulation results.

In this communication, an electronically tunable voltage-mode biquadratic filter with three input and one output terminals (TISO) consisting of only single active IC package LT1228, one resistor and two capacitors is introduced. The proposed TISO filter can realize the five standard biquadratic filtering responses, namely low-pass (LP), band-pass (BP), high-pass (HP), band-stop (BS) and all-pass (AP), all at a single output terminal without modifying a circuit structure. It also provides an electronic adjustability of its pole angular frequency (ω_o) and quality factor (Q) via the external bias current of the LT1228 IC chip. The theoretical propositions are confirmed by PSPICE simulations with LT1228's model parameters, and the simulated results corroborate the theory. In addition, all conclusions discussed in this work are also verified by

*Corresponding Author: Worapong Tangsrirat, Email: worapong.ta@kmitl.ac.th
This paper is an extended version from the proceedings of 2020 8th International Electrical Engineering Congress (iEECON) [45]

https://dx.doi.org/10.25046/aj0601125

the measurement results of an experimentally test circuit with a single IC package LT1228, and the experimental findings are found to be in agreement with the theoretical values.

2. Description of IC Package LT1228

Our design utilizes only one active cell of a commercially available IC LT1228 from Linear Technology Company [46]. An active cell LT1228 is internally a combination of an operational transconductance amplifier (OTA) and a current feedback operational amplifier (CFA) in 8-pin IC package, as demonstrated in Figure 1. This device has three high impedance input terminals (p, n, and z), and one low impedance output terminal (o). It provides the output current i_z at intermediate terminal z which is the difference of two input voltages v_p and v_n ($v_p - v_n$) multiplied by transconductance gain (g_m). An external impedance Z_z is connected to the terminal z, and the potential v_z developed across Z_z will transfer to the output voltage v_o at the terminal o by the CFA. Its ideal terminal characteristics can be described as:

$$\begin{bmatrix} i_p \\ i_n \\ i_z \\ v_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ g_m & -g_m & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_p \\ v_n \\ v_z \\ i_o \end{bmatrix} . \quad (1)$$

Thanks to the LT1228 manufacturing, the g_m-value can be altered to the desired value through the external DC bias current I_B by the following relation: [46]

$$g_m = 10 I_B . \quad (2)$$

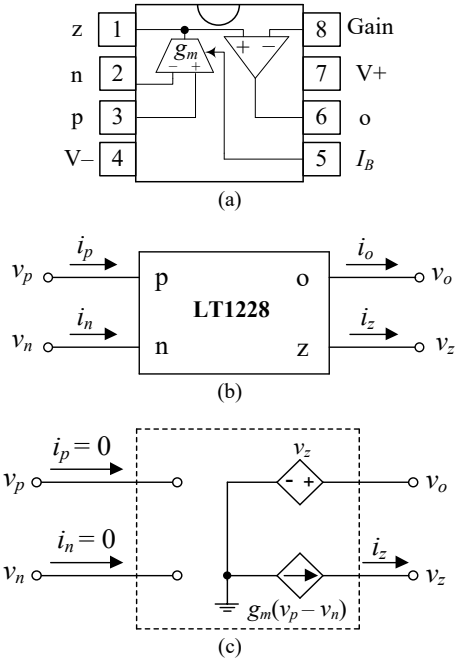


Figure 1: IC device LT1228.
(a) active elements in LT1228 (b) schematic representation
(c) equivalent circuit.

3. Proposed TISO Biquadratic Filter

The realization of an electronically tunable TISO voltage-mode biquadratic filter is given in Figure 2. The proposed TISO filter is implemented with a single LT1228 together with one resistor and two capacitors. A straightforward analysis of the proposed TISO filter reveals the following output voltage function:

$$V_{out}(s) = \frac{s^2 R_1 C_1 C_2 V_3 + s C_1 V_2 + g_m V_1}{D(s)} , \quad (3)$$

where the denominator $D(s)$ is found to be:

$$D(s) = s^2 R_1 C_1 C_2 + s C_1 + g_m . \quad (4)$$

From an inspection of Equations (3)-(4), it appears the five standard biquadratic filter functions can be obtained all at the terminal v_{out} of the proposed circuit by the following conditions.

(i) The LP response is obtained by setting $v_{in} = v_1$ (input voltage signal) and $v_2 = v_3 = 0$ (grounded).

(ii) The BP response is obtained by setting $v_{in} = v_2$ and $v_1 = v_3 = 0$.

(iii) The HP response is obtained by setting $v_{in} = v_3$ and $v_1 = v_2 = 0$.

(iv) The BS response is obtained by setting $v_{in} = v_1 = v_3$ and $v_2 = 0$.

(v) The AP response is obtained by setting $v_{in} = v_1 = -v_2 = v_3$.

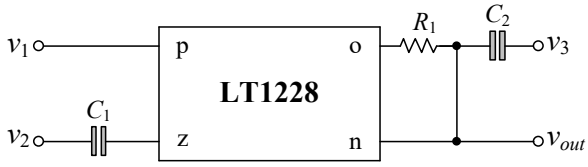


Figure 2: Proposed electronically tunable TISO biquad implementation employing single LT1228.

Therefore, the proposed TISO filter of Figure 2 does not require any element matching conditions or equality constraints for the desired filter function realizations. In all types, the important characteristics ω_o, and Q are respectively found as:

$$\omega_o = 2\pi f_o = \sqrt{\frac{g_m}{R_1 C_1 C_2}} , \quad (5)$$

and

$$Q = \sqrt{\frac{g_m R_1 C_2}{C_1}} . \quad (6)$$

In case of practical design, if $C = C_1 = C_2$, then the ω_o and Q simplify to:

$$\omega_o = \frac{1}{C}\sqrt{\frac{g_m}{R_1}} , \quad (7)$$

and

$$Q = \sqrt{g_m R_1} . \quad (8)$$

In view of the above expressions, the parameters ω_o and Q of the proposed TISO filter can be altered electronically by means of g_m-value. According to Equation (2), the g_m variation can be obtained by an adjustment of the bias current. Also note that since the major contribution of this work is to design a compact and

minimum configuration voltage-mode TISO filter with electronic tunability, an orthogonal control of ω_o or Q is not expected.

4. Non-Ideal Analysis and Sensitivity Performance

In consideration of the non-ideal behavior, the terminal behaviors of LT1228 can be rewritten as:

$$\begin{bmatrix} i_p \\ i_n \\ i_z \\ v_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \alpha g_m & -\alpha g_m & 0 & 0 \\ 0 & 0 & \beta & 0 \end{bmatrix} \cdot \begin{bmatrix} v_p \\ v_n \\ v_z \\ i_o \end{bmatrix} , \quad (9)$$

where $\alpha = (1 - \varepsilon_{gm})$ and $\beta = (1 - \varepsilon_v)$, where $|\varepsilon_{gm}| << 1$ and $|\varepsilon_v| << 1$ are the transconductance inaccuracy and the voltage transfer error, respectively. Taking this effect into account, the characteristics ω_o and Q given in Equations (5) and (6) are modified to:

$$\omega_o = \sqrt{\frac{\alpha\beta g_m}{R_1 C_1 C_2}} , \quad (10)$$

and

$$Q = \sqrt{\frac{\alpha\beta g_m R_1 C_2}{C_1}} . \quad (11)$$

In this case, all sensitivity coefficients of ω_o and Q with respect to the active and passive components are derived and found to be as follows:

$$S_\alpha^{\omega_o} = S_\beta^{\omega_o} = S_{g_m}^{\omega_o} = \frac{1}{2} , \quad (12)$$

$$S_{R_1}^{\omega_o} = S_{C_1}^{\omega_o} = S_{C_2}^{\omega_o} = -\frac{1}{2} , \quad (13)$$

$$S_\alpha^Q = S_\beta^Q = S_{g_m}^Q = S_{R_1}^Q = S_{C_2}^Q = \frac{1}{2} , \quad (14)$$

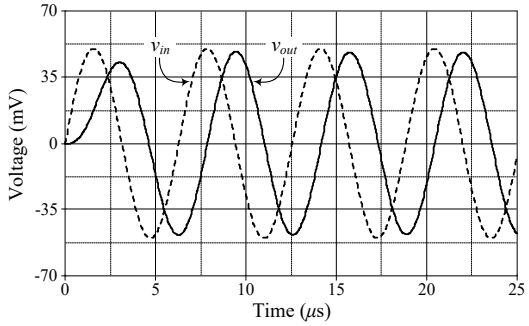
and

$$S_{C_1}^Q = -\frac{1}{2} . \quad (15)$$

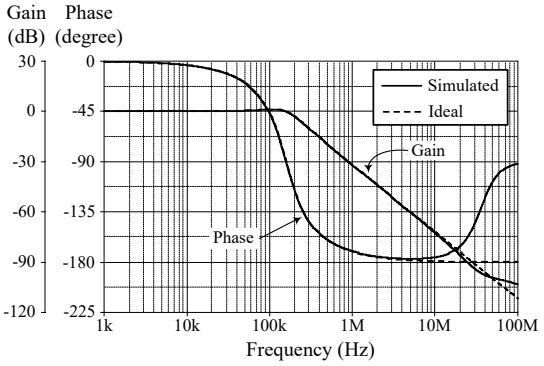
It is clear from Equations (12)-(15) that the absolute values of the ω_o- and Q-sensitivities are all equal to 0.5. These values ensure that the sensitivity performance of the circuit is to be of low value.

5. Simulation Results

In this section, the proposed circuit and its filtering responses are simulated and discussed through the PSPICE simulation program. For ideal simulation, the LT1228 macro-model parameters obtained from Linear Technology Company and DC supply voltages of ±5V were employed. To demonstrate the functionality of the proposed filter, the circuit is designed for f_o = 159.15 kHz and Q = 1. In this case, the various component values have been set as I_B = 100 μA for g_m = 1 mA/V, R_1 = 1 kΩ and C_1 = C_2 = 1 nF. The simulation results for all filter responses are shown in Figures 3-7, which demonstrates very close agreement with the theoretical responses. For time-domain responses, a 159-kHz sine-wave input voltage with 50 mV peak amplitude was applied to the filter. The simulation results show that the error in f_o-value was found to be less than 1%.

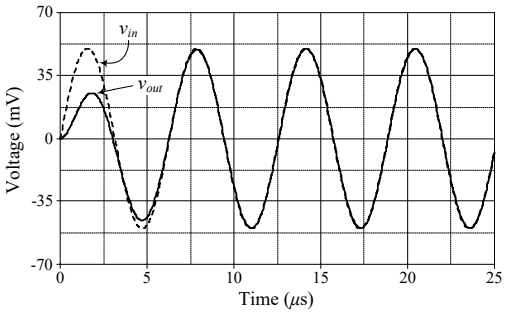


(a)

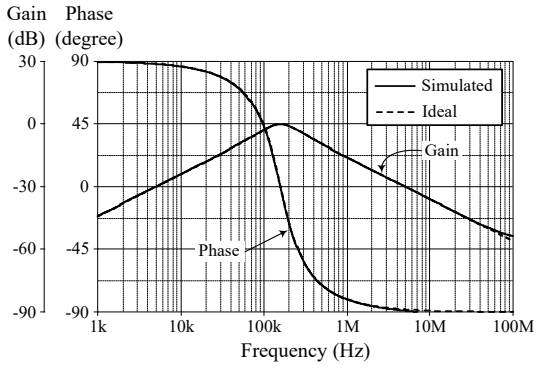


(b)

Figure 3: Ideal and simulated LP characteristics
(a) time-domain responses (b) frequency responses

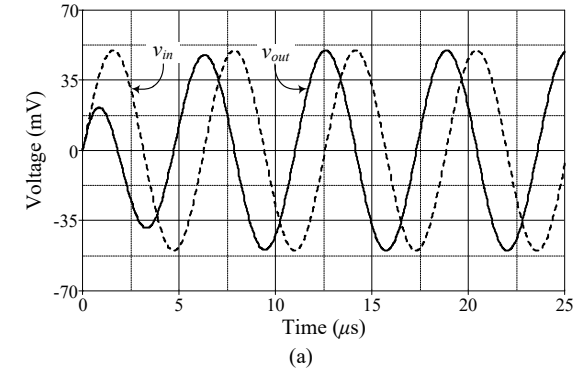


(a)

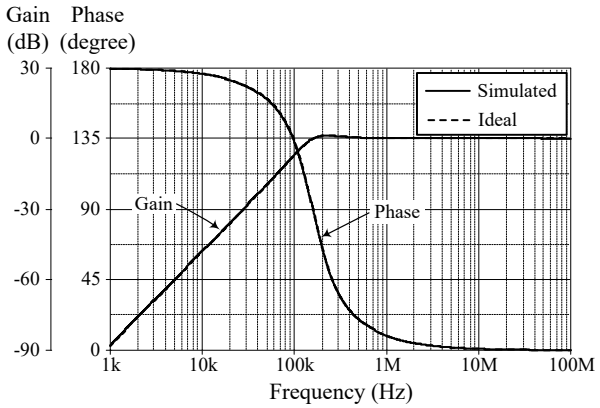


(b)

Figure 4: Ideal and simulated BP characteristics
(a) time-domain responses (b) frequency responses

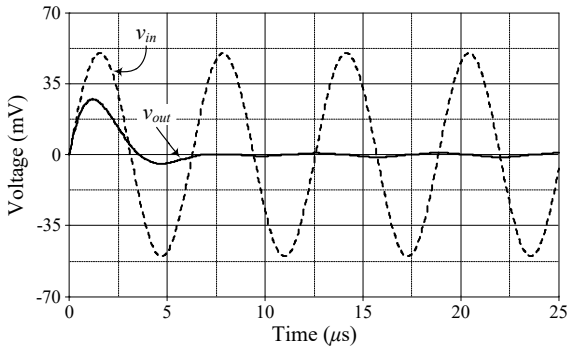


(a)

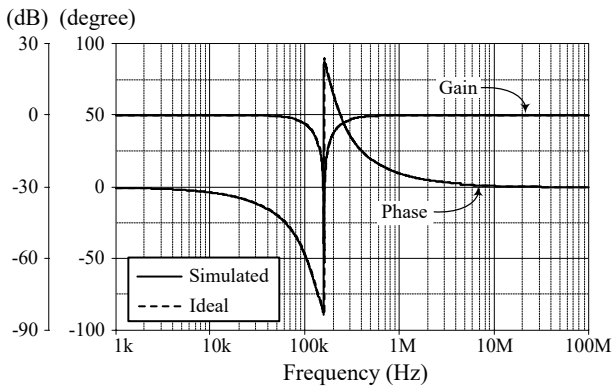


(b)

Figure 5: Ideal and simulated HP characteristics
(a) time-domain responses (b) frequency responses

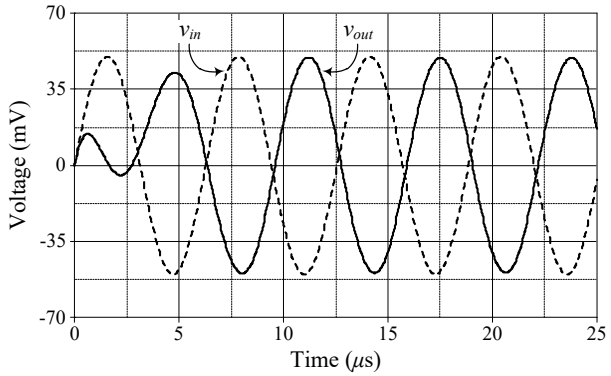


(a)

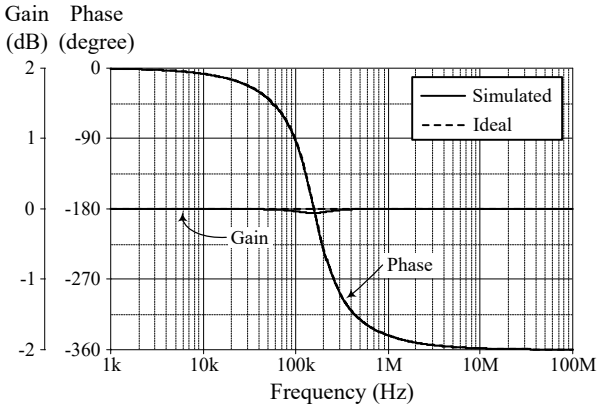


(b)

Figure 6: Ideal and simulated BS characteristics
(a) time-domain responses (b) frequency responses



(a)



(b)

Figure 7: Ideal and simulated AP characteristics
(a) time-domain responses (b) frequency responses

Furthermore, the electronic tuning of gain characteristic for BP filter concerning I_B is observed. The related gain expressions of the proposed BP filter, as shown in Figure 8, are plotted for I_B = 50 µA, 200 µA, and 500 µA, which resulted in g_m = 0.5 mA/V, 2 mA/V, and 5 mA/V, respectively. From Figure 8, the simulation conditions, and corresponding theoretical and simulated f_o and Q are summarized in Table 1.

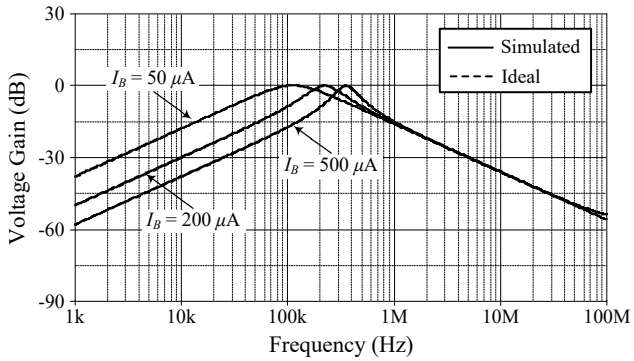


Figure 8: Ideal and simulated frequency responses of the proposed BP filter with an adjustment of I_B.

6. Experimental Results

To further validate the practical workability of the TISO biquadratic filter in Figure 2, the prototype circuit built with readily available IC element LT1228 and discrete passive elements were used to execute experimentally laboratory tests. The circuit was measured using Keysight EDUX1002G digital storage oscilloscope. All of the measured results were performed

Table 1: f_o and Q adjustment of the proposed filter by varying I_B

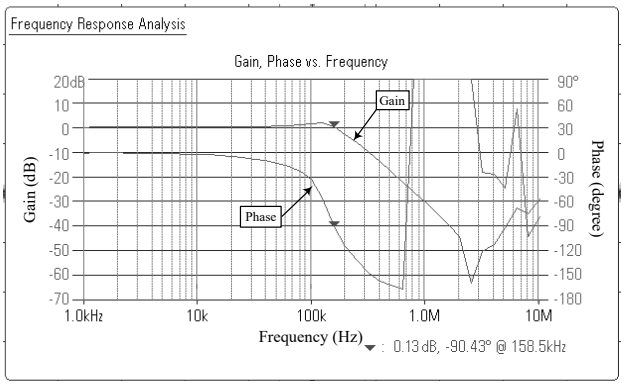
I_B (μA)	g_m (mA/V)	R_1 (kΩ)	C (nF)	Q	f_o (kHz)		% f_o deviation
					Simulated	Ideal	
50	0.5	1	1	0.7	111.43	112.54	0.99
100	1	1	1	1	157.70	159.15	0.91
200	2	1	1	1.4	222.84	225.08	0.99
500	5	1	1	2.2	352.37	355.88	0.99

at symmetrical supply voltages of ±5 V, and I_B = 100 μA (g_m = 1 mA/V), R_1 = 1 kΩ, and $C_1 = C_2$ = 1 nF. This results in f_o = 159.15 kHz and Q = 1. To observed transient response, the measurement was carried out with a 159-kHz sine-wave signal input of 50 mV peak amplitude. The experimental results for the transient and frequency responses as well as the associated frequency spectrums are displayed in Figures 9-13. Also from Figures 9(c)-13(c), the measured results of the percentage total harmonic distortion (%THD) of the v_{out} for each filtering responses are noted in Table 2. It can be concluded that the measured results are close to the theoretical analysis, and also verify the functionality of the proposed circuit.

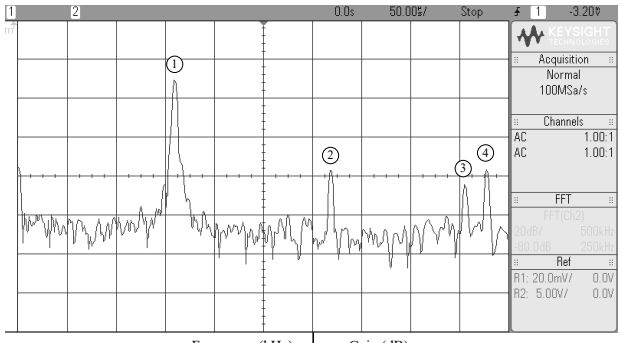
Table 2: Total harmonic distortions of v_{out} in Figure 2.

Filter	THD (%)
LP	0.67
BP	4.47
HP	0.73
BS	2.4
AP	0.32

Another set of measurements have been carried out to examine the electronic adjustability of the proposed TISO filter. BP filter response is used for illustrative purposes. Figure 14 illustrates the measured BP frequency responses for various bias current I_B. The g_m-values of the considered filter have been set as 0.5 mA/V, 2 mA/V, and 5 mA/V, for I_B = 50 µA, 200 µA, and 500 µA, respectively. As follows from Equations (5) and (6), the f_o values have been obtained as 112.54 kHz, 225.08 kHz, and 355.88 kHz, while the Q values have been obtained as 0.7, 1.4, and 2.2, respectively.



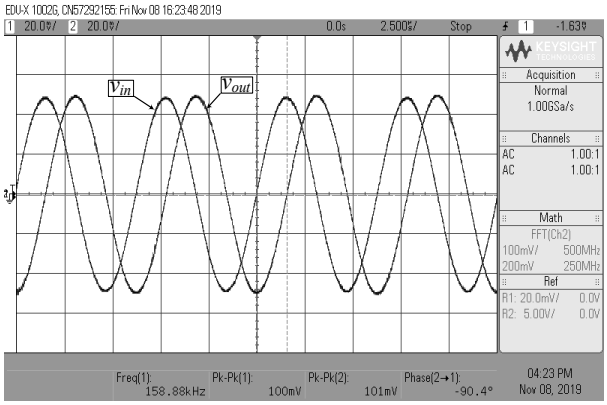
(b)



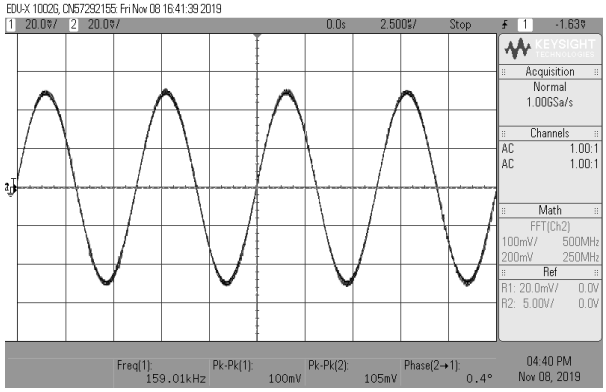
	Frequency (kHz)	Gain (dB)
①	159.1	-30.625
②	318.5	-78.875
③	455.8	-84.375
④	477.7	-76.250

(c)

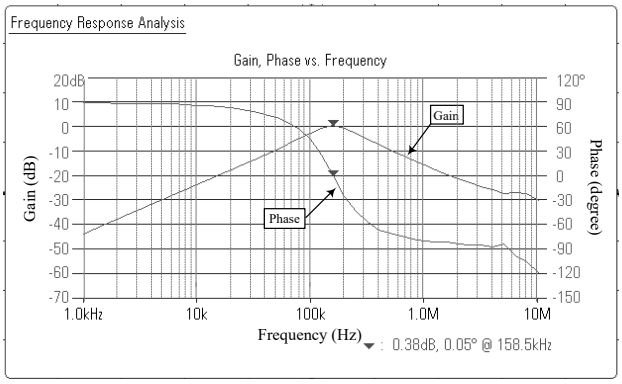
Figure 9: Experimental results of the proposed LP filter.
(a) time-domain responses (b) frequency responses
(c) frequency spectrum



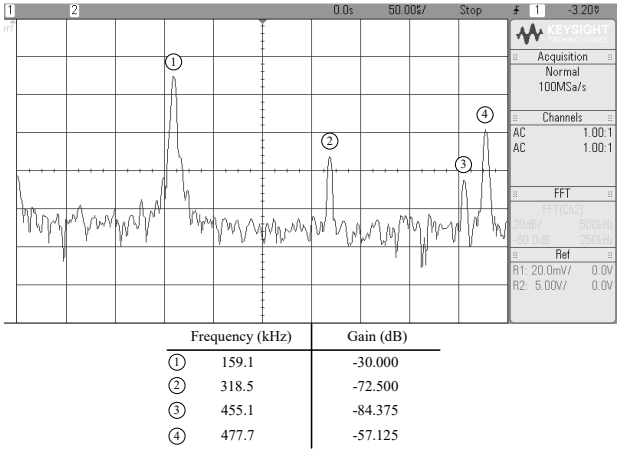
(a)



(a)



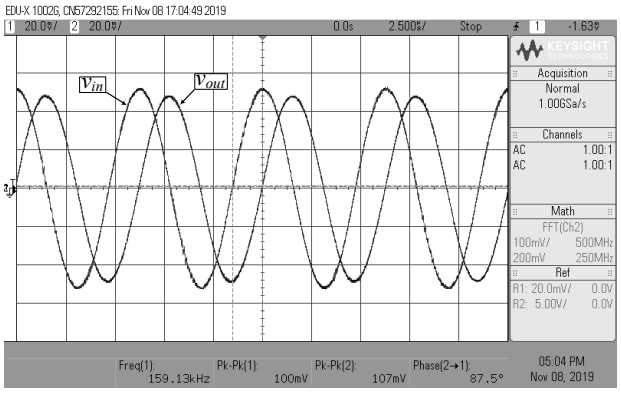
(b)



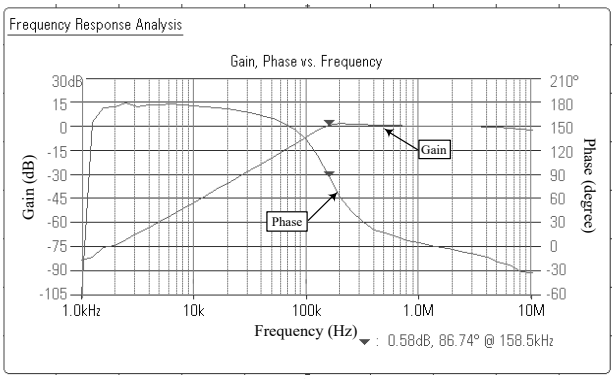
	Frequency (kHz)	Gain (dB)
①	159.1	-30.000
②	318.5	-72.500
③	455.1	-84.375
④	477.7	-57.125

(c)

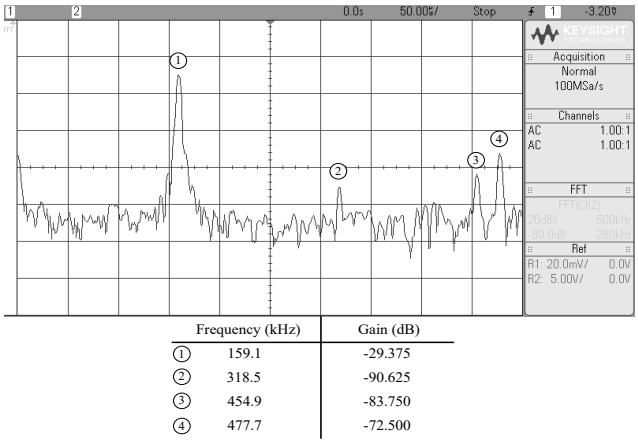
Figure 10: Experimental results of the proposed BP filter
(a) time-domain responses (b) frequency responses
(c) frequency spectrum



(a)



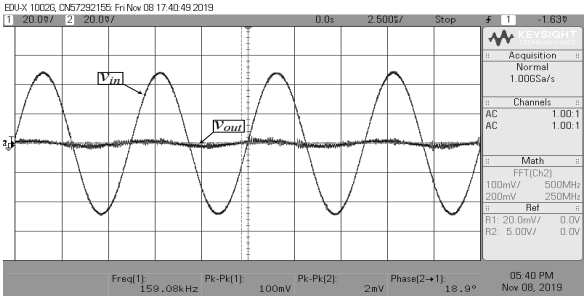
(b)



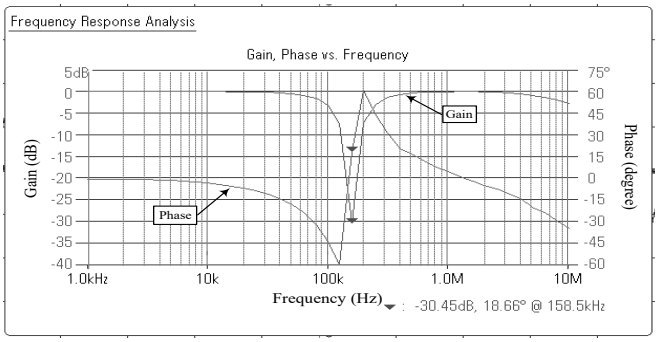
	Frequency (kHz)	Gain (dB)
①	159.1	-29.375
②	318.5	-90.625
③	454.9	-83.750
④	477.7	-72.500

(c)

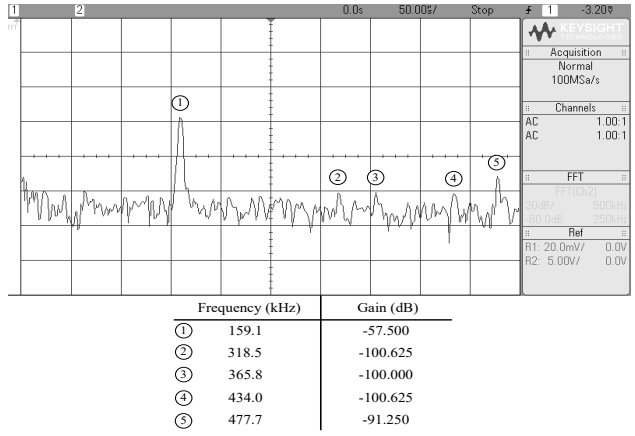
Figure 11: Experimental results of the proposed HP filter
(a) time-domain responses (b) frequency responses
(c) frequency spectrum



(a)



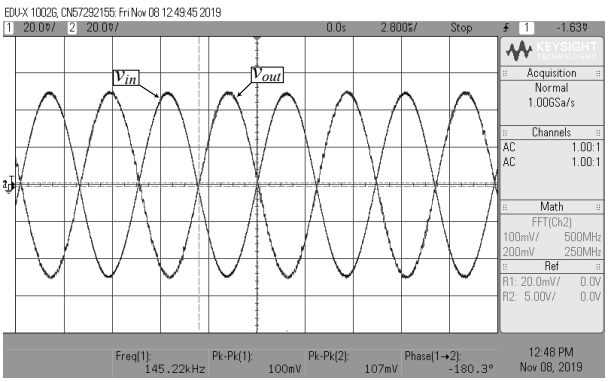
(b)



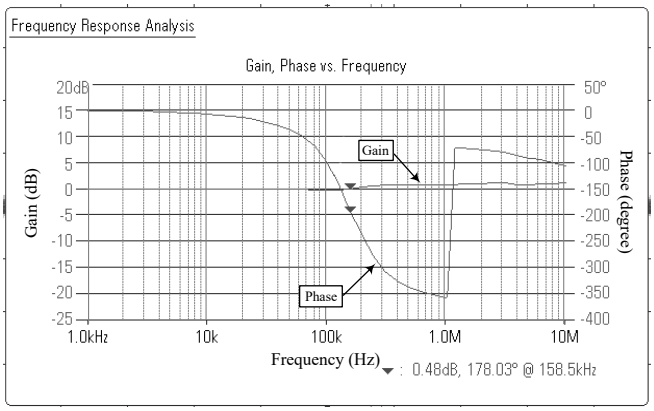
	Frequency (kHz)	Gain (dB)
①	159.1	-57.500
②	318.5	-100.625
③	365.8	-100.000
④	434.0	-100.625
⑤	477.7	-91.250

(c)

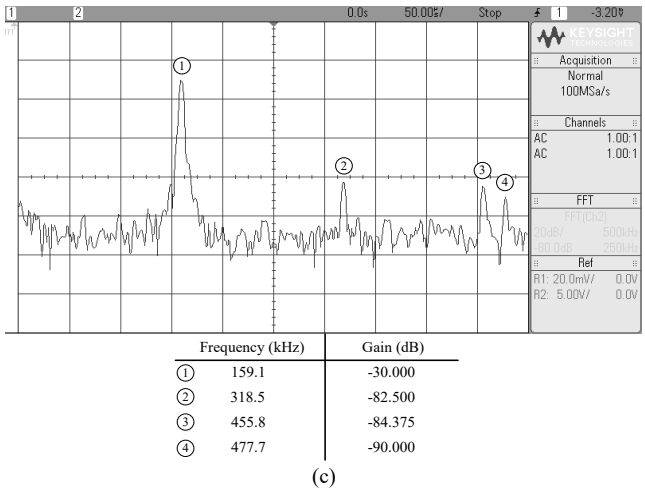
Figure 12: Experimental results of the proposed BS filter
(a) time-domain responses (b) frequency responses
(c) frequency spectrum



(a)



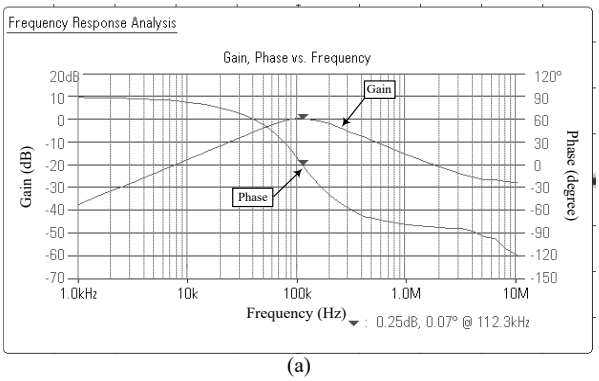
(b)



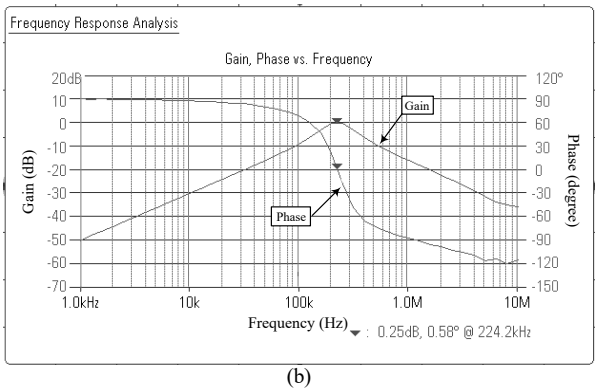
	Frequency (kHz)	Gain (dB)
①	159.1	-30.000
②	318.5	-82.500
③	455.8	-84.375
④	477.7	-90.000

(c)

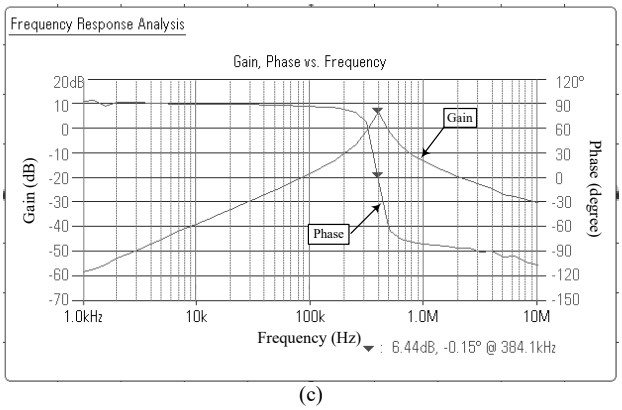
Figure 13: Experimental results of the proposed AP filter
(a) time-domain responses (b) frequency responses
(c) frequency spectrum



(a)



(b)



(c)

Figure 14: Measured gain responses for the proposed BP filter with an adjustment of I_B.
(a) $I_B = 50\ \mu A$ (b) $I_B = 200\ \mu A$ (c) $I_B = 500\ \mu A$

7. Conclusions

This contribution describes the practical implementation of an electronically tunable voltage-mode biquadratic filter with triple input terminals and single output terminal. The proposed filter employs only a single commercially available IC LT1228 together with one resistor and two capacitors. The filter can realize all five standard biquadratic filtering functions all at a single output terminal by an appropriate input signal selection. The characteristics of ω_o and Q can be controlled electronically and linearly in an electronic manner via the external bias current. Simulation results obtained from the PSPICE macro-model of the LT1228 by Linear Technology as well as constructed in prototype hardware using commercially available IC LT1228 are performed to confirm the properties of the proposed circuit.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgment

This work was supported by the Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL) under the contract number 2563-02-01-002.

References

[1] M. Higashimura, "Realisation of voltage-mode biquads using CCIIs" Electron. Lett., **27**(15), 1345–1346, 1991. https://doi.org/10.1049/el:19910847

[2] A.M. Soliman, "Kerwin-Huelsman-Newcomb circuit using current conveyors" Electron. Lett., **30**(24), 2019–2020, 1994. https://doi.org/10.1049/el:19941368

[3] M. Higashimura, Y. Fukui, "Universal filter using plus-type CCIIs" Electron. Lett., **32**(9), 810–811, 1996. https://doi.org/10.1049/el:19960518

[4] J. W. Horng, J. R. Lay, C. W. Chang, M. H. Lee, "High input impedance voltage-mode multifunction filters using plus-type CCIIs" Electron. Lett., **33**(6), 472–473, 1997. https://doi.org/10.1049/el:19970297

[5] C. M. Chang, M. J. Lee, "Voltage-mode multifunction filter with single input and three outputs using two compound current conveyors" IEEE Trans. Circuits Syst.-I: Fundamental Theory and Applications, **46**(11), 1364–1365, 1999. https://doi.org/10.1109/81.802827

[6] J. W. Horng, "Voltage/current-mode universal biquadratic filter using single CCII+" Indian J. Pure & Appl. Phys., **48**(10), 749–756, 2010.

[7] J. W. Horng, Z. R. Wang, C. C. Liu, "Voltage-mode lowpass, bandpass and notch filters using three plus-type CCIIs" Circuits and Systems, **2**(1), 34–37, 2011. https://doi.org/ 10.4236/cs.2011.21006

[8] J. W. Horng, W. Y. Chiu, H. Y. Wei, "Voltage-mode highpass, bandpass and lowpass filters using two DDCCs" Int. J. Electronics, **91(8)**, 461–464, 2004. https://doi.org/10.1080/00207210412331294603

[9] M. A. Ibrahim, H. Kuntman, O. Cicekoglu, "Single DDCC biquads with high input impedance and minimum number of passive elements" Analog Integr. Circ. Sig. Process., **43**, 71–79, 2005. https://doi.org/10.1007/s10470-005-6572-0

[10] W. Y. Chiu, J. W. Horng, "High-input and low-output impedance voltage-mode universal biquadratic filter using DDCCs" IEEE Trans. Circuits Syst.-II: Express Briefs, **54**(8), 649–652, 2007. https://doi.org/10.1109/TCSII.2007.899460

[11] H. P. Chen, "Universal voltage-mode filter using only plus-type DDCCs" Analog Integr. Circ. Sig. Process., **50**, 137–138, 2007. https://doi.org/10.1007/s10470-006-9005-9

[12] W. Y. Chiu, J. W. Horng, "Voltage-mode highpass, bandpass, lowpass and notch biquadratic filters using single DDCC" Radioengineering, **21**(1), 297–303, 2012.

[13] H. P. Chen, S. S. Shen, "A versatile universal capacitor-grounded voltage-mode filter using DVCCs" ETRI Journal, **29**(4), 470–476, 2007. https://doi.org/10.4218/etrij.07.0106.0335

[14] E. Yuce, "Voltage-mode multifunction filters employing a single DVCC and grounded capacitors", IEEE Trans. Instrum. Meas., **58**(7), 2216–2221, 2009. https://doi.org/10.1109/tim.2009.2013671

[15] W. Tangsrirat, O. Channumsin, "Voltage-mode multifunctional biquadratic filter using single DVCC and minimum number of passive elements" Indian Journal Pure & Applied Physics, **49**(10), 703–707, 2011.

[16] J. W. Horng, "Voltage-mode multifunction bi-quadratic filter employing single DVCC", Indian Journal Pure & Applied Physics, **99**(2), 153–162, 2012. https://doi.org/10.1080/00207217.2011.623268

[17] F. Kaçar, A. Yeşil, "Voltage mode universal filters employing single FDCCII" Analog Integr. Circ. Sig. Process., **63**, 137–142, 2010. https://doi.org/10.1007/s10470-009-9440-5

[18] K. N. Salama, A. M. Soliman, "Voltage mode Kerwin-Huelsman-Newcomb circuit using CDBAs" Frequenz, **54**(3–4), 90–93, 2000. https://doi.org/10.1515/FREQ.2000.54.3-4.90

[19] W. Tangsrirat, T. Pukkalanun, W. Surakampontorn, "CDBA-based universal biquad filter and quadrature oscillator" Active and Passive Electron. Components, **2008**, (247171, 6 pages), 2008. https://doi.org/10.1155/2008/247171

[20] J. Pathak, A. K. Singh, R. Senani, "New voltage mode universal filters using only two CDBAs" ISRN Electronics, **2013**(12), (987867, 6 pages), 2013. https://doi.org/10.1155/2013/987867

[21] S. I. Liu, D. S. Wu, "New current-feedback amplifier-based universal biquadratic filter" IEEE Trans. Instrum. Meas., **44**(4), 915–917, 1995. https://doi.org/10.1109/19.392891

[22] J. W. Horng, M. H. Lee, "High input impedance voltage-mode lowpass, bandpass and highpass filter using current-feedback amplifiers" Electron. Lett., **33**(11), 947–948, 1997. https://doi.org/10.1049/el:19970618

[23] J. W. Horng, "New configuration for realizing universal voltage-mode filter using two current feedback amplifiers" IEEE Trans. Instrum. Meas., **49**(5), 1043–1045, 2000. https://doi.org/10.1109/19.872927

[24] N. Shah, M. F. Rather, S. Z. Iqbal, "CFA-based three input and two outputs voltage-mode universal filter" Indian J. Pure & Appl. Phys., **43**(8), 636–639, 2005.

[25] S. Topaloglu, M.Sagbas, F.Anday, "Three-input single-output second-order filters using current-feedback amplifiers" Int. J. Electron. Commun. (AEÜ), **66**(8), 683–686, 2012. https://doi.org/10.1016/j.aeue.2011.12.009

[26] V. K. Singh, A. K. Singh, D. R. Bhaskar, R. Senani, "New universal biquads employing CFOAs" IEEE Trans. Circuits Syst.-II: Express Briefs, **53**(11), 1299–1303, 2006. https://doi.org/10.1109/TCSII.2006.882345

[27] S. F. Wang, H. P. Chen, Y. Ku, P. Y. Chen, "A CFOA-based voltage-mode multifunction biquadratic filter and a quadrature oscillator using the CFOA-based biquadratic filter" Appl. Sci., **9**(11), 2019. https://doi.org/10.3390/app9112304

[28] W. Tangsrirat, "Novel Current-mode and voltage-mode universal biquad filters using single CFTA" Indian Journal Pure & Applied Physics., **17**(2), 90–104, 2010.

[29] J. Sirirat, W. Tangsrirat, W. Surakampontorn, "Voltage-mode electronically tunable universal filter employing single CFTA" in 2010 International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 2010.

[30] R. Nawrocki, U. Klein, "New OTA-capacitor realisation of a universal biquad" Electron. Lett., **22**(1), 50–51, 1986. https://doi.org/10.1049/el:19860034

[31] J. Wu, E. I. E. Masry, "Universal voltage- and current-mode OTAs based biquads" International Journal Electronics, **85**(5), 553–560, 1998. https://doi.org/10.1080/002072198133842

[32] J. W. Horng, "Voltage-mode universal biquadratic filter using two OTAs" Active and Passive Elec. Comp., **27**, 85–89, 2004. https://doi.org/10.1080/0882751031000116160

[33] M. Kumngern, M. Somdunyakanok, P. Prommee, "High-input impedance voltage-mode multifunction filter with three-input single-output based on simple CMOS OTAs" in 2008 International Symposium on Communications and Information Technologies (ISCIT), Lao, China, 2008. https://doi.org/10.1109/ISCIT.2008.4700228

[34] J. Sarasri, B. Knobnob, M. Kumngern, "Electronically tunable voltage-mode universal filter using simple OTAs" in 2011 IEEE International Conference on Computer Science and Automation Engineering (CSAE), Shanghai, China, 2011. https://doi.org/10.1109/CSAE.2011.5952912

[35] F. Kaçar, A. Yesil, A. Noori, "New CMOS realization of voltage differencing buffered amplifier and its biquad filter applications" Radioengineering, **21**(1), 333–339, 2012.

[36] J. Pimpol, N. Roongmuanpha, W. Tangsrirat, "Low-output-impedance electronically adjustable universal filter using voltage differencing buffered amplifiers" in 2019 International Conference on Informatics, Environment, Energy and Applications (IEEA), Osaka, Japan, 2019. https://doi.org/10.1145/3323716.3323738

[37] O. G. Sokmen, S. A. Tekin, H. Ercan, M. Alci, "A novel design of low-voltage VDIBA and filter application" Elektronika Ir Elektrotechnika, **22**(6), 51–56, 2016. https://doi.org/10.5755/j01.eie.22.6.17224

[38] V. Biolkova, Z. Kolka, D. Biolek, "Fully balanced voltage differencing Buffered amplifier and its applications" in 2009 52nd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), Cancun, Mexico, 2009. https://doi.org/10.1109/MWSCAS.2009.5236157

[39] J. Satansup, T. Pukkalanun, W. Tangsrirat, "Electronically tunable single-input five-output voltage-mode universal filter using VDTAs and grounded passive elements" Circuits Syst. Signal Process., **32**, 945–957, 2013. https://doi.org/10.1007/s00034-012-9492-0

[40] D. Prasad, D. R. Bhaskar, M. Srivastava, "Universal voltage-mode biquad filter using voltage differencing transconductance amplifier" Indian J. Pure & Appl. Phys., **51**, 864–868, 2013.

[41] W. Tangsrirat, "Linearly tunable CMOS voltage differencing transconductance amplifier (VDTA)" Informacije MIDEM, **49**(2), 61–68, 2019. https://doi.org/10.33180/InfMIDEM2019.202

[42] O. Channumsin, W. Tangsrirat, "SITO-type high-input impedance voltage-mode multifunction filter using single active element" in 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST), Luang Prabang, Laos, 2019. https://doi.org/10.1109/ICEAST.2019.8802603

[43] N. Roongmuanpha, T. Dumawipata, W. Tangsrirat, "Triple-input single-output electronically controlled voltage-mode biquadratic filter" in 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST), Luang Prabang, Laos, 2019. https://doi.org/10.1109/ICEAST.2019.8802540

[44] P. Moonmuang, T. Pukkalanun, W. Tangsrirat, "Voltage differencing gain amplifier-based shadow filter: a comparison study" in 2020 6th International Conference on Engineering, Applied Sciences and Technology (ICEAST), Chiang Mai, Thailand, 2020. https://doi.org/10.1109/ICEAST50382.2020.9165352

[45] N. Roongmuanpha, T. Pukkalanun, W. Tangsrirat, "Three-input one-output voltage-mode biquadratic filter using single VDBA" in 2020 8th International Electrical Engineering Congress (iEECON), Chiang Mai, Thailand, 2020. https://doi.org/10.1109/iEECON48109.2020.229503

[46] Linear Technology, "100MHz current feedback amplifier with DC gain control", LT1228 datasheet, 1994.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1128-1136 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Performance of Project Teams Selected Based on Student Personality Types: A Longitudinal Study

Svitlana Ivanova[1], Lubomir Dimitrov[2], Viktor Ivanov[*,3], Galyna Naleva[4]

[1]*South Ukrainian National Pedagogical University named after K.D. Ushynsky, Department Mathematics and its teaching methods, Odesa, 65020, Ukraine*

[2]*Technical University of Sofia, Department Mechanical Engineering, Sofia, 1000, Bulgaria*

[3]*Odessa National Polytechnic University, Department Mechanical Engineering and elements of machine, Odesa, 65044, Ukraine*

[4]*Ukraine National University "Odesa Maritime Academy", Department of Higher Mathematics, Odesa, 65029, Ukraine*

ARTICLE INFO

Article history:
Received: 28 November, 2020
Accepted: 05 February, 2021
Online: 25 February, 2021

Keywords:
Project Method
Heuristic Method
Personality Types

ABSTRACT

The use of heuristic methods in teaching is not possible without the cooperation of all or part of the students' academic group. That is, the teacher who made the decision to apply the heuristic method de facto solves the issues of organizing project method in teaching. There are a large number of indications on the relationship between the effectiveness of the use of heuristic methods, taking into account the personality differences of students. As well as the importance of taking into account the personality differences in the project method. However, there is a lack of information about experimental studies in which three components: the project method, the heuristic method and the Myers Briggs personality types methodology, would be considered simultaneously. This prompted us to conduct this study. As part of the project method, a tournament among students of prospective mathematics teachers was held during 2014-2020. Teams of three types to participate in the competition were formed. There was a team whose members were not previously trained. The team whose members studied the heuristic method - "Creativity enhancement method". And also a team whose members, along with the study of the heuristic method, were selected in a special way. Students included in this group had personality types most suitable for performing heuristic techniques, which are components of the heuristic method. The task of the tournament was to compile a set of educational problems in geometry that can be used in the school curriculum. The problems developed by the team were evaluated by the panel. Members of other teams acted as opponents and reviewers. Using the heuristic method allowed teams to prepare more problems and systematize them. The best results in the use of the heuristic method showed the team, the composition of which was selected in a special way. The survey conducted according to the results of the tournament showed an increase in students' interest both in the studied discipline and in the project method, as well as a willingness to use the project method in their future work.

1. Literature review

1.1. Problem-based learning

The experiment on the application of the project method in the study of the discipline "Mathematics training methodology" was carried out from 2014 to 2020 year. This article is a generalization and continuation of the work originally presented at the International Conference on High Technologies for Sustainable Development, which describes the 2014-2018 experiment [1].

Problem-based learning (PBL) is a widely spread teaching method. This method has the following features: using collaboratively work in groups or teams, using principle "learning by doing", increased motivation and interest in learning outcomes [2].

*Corresponding Author: Viktor Ivanov, Email: ivv@opu.ua

Creative techniques are also used [2]. In its turn, principle "learning by doing" include such techniques: learning by trial and error, evidence in practice, experience versus book learning, discovery versus instruction [3]. Another point of view on the main features of PBL in [4] is as follows: the method involves solving a problem posed by the student, initiative of a group of students by the form of study, results of research and activities in the form of the final product, long-term project, and teachers are used only as consultants. The PBL is considered as three principal approaches: learning, contents and social [5]. The term learning means cognitive learning, which have follows peculiarities: the problem is focused on learning, learning versus experiment, enhancement motivation and study within the framework of the project. These peculiarities are the same that point out above. The content approach first of all implies interdisciplinary learning. The social approach is mean that student organize in project team and get project management skills.

At all source mentioned above indicate such features of PBL; cognitive character of learning (including heuristic methods [5]); student should be organized in group; the goal is to solve the problem (this solution to the problem is carried out within the framework of the project [4, 5]). One author is using term a group [4], another is using both term group and team [2]. It is more clearly stated in [5] - the project team is working on the project. Direct link to project management is here.

1.2. Elements of role-playing projects in PBL

In this and other publications, the project method is equated to PBL. But, according to [6], there are four types of projects: construction, enjoyment, problem and specific learning. Thus PBL is only a third type of project. The "construction" project constructs the interaction of the project team members according to certain circumstances. This means that role projects, as well as their well-known variety - a business game - are construction projects. We believe that all types of projects include elements of construction project and PBL. When considering PBL projects, at least the role of the teacher and the role of the team leader are highlighted. Also, the size of the team is usually taken into account.

In [2] the team contract is considered to be an interesting PBL tool. This contract contains the project rules namely: objective of the project, communication between team members, rules and procedures for meetings, etc. At this contract experience of project method is used. As example, a resource and time management elements are include. The team role distribution is considered at item - define the responsibilities of each team member. The role of the teacher is described as a facilitator of the learning process in [2]. Group size guidelines are given. Students which study engineering in framework of PBL perform project tasks in industry. The teacher acted as a mediator among the student group and employees of the firm. The number of assistant lecturers should be limited [7].

In the case of real project, special employee of the company − an instructor can be assigned. He selects a project topic, prepares the presentation of the project for a group of students with a statement of the problem and meets regularly with the members of the group [7]. But inspector should not execute the project instead of students or impose their opinion on them. It is also considered necessary to have an instructor who must prepare important data and organize information for students [2]. The instructor ensures a link between educational tasks with real problems. Another problem is that the use of the project method to solve real problems presupposes the existence of a core of good academic standing and conscientious students. Lack such of core creates an incredible administrative struggle to compensate for lack of student's motivation and to maintain the quality of the final product for the client [7]. Also, the potential problem for the implementation of real projects should be limited in scope so that the task does not exceed academic terms - a quarter or a semester.

In [8] example of student's group from 5 to 45 is provided. But note that more often groups of 5-7 are applied. It is point out that two faculty advisers should provide the work of student's group [8]. In [9] teaching roles in problem-based learning is described as a process-oriented supervisor.

PBL uses the results of research in the field of cognitive psychology to enhance creativity and increase interest [10]. The size of group is considered; note that for most purposes 3-5 members feasible is recommended. But this recommendation is not based on scientific analysis, but only on experience. Regarding the long-term project, it is indicated that most of the project terms are four years, when the whole process of studying at the university is performed on one project [10]. It should be noted that the work of the project team is provided by one or two teachers [10].

1.3. The role of teacher

The role of teachers is advisory, rather than authoritarian. Teachers used teaching assistant as mentors. On a final stage the report to the client is prepared with help a research assistant [10]. In [11] it was concluded that the group size and team tenure don't influence at the team performance. Note that both the term group and command are used in [10] and [11].

At the each publication mention above the role of a teacher is discussed. In detail the teacher's role is investigated in [12]. This role changes according to stage of a project. At the preparatory stage, the teacher motivates the students, helps in setting the goals of the project. When performing the main stage teacher directs the process. At the final stage the teacher is involved in summarizing works as an independent expert [12]. There are notes that it is very difficult for a teacher to plan or execute projects with students and control them at the same time. In [10] for decided this problem the senior seminar students as mentors are used. In order to avoid the monitoring problems, the teacher may assign a leader for each group of students. Collective execution of the project may lead to the fact that the poor academic standing students avoid academic work. Therefore, the teacher can use individual evaluations, along with a general evaluation of the performance of the group [7]. Such an evaluation can also be given by the instructor (research assistant) and each member of the group can evaluate the contribution of his colleagues during the various stages of the project [7]. In [13] identifies various ways of organizing the work of the teacher: with open (explicit) coordination or with hidden coordination (including the coordinator in the work on the project). In [14] refer different options for the distribution of roles between the teacher and students. In the case that the total amount of work can be divided into independent parts for each student, the role of the co-ordinator falls upon the teacher. In the case when it is possible to identify the main and auxiliary tasks, and also when the

final distribution of tasks between students occurs in the process of performing assignments, the role of the co-ordinator falls upon the teacher and one of the students. Such a distribution of roles exists when performing of integrated master's theses when several students develop one topic [14].

The main sources mentioned above are connected project approach with creating a group. In [15] note that obviously, the group is fundamentally and, when possible, students should be grouped. It is important to give students time to study group behaviour and performance. A study of the work in the student group showed that although the teacher's intervention and management led to improved student satisfaction with the results of the project, this negatively affected group unity [15]. An interesting variant in which the function of the co-ordinator is entrust to one of the students. This student co-ordinator is usually chosen from among the most successful students [16]. In general, many researchers distribute students into "strong" and "week", meaning only their academic achievements. For election of co-ordinator, such a feature as a high sense of responsibility, can take into account, for example [15].

In [11] pays attention to the need to take into account the personal differences of students while monitoring their academic achievements. So, students in different ways can show themselves in oral or written evaluation. To take account of these differences is proposed to use socionics typology [17]. Refer that the project method is also assumed control of academic achievements. In [18] insists that the same teaching method gives different results for students with different personality types. He invited to divide students into small groups so that each group included representatives of only one personality types. [18].

1.4. Student competition

The foregoing article consider different aspect of project method at education; these are: increasing creativity, “learning by doing”, increased motivation, results of research and activities in the form of the final product, teacher's role, size of group and selection team members in according there personality types. Depending on the type of project, this or that moment turns out to be very important. It is consider the projects, which included this entire item. These projects are based on student's competitions.

The Formula Student team includes from 40 to 50 members typically [19]. This large team is parted into sub-teams of 3 or 4 participants. A large project team has a leader called the captain. Usually this role is performed by a graduate student. The formalized roles are systems engineer and managers responsible for marketing and financial matters [19]. The captain and faculty member is controlled financial issues, also. An adviser and industry mentor supported each large team. In this case is realized problem learning study: team solve problem of automobile design with final product - racing car. As well role-playing project is realized. The students are performed role of design division of automobile company.

The same examples when there is a PBL project and a role project and also the final product are World Robot Olympiad and Robot Soccer World Cup. Another example is student tournaments. International Natural Sciences Tournament (INST) is including following specializations: Chemistry, Material Science, Geography, and Engineering. University students can participate in the team. The researcher or doctoral student acts as a trainer. Separate tournaments in mathematics (International Tournament of Young Mathematicians) and physics (International Physicists’ Tournament) are held.

The International Physicists’ Tournament (IPT) has the most established and regulated rules for conducting [20]. The IPT preparatory work on the formation of tasks carried out group of teachers, which is then not in contact with students. The other teachers, usually four of them, are panel members. Each team has the adviser - PhD student. Team is headed by the captain; he is one of a student. Teams of six students receive a list of tasks that can be solved experimentally or theoretically. In a round of a tournament, each team plays one of the following roles: speaker, opponent, and reviewer. Each task consists of three rounds, i.e. each team must play all three roles. Along with the team competition, the participant's personal contribution is assessed. Speakers and Opponents who receive an above average grade for their performances receive personal points. Individual winners of the Tournament are determined by the total amount of personal points.

Point out that in a number of cases, the work result of the project team may be not only in obtaining certain knowledge and skills, but it have practical result. Practical result is important when:

- A project is real and results of which will be used in practice;
- Competition between projects groups, for example - Formula Student, IPT, INST, etc.;
- Business games, when the insufficient level of results breaks the plot of the game and does not allow achieving the necessary knowledge and skills;
- This is the study of heuristic methods, when the absence of a result will not allow to fully revealing the content of the topic.

1.5. The results of the review and the objectives of the study

The effectiveness of the application of project method, as reviewed by the literature sources, is related to the features of the formation and work of project teams. Recommendations for the size of the project team are very different - from 2 to 12 members. Note that there are usually 15-25 students in an academic group. The size of the project team depends on the size of the academic group, due to organizational factors. Conveniently if all students of an academic group are included in the project team or two teams are formed from this academic group. The most common recommendation is that the project team should be small, but for more teams, more teachers should be involved than a lecturer and assistant.

In two cases, it is important to distribute roles between students in the project team - in the implementation of long-term projects and in student team tournaments. Long-term projects can be implemented for execution of a bachelor's degree project, a master's degree project and a term paper formed to work for a long period of time. The project team can be formed for a year or more for implementing collaborative interdisciplinary or research projects. In [11], it was concluded that group size does not affect team performance. We assumed this was due to the fact that team members were selected without regard to their personal

differences. Extensive experience in the formation and organization of project team interaction has been accumulated in project management. As a rule, the distribution of roles between students of the team members is spontaneous: someone is chosen for the role of leader, someone is a generator of ideas, and someone is responsible for processing the results of the work. Sometimes students themselves are called to perform certain functions in the team. However, there are various scientifically valid typologies of team roles. This distribution can be performed on the basis of the project management theory. The terminology accepted in project management, namely: project team, teamwork, team roles, is also expedient to use. The heuristic nature of the design method is often noted. The content of the project method can be the work of the team to study and apply the heuristic method [21]. In this case the distribution of team roles should be performed on the basis of personality type's methodologies [21]. The important question is about the role of the teacher in the project method. The roles of teachers vary. Teachers prepare material and project rules. They are advisers and evaluate students. Is teacher outside the project team or its member? Usually it depends on which project we are considering. In a short-term project, the teacher is out of the project. If the subject of the study is a long-term project, then the teacher can be an advisor and even be a member of the project team.

The study includes a preparatory stage, which consists of building project teams, teaching students to use the heuristic method and surveys to determine their personality type. The experiment itself included the organization of a tournament, work with the use of a heuristic method, including taking into account the personality type of students. As well as the panel's assessment of the teams' work results and a survey of the participants about the experiment results. At the analysis stage, the performance of the project teams, and the effectiveness of using individual techniques of the heuristic method by students with certain personality types were compared.

2. The methodology of experiment

2.1. The hypothesis of the study

The project method is widely used in education. The scientific basis of this method was developed in project management. The guidelines for team size and selection of team members were generally developed based on the team role methodology. Often the purpose of forming a project team is to ensure teamwork within the heuristic method. The methodology of personality types is more suitable for selecting group members and organizing the interaction of team members in the case when the application of the heuristic method is the reason for the formation of the project team. The lack of data on the combination of personality type's methodology, heuristic method and project method should make up for this experiment.

The following statement is accepted as hypotheses of a study:

- H_a - The use of the heuristic method, allows to improve the performance of the team.
- H_b - The formation of a project team, taking into account the personality types of students, allows to improve the performance of the team.
- H_c - The use of the project method in the form of team competition increases the interest of team members in the application of this method.

2.2. Tools are used to test these hypotheses of a study

The Keirsey Temperament Sorter. Student's personality types were determined using the Keirsey questionnaire [22]. The first test was conducted six months before the tournament. The second test was performed immediately before the tournament. Along with personality types, their brightness was determined. The test was carried out on the basis of an adapted version of the test in Ukrainian (http://ibib.ltd.ua/2211-test-keyrsi-tekst-35784.html). Two academic groups consisting of 15-20 students in each took part in testing. The coincidence of the results of the first and second tests was in the range of 76–81%, which is within the acceptable range (75–90%) in accordance with [23]. Students with personality types necessary for the execution of the heuristic method algorithm were selected to the first group. Of the students with the same personality types, preference was given to the participants with the brighter type.

The expert evaluation. The teacher and two PhD students prepared tasks for the tournament and became members of the panel. The panel members evaluate each problem based on a report and discussion. The maximum score is 100 points. The problem was evaluated on four items: correctness of composing and solving the problem (40), completeness of methodological recommendations (20), compliance with the requirements of the educational set of problems (20), consistency, clarity and validity of each stage in the presentation of the problem (20).The maximum number of points for each item is indicated in parentheses. The individual contribution of each participant was also assessed. The maximum score also is 100 points. The problem was evaluated on five items: ability to report, ability to debate, quickness in correcting your mistakes, finding mistakes in opponents, the ability to analyze the proposed sets of problems from a methodological point of view. The maximum score of each item is 20 points.

The survey. A survey was conducted following the experiment. All students answer the questions "Do you like the project method?" and "Will you use the project method in your future work?". The members of first and second team answer the questions "Do you like the heuristic method?" and "Will you use the heuristic method in your future work?". The members only of first team answer the questions "Do you like the distribution team role due personality type?" and "Will you use the distribution team role due personality type in your future work?".

2.3. The heuristic method

It was decided to use as a heuristic method "creativity enhancement method" [21]. This method included several heuristic techniques, namely: collective discussion, pause between the presentation of ideas and their criticism, random associations, analogy, expert evaluation, using a matrix. The sequence of heuristic techniques is follow. At first is applied three techniques collective discussion, pause between the presentation of ideas and their criticism, expert evaluation. This three techniques is represented a all known brainstorming. The main difference of a creativity enhancement method in contradistinction to another

method based on brainstorming is using matrix. The purpose of the first three heuristics is to determine what sense should be given to the row and column. Second part of heuristic techniques follows. These are random associations, analogy and expert evaluation. The goal of techniques random associations and analogy is to generate ideas about the semantic content of the first row and first column elements. Expert evaluation technique is used to select better variant of ideas. Third part of heuristic techniques grouped around the matrix. These are using a matrix, collective discussion and expert evaluation. The combination of elements first row and first column give a content of each matrix cells [24]. Each cells can contain a possible solve of problem. But in fact, many combinations give impossible options or options that are not implemented under given restrictions. The cells that are have sense the material of collective discussion. The final selection is performed using expert evaluation techniques.

2.4. The personality types selection

The preparation of students to use creativity enhancement method one lecture session and one practical session were provide.

According to the method especially selected students performed determination techniques. The moderator of discussion should be with personality type ENTJ or ESTJ [21, 22]. For collective discussion the ENTJ, ENTP, ESTJ and INTJ personality types are recommended [22, 25, 26]. These types are arranged in the order in which they are best suited for this technique. For criticism the types sequences in order of their degree of preferable: ISTP, INTP, ISTJ, ESTJ [25, 27, 28]. For expert evaluation the same type can be used only at another sequences their degree of preferable. For technique of random associations the types ENTP and ENTJ should be added to discussion group [21, 26]. For technique of analogy a type INTJ should be used [22, 25].

3. The results of the experiment

The experiment was attended by students of SUNPU - prospective teachers of mathematics. Every year, the subject studied two academic groups. The experiment was carried out during 2014-2020. Students were assigned to two or three project teams. The academic group was reorganized into a project team if the number of students was less than 15 or two project teams if the number of students was more than 15. Thus, there were two or three project teams. The teams were of three types. The third project team were participants, who were doing methodical task without special preparatory training for this task. The members of the first project team took additional sessions to learn the creativity enhancement method. The participant of this team especially selected on the base of their personality type. The types, which corresponded with personality type to need a heuristic technique of method, were included in team. Members of the second project team are also studying the heuristic method. Students of the third project team didn't have additional classes. In total, sixteen project teams were created: a fifth team of the first type, three teams of the second type, and three teams of the third type (Table 1).

The total number of problems prepared by the teams is presented in Figure 2. The tournament was organized for project teams as part of the study of the discipline «Mathematics training methodology». IPT rules were used for the competition. If there were three teams, then each of the teams consistently acted as a speaker, opponent and reviewer. In the competition between the two teams, there were only speakers and opponents. The teams were tasked with developing a set of geometric problems, namely, a planimetry subsection. The topic of these problems set is established the same for all teams. This set of problems can be used in the school curricula. Specific requirements for this set are established: definite purpose of using, the arrangement of the problems in increasing order of difficulty and the availability of the alternate problems (with the same level of difficulty as well as methodical recommendations).

Table 1: Types of team per year

Types of team \ Year	2014	2015	2016	2017	2018	2019	2020
First type							
Second type							
Third type							

The work of teams of the first and second types was organized in accordance with the heuristic method algorithm. The collective discussion group included from 5 to 10 people, depending on the number of project team members. The ideas put forward by them were discussed by a group of critics of 3-5 people, together with some of discussion group participants. The final decision is made by an expert evaluation group of three people, one of whom is a moderator. Then the team was divided into two parts, one discussed the specific content of the rows, the second columns. Groups is performed this task using free association and analogy techniques. Based on the results of the discussion, a matrix was drawn up. Schemes of problems were depicted in the corresponding cells of the matrix. The discussion group is selected problems for further development. Then, the team, to increase productivity, split into parts and worked on selected problems. For each problem it was necessary to find several solutions and develop methodological recommendations. The expert evaluation group with high-achieving students selected problems that could be presented.

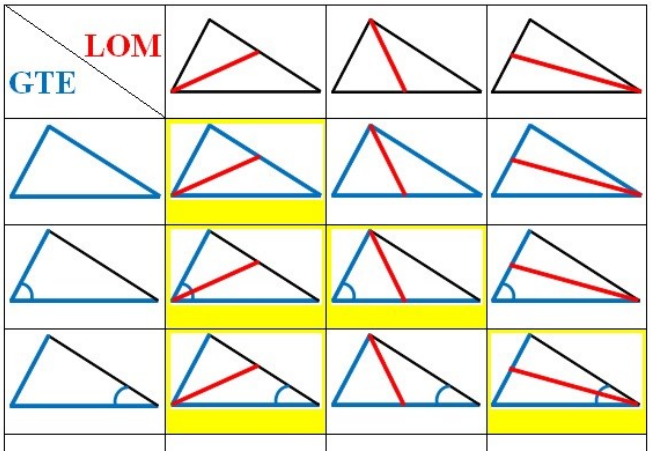


Figure 1: Example of a table with options for problems

An example of a fragment of one of the matrices obtained by the command of the second team type is presented in Figure 1. The

elements that need to be found are marked in red. These are the medians drawn to the three sides of the triangle. The first row of the matrix is location options of medians (LOM). The first column contains the given triangle elements (GTE). Each cell of the second row corresponds to a triangle with all sides are given.

Of the three options for the location of the median, only one was chosen, since in all three cases the solution algorithm is the same. In the third row, there are triangles in which two sides and the angle between them are known. Two options are selected when the median is drawn from a known angle and when it is drawn from an unknown angle. In the fourth row, there are triangles in which two sides and angle are known but the angle is not located between them. Two variants are also chosen when the median is drawn from a known angle and when it is drawn from an unknown angle. Further, in the matrix considered triangles in which one side is known. These rows are not shown. For each of the triangles marked with a yellow background, the team developed several solutions with explanations, as well as recommendations for the teacher on how to apply these problems in the school curricula.

The total number of problems prepared by the teams is presented in Figure 2 and Table 2. The team reported the results, usually there were several speakers. Students preparing the problem presented it. The panel evaluated the team members who participated in the presentation individually. The members of the opponent's team who participated in the discussion were also individually assessed. The participants of the opposing team looked for inaccuracies in the solution and deficiencies in the explanation. They raised objections to methodological recommendations and to the possibility of using the problem in the school course. On the basis of the report and discussion, the panel evaluated each of a set of problems. Problems with more than 100 points were selected for use in the school course. The number of problems selected by the panel is shown in Figure 3 and Table 2.

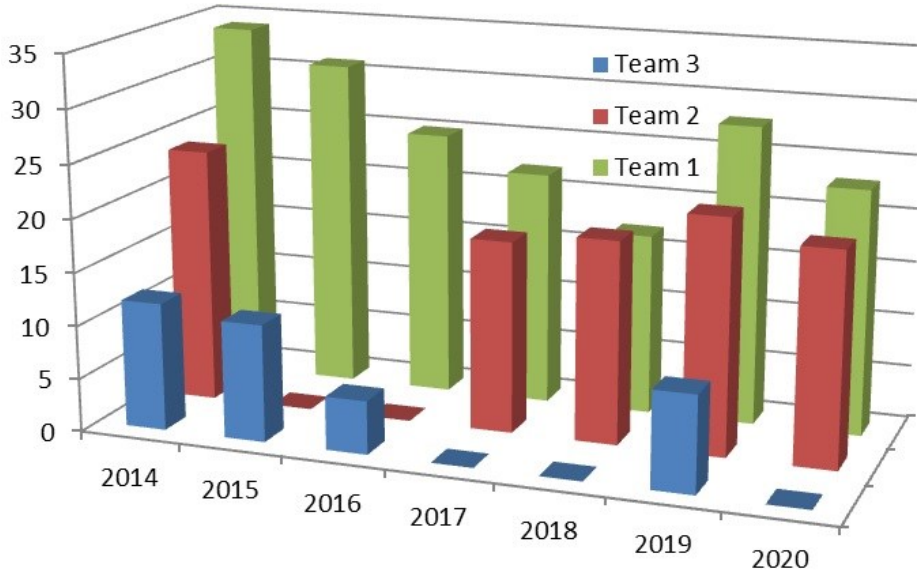


Figure 2: Total number of problems by year

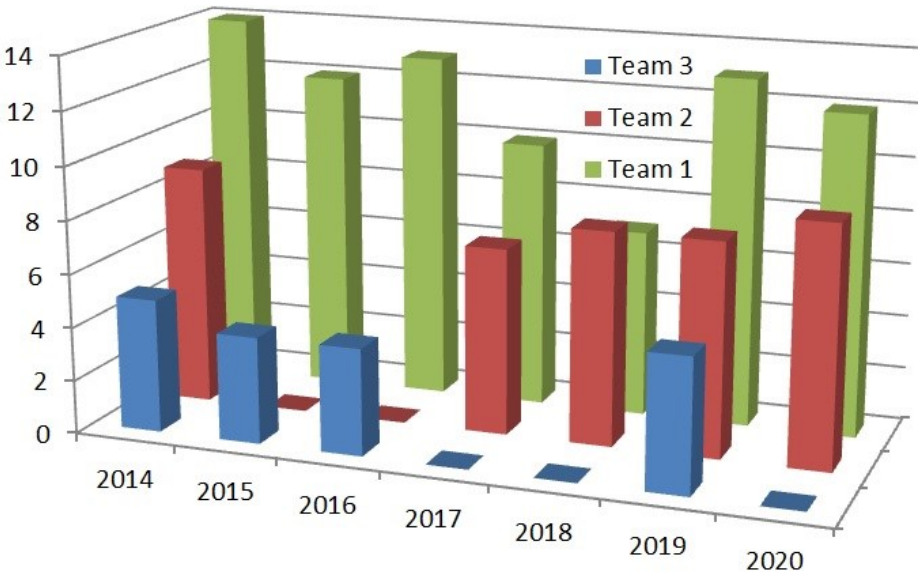


Figure 3: Selected number of problems per year

Table 2: Team performance

Team No / Year	Team 1		Team 2		Team 3	
	Total	Selected	Total	Selected	Total	Selected
2014	34	14	24	9	12	5
2015	31	12	-	-	11	4
2016	25	13	-	-	5	4
2017	22	10	18	7	-	-
2018	17	7	19	8	-	-
2019	29	13	22	8	9	6
2020	24	12	20	10	-	-

The results of the student survey showed an increase in the interest of team's members of all types in the project method and the heuristic method. The students of the first type of team showed enthusiasm for defining their personality types and organizing the team work based on this data. As for the prospects for using the acquired skills in further work, everything is not so unequivocally. Most of the students of the first type team are not going to use the knowledge gained about personality types in their further work. Almost half of them are ready to use the heuristic method in the future, in contrast to participants from teams of the second type. The project method of work was assessed by the students of teams of three types as the most deserving of use in the work of a teacher (Table 3).

Table 3: Types of team per year

Survey questions / Team types	First team	Second team	Third team
Do you like the project method?	89%	84%	73%
Do you like the heuristic method?	77%	69%	-
Do you like the distribution team role due personality type?	82%	-	-
Will you use the project method in your future work?	65%	71%	58%
Will you use the heuristic method in your future work?	44%	24%	-
Will you use the distribution team role due personality type in your future work?	23%	-	-
Has interest in studying the discipline increased due to innovations in the form of education?	81%	73%	67%

4. Analysis of experiment result and discussion

The performance of the second type team, which used the heuristic method, is higher than the performance of the third type team. Teams of the second type developed 24 problems. The average number of problems in a set is 8. The third type team developed 13 problems. The average number of problems in a set is 4,3. Thus, the results of the experiment confirmed the hypothesis H_a.

The application of the heuristic method requires special training. One lecture and one practical session are the minimum time to explain the essence of the method. It would be desirable to have more time for this. But there is a problem associated with the allocation of even this short time. If the heuristic method were used in teaching several disciplines, then there would be an opportunity for additional training. The method is quite complex, and several times the teams invented the contents of the matrix during brainstorming and heuristic techniques of random associations and analogies are not used. Heuristic techniques of random associations and analogies can be used to search for the contents of rows and columns. If one technique does not work, the other can be used additionally. It is promising to use the heuristic method in long-term projects carried out during a semester or longer.

The performance of the first type team, which was formed taking into account the personality types of students, is higher than the performance of the second type team and much higher than the performance of the third type team. The first type of team developed sets of 56 problems. The average number of problems in a set is 11,2. The experimental results confirmed the hypothesis of H_b. This means that the selection of performers for the execution of heuristic techniques was performed correctly.

It is very difficult to select the necessary personality types for the project teams, considering the level of knowledge, motivation and the personal interaction and relationships due to the small number of students learning a certain subject. For to form first type team in some cases it was necessary to include the students that had rare personality types ENTJ (moderator) ENTP (discussion group) from another academic group. This reduced the performance of teams of the second and third types. Experience has shown the academic achievements of the students can be more important than the exact matching of their personality types with their role in the creativity enhancement method. Sometimes we had to include the students from different academic groups into the project teams to strike a balance between the levels of their knowledge. The correct selection of team members of the first type allows students to understand the advantages of the method through a successful result. The project team included not only students with a personality type that is suitable for performers of heuristic techniques. Thus, it was necessary to find a model of participation for those who have a different personality type. Experience shows that the ENFP can act as a moderator, facilitating discussion through one's own imagination and enthusiasm. The personality type ENFJ can substitute INTJ in the discussion group because of their ability to generate the ideas. The personality type ENFJ can also fulfil the role of the teacher assistant when teaching creativity enhancement method. Having the big number of the representatives of personality types ISTJ, INTJ, INTP, ISTP, in the discussion group, it makes sense to give the students time to make the record of their ideas before the beginning of the discussion. A higher level of students' knowledge is created the prerequisites for the effective execution of the chosen role when working in a project team. However, this students, who did not have high grades earlier with the matching of their personality type with the appointed role surprisingly enough showed a better efficiency. For example, a student of personality type ISTJ, with an average level of academic achievement, working as a critic in a project team, enthusiastically is joined the work in detail analyzing the proposals of other team members. This

is allowed him to systematize and deepen his knowledge and skills appropriate way for him. The personality types of students, who show high results in the experiment, usually coincide with the typical personality types for STEM specialties: ISTJ, INTP, ESTJ, INTJ, ENTP, and ENTJ.

The survey showed that the members of the three types of teams were enthusiastic about the project method implemented in the form of competition. 82% of the total number of students liked the participation in the project. Students, who are going to use the project method in their profession - 64,7%. Thus, hypothesis H_c is confirmed.

The number of students in project teams is higher than recommended by the majority of research. This is due to the lack of teachers to conduct the research. Another problem of the project method was the organization of teamwork in such a way that all team members took an active part in it. We did not have this problem with the work of the first type of teams, unlike the second and third type teams. This serves as an indirect confirmation of the success of team interaction based on personal types. When using project method, there has always been the problem of assessing the personal contribution of team members and, accordingly, evaluating their academic achievements. Not all students presented solutions to problems and participated in the discussion. In the first type team, the students, in particular, the moderator, were involved in the assessment, which contributed to its objectivity and validity. The fact that students who solved the problem represent it is good for their self-realization, but it can be bad for the team score. From the point of view of higher score of the team, it is advisable to entrust the presentation of problems to students with high academic achievements.

5. Conclusion

The experiment was conducted in the form of a tournament, which made it possible to implement the PBL project and at the same time the role-playing project. The project had the final product - a set of geometry problems. Students studied the creativity enhancement method using the principle of "learning by doing". Thus, the project was comprehensive. Three types of project team took part at competition. Students of first and second project team are learned the heuristic method. Students of third project team haven't special training. It was found that the use of the project method in the form of a team competition aroused increased interest among students. It has been established also that the use of the heuristic method improves the performance of the project team. Members of first project team were selected according to their personality types. It was proven that the selection of students with certain personality types to the project team, the most suitable for the implementation of the heuristic techniques of the method, allowed increasing the performance of the project team. This confirms the correctness of the recommendations for the selection of personality types corresponding to a certain heuristic techniques. It is best to choose students with the ENTJ personality type for the role of moderator. It is recommended to choose personality types ENTP, ENTJ, INTJ for discussion and generation of ideas. The ISTP, INTP and ISTJ personality types should be used as criticisms. Several members of a collective discussion group must necessarily be at the group of critics. However, among the critics, there should be several members of the project team who did not participate in the discussion, for unbiased analysis of the proposed solutions. Academic achievement affects the performance of teams along with the personality types of students. This is important to consider when choosing a moderator and discussion group members. For this reason, the preferred ENTJ types can be replaced with ESTJ as the moderator, and ENTP can be replaced with INTP in the discussion group.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments:

This work has been accomplished with financial support by the Grant No BG05M2OP001-1.002-0011 "MIRACle (Mechatronics, Innovation, Robotics, Automation, Clean technologies)", financed by the Science and Education for Smart Growth Operational Program (2014-2020) and co-financed by the European Union through the European structural and Investment funds.

References

[1] S. Ivanova, L. Dimitrov, V. Ivanov and G. Naleva, "An Experiment on the Joint use of the Heuristic and Project Methods at the University," in 2019 IEEE International Conference on High Technology for Sustainable Development (HiTech), 1-5, doi: 10.1109/HiTech48507.2019.9128248.

[2] R. Seidel, E. Godfrey, "Project and team based learning: An integrated approach to engineering education," In 4th ASEE/AaeE Global Colloquium on Engineering Education, №146, 2005.

[3] H. W. Reese, "The learning-by-doing principle. Behavioral development bulletin, " **17(1)**, 1-19, 2011. doi.org/10.1037/h0100597.

[4] J. Adderley, Project Methods in Higher Education, Research into Higher Education Monograph 24 London: Society for Research into Higher Education, 1975.

[5] A. Kolmos, E. De Graaff, X. Du, "Diversity of PBL–PBL learning principles and models," In Research on PBL practice in engineering education, 9-21, 2009, doi.org/10.1163/9789087909321_003.

[6] W.H. Kilpatrick, "The project method," Teachers college record, **19(4)**, 319-335. 1918.

[7] N.K.Malhotra, A. Tashchian, A.K. Jam, "The project method approach: An integrated teaching tool in marketing research," Journal of Marketing Education, **11(2)**, 32-40, 1989, doi.org/10.1177/027347538901100206.

[8] J. E. Mills, D. F. Treagust, "Engineering education—Is problem-based or project-based learning the answer," Australasian journal of engineering education, **3(2)**, 2-16, 2003.

[9] A. Kolmos, "Reflections on project work and problem-based learning,: European journal of engineering education, **21(2)**, 141-148, 1996, doi.org/10.1080/03043799608923397.

[10] L. Helle, P. Tynjälä, E. Olkinuora, "Project-Based Learning in Post-Secondary Education – Theory, Practice and Rubber Sling Shots," High Educ., **51**, 287–314, 2006, https://doi.org/10.1007/s10734-004-6386-5.

[11] R. Harms, "Self-regulated learning, team learning and project performance in entrepreneurship education: Learning in a lean startup environment," Technological Forecasting and Social Change, Elsevier, **100**, 21-28, 2015, doi: 10.1016/j.techfore.2015.02.007.

[12] Е.С. Полат, Новые педагогические и информационные технологии в системе образования. Академия. 2009.

[13] Т.И. Гречухина, А.В. Меренков, С.В. Куньщиков, И.Ю. Вороткова, А.В. Усачева, Самостоятельная работа студентов: виды, формы, критерии оценки, Екатеринбург: Издательство Уральского университет, 2016.

[14] А.Т. Ашеров, В.І. Шеховцова, "Проектная культура будущих инженеров–педагогов компьютерного профиля: сущность понятия," Теорія і практика управління соціальними системами, **4**, 70-79, 2007.

[15] G. Kirkpatrick, "Online 'chat' facilities as pedagogic tools: a case study," Active Learning in Higher Education, **6(2)**, 145-159, 2005, doi.org/10.1177/1469787405054239.

[16] S.A Barab, T.M. Duffy, "From practice fields to communities of practice," Theoretical foundations of learning environments, 25–55, 2000.
[17] О.П. Кальнік, "Індивідуалізація контролю навчальних досягнень студентів вищих навчальних закладів на основі використання соціонічних типів та установки на вид діяльності. Педагогіка та психологія." **53**, 66-74, 2016.
[18] D. Prorak, T. Gottschalk, M. Pollastro, "Teaching method and psychological type in bibliographic instruction: effect on student learning and confidence," RQ, **33(4)**. 484-495. 1994.
[19] I. Talmi, O. Hazzan, R. Katz, "Intrinsic Motivation and 21st-Century Skills in an Undergraduate Engineering Project: The Formula Student Project," Higher Education Studies, **8(4)**, 46-58, 2018.
[20] V. Vanovskiy, "International Physicists' Tournament - the team competition in physics for university students," European Journal of Physics, **35(6)**, 064003, 2014. doi: 10.1088/0143-0807/35/6/064003.
[21] V. Ivanov, L. Dimitrov, S. Ivanova and O. Olefir, "Creativity enhancement method for STEM education," 2019 II International Conference on High Technology for Sustainable Development (HiTech), Sofia, Bulgaria, 1-5, 2019, doi: 10.1109/HiTech48507.2019.9128255.
[22] D. Keirsey, M. Bates, Please understand me. Del Mar, CA: Prometheus Nemesis, 1978.
[23] R.M. Capraro, M.M. Capraro, "Myers-Briggs Type Indicator score reliability across studies: A metaanalytic reliability generalization study," Educational and Psychological Measurement, **62**(4), 590-602, 2002, doi.org/10.1177/0013164402062004004.
[24] V. Ivanov, G. Urum, S. Ivanova, G. Naleva, "Analysis of Matrix and Graph Models of Transmissions for Optimization Their Design." Eastern-European Journal of Enterprise Technologies, **4(1)**, 11-17, 2017, doi:10.15587/1729-4061.2017.107182
[25] M.H. McCaulley, "Myers-Briggs Type Indicator: A bridge between counseling and consulting". Consulting Psychology Journal: Practice and Research, 2000. https://dx.doi.org/10.1037/1061-4087.52.2.117]
[26] L. Moller, C. Soles, "Myers Briggs type preferences in distance learning education", International Journal of Educational Technology, **2(2)**, 2001.
[27] C.F. Yokomoto, R. Ware, "Applications of the Myers-Briggs Type Indicator in engineering and technology" Education--part II. Indiana University-Purdue University Indianapolis. Session, **2230**, 4.88.1-4.88.10, 1999.
[28] D.W. Salter, N.J. Evans, D.S. Forney, "A longitudinal study of learning style preferences on the Myers-Briggs type indicator and learning style inventory", Journal of College Student Development, **47(2)**, 173-184, 2006.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1137-1144 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Artificial Neural Network Approach using Mobile Agent for Localization in Wireless Sensor Networks

Basavaraj Madagouda[1,*], R. Sumathi[2]

[1]*Department of Computer Science and Engineering, Jain College of Engineering and Research Belagavi, VTU Belagavi, 590001, India*

[2]*Department of Computer Science and Engineering, Siddaganga Institute of Technology Tumkur, VTU Belagavi, 590001, India*

ARTICLE INFO

Article history:
Received: 15 December, 2020
Accepted: 31 January, 2021
Online: 25 February, 2021

Keywords:
Artificial Neural Network
Localization
Mobile Anchor Node
Wireless Sensor Network

ABSTRACT

Wireless sensor networks (WSNs) are having large demands in enormous applications for the decade. The main issue in WSNs is estimating the exact location of unknown nodes. All applications are dependent on the location information of unknown nodes in WSNs. Location information of mobile anchor node is used to estimate the location of unknown nodes. A new approach is implemented in this paper for the localization of unknown nodes using Artificial Neural Networks. Specifically, a neural feed network is used for the indoor position process. Also several neural network configuration sets have been tested, which includes Bayesian regularisation (BR), Levenberg-Marquardt (LM), resilient back propagation (RP), Scaled Conjugate Gradient (SCG) and Degree Descent (SCG),etc. At the end results are simulated using MATLAB and Mean Square Error is calculated and compared with other existing *approaches. The proposed approach is energy efficient and uses only a two-way message to obtain inputs for the localization. Even the cost is minimized as in the proposed system only one mobile anchor node is used.*

1. Introduction

With the development of new applications in numerous areas, wireless sensor networks (WSNs) have gained a vast reaction from academia and industry. Applications for the future include monitoring of vegetation and the environment, search and rescue processes, object tracking in hospitals, monitoring of patients, and military applications. Because of technological development, small, cheaper sensor nodes have been developed that consume less power and feature capabilities like sensing, data storage, computer technology, and wireless communication.

Localization among others, architecture, deployment, synchronization, calibration, quality of service, and security are the key and difficult part of many WSN applications. The sensor data has to be attached to measured data in order to make this important for location since it is essential to monitor and record broad information like acoustic, visual, thermal, and seismic or any other type of measured observation. For example, the place or part of the farm from which the data is collected would be required for the person monitoring a big vegetation farm, with a wide variety of vegetables, since various vegetables have different special needs. Reporting Object origin, node location is required so as to support or answer group querying, routing and network coverage questions. Research into the problem of node location has provided a number of solutions over the course of the decade. Overall, most of the time there is a compromise between localization accuracy, computer complexity, and technology energy efficiency according to their application needs.

A large number of space-dispersed sensor nodes is usually used in a WSN such that the location of the sensor node when the nodes are used is unable and inefficient. Sensor nodes may also be relocated from their original position. An algorithm is therefore necessary for an independent assessment of the location of the node. In this research paper, a new location scheme for neural network-based nodes is proposed. In order to address the node location problem, the Received Signal Strength Indicator (RSSI), is utilized to evaluate the coordinates and location of WSN nodes. Contributions are as follows:

Is RSSI an appropriate component for WSN node location? We use RSSI to determine where the sensor nodes are located and explain how RSSI is suitable for locations. There will not be a need for additional hardware because sensor nodes are openly outfitted with RF modules for wireless communication. Multi track fading and noise, however, can contaminate the RF signal and thus affect the RSSI value through a noise-to-signal correction.

Also, of the utmost importance is the choice of the evaluation model. A 2D indoor environment that is determined by the x and y co-ordinates is considered in this research. A single node of anchor requires directional steering antennas, which directly

*Corresponding Author: Basavaraj Madagouda, basavarajkm29@gmail.com

boosts the sensors' cost and electricity consumption. Therefore, the two or more anchor nodes with different configurations explain the best configuration used for the anchor nodes.

A new 2D indoor locating system based on a neural network is introduced. For an accurate localization, the global positioning system can be used. The use of GPS does require the attachment of external hardware to the sensor nodes, thereby increasing cost and electricity consumption. In addition, a GPS viewing line is required and is therefore not applicable to indoor environments.

A neural feed network is used for the proposed indoor position process. A variety of neural network configuration sets have been tested, and Bayesian regularisation (BR), Levenberg-Marquardt (LM), resilient back propagation (RP), Scaled Conjugate Gradient (SCG) and Degree Descent (SCG) are the best neural network training and implementation (GD).

2. Literature survey

In [1], the authors presented a survey on machine learning algorithms for WSNs. They have presented techniques to identify intrusion by merging clustering and deep neural network methods and discussed machine learning-based localization techniques for WSN such as fuzzy logic, k-means + fuzzy-c-means, CNN+SVM, ANN, SVM. Also discussed ML-based congestion control strategies. Logic-based cram detection and control systems have been configured for efficient and active queue management as the basis for minimizing packet loss. It can be carried out in three stages. Second, it senses congestion in the second it modifies congestion control and by balancing the flow it recovers from the congestion.

In [2], the authors have suggested a new distance measuring capability using the Adaptive Neural Fuzzy Inference Scheme to render model, map RSSI to the precise T-R distance and experiments prove feasible.

In [3], the authors has proposed an adaptable model dependent on the neural organization and lattice sensor preparing stage for exact limitation of sensors. Recreation results show that the area exactness can be expanded by expanding the framework sensor thickness and the number of passages. The author proposed and explored a confinement procedure for WSNs utilizing NN. They think about NN for building adaptable planning among RSS and the situation of the sensor hubs. In this method, the NN is prepared to utilize the RSS estimations of the lattice sensors. The positive highlights of the framework are its dependence against the adjustment in the RSS estimations with time and no requirement for the additional course of action of equipment. The organization's preparation is refreshed with a standard timespan to limit area mistake. Simulation experiments demonstrate that the area exactness can be expanded by expanding the matrix sensors' density and APs.

Nowadays for enhancing the accuracy of localization in the Range-Free approach is more focused. In [4], the authors introduced many algorithms to improve accuracy such as BP artificial neural network (BP ANN). Rather than the amorphous algorithm is very easy to develop, and it will generate huge errors in the localization procedure. For resolving the issue and to diminish localization error, presented an improved localization algorithm. A widely utilized model in neural networks (NN) is BP NN. BP algorithm is a NN of three-level. The methodology used for the detection of network weight and the threshold between BP algorithms with minimal error and recursive technology is known as the deepest descent. BP ANN practitioners are incredibly self-organizing and self-learning, error replication, and a minimal theory of error in error optimization.

In [5], authors presented a study on localization in WSN by utilizing neural networks (NN). This study analyses localization algorithms and their advantages and disadvantages. They Discussed localization, localization classification methods, localization classification algorithms, and the pros and cons.

In [6], authors submitted comparable analyses of the localization WSN localization networks of Sigmoidal Feed ANN (SFFANNs) and Radial Base Function (RBF). The system proposed uses the Obtained Signal Force Indicator (RSSI), a static node calculated on a grid of 100 x 100 m2 of the three anchor nodes. The key aim of the SFFANNs' and RBF network system comparative analysis is the development of an efficient WSN localization structure. Both ANN models approximate the position error by using RSSI values. Quality comparison has been done with respect to standard deviations, mean position errors, and statistical parameters of R-square. MLP has the best efficiency, shown by simulation outcomes, where accuracy is considered.

In [7], the RBF network has been reviewed by the authors for the implementation of the WSN position method. For reasons of faster convergence and the lowest measurement rate, the RBF-based localization architecture should be evaluated. The analyses of RBF in this paper have been a probabilistic NN and general regression NN. For designing an efficient localization method, the suggested technique may be used. To minimize position errors, the predictive location algorithm proposed by the authors [8] relies on RBF NN. An algorithm detains moving node standards inherent in it and learns them from them. A field of mobile nodes can be expected by removed traveling highlights. Reproduction findings confirm that these calculations will be more reliable in the visually impaired interval to understand detailed localizations.

In [9], the authors proposed computationally – cheap arrangements rely upon support vector machine (SVM) and twin SVM learning algorithms where network nodes right off wanted signal. At that point, the network is trained to determine nodes in the region of source; subsequently, locale of the function is distinguished. At last, centroid of event region is assessment of source area. Re-enactments show the productivity of the proposed techniques. SVMs are integral tools for regression and classification which depend on the statistical learning hypothesis. Contrasted and other ML strategies, for example, ANNS, SVMs have a superior speculation ensure. SVM separates the data tests of 2 classes by deciding a particular hyperplane in input space. This hyperplane boosts the partition among 2 classes by taking care of a quadratic programming issue whose arrangement is around the globally optimal.

In [10], the author has examine the reasonableness of a portion of these algorithms while investigating diverse compromises. In particular, authors initially figure new method of characterizing many feature vectors for planning localizing issues onto diverse AI models. Instead of regarding localization as a classification issue, as accomplished in the majority of detailed work, authors treat it as a regression problem. Authors have contemplated an effect of varying network parameters, for example, anchor population, network size, transmitted signal power, and remote channel quality, on confining precision of the models. Authors have additionally considered the effect of sending anchor nodes in the framework as opposed to putting the nodes arbitrarily in

deployment areas. Their outcomes have uncovered fascinating experiences while utilizing the multivariate regression model and SVM regression model with an RBF. Various models have been characterized for various ML algorithms. A portion of the normally utilized models incorporates neural networks, multiple linear regression model, linear regression model, SVM with various kinds of kernels, Logistic regressions or classifications, k-nearest neighbours (KNN), and random forest (RF) clustering.

In order to address the Wireless Sensor Networks (WSN) localization accuracy problem in [11], the authors have suggested the latest Methodology for Localization Backpropagation Network (BPNN). At first, node localization calculations are provided with range intervals and signal power, and on BPNN the parameters solve easily. In a simulation experiment, the main factors of NS2 and MATLAB are also evaluated. The results show that this technique is successful in comparison with other localization algorithms and can minimize localization errors effectively.

In [12], the authors proposed ANNs way to deal with localization in WSN through change of ANNs structures utilizing Genetic Algorithms. Populace of feed forward ANNs composing their structure in genetic code is developed more than 20 ages. Every individual is assessed through the preparation of training of ANN and further estimation of its root mean square error for all testing set. The RSSI estimations were utilized as ANN inputs to localize hubs. The methodology is to collect information from the ANNs under an input reactivated static 26 x 26 meter indoor organizational climate, which consists of eight anchor hubs, under a reactivation of the static 26x26 meter indoor organization environment. The methodology has been used. Hereditary calculations of MATLAB and fake neural organizations were used for stashing tools.

In [13], the authors emphasise the comparison and study of MLPNN and RBFNN for the creation of a position system in WSNs. Various sensor node collection were evaluated using both MLPNN and RBFNN along with the mean square error of various input data sets. Their simulation findings effectively suggest that MLPNN is more reliable with respect to localization than RBFNN.Use the "Insert Citation" button to add citations to this document.

In [14], the authors suggest a new algorithm for localizing the sensor nodes using virtual anchor nodes. VAN is located in its system with minimum anchor nodes. The goal of their work is to improve the accuracy of the position by optimizing the location ratio and minimizing the costs of several anchor nodes.

In [15], the author explained the localization based on ANN in WSN. Localization refers to the issue of finding the location or position of sensor nodes in a network. But localization problem is still an open challenge in WSN. Author analyzed and classified existing works on localization algorithms that uses ANN. ANN algorithms were identified based on how they were utilized in the localization processes; noise characterization, model construction, training phase optimization, clustering, and prediction. Author also discussed about the features and related work about it and how to improve the accuracy of previous localization using ANN.

In [16], the authors proposed an approach known as ANN approach for moderating the reaction of mixed noise sources and harsh factory conditions on the localization of the wireless sensors. They have focused on investigating the effect of blockage and ambient conditions on the accuracy of mobile node localization. They have added a simulator for simulating the noisy and dynamic shop conditions of manufacturing environments, also for analyzing the proposed neural network.

In particular, in networks with sparse anchor deployment, cooperative localization improves the accuracy of position estimates. In [17], the authors presented implementation of it in low-cost WSN that are low-power. A WSN hardware platform with ultra-wideband range and communication is shown and the implementation of a cooperative localization algorithm based on belief propagation is addressed. The key sources of error are addressed and indoor tests of the device are performed using a simple channel access scheme and a standard ranging process.

In [18], the authors presented an enhanced centroid localization algorithm for WSN. One of the main technologies in the WSN is node localization. The center position algorithm depends entirely on the size and distribution density of the anchor nodes, but they are distributed randomly and their density is small. Given the conventional centroid localization algorithm, an enhanced centroid localization algorithm is vulnerable to the effect of network deployment status and the current situation of low positioning accuracy. Second, the idea of reference anchor nodes is implemented by bringing weight thought into the measurement method of node coordinates.

This template, modified in MS Word 2007 provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout the conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

3. Methodology

An ANN, also known as the network of neurons, offers a simple way of revealing valued and true functions as examples. ANN uses a supervised curriculum. Inputs and outputs to the ANN are given to learn and create a correct model. The classification and regression dilemma is normally tackled. Typically, the artificial neuron network has 3 layers, input layer, hidden layer(s), and finally output layers.

Neural feedback and feedback networks form two classes of neural networks. In a neural feed network, the outcome is fed from one layer of neurons to the next layer, no layers are skipped during this process, and the system has no feedback. The position is composed of three major neural networks: MLP, the Recurrent Neural Network and, the RBF.

It is to be noted that RSSI values are very unpredictable and vary under ambient noise and sensor mobility. The benefit of a neural network is that prior environmental and noise propagation is not needed. More precision is achieved through a neural network compared to other approaches such as Kalman filters. Comparing other neural network types, for the study the

compromise between MLP neural accuracy and the network memory requirement was chosen.

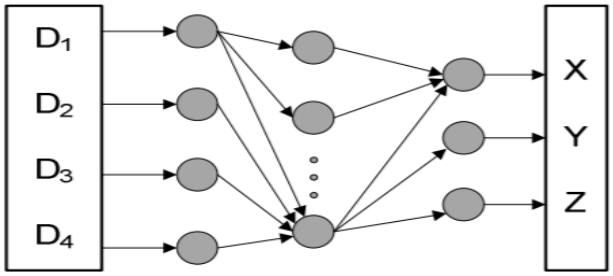


Figure 1: A general feed-forward neural network structure with four inputs and three outputs.

3.1. Network Model

WSN model for region assessment of static sensor nodes. Static nodes are moved to region of recognition and the node of mobile

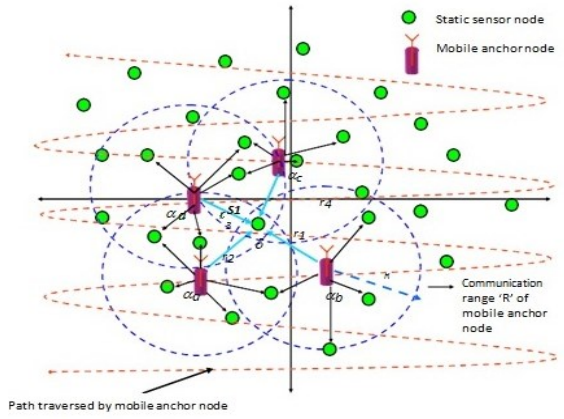


Figure 2: Illustration of broadcasting and reception of beacon signal from a Mobile Anchor and static sensor node respectively.

anchor (MA) ready for intersection. In addition, in the fig, the way MA navigates. 2.

It is assumed that R is a contact range of the MA hub. MA has a GPS recipient and knows its location often. The MA is sending beacon signal at a fixed "Live" interval with an id, transmission radius, current location, transmission power and format shown in figure 2. Before the MA node moves around the sensor area, the administrator determines the 'alternative' intervals.

MA_id	MA current Position (x, y)	Transmitted power	Transmission range 'R'

Figure 3: Packet format of beacon message

Any static sensor node in MA node's "R" range receives a signal from the beacon. A table called Mobil Anchor Location (MAP), as represented in Table 1, is retained in every static sensor Node.

When static sensor node beacon signals are obtained, X-y coordinate values are extracted from signal and stored in the MAP table in corresponding fields. The distance is determined according to the signal power obtained and stored in the Chart Table.

Consider scenario in Figure 2, let $\alpha_a, \alpha_b, \alpha_c$ and α_d be the MA positions throughout its traversal time and having (x_1, y_1), (x_2, y_2), (x_3, y_3) and (x_4, y_4) as their coordinates. Static sensor node S1, at position δ whose co-ordinate is (x_0, y_0), receives beacon signals when MA is at position $\alpha_a, \alpha_b, \alpha_c$ and α_d and stores their x-y co-ordinate values into MAP table. Similarly, S1 receives beacon signals when it comes into MA node proximity. Eventually, x-y co-ordinate values are extracted from received signal and stored into MAP table.

3.2. RSS Measurement

Distance from the MA node at position I to the static node S1 at position (x0, y0) is referred to as a correspondence to r (i), as seen in the following equations; (R2) The noise free RSS or received power at S1, denoted by P_i^r

$$P_i^r = k_i \frac{P_i^t}{d_i^\alpha}, i = 1,2,3, \ldots., m \quad (1)$$

$$d_i^\alpha = k_i \frac{P_i^t}{P_i^r}, i = 1,2,3, \ldots., m \quad (2)$$

If the beacon signals received by the static node i=1, 2... m number, the transmitted power is P_i^t, the power received is P_i^r, and the propagation constant is α, k (i)denotes other environmental variables that influence the received power.
Range related measurement can be modeled, as expressed in (R1)

$$r_i = k_i \frac{P_i^t}{P_i^r} \quad (3)$$

$$r_i = d_i^\alpha + \eta_i \quad (4)$$

Every static node calculates r_i for each received beacon signal and stores into MAP table as demonstrated in table 1.

Table 1: MA Positions stored at static sensor node

X-co-ordinate	Y-co-ordinate	Distance (r)
(x_1)	(y_1)	r_1
(x_2)	(y_2)	r_2
(x_3)	(y_3)	r_3
(x_4)	(y_4)	r_4
•	•	•
•	•	•
(x_m)	(y_m)	r_m

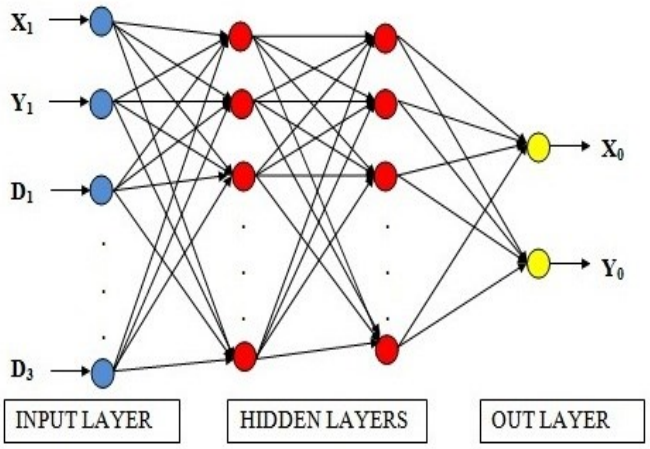


Figure 4: Multi-layer perceptron neural network

3.3. Multi-Layer Perceptron Neural Network Model

Neural network model (framework), initially trained with training dataset. Proposed system consisting of Multilayer Perceptron neural network [MLPNN] model having input layer, one or more hidden layers and output layer.

Nine inputs are passing to MLPNN, consisting of three different positions of MA with respect to distance from sensor node. Hidden layer having various neurons. Each neuron, having transfer function to find out weights of each connection between inputs to hidden layer neurons. Output layer having two neurons produces the location coordinates of unknown sensor node. Activation function (linear or nonlinear) is used to calculate location of sensor nodes.

3.4. Algorithm

The multilayer perceptron algorithm is as follows:

Step 1: Initialization: Arbitrary values are used to trigger all networks and thresholds, which are consistently distributed across a limited range. If these values are 0, gradients will also be computed to 0, if and only if there will be no direct relation between input and output and network. More training trials, with various initial weights, are planned to have the best cost-effectiveness. If the initial values are sufficiently high, however these units appear to be saturated.

Step 2: A new period of formation: an era means all the examples are presented in the training set. Training on the network typically requires further training times. The weights are calibrated for statistical rigour, only after the test vectors are applied on the network. Therefore, after each model in the training set and after the end of the training stage, weights must be memorised and adjusted, the weights changed only once. There is a "online version, simpler, which directly updates weights, in which case there may be the order in which the network vectors are displayed. Both weight gradients and current error have been initialised at 0 (alternative to 0) and E=0).

Step 3: The forward propagation of the signal

An example from the training set is applied to the inputs. Outcomes of neurons from hidden layer are computed:

$$y_j(p) = f\left(\sum_{i=1}^{n} x_i(p) \,.\, w_{ij} - \theta_j\right) \quad (5)$$

Where n is hidden layer number of inputs for neuron j and f is sigmoid activation.

Real network outputs are measured:

$$y_k(p) = f\left(\sum_{i=1}^{n} x_{jk}(p) \,.\, w_{jk} - \theta_k\right) \quad (6)$$

Where m is the input number from the output layer of neuron k.

Error updated by epoch:

$$E = E + \frac{(e_k(p))^2}{2} \quad (7)$$

Step 4: Backward error propagation and weight correction. Neuron error gradients are computed in the output layer:

$$\delta_k(p) = f' e_k(p) \quad (8)$$

where f' is the activation derivative function. A mistake occurs as

$$e_k(p) = y_{d,k}(p) - y_k(p) \quad (9)$$

If we use the single-pole sigmoid (equation 1, its derived is:

$$f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = f(x).(1 - f(x)) \quad (10)$$

Furthermore, the single-pole sigmoid is the function used. Then it becomes equation (10):

$$\delta_k(p) = y_k(p).\left(1 - y_k(p)\right).e_k(p) \quad (11)$$

The weight gradients from the secret layer to the output layer are updated:

$$\Delta w_{jk}(p) = \Delta w_{jk}(p) + y_j(p).\delta_k(p) \quad (12)$$

The gradients of the neuron errors are determined in the secret layer:

$$\delta_j(p) = y_j(p).\left(1 - y_j(p)\right).\sum_{k=1}^{l} \delta_k(p)\, w_{jk}(p) \quad (13)$$

where l is the number of network outputs.
The weight gradients of the input layer are modified with the hidden stage.

$$\Delta w_{ij}(p) = \Delta w_{ij}(p) + x_i(p).\delta_j(p) \quad (14)$$

Step 5: A new version. A new one. If the current training cycle still has test vectors, pass stage 3. If not the weights of all the links are changed according to the weight gradients:

$$w_{ij} = w_{ij} + \eta\,.\Delta\, w_{ij} \quad (15)$$

Where μ is the rate of analysis. When a stage is over, we check if it meets the termination condition ($E < E_{max}$ or a maximum number of periods of training have been achieved). Otherwise, we will move on to phase 2. If so, the MLP algorithm is over.

3.5. Simulations

On a PC with an i5 processor with an 8 GB RAM processing speed of 3.50 GHz, simulations can be brought. Sensor nodes are positioned on the grid intersection points and are first learned to locate as seen in the Figure. 5.

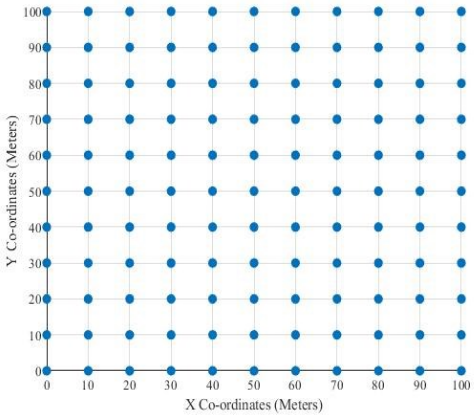


Figure 5: Illustrate the deployment of Sensor nodes and Mobile anchor node

The trained network with numerous training algorithms has placed localised sensor nodes on arbitrary points on the grid. The research is conducted. The node position is determined with the positioning information of the moving anchor node. The network is taught and checked using a handheld anchor node with the initial position to determine network efficacy (0.5, 0.5).

The sensor nodes are initially installed in a fixed location on the grid during testing. The handheld anchor node that has a set network gap has been deployed. Each mobile anchor node interval 't' sends its present position (X and Y coordinates) to sensor nodes that fall within its radio frequency spectrum. Each sensor node shall compute its current position using the proposed position algorithm by obtaining a minimum of three separate location information from mobile anchor node positions.

To determine the position of sensor nodes in the same grid, a further test of the qualified network is carried out. The nodes of the sensor were spread around the test grid uniformly. The testing is seen via the coordinate grid.

4. Results and discussion

For this study 5 different training algorithms BR, LM, RP, GD and SCG have been used for formation of the neural network, requiring supervised learning.

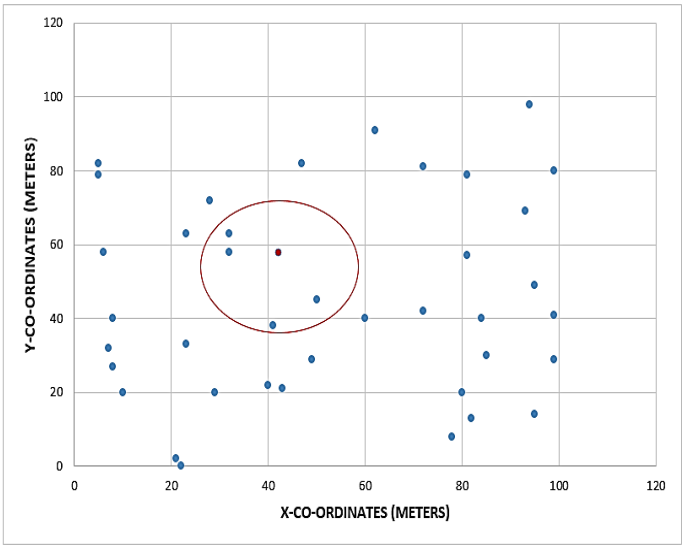


Figure 6: Illustration of Mobile anchor node with random deployment of sensor nodes

A neural network. Several testing functions with a different number of layers and nodes have been conducted. There are optimal results from network of 3 layers with a 9-12-2 structure. There were 9 neurons in input layer, hidden layer with 12 neurons and two in the output layers. Hyperbolic sigmoid tangent activation function for hidden layers, whereas pure line function in output layer has been used. This structure was utilized for the evaluation of the performance of the following anchor node configurations with different numbers of anchors.

An experiment was performed. 5 different BR, LM, RP, GD and SCG training algorithms for the formation of a neural network that require supervised learning have been used to this study. A network of neural components. A number of test functions have been performed with different layers and nodes. A network of three layers with a structure 9-12-2 produces optimal results.

$$e = \sum_{i=1}^{n} \frac{1}{n} \sqrt{(x_i - x_{oi})^2 + (y_i - y_{oi})^2}$$

where n is the overall number of test sets and exact position (x_{oi},y_{oi}) and estimated position of sensor node in i^{th} test dataset is (x_i,y_i).

The above mobile anchor node configurations were used for assessments of the performance of various anchor numbers. The average location error during this phase is illustrated in Figure 7. Figure 8 shows the time required for the network to train using different training algorithms.

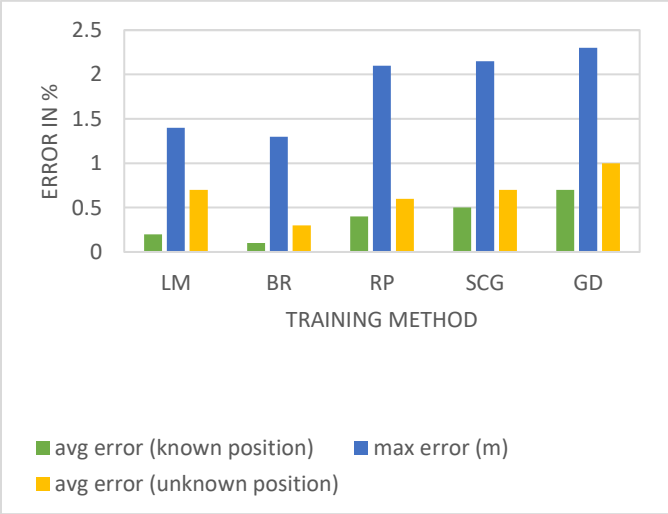


Figure 7: Illustration of localization error of different training algorithm with mobile anchor.

In Figure 7, Depicts different training algorithms have compares with localization error. The BR training algorithm is having less average error as compared to other training algorithms.

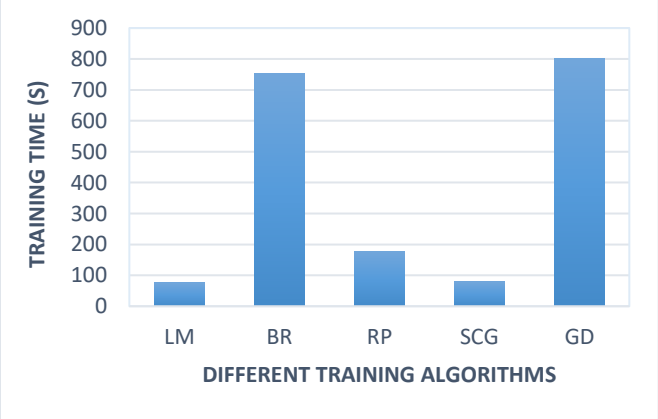


Figure 8: Illustrate time taken to train neural network for different training algorithms with mobile anchor

In Figure 8. Shows the training time taken by the ANNN model for different training algorithms. LM, RP and SCG are taking less training time and BR and GC training algorithms are taking more training time.

Cost of the network will be reduced with using one mobile anchor agent. Estimated Mean Localization Error (Meters) for different number of sensor nodes using BR training algorithm is shown in Figure 9 for the communication range of mobile anchor node that is R=40 meters, 60 meters and 80 meters.

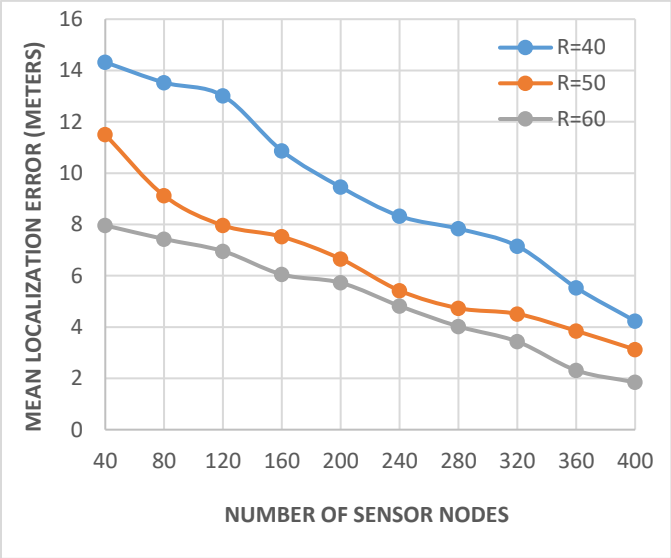


Figure 9: Estimation error when varying communication range.

In Figure 10. Depicts convergence accuracy of mobile anchor node with different time stamps for broadcasting their location information to static sensor nodes using BR training algorithm.

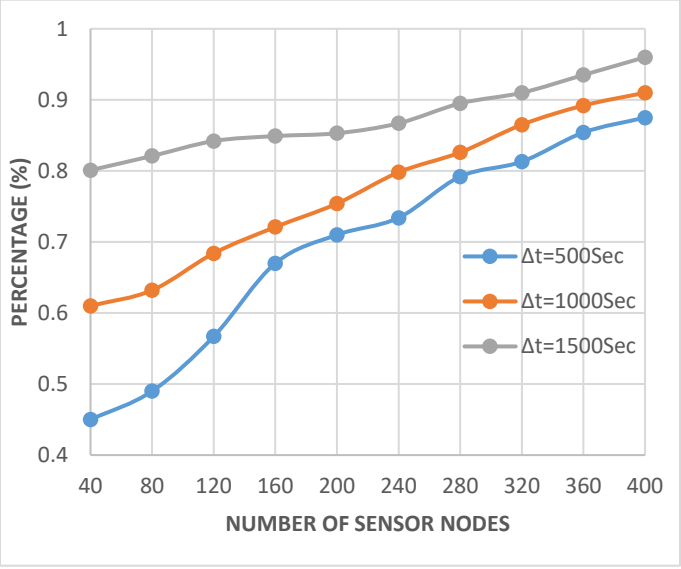


Figure 10: Convergence accuracy for various regular intervals.

5. Conclusions

The use of ANN has been suggested for an effective 2D localization algorithm for WSN. The RSSI signal values from the mobile anchor node are the entry to the proposed process. Since every mobile anchor node is fitted with wireless RF modules, no external hardware is necessary. The approach proposed is therefore , energy-efficient and uses just a two-way message to obtain the location inputs. As in the proposed system, only one mobile anchor node is used, even the cost is minimized. It is however, advisable to use the RSSI value of the mobile-based anchor Node beacon request signals, which may further reduce the location error. The following is recommended.

In order to obtain the finest neural network for localization, the BR training algorithm is evaluated to give the best results. Hence, the training time for this algorithm is quite high compared to other methods such as LM, RP, and SCG. It is therefore recommended that the BR method is used when offline training is performed as in this case. The LM method is recommended for applications requiring online training.

Conflict of Interest

The authors declare no conflict of interest related to this paper.

References

[1] D. Praveen Kumar, T. Amgoth, C. S. R. Annavarapu, "Machine learning algorithms for wireless sensor networks: A survey," Information Fusion, 49, https://doi.org/10.1016/j.inffus.2018.09.013.

[2] X. Feng, Z. Gao, M. Yang, S. Xiong, "Fuzzy distance measuring based on RSSI in Wireless Sensor Network," 2008 3rd International Conference on Intelligent System and Knowledge Engineering, https://doi.org/10.1109/iske.2008.4730962.

[3] M. S. Rahman, Y. Park, K. D. Kim, "Localization of Wireless Sensor Network using artificial neural network," 2009 9th International Symposium on Communications and Information Technology, https://doi.org/10.1109/iscit.2009.5341165.

[4] L. Zhao, X. Wen, D. Li, "Amorphous Localization Algorithm Based on BP Artificial Neural Network," International Journal of Distributed Sensor Networks, 11(7), https://doi.org/10.1155/2015/657241.

[5] K. Madhumathi, Dr. T Suresh, "Localization of Mobile Wireless Sensor Node Using Randomized Feature Sample of Estimated RSSI Value," Journal of Advanced Research in Dynamical and Control Systems, **11**(11-SPECIAL ISSUE), https://doi.org/10.5373/jardcs/v11sp11/20192943.

[6] A. Payal, C. S. Rai, B. V. R. Reddy, "Comparative Analysis of Sigmoidal Feedforward Artificial Neural Networks and Radial Basis Function Networks Approach for Localization in Wireless Sensor Networks," International Scholarly and Scientific Research & Innovation, **10**(7), https://doi.org/doi.org/10.5281/zenodo.1125123.

[7] A. Payal, C. S. Rai, B. V. R. Reddy, "Analysis of Some Feedforward Artificial Neural Network Training Algorithms for Developing Localization Framework in Wireless Sensor Networks," Wireless Personal Communications, **82**(4), https://doi.org/10.1007/s11277-015-2362-x.

[8] C. Xiao, & N. Yu," A predictive localization algorithm based on RBF neural network for wireless sensor networks," 2015 International Conference on Wireless Communications & Signal Processing (WCSP), https://doi.org/10.1109/wcsp.2015.7341145.

[9] S. Javadi, H. Moosaei, D. Ciuonzo, "Learning Wireless Sensor Networks for Source Localization," Sensors, 19(3), https://doi.org/10.3390/s19030635.

[10] G. Bhatti, "Machine Learning Based Localization in Large-Scale Wireless Sensor Networks," Sensors, **18**(12), https://doi.org/10.3390/s18124179.

[11] C. Zhou, L. Wang, L. Zhengqiu, "The Study of WSN Node Localization Method Based on Back Propagation Neural Network," Advances in Intelligent Systems and Computing, https://doi.org/10.1007/978-3-319-67071-3_54.

[12] S. H. Chagas, J. B. Martins, L. L. De Oliveira," An approach to localization scheme of wireless sensor networks based on artificial neural networks and Genetic Algorithms," 10th IEEE International NEWCAS Conference, https://doi.org/10.1109/newcas.2012.6328975.

[13] B.K. Madagouda & R. Sumathi, "Analysis of Localization Using ANN Models in Wireless Sensor Networks," 2019 IEEE Pune Section

International Conference (PuneCon), https://doi.org/10.1109/punecon46936.2019.9105871.

[14] B.K. Madagouda, R. Sumathi, A. H. Shanthakumara, "Localization of Sensor Nodes Using Flooding in Wireless Sensor Networks," Communications in Computer and Information Science, https://doi.org/10.1007/978-3-642-29219-4_72.

[15] R. Dela Cruz, "Artificial Neural Network-based Localization in Wireless Sensor Networks".

[16] M. Gholami, N. Cai, R. W. Brennan, "An artificial neural network approach to the problem of wireless sensors network localization," Robotics and Computer-Integrated Manufacturing, **29**(1), https://doi.org/10.1016/j.rcim.2012.07.006.

[17] B. Etzlinger, A. Ganhor, J. Karoliny, R. Huttner, A. Springer, "WSN Implementation of Cooperative Localization," 2020 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), https://doi.org/10.1109/icmim48759.2020.9299096

[18] X. Yang, X. Wang, W. Wang, "An Improved Centroid Localization Algorithm for WSN," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), https://doi.org/10.1109/itoec.2018.8740471.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1145-1150 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Allocation of Total Congestion Cost and load participation to Generators for a PoolCo Market in Deregulated Power System

Yashvant Bhavsar[1,*], Saurabh Pandya[2]

[1]*Electrical Engineering Department, Vishwakarma Government Engineering College, Ahmedabad, 382424, India*

[2]*Electrical Engineering Department, Lukhdhirji Engineering College, Morbi, 363641, India*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 05 February, 2021
Online: 25 February, 2021

Keywords:
Open access
Transmission congestion cost
Congestion cost allocation
Bialek' up-stream theorem
Poolco market

ABSTRACT

The objective of this paper is to allocate transmission congestion cost to responsible generators using a novel method. Deregulation of the electrical power system leads to the compulsion of open access to the transmission system for all entities of the power system. There is a trend to utilize cheaper generators by all loads. This leads to a violation of the operational and physical constraints of transmission corridors connected to those generators. It is not possible to utilize cheaper generators all the time due to the operational and physical constraints of the transmission lines. Hence there is an increase in the cost of energy produced. This increase in energy cost is taken into account as total congestion cost. Allocation of total congestion costs among various entities is always a complex task. Here, generators liable for the increase in total congestion cost identified using Bialek's algorithm. Bialek's upstream algorithm was applied to allocate congestion costs to generators. Results are obtained on IEEE-14 bus and IEEE-30 bus standard test systems.

1. Introduction

The electric power industry was operated as a *"regulated monopoly"* up to the nineties decade. Then after the structure of the electric power industry took a major shift and all its operations were unbundled, i.e. de-regulated. Because of the restructuring of the power utilities, the power industry is becoming turbulently competitive, and going through technological and regulatory changes, which affect its planning, operation, control, and services to customers. It is important to spot the consequences and impacts of those changes on the planning, operation, control and cost of the power system. In a deregulated power system, the transmission system must be available to all users without any discrimination. Due to physical and operating constraints of transmission corridors, not all the load be served by the cheaper generators, leading to the increased generation cost. The system operator (SO) must identify the responsible entity for this increase in cost. Determination of congestion cost and distribution of this among all participants is a crucial issue with the operation of a deregulated power grid.

*Corresponding Author: Yashvant Bhavsar, yashvantbhavsar@gmail.com

Market-based methods are methods that used either dispatch methods, generator rescheduling, load management or nodal pricing based methods. Pool and contract dispatch models with a priority of load curtailment applied using social welfare and contracts between generators and load [1, 2]. Optimal bus price-based simple methods were proposed [3] to calculate congestion cost supported the very fact that additional flow increases congestion. Various dispatch methodologies introduce [4] for open access of transmission line by representing a conceptual model of pool dispatch, bilateral dispatch and multilateral dispatch alongside the need for dispatch coordination between these models. Effectiveness of congestion clusters method for congestion management discussed for cost and loss minimization [5] with the definition of transmission congestion distribution factors (TCDFs) which used to evaluate the change in the flow of a rise in the injection of power at any bus. Two sets of sensitivity indices: Real Power transmission Congestion distribution Factors (PTCDFs) and Reactive Power transmission Congestion Distribution Factors (QTCDFs) used for congestion management [6] during which selection and participation of generators depend upon both their reactive sensitivity and bid price for up/downregulation. A simplified approach proposed [7] for

security-oriented power system operation in which the first contribution of each generator for a particular overloaded line is identified and then based on the relative electric distance (RED) concept the specified proportion of generation for the specified overload relieving obtained. A combined objective function [8] was obtained using the generation cost function and congestion management cost function incorporating additional penalty factors that were further solved by a bacterial foraging algorithm. A multi-objective congestion management framework with optimization of congestion management cost, voltage security and dynamic security called augmented ε-constraint technique discussed in which the first objective function is that the cost of congestion management while the second and third objective function is of voltage stability margin (VSM) and Corrected Transient Energy Margin (CTEM) [9]. Simple indices introduced [10] for effective and agreeable load curtailment in congested lines because of the sensitivity of load to a congested branch, economic incentives to reduce consumption and customer willingness to curtail load during congestion. These indices are achieved by mathematical formulation for maximizing social welfare, the overall index for possible load management and generator re-dispatch. Zonal congestion management approach discussed in[11,12], using real and reactive power sensitivity index as well as balancing energy up/down service settlement and other similar factors. Based on the calculation of the congestion cost index (CCI) total congestion cost index (TCCI) is minimized for a given number of buses [13]. CCI is obtained by difference of bidding rate and power during the pre-dispatch and re-dispatch era. Security constraint optimal power flow (SCOPF) with and without line flow constraints [14] used to determine total congestion cost which further distributed into line-wise congestion cost, whereas same SCOPF with inputs from energy management system (EMS) and state estimator applied iteratively [15] with contingency analysis used to estimate the actual cost of congestion. The participation of generators and loads within the use of the transmission network is decided employing a modified z-bus matrix [16]. The modified z-bus matrix is formed by modeling both generators and loads as constant admittances in the network considering them separately. A novel topological approach to MW-Mile proposed based on electricity tracing method [17]. The equation for topological generation distribution factor almost like that of generalized generation distribution factor defined, which further used for formation of supplement transmission charge of any generator in the network. Distributed energy sources (DERs) plays important role for alleviating transmission congestion cost. The usage based transmission cost division method is introduced in which transmission capacity is divided into four capacities to gauge the contribution of DERs to transmission cost [18]. A modified inage domain method for power flow tracing [19] used to allocate transmission cost and advance congestion management. Power tracking coefficient developed [20]with the help of kirchof's law for the active power flow tracking and revenue collected by individual generator determined. A method [21] suggested allocating the transmission fixed charge based on the modification of the impedance matrix in which real power flow through the line expressed in terms of load current also as electrical distance and voltage injection at the bus. Iteratively one by one impact of every generator selected on the network using the z-bus of the network [22] and transmission congestion cost determined based on the current decomposition method. A DC power flow based problem [23] using generation shift factor is defined as a lossless system to work out contribution-based congestion cost allocation methods in a bilateral market.

A fair allocation of transmission congestion cost to the generation companies is required in poolco market where generator dispatch pattern changes due to congestion. Here in this work transmission line congestion is identified by observing the value of the dual variable associated with constraints [14]. Total congestion cost (TCC) evaluated by generator re-dispatch method with and without constraints. The dual variable associated with the congested line is utilized to find the line allocation factor. Line-wise congestion cost (LWCC) of each congested line is evaluated with help of the line allocation factor. For power flow tracing, results obtained from common methods like graph method and node method are not accurate and it takes more computational time. To allocate congestion cost, the cost associated with congested lines should only be considered. A novel method discussed here using bialek's upstream algorithm [17,24] to allocate this TCC to all generators as well as the generator's contribution to supplying the load of the system.

2. Problem formulation

2.1. Total Congestion Cost (TCC) and Allocation to Line

A simple method to evaluate TCC is the generator re-dispatch method. In this method, optimal power flow (OPF) is applied with and without line flow constraints. The difference in evaluated costs of both cases is identified as TCC. This TCC is allocated to the lines under congestion. Generalized optimal power flow problem with constraints is formulated as:

$$C(P_{Gi}) = \sum_{i \in N_{gi}} (a_i P_{Gi}^2 + b_i P_{Gi} + C_i) \quad (1)$$

where,

N_{Gi} = Number of generators in the system

The above optimized problem is with several constraints as follows:

$$P_{Gi} - P_{Di} - P_{loss} = 0 \quad (2)$$

$$P_{Gimin} \leq P_{Gi} \leq P_{Gimax} \quad (3)$$

$$Q_{Gimin} \leq Q_{Gi} \leq Q_{Gimax} \quad (4)$$

$$V_{imin} \leq V_i \leq V_{imax} \quad (5)$$

$$P_{ij} \leq P_{ijmax} \quad (6)$$

where,

P_{Gi} = Active power generation in *MW* at bus *i.*

P_{Di} = Load in *MW* at bus *i.*

P_{Gimin} = Minimum limit of active power generation at bus *i.*

P_{Gimax} = Maximum limit of active power generation at bus *i.*

Q_{Gimin} = Minimum limit of reactive power generation at bus *i.*

Q_{Gimax} = Maximum limit of reactive power generation at bus *i*

V_{Gi} = Voltage at bus *i*

V_{Gimax} = Maximum voltage limit at bus *i*

V_{Gimin} = Minimum voltage limit at bus *i*

P_{ij} = Power flow in MW between bus *i* and *j*

P_{ijmax} = Power flow limit in MW between bus *i* and *j*

In the above constraints (2) is power balance constraint whereas (3) to (6) are inequality constraints.

Following are steps to evaluated TCC and its allocation to the congested line:

1. Run optimal line flow without and line flow constraints and obtain total generation cost (*TGC*) as per (1).
2. With line flow constraints, repeat optimal power flow for obtaining a revised total generation cost (*TGC'*).
3. Total congestion cost (*TCC*) is evaluated as:

$$TCC = TGC' - TGC \tag{7}$$

4. The dual variable associated with the constraint of (6) is utilized to obtain the line allocation factor L_{ij} as [14],

$$L_{ij} = \frac{\mu_{ij}(S_{ijmax} - S_{ij})}{\sum \mu_{ij}\,(S_{ijmax} - S_{ij})} \tag{8}$$

where,

μ_{ij} = Dual variable associated with constrained line

S_{ij} = Actual line flow between bus *i* and *j*

S_{ijmax} = Maximum line flow limit between bus *i* and *j*

5. Congested cost allocated to the constrained line is evaluated as:

$$LWCC_{ij} = TCC * L_{ij} \tag{9}$$

Equation (9) determines the line-wise congestion cost allocated to the constrained line from that of TCC.

2.2. Generator share to line congestion cost

In poolco model of the deregulated power system, it requires to identify generators (suppliers) liable for line congestion as well as the allocation of congestion cost to those generators is equally important. For fair allocation of congestion cost to the generators, generator share to line flow is to be determined. Here bialek's upstream algorithm [17,24] is employed to spot generators participation for the corresponding line flows from congested lines. The subsequent steps demonstrate to determine the sharing of line congestion cost to generators:

1. After carrying out OPF, convert net real power into lossless line flows $|P_{i-j}|$.
2. Next to determine node injection power as,

$$P_i = \sum_j |P_{i-j}| + P_{Gi} \tag{10}$$

For *i = 1, 2 ...n* *n* = number of buses

where,

n = number of buses
j = node directly supplying to the node *I*

3. The ratio of line flow to the node injection power, Generation cost coefficient (GCC), determined as,

$$C_{ji} = \frac{|P_{j-i}|}{P_i} \tag{11}$$

4. An up-stream distribution matrix A_u evaluated whose elements are determined as,

$$[A_u]_{ji} = \begin{cases} 1 & for\ i = j \\ -C_{ji} & for\ j\ supplyinng\ to\ node\ i \\ 0 & other\ wies \end{cases} \tag{12}$$

5. Take the inverse of the upstream distribution matrix evaluated with help of (12).
6. Determine the topological distribution factor by,

$$D_{i-l,k}^{g\ fact} = \frac{|P_{i-l}|[A_u^{-1}]_{ik}}{P_i} \tag{13}$$

7. Generator share to line flow is determined as,

$$|P_{i-l}| = \sum_{k=1}^{n} \left(D_{i-l,k}^{g\ fact}\right) P_{GK} \tag{14}$$

8. Line congestion cost allocated to generators as,

$$P_{gcc} = \frac{|P_{i-l}|}{P_{ij}} LWCC_{ij} \tag{15}$$

Distribution of line congestion cost to the generators on the proportional sharing base is obtained by (15).

2.3. *Generator participation to supply each load*

With help of the same upstream algorithm suggested by bialek, generator participation to supply the load can be determined. Following are steps to be followed for evaluating the same.

1. Power outflow from each node is determined as,

$$P_i^{ndout} = \sum_j P_{i-j} \tag{16}$$

For every node *j* connected to node *i*.

2. The ratio of net power flow to the node injection power P_i is given as,

$$P_i^{ndration} = \frac{P_i - P_i^{ndout}}{P_i} \tag{17}$$

3. Generator at node I participation to each load at bus j is calculated as,

$$P_{i-j}^{GL} = [A_u]_{j,i}^{-1}\, P_j^{ndratio} P_{Gi} \tag{18}$$

where,

P_{Gi} = Power generation at bus i
$[A_u]_{j,i}^{-1}$ = Upstream distribution matrix evaluated in section 2.2

3. Results and Analysis

The problem formulated during the previous section was solved with help of MATLAB programming. Initially, test results are obtained on small test bus systems like 6-bus and 9-bus test systems. Then after standard IEEE-14 and IEEE-30 test bus systems are used for obtaining results. In this section, the results of higher test bus systems are discussed. These test results are obtained for TCC, line congestion cost, and share of generators to TCC as well as line congestion cost.

3.1. IEEE-14 bus test results

IEEE-14 bus test system is with 20 lines and 5 generators. Figure 1 shows the line diagram of IEEE-14 bus system. OPF is

carried out without line flow constraints then after imposing line flow on line number 1 (connecting bus 1 and 2) and 4 (connecting bus 2 and 4). Generators' cost function is of 3rd order polynomial function. Bus number 1 is considered as slack bus for OPF. After that TCC, line congestion cost and generator share to the TCC as well as generator participation to each load are obtained as described in the previous section.

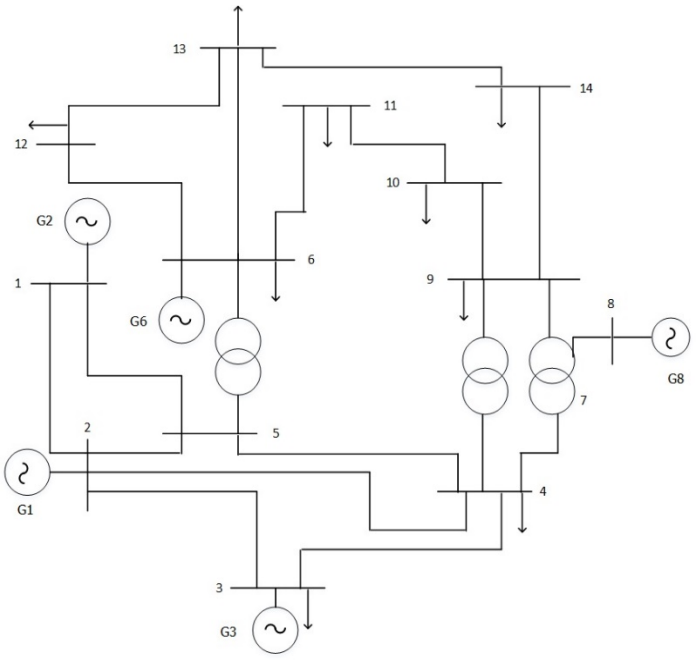


Figure 1: IEEE-14 bus test system

Table 1 shows the results obtained for the line congestion cost. Power flow without line flow constraints is 129.667 MW through line 1. With an imposing line flow limit of 110 MW, it is reduced to 109.94 MW. Similarly, for line number 4, after imposing a line flow limit of 40 MW, line flow is reduced to 38.392 MW from that of 55.3137 MW.

Table 1: IEEE-14 bus results for line congestion cost

Line no	Fbus	Tbus	Pflow w/o constraints in MW	Pflow with constraints in MW	Max line flow in MW	Line constraints allocation factor	Line congestion cost in $/hr	Total congestion cost in $/hr
1	1	2	129.667	109.94	110	0.493	28.53	57.88
4	2	4	55.3137	38.392	40	0.507	29.34	

Total congestion cost is 57.88 $/hr, further allocated to the congested line from line constraints allocation factor is 28.53 $/hr and 29.35 $/hr. Equation (14) is for identifying generator share to line flows. For IEEE-14 bus system, the generator share to the line flows is shown in Figure 2.

As shown in Figure 2, generator 1 at bus 1 is responsible for the major contribution to the line flow because the generation of the generator is remarkably high compared to all other generators. Subsequent contributor to the line flow is generator 2 connected at bus 2, whereas other generators are contributing very less to the line flow. Table 2 shows the allocation of congestion cost to the generators from that of line congestion cost.

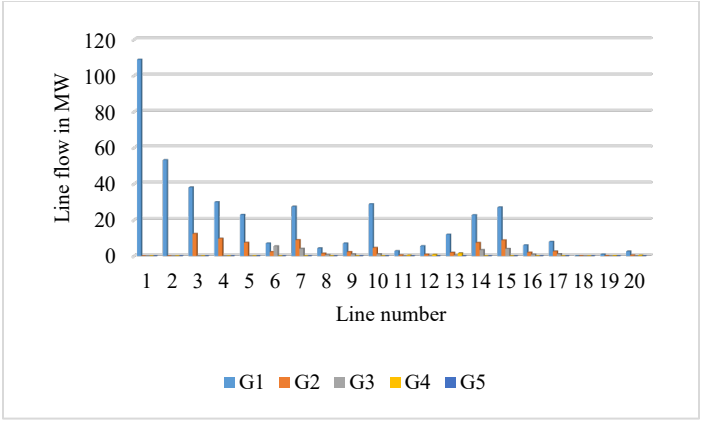


Figure 2: IEEE-14 bus generator share to line flows

Table 2: IEEE-14 bus line congestion cost allocated to generators

From bus	To bus	G1	G2	G3	G6	G8	Total	Line Cong Cost
1	2	28.532	0	0	0	0	28.532	28.532
2	4	22.175	7.173	0	0	0	29.348	29.348
Total Cost		50.707	7.173	0	0	0	57.880	57.880

All congestion costs are in $/hr in Table 2. As described in Table 2, generator 1 is sharing most of the line congestion cost as it is a major supplier to the load. It is observed that the summation of congestion cost shared by generators is equal to the total of the line congestion cost as well as that of TCC. Hence results of TCC and line congestion costs allocated to generators are self validated.

Generation participation to the system load is identified with help of bialek's upstream algorithm described in section 2.3. . The total load of the IEEE-14 bus system is 259 MW and the total generation of the system is 264.13 MW. Following Figure 3 is depicting how each load of the system is shared by individual generators.

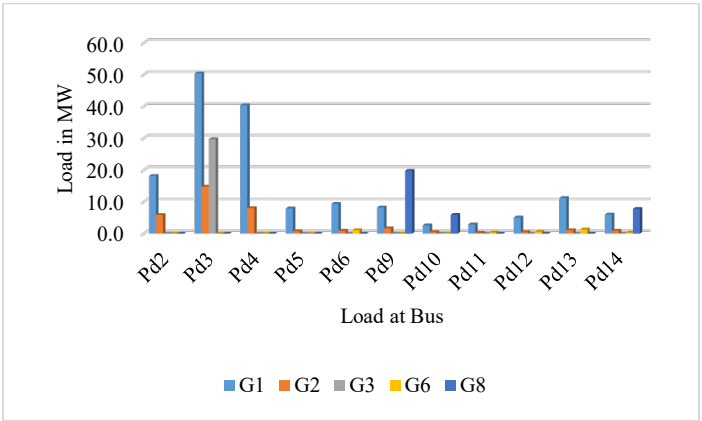


Figure 3: IEEE-14 bus load share by individual generators

Pd2, Pd3… Pd14 are the power demands (loads) at corresponding buses of the system. It is observed that generator 1 is sharing most load of the system followed by generator 2. This is in line with the sharing of line flow by all generators. Generation of generator 1 is 162.15 MW and so it serves a total load of 162.15 MW of the system with maximum sharing of 50 MW load connected at bus 2. Generators connected at bus numbers 1, 2 and

3 are supplying a load of 227.03 MW load, almost 85% of the system load.

3.2. IEEE-30 bus test results

Another standard test bus system, IEEE-30 bus system is used to get similar results as obtained on IEEE-14 bus system. As shown in Figure 4, IEEE-30 bus system having 41 lines along with six generators. The entire load connected to the system is 283.4 MW.

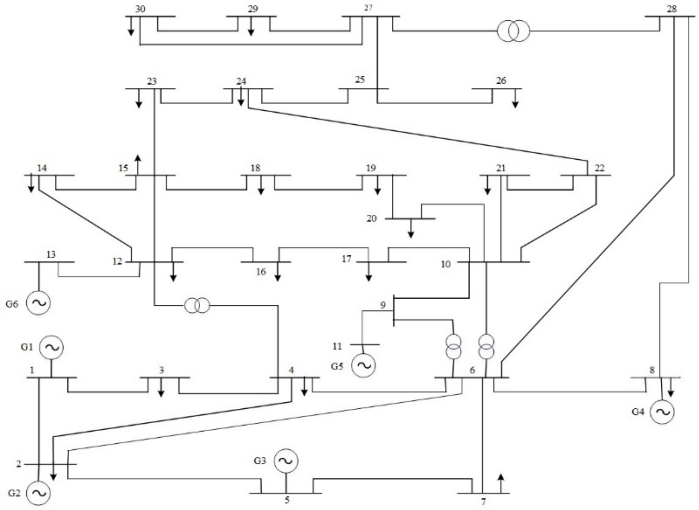


Figure 4: IEEE-30 bus system

TCC and line congestion costs are evaluated for the IEEE-30 bus system and the result obtained are tabulated in Table 3. Here, OPF is carried out without line flow constraints and then imposing line flow constraints on line one (connecting bus 1 and 2) and seven (connecting bus 4 and 6) to identify TCC. Allocation of TCC to line congestion cost is shown in Table 3.

Table 3: IEEE-30 bus results for line congestion cost

Line no	Fbus	Tbus	Pflow w/o constraints in MW	Pflow with constraints in MW	Max line flow in MW	Line constraints allocation factor	Line congestion cost in $/hr	Total congestion cost in $/hr
1	1	2	139.115	99.997	100	0.0511	8.16	159.48
7	4	6	55.314	38.393	40	0.9489	151.32	

Here it is revealed that after imposing line flow limits, power flow reduced to 99.99 MW and 38.39 MW through line numbers one and seven. TCC evaluated is 159.48 $/hr due to congestion. With help of line allocation factors of every congested line, this TCC was further allocated as 8.16 $/hr for line number 1 and 151.32 $/hr for line number 7.

Figure 5 shows the line flow share of generators up to line number 20. Again in this case also generator 1 is sharing the remarkable load of the system, hence the line flow share of generator 1 is high compared to all other generators. For the remainder of the lines, it is noticed that only except that of generator 2, a generator connected at bus 4 (generator 4) is sharing somewhat of line flow particularly of line number 36 and 40.

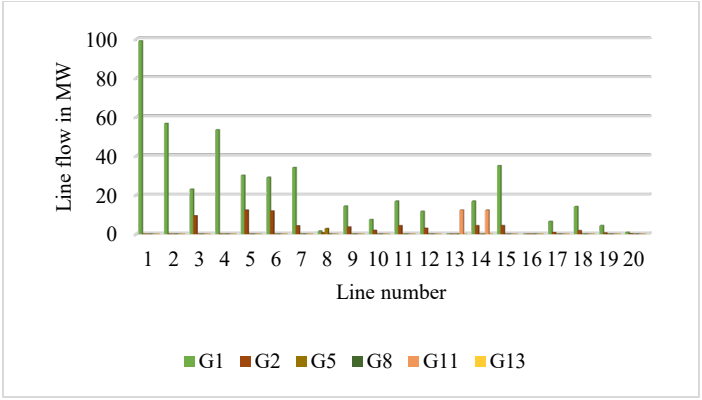


Figure 5: IEEE-30 bus generator share to line flows (up to line 20)

Table 4: IEEE-30 bus line congestion cost allocated to generators

From bus	To bus	G1	G2	G3	G4	G8	Total	Line Cong Cost
1	2	8.16	0	0	0	0	8.16	8.16
4	6	134.84	16.48	0	0	0	151.32	151.32
Total Cost		143.0	16.48	0	0	0	159.48	159.48

Table 4 states line flow congestion cost allocation to individual generators. Although IEEE-30 bus system contains a total of six generators, Table 4 shows five generators as only two generators (at bus numbers 1 and 2) are sharing line congestion cost. Here also the summation of individual sharing of line congestion cost by generators is equal to TTC and summation of line congestion cost.

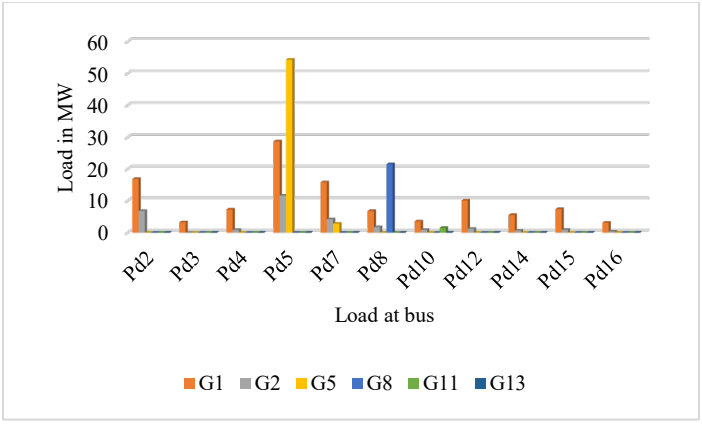


Figure 6: IEEE-30 bus load share by individual generators (first 11 loads)

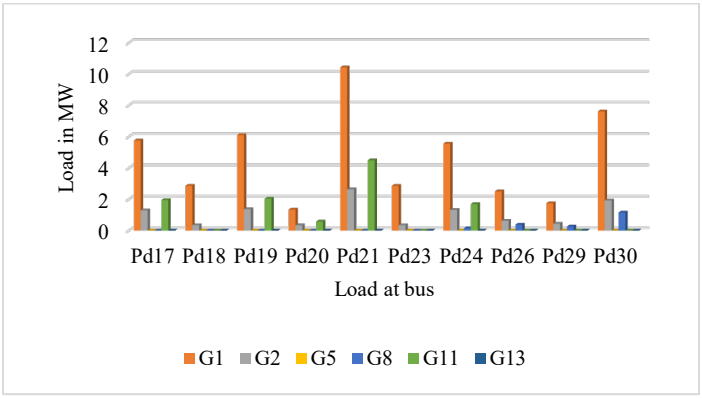


Figure 7: IEEE-30 bus load share by individual generators (Rest10 loads)

Figure 6 and Figure 7, show the load participating by all generators of IEEE-30 bus system. There are a total of 21 loads connected to various buses as shown in Figure 5 and Figure 6. Out of a total system load of 283.4 MW, generator 1 is sharing 155.2 MW of load, and generator 3 is connected at bus 5 sharing a load of 57.13 MW.

4. Conclusion

In this paper, a method is proposed to allocate TCC to congested lines with help of a dual variable associated with these congested lines. Further, a novel method discussed to allocate these line congestion costs to generators with help of bialek's upstream algorithm of tracing the flow of electricity. Also, with help of the same algorithm formulation is done for the participation of system generators to each load connected to the system. Results are obtained for IEEE-14 bus and IEEE-30 bus test systems. Results obtained are self-validictory, and reveal that TCC is equal to the summation of individual line congestion cost as well as equal to the summation of individual congestion cost shared by all accountable generators.

References

[1] A.K. David, R.S. Fang, "Congestion management of electric power systems under open transmission access," IEE Conference Publication, (450), 469–474, 1998, doi:10.1049/cp:19971879.

[2] R.S. Fang, A.K. David, "Transmission congestion management in an electricity market," IEEE Transactions on Power Systems, **14**(3), 877–883, 1999, doi:10.1109/59.780898.

[3] T.W. Gedra, "On transmission congestion and pricing," IEEE Transactions on Power Systems, **14**(1), 241–248, 1999, doi:10.1109/59.744539.

[4] A.K. David, "Dispatch methodologies for open access transmission systems," IEEE Transactions on Power Systems, **13**(1), 46–53, 1998, doi:10.1109/59.651612.

[5] S. Charles Raja, P. Venkatesh, B. V. Manikandan, "Transmission Congestion Management in restructured power systems," 2011 International Conference on Emerging Trends in Electrical and Computer Technology, ICETECT 2011, 23–28, 2011, doi:10.1109/ICETECT.2011.5760085.

[6] E. Muneender, D.M. Vinod Kumar, "Optimal rescheduling of real and reactive powers of generators for zonal congestion management based on FDR PSO," Transmission and Distribution Conference and Exposition: Asia and Pacific, T and D Asia 2009, 1–6, 2009, doi:10.1109/TD-ASIA.2009.5356989.

[7] G. Yesuratnam, M. Pushpa, "Congestion management for security oriented power system operation using generation rescheduling," 2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2010, 287–292, 2010, doi:10.1109/PMAPS.2010.5528719.

[8] A. Ramesh Kumar, L. Premalatha, "Security constrained multi-objective congestion management in transactional based restructured electrical network using bacterial foraging algorithm," Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2013, (1), 63–67, 2013, doi:10.1109/ICCPCT.2013.6528913.

[9] M. Esmaili, N. Amjady, H.A. Shayanfar, "Multi-objective congestion management by modified augmented ε-constraint method," Applied Energy, **88**(3), 755–766, 2011, doi:10.1016/j.apenergy.2010.09.014.

[10] Y. Niu, Y.L. Cong, T. Niimura, "Transmission congestion relief solutions by load management," Canadian Conference on Electrical and Computer Engineering, **1**, 18–23, 2002, doi:10.1109/ccece.2002.1015168.

[11] J. Yu, J. Galvin, "Zonal congestion management and settlement," Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference, **1**(SUMMER), 249–254, 2001, doi:10.1109/pess.2001.970021.

[12] A. Kumar, S.C. Srivastava, S.N. Singh, "Congestion management in competitive power market: A bibliographical survey," Electric Power Systems Research, **76**(1–3), 153–164, 2005, doi:10.1016/j.epsr.2005.05.001.

[13] K.Y. Lee, S. Manuspiya, M. Choi, M. Shin, "Network congestion assessment for short-term transmission planning under deregulated environment," Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference, **3**(WINTER MEETING), 1266–1271, 2001, doi:10.1109/pesw.2001.917258.

[14] A. Jana, "A new approach to transmission congestion cost calculation," 2008 Joint International Conference on Power System Technology POWERCON and IEEE Power India Conference, POWERCON 2008, 9–12, 2008, doi:10.1109/ICPST.2008.4745363.

[15] T.J. Overbye, "Estimating the actual cost of transmission system congestion," Proceedings of the 36th Annual Hawaii International Conference on System Sciences, HICSS 2003, 2003, doi:10.1109/HICSS.2003.1173857.

[16] Y.P. Molina, O.R. Saavedra, C. Portugal, "Allocation of transmission network cost using modified Zbus matrix," International Journal of Electrical Power and Energy Systems, **63**, 323–330, 2014, doi:10.1016/j.ijepes.2014.06.007.

[17] J. Bialek, "Topological generation and load distribution factors for supplement charge allocation in transmission open access," IEEE Transactions on Power Systems, **12**(3), 1185–1193, 1997, doi:10.1109/59.630460.

[18] J.X. Wang, H.W. Zhong, Q. Xia, "A usage-based transmission cost allocation scheme considering impacts of DERs," IET Conference Publications, **2016**(CP705), 2016, doi:10.1049/cp.2016.1184.

[19] A. Baczynska, W. Niewiadomski, "Power flow tracing for active congestion management in modern power systems," Energies, **13**(18), 2020, doi:10.3390/en13184860.

[20] D. Bhowmik, A.K. Sinha, "An efficient cost based allocation approach for individual generators associated with the system," Computers and Electrical Engineering, **70**, 212–228, 2018, doi:10.1016/j.compeleceng.2017.12.040.

[21] J. Nikoukar, M.R. Haghifam, A. Parastar, "Transmission cost allocation based on the modified Z-bus," International Journal of Electrical Power and Energy Systems, **42**(1), 31–37, 2012, doi:10.1016/j.ijepes.2012.03.002.

[22] H. Zein, "Cost allocation of transmission usage based on current magnitude," 2013 International Conference on Quality in Research, QiR 2013 - In Conjunction with ICCS 2013: The 2nd International Conference on Civic Space, 175–179, 2013, doi:10.1109/QiR.2013.6632560.

[23] M.P. Abdullah, M.Y. Hassan, F. Hussin, "Assessment of contribution-based congestion cost allocation using AC and DC for bilateral market," PECon2010 - 2010 IEEE International Conference on Power and Energy, (2), 897–901, 2010, doi:10.1109/PECON.2010.5697706.

[24] J. Bialek, Tracing the flow of electricity, IEE Proceedings: Generation, Transmission and Distribution, **143**(4), 313–320, 1996, doi:10.1049/ip-gtd:19960461.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1151-1156 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Parameters Degradation Analysis of a Silicon Solar Cell in Dark/Light Condition using Measured I-V Data

Dominique Bonkoungou[1,*], Toussaint Guingane[1], Eric Korsaga[1], Sosthène Tassembedo[1], Zacharie Koalaga[1], Arouna Darga[2], François Zougmore[1]

[1]*Laboratory of Materials and Environment (LAME), University of Ouagadougou, Ouagadougou, 7021 Burkina-Faso*

[2]*Group of electrical engineering, Paris (LGEP) 11, rue Joliot Curie, Plateau de Moulon, 91192, Gif sur Yvette*

ARTICLE INFO

Article history:
Received: 26 December, 2020
Accepted: 31 January, 2021
Online: 25 February, 2021

Keywords:
Parameters degradation
Dark condition
Light condition
Efficiency
Shade condition

ABSTRACT

In this paper, we investigate and analyze parameters degradation in a typical photovoltaic (PV) cell, which lead to power loss under dark as well as light condition using measured current-voltage (I-V) data. A nonlinear least squares method to extract the parameters such as the reverse saturation currents, the ideality factors, the series and shunt resistances of the cell from the dark current-voltage (I-V) curves is used. In order to analysis the sensitivity of the dark current-voltage (I-V) measurement to each of the six extracted parameters as a function of the voltage as well as the temperature and the density current, we simulate the operation of a silicon solar cell (KXB0022-12X1F). The analysis of the dark current-voltage (I-V) curves permit us to detect variation as small as 15% in the series resistance. We also extends the use of dark as well as light current-voltage (I-V) measurements to modules configurations of cells and uses a nonlinear least squares method to evaluate the cell efficiency parameters in the modules. Results obtained show a degradation of the values of the maximum power (Pmax) as compared to initial values by about 12, 3%, 12, 06% and 10, 21 % respectively in Total-Cross-Tied (TCT), Bridge-Link (BL) and Honey-Comb (HC) configurations.

1. Introduction

The analysis of the current-voltage curves of solar cells parameters is an important tool for quality control and evaluation of their efficiency. Though photovoltaic (PV) cells are a very important source of electrical energy, some studies [1-5] show that in a number of ways, the photovoltaic (PV) cells/modules can fail. In the circumstances; it is very important to understand the origin and behavior of these faults, before trying to improve them. However, the analysis approach required for determining the source of power loss within a solar cell depend on the fault itself, and may necessitate an analysis process which adds to system operation costs and may not always be feasible [6-8]. Authors in [9] have classified the degradation of PV cells into groups such as early degradation and long-term degradation. Early degradation are defined as encapsulant discoloration and delamination; cracks in the cell [7, 8], burnt cell caused by hot spot; soiling and shading cell, defect in anti-reflective [9]. The early degradation are also inspected in long-term field-aged PV cells/modules. Long-term degradation can be defined as corrosion of solar joint or crack in solder joint of the PV cells connection [9]. Understanding the cause of PV cell fault and how they affect the efficiency of the PV cells is essential to improve their dependability.

In this paper, we will focus on investigating and analyzing some parameters degradation that lead to power loss in PV cells under dark/light condition using measured I-V data. The 2-diodes equivalent circuit is used to describe the electronic properties of the solar cell. A set of analyzed parameters are determined from the dark/light I-V characteristics that can identify some fault of the PV modules aggregation of electrical components and solar cells, as well as shunting and recombination losses occurring due to cell cracks, shading cell [8, 10]. An experimental setup for measuring both dark and light I-V characteristics of solar cells is performed for this purpose. It is a complementary study of a previous work in which parameters of a typical solar cell named KX0B22-12X1F have been extracted under dark condition. This work also extends the use of dark as well as light I-V measurements to modules configurations of cells and uses a nonlinear least squares method [11] to evaluate the cell efficiency parameters in the modules. Results obtained are used to analyze

*Corresponding Author D. Bonkoungou, P.O. Box 243 Ouaga 10, contact: 00226 78 61 52 85 & Email: dominique.bonkoung@gmail.com

the shade impact in the PV cell efficiency and the effect of the cell imperfection on PV module efficiency.

2. Experiment

We have made measurement on 270 mm² silicon solar cell (KXB0022-12X1F). The I-V characteristics of the cell have been measured with a system according to international standard consisting of a Keithley 2400 multimeters; a sol-UV simulator (Oriel 81192/ 1000 W); an arc lamp source (Oriel -68920). A desktop computer is used to processes the measured data using LABVIEW. To measure the I-V curve under dark we used the procedure described into [11]. For light I-V curve measurements, the cell is illuminated by a halogen (ORIEL) .This last is fed by a halogen power supply. We performed two experiments for the measurements with homogeneous as well as heterogeneous impacts on light I-V curves. Pictures of the experiment approach are displayed in figure 1. Under dark condition, the data are measured in the forward and reverse direction using a power supply. For the power supply, the voltage is measured from – 5 Volts to 1 Volt with a resolution of 0.01 Volt. The maximum current was 1 A, which corresponds to 42.4 mA/cm².



Figure 1: Experimental setup

3. Parameters extraction and fault analysis

3.1. I-V characteristics under dark condition

The 2-diodes model is used to describe the dark I-V curves of the PV cell [11].

$$I = \frac{V - R_S I}{R_p} + I_{02}\left[\exp\left(\frac{V - R_S I}{n_2 U_{th}}\right) - 1\right] + I_{01}\left[\exp\left(\frac{V - R_S I}{n_1 U_{th}}\right) - 1\right] \quad (1)$$

where $U_{th} = \frac{KT}{q}$ is the thermal voltage, q is the electron charge, K is the Boltzmann's constant, T is the temperature; R_s is the series resistance; $I_{0i}(i = 1,2)$ is the dark saturation current of the diode; R_p is the shunt resistance; $n_i(i = 1,2)$ is the ideality factor. In our previous work [11], the six parameters like $\left[I_{0i}(i = 1,2); n_i(i = 1,2); R_s; R_p\right]$ are obtained by fitting (1) to a set of measured data using a nonlinear squares method of dark I-V measurement data. The standard deviation (SD) between the computed and measured data is less than 3% [11].

3.2. I-V characteristics under illuminated condition

To analyze the results obtained for the various parameters $\left[I_{0i}(i = 1,2); n_i(i = 1,2); R_s; R_p\right]$ under dark condition [11], we have made additional measurements in light condition. Under illuminated condition, the PV cell generate a photocurrent I_{light} and this last has to be added to (1) expressed by (2):

$$I + I_{light} = \frac{V - R_S I}{R_p} + I_{02}\left[\exp\left(\frac{V - R_S I}{n_2 U_{th}}\right) - 1\right] + I_{01}\left[\exp\left(\frac{V - R_S I}{n_1 U_{th}}\right) - 1\right] \quad (2)$$

All the parameters are defined previously. To study the efficiency of the PV cell in light condition, it is important to analyze the following parameters such as: the open circuit voltage (V_{OC}), the short circuit photo-generated current (I_{SC}) and the fill factor (FF) [3]. For the determination of the cell parameters under illumination, we proceeded as follows. We have determined the I_{SC} value by assuming that the short circuit condition is verified i.e. for $V = 0$ and $I = I_{SC}$. We get the following equation:

$$I_{SC} + I_{light} = \frac{0 - R_S I_{SC}}{R_p} + I_{02}\left[\exp\left(\frac{0 - R_S I_{SC}}{n_2 U_{th}}\right) - 1\right] + I_{01}\left[\exp\left(\frac{0 - R_S I_{SC}}{n_1 U_{th}}\right) - 1\right] \quad (3)$$

Likewise, for $V = V_{OC}$ and $I = 0$ at open circuit condition, we obtain:

$$I_{light} = \frac{V_{OC}}{R_p} + I_{02}\left[\exp\left(\frac{V_{OC}}{n_2 U_{th}}\right) - 1\right] + I_{01}\left[\exp\left(\frac{V_{OC}}{n_1 U_{th}}\right) - 1\right] \quad (4)$$

Using (3) and (4), we yield a relationship between V_{OC} and I_{SC} given by equation 5:

$$V_{OC} + R_S I_{SC} = R_P\left[I_{01}\left(\exp(-\beta_1 R_S I_{SC}) - \exp(\beta_1 V_{OC})\right) + I_{02}\left(\exp(-\beta_2 R_S I_{SC}) - \exp(\beta_2 V_{OC})\right)\right] \quad (5)$$

where $\beta_i = \frac{1}{n_i U_{th}}$ $(i = 1,2)$ .To fitting the illuminated I-V characteristics, we inserted the parameters values obtained in dark condition into (2) and determined V_{OC} as well as I_{SC}. For I_{light}, we have exploited the value of I given by (1) whose expansion for low voltages allows us to establish the following relation:

$$I_{Light} = I_{SC}\left(1 + \frac{R_S}{R_P}\right) \quad (6)$$

The fill factor (FF) can be calculated by means of the definition:

$$FF = \left(V_{mp} I_{mp}\right)/\left(V_{OC} I_{SC}\right) \quad (7)$$

where $(V_{mp}.I_{mp})$ is the upper value of the voltage-current (V.I) product.

4. Results and Discussion

4.1. PV cell parameters sensitivity analysis both in dark and light condition

In order to analysis the sensitiveness of the dark current-voltage measurement to each of the six parameters as a function

R_P, R_S are relatively neglected. We analyze the different power losses by progressively degrading the parameters of the cell, one at a time. We varied the parameters such as I_{01}, I_{02}, R_S and n_2 by increasing their value. During this time, we decreased the value of R_P and n_1. It can be seen in Figure.2 that increased I_{02} impact the low value voltage as well as the medium voltage region of the dark I-V curve. This phenomenon can be explained by the increase of the number of fault in the cell junction. Unlike I_{02}, the increase in I_{01} affects both the medium and the high voltage ranges of the dark I-V curve. The sensitivity to R_P and n_2 are important at low and medium voltage ranges. We can conclude that the increase in I_{01}, I_{02}, R_P and n_2 affect and present distinct curvatures in the low and medium voltage ranges of the curve.

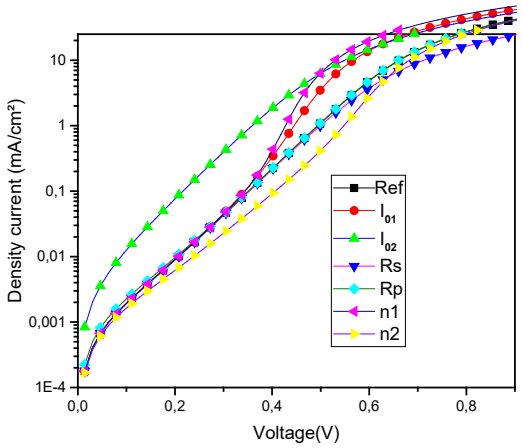


Figure 2: Dark I-V curves of a particular solar cell (KX0B22-12X1F) affected by variant faults

Then, an exhaustive analysis of these different variations (see Figure 1) can make it possible to detect operating faults. Using the measured dark I-V characteristics, we analyzed the sensitivity of each extracted parameter as a function of the current, shown in Figure 3. The parameter β is defined as being the sensitivity ratio described as follow: $\beta = \frac{\delta V}{\delta \alpha_i} * \alpha_i$ where $\alpha_i (i = 1,2,3,4,5,6)$ represents the different parameters i.e. $\alpha_1 = I_{01}$, $\alpha_2 = I_{02}$, $\alpha_3 = R_P$, $\alpha_4 = R_S$, $\alpha_5 = n_1$ and $\alpha_6 = n_2$. These sensitivity coefficients show the dependencies of the solar cell parameters on the current level. As can be observed that for currents smaller than 400 mA/cm², the sensitivity coefficient to R_P rises to 0, 3% as well as 0, 45% for n_2. And at high current ranges, the sensitivity to R_S increases from 0% to 0, 15% in the studied interval.

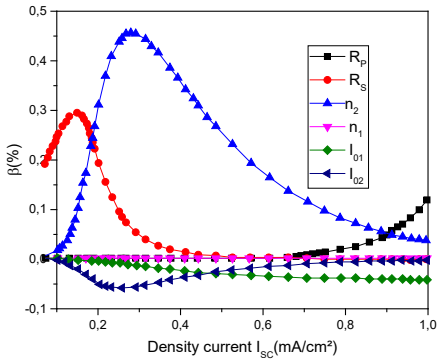


Figure 3: Sensitivity analysis of each extracted parameter $\left[I_{0i}(i=1,2); n_i(i=1,2); R_s; R_p\right]$ as a function of density current

The increase of the value of R_S drastically affect the high current and voltage regions shown in Figure 4. This shows that the analysis of the dark I-V characteristics of a solar cell is a very important tool which permit us to detect variations as small as 15% in R_S. Besides, the disturbance of R_S under dark condition made it possible to identify the faults which hinder the proper functioning of the PV cell, when we realize that the R_S in a particular PV cell as well as module would have to double before a 5% drop in power resulted [12]. The analysis of disturbances in the other parameters such as $I_{0i}(i=1,2); n_i(i=1,2); R_p$ can allow us to identify certain operating faults. And it could make easier to modify the manufacturing process to improve cell performance.

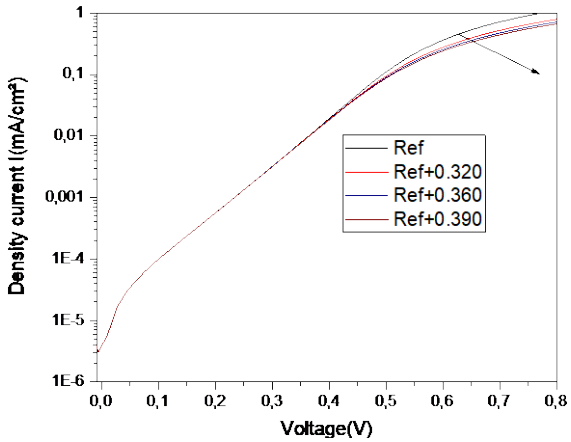


Figure 4: Effects of series resistance on the semilog plot of the dark I-V curves

Figure 5 presents the fitted and measured I-V characteristics of the PV cell in light condition. As can be noted that the fitted I-V curve agreed very well to the measured I-V characteristic. From this figure, one has determined the values of I_{SC} and V_{OC} by extrapolation to $V = 0$ axis respectively to $I = 0$ and FF by using equation 7.

The different parameters V_{OC}, I_{SC} and FF are directly obtained from the light I-V curve. The values of these parameters are given in Table1. In this table, we compares the fitted and

measured data at room temperature. We can remark that the results obtained reveal a good approximation with the two I_{sc}-V_{oc} characteristics.

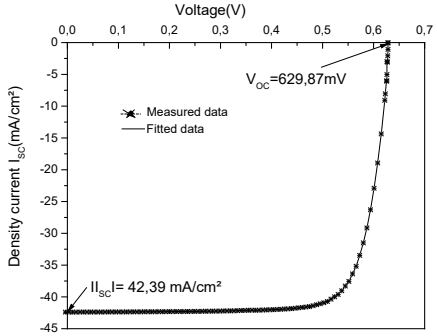


Figure 5. Fitted and measured light I-V curves at 300K.

Table 1: Measured and fitted parameters of the studied PV cell under illumination condition

Parameters	I_{SC} (mA/cm²)	V_{OC} (mV)	FF (%)
Measured data	-42,39	629,87	72, 21
Fitted data	-42,39	628,76	71,32
Deviation (%)	0	-0,27	-1,2

The plots in Figure 6 present the effects of temperature on the light I-V characteristics of a particular silicon solar cell. The plot clearly demonstrates that the open circuit voltage (V_{OC}) decreases with increasing temperature; whereas the short circuit current (I_{SC}) remains unchanged throughout the temperature range studied i.e. $I_{SC} \cong 42{,}40$ mA/cm². This shows that V_{OC} considerably depends on the temperature.

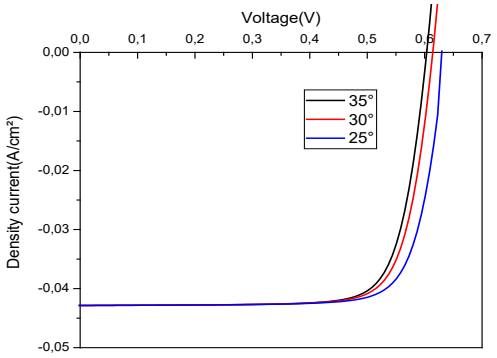


Figure 6: Density current I_{SC} versus voltage V_{OC} for variant temperatures

In order to analyze the effect of partial or total shading on the studied PV cell, we performed two experiments for the measurements with homogeneous as well as heterogeneous impacts on light I-V curves. Firstly, we measured the light I-V curves of the PV cell with uniform irradiances at Standard Test Condition (STC: 1000 W/², 25° C; AM.1.5) shown in Figure 7; these data are used as a baseline for the future analysis.

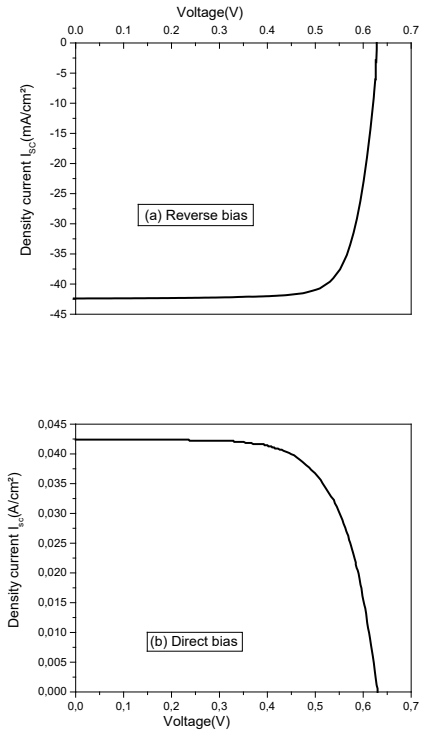


Figure 7: measured I-V characteristic of the studied PV cell under illumination in direct (a) and reverse (b) bias at STC

Secondly, we covered 25% (respectively 50% and 75%) of the area of the PV cell with a light-tight lid. Afterwards, the PV cell was illuminated over about 25% (respectively 50% and 75%) of its area. Results can be observed in Figure 8. It can be noted that the increase in the rate of shading of the cell produces a higher reduction of short circuit current I_{SC}, which causes a reduction in maximum power. It clearly demonstrate that the cell generated current is a function of the rate of light transmittance.

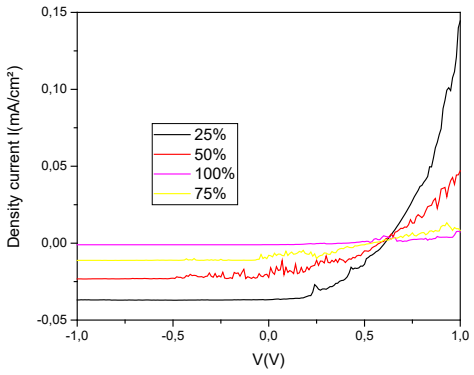
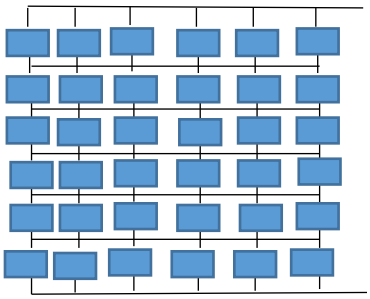


Figure 8: effect of partial shading of the studied PV cell on light I-V characteristics

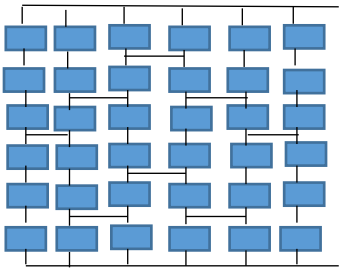
4.2. Analysis of three (3) types cell interconnections under partial shade conditions

This section offers an extensive study of different reconfigurations of PV cells operating under partial shade conditions. For this purpose, we have studied the shading rate of 3 types of configurations such as Total-Cross-Tied (TCT),

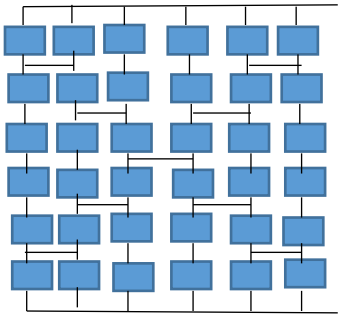
Bridge-Link (BL) and Honey-Comb (HC) of PV cells shown in Figure 9.



a) TCT



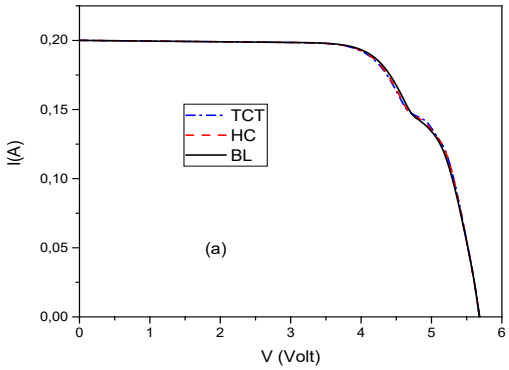
b) BL



c) HC

Figure 9: Types of 36 PV cells interconnection in a typical PV module

The results of shading rate of these configurations have been analyzed by shading completely selected cells in the PV module. Two experiments are made: firstly, one of the 6 X 6 PV cells is 100% shade for each type of configuration; secondly, nine (9) cells are 100% shaded and twenty-seven (27) working properly. I-V and P-V characteristics of these conditions are shown in Figures 10 and 11. First thing that can be observed is that there is a degradation of the power produced depending of the type of configuration. For the first experiment, shading one cell involve an efficiency decrease; for the second case, shading nine (9) cells involve efficiency to decrease further. We can note that in Figure 10, the values of the maximum power (Pmax) decreased as compared to initial values by about 12, 3%, 12, 06% and 10, 21 % respectively in TCT, HC and BL interconnections. Figure 11 proves that the decrease of Pmax can be imputed to the decrease of the short circuit current I_{SC}. The reduction in I_{SC} is explained by the rate of shading cells which limit the light transmittance. The influence of the number of shaded cell is presented in bar chart form in Figure 12. The TCT and BL configurations exhibit themself as the most



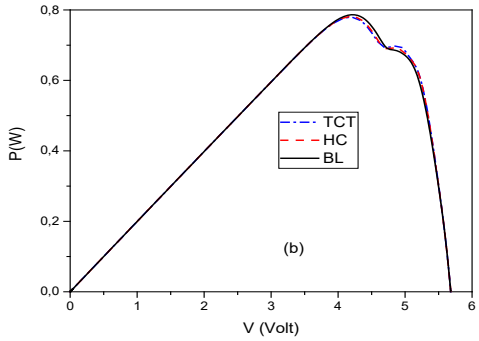
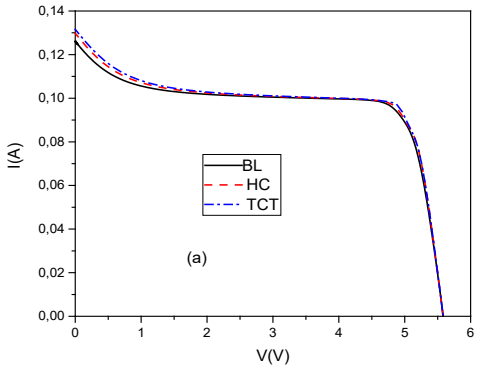


Figure 10: effect of shading one cell in a typical PV module (36cells) for three various configurations: (a) I-V curve and (b) P-V curve



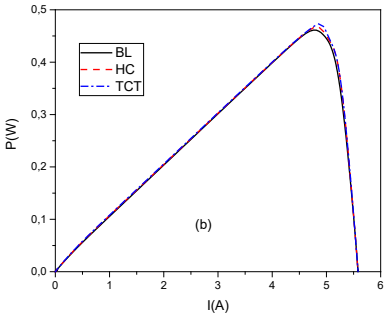


Figure 11: effect of shading nine cells in a typical PV module (36cells) for three various configurations: (a) I-V curve and (b) P-V curve

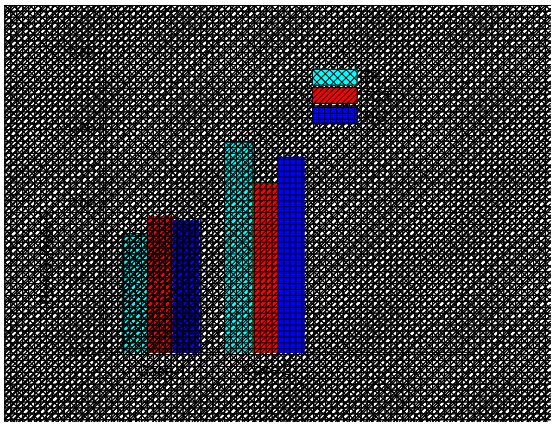


Figure 12: effect of the number of shaded cells on power losses

5. Conclusion

In this paper, an analysis parameters degradation in a typical PV cell, which lead to power loss under dark as well as light condition using measured I-V data was made. In order to analysis the sensitivity of the dark I-V measurement to each of the six extracted parameters as a function of the voltage as well as the temperature and the density current, we simulated the operation of a silicon solar cell (KXB0022-12X1F). The analysis of the dark I-V curves permit us to detect variation as small as 15% in the series resistance. We have also extended the use of dark as well as light I-V measurements to modules configurations of cells and uses a nonlinear least squares method to evaluate the cell efficiency parameters in the modules. Results obtained have shown a degradation of the values of the maximum power (Pmax) as compared to initial values by about 12, 3%, 12, 06% and 10, 21 % respectively in Total-Cross-Tied (TCT), Bridge-Link (BL) and Honey-Comb (HC) configurations. The decrease of Pmax could be imputed to the decrease of the short circuit current Isc. The reduction in Isc is explained by the rate of shading cells which limit the light transmittance.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors wish to express their gratitude to the Laboratory of Materials and Environment (LA.M.E), Burkina Faso and the Group of electrical engineering, Paris (GeePs) who permit to realize the experimental setup.

References

[1] S. K. Firth; K.J. Omas, S.J. Rees," A simple model of PV system performance and its use in fault detection " Solar Energy, **84** (4), 2009, 624-635, doi:10.1016/j.solener.2009.08.00

[2] M.A. Munoz, M.C Alonso-Gracia, N. Vela, F. Chenlo.," Early degradation of silicon PV modules and guaranty conditions " Solar Energy, **85**(9), 2011, 2264- 2274 doi:10.1016/j.solener.2011.06.011

[3] J.I Van Molken, U.A. Yusufoglu, A. Safiei, H. Windgassen, R. Khandelwal, T. M. Pletzer, H. Kurz," Impact of micro-cracks on the degradation of solar cell performance based on two-two model parameters" SiliconPV: April 03-05, Leuven, Belgium, Energy Procedia 27, 2012, 167- 172 doi:10.1016/j.egypro.2012.06.046

[4] E.L. Meyer and E.E. Van Dyk,"Assessing the reliability and degradation of photovoltaic module performance parameters", **53**(1), 83-93, 2004, doi:10.1109/TR.2004.824831.

[5] P. Hacke, R. Smith, K. Terwiliger, S. Glick, D. Jordan, S. Johnson, M. Kempe and S. Kurtz," Testing and analysis for lifetime prediction of crystalline silicon PV module undergoing degradation by system voltage stress", 38 th IEEE Photvoltaic specialists conference, Austin, TX, 2012, 001750-001755 doi:10.1109/PVSC.2012.6317933.

[6] D. Sera, S. Spataru, L. Mathe, T. Kerekes, R. Teodorescu," Sensorless PV array Diagnostic method for residential PV system" 26 th European photovoltaic Solar Energy conference and exhibition, 2011, 3776- 3782, doi:10.4229/26thEUPVSC.2011-4AV.3.37

[7] A. Drews, A. C. de Keizer, H.G. Beyer, E. Lorenz, J. Betcke, W. G. J.H. van Sark, W. Heydenreich, E. Wiemken, S. Stettler, P. Toggweiler, S. Bofinger, M. Schneider, G. Heilscher, D. Heinemann," Monitoring and remote failure detection of grid connected PV systems based on satellite observations " Solar Energy, **81**(9), 2007, 548- 564, doi:10.1016/j.solener.2011.06.011

[8] S. V. Spataru, D. Sera, P. Hacker, T. Kerekes, R. Teodorescu " Fault identification in crystalline silicon PV modules by complementary analysis of light and dark current-voltage characteristics " Progress in photovoltaics: research and applications., 2015, doi:10.1002/pip.2571.

[9] N.C Park, J.S. Jeong, B.J. Kang, D.H. Kim, "The effect of encapsulant discoloration and delimination on the electrical characteristics of photvoltaics module," Microelectronics reliability **53**(2013), , 1818-1822, doi:10.1016/j.microrel.2013.07.062.

[10] P. Hacke, K. terwilliger, R. smith, S. Glick, J. Pankow, M. Kempe, S. Kurtz, I. bennett, M. Kloos," System voltage potential-induced degradation mechanisms in PV modules an methods for test" 37 th IEEE Photovoltaic Specialists Conference, Seatle WA, 2011, 000814-000820,, doi:10.1109/PVSC.2011.6186079.

[11] D. Bonkoungou, T. Guingane, E. Korsaga, S. Tassembedo, K. Zacharie, A. Darga, F. Zougmore" Measurement and analysis of the I-V-T characteristics of a photovoltaic cell: KX0B22-12X1F", 8 International conference on smart Grid (icSmartGrid), Paris, France, 2020, 157-162 doi:10.1109/icSmartGrid49881.2020.9144904

[12] D.L.King, B. R. Hansen, J.A. Kratochvil, and M. A. Quintana, "Dark current-voltage measurement on photovoltaic modules as a diagnostic or manufacturing tool". conference record of the twenty sixth IEEE Photvoltaic specialists conference, Anaheim, CA, USA, 1997, 1125-1128 doi:10.1109/PVSC.1997.654286.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1157-1163 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Operational Responsibility and Task Monitoring Method: A Data Breach Case Study

Saliha Assoul[1], Anass Rabii[*, 2], Ounsa Roudiès[2]

[1]*ENSMR, Siweb, E3S, Mohammed V University, Rabat, 10010, Morocco*

[2]*EMI, Siweb, E3S, Mohammed V University, Rabat, 10010, Morocco*

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 09 February, 2021
Online: 25 February, 2021

Keywords:
Information Security
Security Breach
Supervision
Monitoring Method
SysML
Capital One

ABSTRACT

As a result of digitalization, services become highly dependent on information systems thus increasing the criticality of security management. However, with system complexity and the involvement of more human resources, it becomes more arduous to monitor and track tasks and responsibilities. This creates a lack of visibility hindering decision making. To support operational monitoring, we propose a method composed of i) a core of security concepts from International Standard Organization (ISO) standards ii) a graphical modeling language iii) a guiding process and iv) a tool that provides verification through formal Object Constraint Language (OCL) queries. Applying this method to the case of the Capital One data breach showcases incident prevention through task supervision. The resulting work product is a formal comprehensive map of assets, actors, tasks and responsibilities. The SysML formalism allows different actors to extract information from the map using OCL queries. This allows for regular task and responsibility verification thus closing any window of attack possible.

1. Introduction

With the advent of digitalization era, the widespread dependence on information systems (IS) rose as well. This new age was kick started by technological developments especially in IoT, Big Data and Cloud, paving the way for newer services and spreading IS use to more diverse fields (Industry-science research alliance). However, this dependence translated into an increase in criticality levels of assets and tasks. Fields such as health, energy, or e-Government, which were already dealing with critical information or physical assets, are facing greater challenges in security management. As a response, practices of security management had been thoroughly addressed by corporations, standardizing organizations and academia. The quest for security approaches and methods to apply systematically to specific situations fueled the creation of best practices, standards and maturity models. The popular ISO 27000 standard series stood out for the creation and management of Information Security Management Systems (ISMS) (ISO 27000). ISO provides baseline requirements in ISO 27001, best practices in ISO 27002, security risk management in ISO 27005 or even specialized recommendation for network security in ISO 27033. Additionally, a plethora of security maturity models were created between 2007 and 2018 to support continuous improvement of security in diverse fields. However, our systematic review [1] of the literature showed that even if many standards and approaches are available, implementation remains problematic. The lack of implementation feedback further makes the study of a real case such as the Capital One Financial Corporation breach [2] more appealing to discuss. We postulate that this issue is due to two major aspects: system complexity and human behavior.

Dealing with increased complexity created between interacting systems, we find it necessary to adhere to a high level of formalism to ensure rigor and precision to tackle one of the four dimensions of human factors in information security: responsibility [3]. Taking root in the system engineering paradigm, our solution is based on the standard system formalization language SysML. We propose a scalable formal representation of security responsibilities in accordance with complex system composition that supports traceability and management. This formalism provided by extended SysML profiles, enables the use of formal verification languages such as OCL to support supervision. We also provide a guiding procedure based on continuous improvement for iterative applications as well as a supporting tool. This method would provide security managers with proper knowledge to track fundamental causes, such as asset vulnerabilities, unaccomplished tasks and accountabilities. In our previous studies, we described

[*]Corresponding Author: Anass Rabii, anassrabii@gmail.com

our generic framework dedicated to information security management, extracted security concepts for SysML profiles [4] as well their application in security requirement management [5]. In this paper, we examine the responsibility issue and illustrate it using a case study.

We present and discuss the case of a data breach that occurred in Capital One in order to analyze the consequences of an informal specification of responsibilities and therefore prove the relevance of the problem we tackle. Next, we apply our method to the case to highlight its advantages. Through the Capital One case study we will address the following research questions:

- RQ1: How did Capital One track the chain of responsibility for their data breach?
- RQ2: How did Capital One security managers verify the implementation of security controls?
- RQ3: How to model tasks and responsibilities in a multi-scale information system?
- RQ4: How to formalize operational monitoring in a multi-scale information system?

In the following sections, we detail the context of our study highlighting the problem that our method aims to solve and how it fits within the existing body of work. Next, we present the components of our proposed method. Lastly, we present the Capital One Case followed by the application of our solution to highlight its applicability and its ability to prevent similar cases.

2. Related Works & Problematic

Information system complexity introduced heterogeneous components with a high connectivity level. In the complex system paradigm [6,7], systems can be decomposed into smaller autonomous collaborating sub-systems. This entails that security breaches and by consequence security management become pervasive. This means that measures taken to address security matters must be propagated to other collaborating systems. Similarly, this complexity makes responsibility harder to pinpoint and necessary tasks less apparent. On the other hand, the transition from theoretical and generic standard recommendations or requirements to practical application by human resources reveals a substantial gap. In 1996, including human behavior was considered as a novel conceptual stance within security management [8]. Since then, both academia and standardizing organizations have provided a plethora of references for security management that take human behavior as an important variable.

On one hand, standardizing organizations such as NIST and ISO provide thorough and well recognized references like the NIST Cybersecurity Framework, ISO 27001 and ISO 21827. These standards have either stood the test of time or have been periodically updated. We consider these standards as an important body of knowledge and procedures for our studies. We position ourselves as an extension of these standards providing the method to guide implementation that is out of their scope. On the other hand, academia has also produced countless information security maturity models such as SOASMM for SOA architecture [9], MMISS-SME [10] for small and medium enterprises, CCSMM for American governmental entities [11] or ISMM-PCI [12] for the payment card industry. Through our systematic literature review [1], we became aware of the lack of implementation and validation results for academic security maturity models. We also perceived that the academic shift towards specialization is also related to the implementation issue, hence why our solution intervenes at the implementation phase.

This implementation issue was due to the impact of human behavior on information security as asserted by both academia and standardizing organizations. According to research [13], human resources are considered the weakest link in securing information systems. In studying the impact of habits in following security policy [14], the authors highlight the effect of individual beliefs, thoughts, actions, attitudes, awareness, and training among others on policy compliance. ISO 27001 and 21827 standards also emphasize the importance of awareness building and skill development for all interested parties. In addition, as system complexity also entails more stakeholders, responsibilities become pervasive and contain multiple levels of accountability. In this context, the problem, firstly, is that actors have difficulties proficiently managing their responsibilities within their own scope of action. While the standard recommendations are clear, organizations have difficulties down-scaling them into actionable tasks and responsibilities [15]. Secondly, managers don't have a multi-scale visibility over the system to insure monitoring adapted to scalability. In fact, we currently rely on individual manager capability for tracking tasks using natural language management tools and renowned methods such as responsibility assignment matrices [16].

As the Design Science [17] method entails, the next step is the production of artifacts to put this knowledge to use. Our method aims to bridge this gap by addressing operational monitoring for security management with a scalable vision, focusing on human behavior. In their study of non-malicious security violations, the authors [18] clear up that while employees are goal centered, job performance is the end goal rather than security. Sharing this view, we perceive that providing managers with the adequate tool to operate is the way forward. Also, studies on the social impact on information security address the influence of external systems in their culture models [19] demonstrating the need for the multi-scale visibility our solution provides. Finally, formalizing human involvement in applying security processes as well as supporting diversity of security approaches through genericity are essential. In order to achieve this, we rely on the prevalent system modeling language SysML. As a matter of fact, several studies use SysML as their basis for security risk assessment [20], secure system design [21] for model transformations [22]. We share the same system engineering vision for security management as these studies to address a different aspect of security that is operational management. Other modeling languages such as KAOS [23] offer the possibility for responsibility modeling but lacks the formalism necessary for supervision. Whereas, our solution enables a formal and complete representation of security related information. As a result, it will help test and verification using OCL, allowing for day to day monitoring, the cornerstone of Agile's [24] success in software engineering.

3. The operational responsibility and task monitoring method

As information systems' complexity blurs the lines of responsibility, we provide a generic method for responsibility, task

and vulnerability traceability. By definition, a method is composed of a set of concepts, a language, a procedure and a tool to support it all. In the following sub-sections, we will detail each component as well as their provenance.

3.1. Core concept

In order to populate our modeling language, we need to introduce information security concepts. We can rely on existing references such as standards or prominent security maturity models in order to extract these concepts. In fact, we analyzed and compared the main security concepts used in the security maturity models we found through our systematic literature review [1]. We have found that these models are highly connected to the ISO 27001 and ISO 27002 standards [4]. Seeing that the ISO standards are aligned, we compared the security concepts used in ISO 21827 to prominent academic and governmental security maturity models. We concluded that the ISO standards are comprehensive sources for information security concepts. The concepts that we put to use in this monitoring method are listed in Table 1.

Table 1: Core of Information security concept list

Concept Name	Semantic Meaning	Example
System	Smart building, Energy grid, Network	Smart building, Energy grid, Network
Asset	Anything of value to the organization	Software, Building, Server, Information, Process
Human Resource	An actor implicated with the system	Client, Chief of security, System Owner
Task	A set of actions assigned to a human resource that must be executed	Log data analysis, Configuration update
Responsibility	The obligation to oversee, take care of a specific asset	Security responsibility, Maintenance responsibility, Access responsibility
Vulnerability	The state of exposure to the possibility of being attacked, damaged or tampered with	ragility, Open port, Plain text storage, Weak password

3.2. Language

In order to provide a graphical modeling language tailored to our concepts, we chose to extend SysML, the standard system modeling language. It supports specification, analysis, modeling, verification and validation of complex systems. It is defined as an extension of the well-known UML language used in software engineering. Another upside for choosing UML is the Object Constraint Language (OCL). OCL is a formal textual language that allows constraint and object query expressions. It is a key point to the testing needed for monitoring. In addition, we can extend the SysML basic concepts by adding our security core concepts by the mean of profiles. This leads to define a new block diagram that includes security concepts. It allows us to define newer associations between these concepts for the necessary interactions. This ensures that we profit fully from the richness of the predefined SysML block diagram concepts. Our profile consists of the security concepts in Table 1 coupled with five new associations added to define new relationships between them. In Table 2 we specify the composition of our profile as well as the properties required for traceability.

Table 2: Security Responsibility Profile Elements

Concept Name	Relationship	Properties
System	Systems contain one or many Assets. Systems can interact with one another.	Owner Purpose
Interact With	Association linking two systems	Influence StartDate EndDate
Asset	Assets collaborate with one another. Assets Belong to one or many systems. Assets have one or many responsibles. Assets have one or many vulnerabilities.	ID Description IsActive LastActiveDate
Belong To	Association between assets and systems	Criticality level
Collaborate With	Association between assets	Nature Description CollabDuration CollabStartDate
Vulnerability	A Vulnerability belongs to one or many assets	Existence likelihood Description Recommendation
Human Resource	A human resource can be responsible for one or many assets. A human resource is a specialization of an asset. A human resource can execute one or many tasks.	Organizational unit
Responsible for	Association linking a human resource to one or many assets.	RespDuration RespStart LastResp RespNature

Execute	Association linking a task to a human resource.	StartDate EndDate
Task	A task can be executed by a human resource. An influence from an external system can impact a task.	ID Description IsExecuted Deadline Nature Criticality level

3.3. Procedure

The third part of our operational monitoring method is a procedure to guide step by step the usage of the modeling language. Drawing from the existing base of knowledge, continuous improvement is a central aspect of the ISO security standards. This translates into a cyclical security management procedure [5]. Using a PDCA cycle [25], operational monitoring would take place during the "Do" and "Check" phases. However, with the facilities enabled by the formal language, we can have a higher frequency of verification.

This is reflected in more frequent cycles of verification and testing. In Figure 1, we zoom into the "Do" and "Check" phase of the overall security management in order to highlight the usage of our method in operational monitoring.

First Step: The system definition resembles that of a general security management approach. It allows us to determine the scope of the system we plan to assess and aim to protect. This also clarifies the borders of the system properly defining the flow of interactions with other systems. The resulting work product is a mapped out system through a security block diagram. This diagram would present assets including their characteristics and collaborations, human resources involved and systems with which the system in question interacts.

Second Step: The organizations must devise an action plan that aligns with their internal goals and security requirements. That action plan is then adapted depending on existing resources and human resources capabilities.

Third Step: The action plan is down-scaled into actionable assignments. These duties and responsibilities are then allocated to the different actors within the system bounds with specific deadlines. This is where, traditionally, managers would use classic management tools defining tasks using User Stories that can be translated into model elements: "As a *Responsibility* I must *Task* by *Deadline*". This information is used to complete the previously mentioned security diagram.

Fourth Step: Human resources execute the tasks allocated to them. Task execution duration is task dependent and supervisions needs to be adapted accordingly.

Fifth Step: Supervisors and managers can design OCL tests to monitor task execution or do bulk monitoring. These tests can immediately yield undone tasks, return human resource work load or other tasks where manual verification would be arduous.

Sixth Step: These managers can determine, through multiple pre-designed tests, the current operational posture for their responsibilities within the system. The cycle then restarts, devising a new action plan based on previous results and future goals.

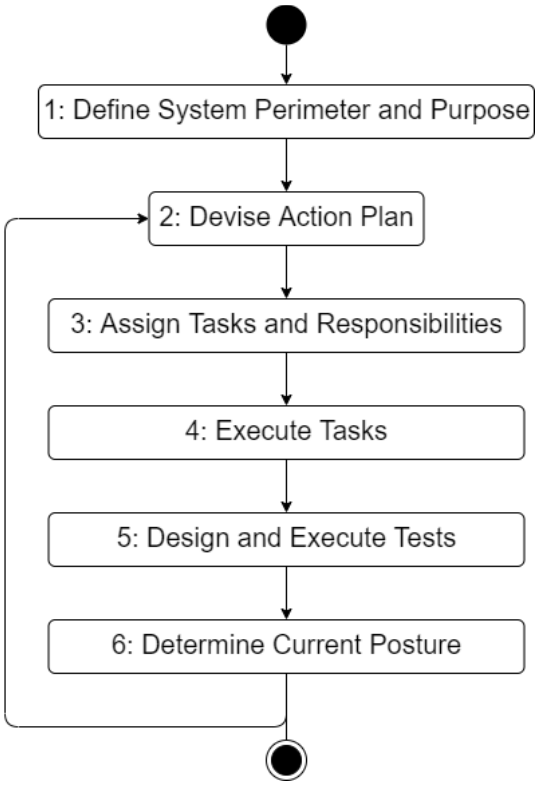


Figure 1: Operational monitoring procedure steps

3.4. Tool

In order to produce a functioning graphic editor for our language to support our method, we set out to find an existing well renowned modeling tool to extend. In Table 3, we set out to find a free and open source modeling solution.

Table 3: Renowned Modeling Tool Comparison

Modeling Tool	Author	OpenSource	Free
Modelio	Modelio Corp.	Yes	No
Entreprise Architect	Sparx Systems	No	No
MagicDraw	NoMagic	No	No
Eclipse Papyrus	Eclipse Foundation	Yes	Yes
Eclipse Sirius	Eclipse Foundation	Yes	Yes

Both Sirius and Papyrus are adequate choices, the scale tipped in favor of Papyrus for the availability of support and documentation. After this selection, the profile for the security SysML extension is created in a separate project. It is meant to act as a new extension point to be added to the profile list. This gives us, through the option of palette configuration, the possibility to add the profile elements. We can also define their graphical syntax

i.e. new shapes if needed, we elected to leave the original block shape with their respective stereotype. The resulting palette is shown in Figure 2. As it is for tests through OCL, Eclipse Papyrus supports the addition and execution of any OCL queries written in all Papyrus projects. By right clicking within the project, Eclipse allows the addition of OCL files, their execution and validation through the drop-down menu in Figure 3.

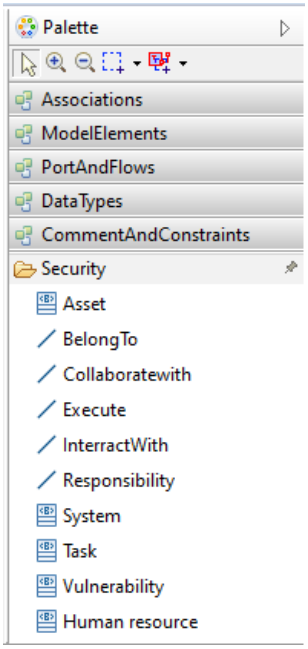


Figure 2: Security Responsibility Palette

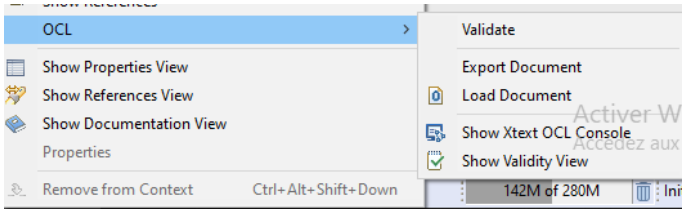


Figure 3: Dropdown Menu for OCL

4. Capital One Breach Case Study

In this section, we follow the operational responsibility and task monitoring method in analyzing the Capital One data breach. At first, we analyze the impact of Capital One's monitoring practices. Secondly, we use the data breach as a case study to apply the steps of this method. Finally, we highlight the lessons learned from this case study.

4.1. Methodological Consideration

Because of the rarity of details concerning information security management, analyzing the detailed study [2] would yield interesting insights. Their study aligns with our vision since their initial hypothesis: Renowned standards do not provide enough guidance for organizations for implementation nor incident management. Also, organizations aren't capable of implementing and maintaining security controls. One of the recommendations made was to find a way to manage the time frame between control implementation, evaluation and audit. This is precisely where our study intervenes analyzing task and responsibility operational monitoring. As per the typology provided in [26] and [27], our units of analysis in this "Descriptive Case Study" are the tasks or responsibilities of a single human resource. Our conceptual framework is that of our proposed solution, detailed in section 3. 1., as we will be studying the different human resources and assets intervening in Capital One's systems as well as the different responsibilities and tasks. We will address them through the research questions previously presented in the introduction:

- RQ1: How did Capital One track the chain of responsibility in their data breach?
- RQ2: How did Capital One security managers verify the implementation of security controls?
- RQ3: How to model tasks and responsibilities in a multi-scale information system?
- RQ4: How to formalize operational monitoring in a multi-scale information system?

4.2. Procedure

Capital One is ranked eighth largest bank overall with a revenue of around 28 billion dollars in 2018. Technology implementation and consistent progress are considered driving ideologies within Capital One. As it is the case in banking, they abide by several security standards such as the New York Stock Exchange corporate governance rules, as well as being one of the participants in supplementing the NIST security standards. Capital One are considered pioneers in migrating their data centers to a cloud environment, environment from which the breach stemmed later. In spite of all their efforts, a breach disclosing the personal data of around 106 million individuals was discovered in July the 19th, 2019. According to the investigations that followed, an employee for the cloud storage service provider Amazon Web Service (AWS) created a scanning tool that allowed the culprit to recognize servers with firewall misconfigurations allowing access to buckets of data. This permitted the later execution of a set of scripts to retrieve access credentials then copy the now available data.

4.3. Analysis

Following the steps of our procedure described in section 3.3, we address primarily the failure in the implementation of technical security controls from Capital One's side. In the ***First step***, we describe the contents of the Capital One Security Management System. They have invested in hiring a renowned Chief Information Security Officer (CISO), talented security engineers as well as on other security related investments for tool development.

Secondly, the bank's action plan allegedly revolves around implementing information security standards such as the NYSE requirement or NIST security controls. This meant that the CISO had the duty of choosing the security controls he deems adequate to protect his system. He then had to specify the time and resources allocated for each security control as well as defining the evaluation and audit periodicity. This is also when he's supposed to decide the means by which the controls are to be satisfied. First of all, a Web Application Firewall (WAF) needs to be implemented and configured to block any entry from malicious

proxy servers or from TOR exit nodes. Secondly, set up proper periodic vulnerability scans for the WAF. Lastly, revoke administrative account access credentials in case of internal changes. In the ***third step***, the CISO should assign these duties to the security engineer and the security manager. In the case as we understand it, there was time window in which the attack occurred and that is between task execution in ***step four*** and the monitoring test that would have identified the failed NIST controls in ***step five***.

Research Question 1: In an article by The Wall Street Journal, Capital One attributes the problem to an error in its own infrastructure. However, in a second article in the same journal, interviewing Capital One employees reported that multiple issues are at the root of this breach. Prior to the incident, there have been complaints that employees have raised concerns that they required more software to spot breaches. This delay is caused by the difference in skillset between IT professionals and governance professionals. As systems become more interconnected, all involved parties must acquire multidisciplinary skills so that governance professionals understand the necessary requirements in terms of technology to be able to provide them within a smaller timeframe. This could also be linked to failures to stay within budget lines despite hefty investments. This shows that Capital One failed to trace back the source of their infrastructural issues that is managerial and financial at its core.

Research Question 2: A second issue voiced by employees was low morale due to an increase in number of employee terminations. This is also due to budget issues that later became detrimental to routine security control implementation. Employees also reported an internal practice giving liberties to programmers the freedom to code in whatever language they choose if it would help complicate breaking into their systems in turn complicating pen-testing. These two incidents reflect that Capital One also lacked visibility as well as control over the tasks executed inside its scope. The Capital One security managers knew the necessary security controls to implement, however, they lacked regular verification of their implementation. This is what created the timeframe the culprit needed to execute his attack.

4.4. Lessons Learned

In the previous subsection, we clarified that Capital One had issues regarding task and responsibility tracking. Now, we will address these issues through discussing the second set of research questions.

Research Question 3: In our previous analysis, the first issue was the lack of multi-scale visibility coupled with lack of understanding between engineers in the security management system and their higher-ups in top management. This an aspect that our solution addresses via a complex system vision, where multiple systems interact with one another and therefore influencing one another. Using our method, in the first step, we can model Capital One as two systems "Governance System" and "Security Management System" each containing their respective human resources. As is shown in Figure 4, we can model the budget that the first system provides to the next or how the financial issues can influence security management. In this case, any employee dismissed must be removed from the model. Carrying on, we add to each human resource their set of tasks to which we assign deadlines and the attribute *Isexecuted* to be set to "True" when it's done. Figure 5 shows the different security tasks and their responsible. Each task is the implementation of a necessary NIST security control [2]. In the case of the layoffs and human resource removal, the tasks would remain and it would be apparent that they need to be reassigned in order to maintain security controls. Verification for human resource task load and tasks with no *responsible* association can be done through OCL tests. In addition, each human resource can find her own tasks by executing a script. This script would return all tasks that are assigned to the human resource with a *responsible* association.

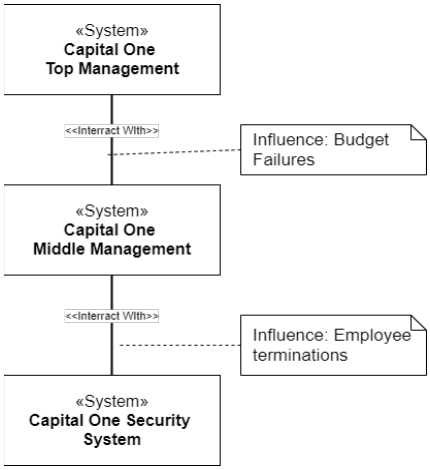


Figure 4: System Influence Diagram

Research Question 4: Following task modeling and responsibility assignment, it is the managers' duties to design tests for task supervision. This would normally be achieved by individually verifying each task accomplishment manually. Through the formal modeling of each task and responsibility, we now have the possibility of designing tests where we can list all tasks that have the attribute *IsExecuted* set to "False" and compare the current date to the deadlines. This would also link us immediately to the person responsible for this task execution. Since the day to day supervision tests don't have to be re-written, this makes monitoring less time consuming and therefore more efficient opening the possibility for day to day supervision. In the case of Capital One, this would have resulted in the revocation of the culprit's access credentials earlier as well as reassigning other employee's tasks in time. In turn, this would have closed any timeframe required to conduct the attack.

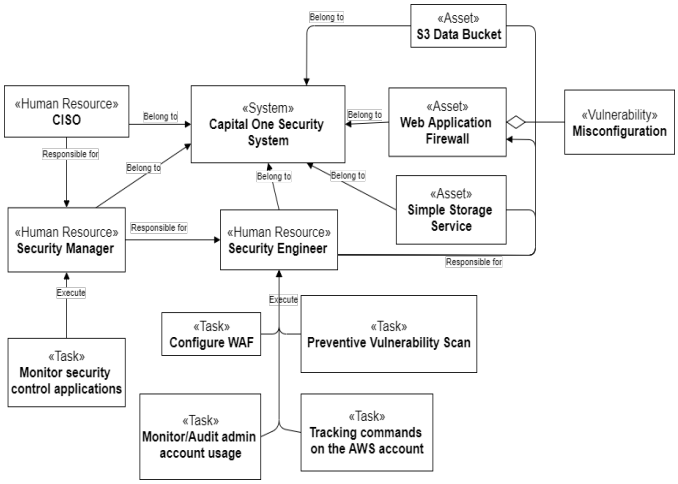


Figure 5: Operational Traceability diagram for Capital One Case

5. Conclusion

This paper presents a method for operational security monitoring applied to a security breach case. It aims to bridge the gap identified between security control identification and implementation. We have presented the components of our method that provide a much-needed formalism in security management. We have built this method using prominent modeling and security management standards in the base of knowledge. We have chosen to extend the SysML modeling standard with security monitoring concepts. This allows the creation of specialized block diagrams that model the actors and assets within a system as well as the tasks and responsibilities. The resulting diagram is a map capable of capturing complex system composition and interaction. Furthermore, this formalism provides the ability to use OCL for proper monitoring and supervision even for complex systems. In fact, the information depicted is of different use to different actors. OCL queries allow managers to regularly verify task execution and manage task loads. They also allow security engineers to see which new tasks were assigned to them or whether undone tasks still remain. Moreover, we have highlighted through our procedure the issues usually faced during the "Do" and "Check" phases of the Deming wheel. This allowed us to showcase the importance and applicability of our method for security management onto a high profile case. The described case study provided a high level of detail that allowed us to properly study the applicability and utility of our proposed method. Our method does not circumvent the complexity of the systems studied, it retains all the necessary information for operational monitoring. It aims to provide the tools that different actors need to carry out their roles in security management.

For future studies, we intend to include this study within a bigger whole designed for security management to reach "security by design". We also plan to further improve the modeling tool facilitating tests and creating a plugin to be featured in Eclipse Marketplace for public usage.

References

[1] A. Rabii, S. Assoul, K. Ouazzani Touhami, O. Roudies, "Information and cyber security maturity models: a systematic literature review," Information and Computer Security, **28**(4), 627–644, 2020, doi:10.1108/ICS-03-2019-0039.

[2] N. Novaes Neto, S.E. Madnick, A. Moraes G. de Paula, N. Malara Borges, "A Case Study of the Capital One Data Breach," SSRN Electronic Journal, (January), 0–24, 2020, doi:10.2139/ssrn.3542567.

[3] A. Alhogail, "Design and validation of information security culture framework," Computers in Human Behavior, **49**, 567–575, 2015, doi:10.1016/j.chb.2015.03.054.

[4] R. Anass, A. Saliha, R. Ounsa, "A Concept & Compliance Study of Security Maturity Models with ISO 21827," in Proceedings of the 22nd International Conference on Enterprise Information Systems, SCITEPRESS - Science and Technology Publications: 385–392, 2020, doi:10.5220/0009569703850392.

[5] A. Rabii, S. Assoul, O. Roudies, "Security requirements elicitation: A smart health case," Proceedings of the World Conference on Smart Trends in Systems, Security and Sustainability, WS4 2020, 776–781, 2020, doi:10.1109/WorldS450073.2020.9210330.

[6] F. Lahboube, O. Roudies, N. Souissi, "Extending supervision metamodel to complex system context," in 2015 Third World Conference on Complex Systems (WCCS), IEEE: 1–7, 2015, doi:10.1109/ICoCS.2015.7483242.

[7] Y. Bar-Yam, "General Features of Complex Systems," Knowledge Management, Organisational Intelligence and Learning and Complexity, **I**(1), 1–10, 1997.

[8] J. Backhouse, G. Dhillon, "Structures of responsibility and security of information systems," European Journal of Information Systems, **5**(1), 2–9, 1996, doi:10.1057/ejis.1996.7.

[9] M. Kassou, L. Kjiri, "SOASMM: A novel service oriented architecture Security Maturity Model," in 2012 International Conference on Multimedia Computing and Systems, IEEE: 912–918, 2012, doi:10.1109/ICMCS.2012.6320279.

[10] MMISS-SME Practical Development: Maturity Model for Information Systems Security Management in SMEs, in Proceedings of the 5th International Workshop on Security in Information Systems, SciTePress - Science and and Technology Publications: 233–244, 2007, doi:10.5220/0002430402330244.

[11] G.B. White, "The community cyber security maturity model," in 2011 IEEE International Conference on Technologies for Homeland Security (HST), IEEE: 173–178, 2011, doi:10.1109/THS.2011.6107866.

[12] S. Yulianto, C. Lim, B. Soewito, "Information security maturity model: A best practice driven approach to PCI DSS compliance," in 2016 IEEE Region 10 Symposium (TENSYMP), IEEE: 65–70, 2016, doi:10.1109/TENCONSpring.2016.7519379.

[13] K.D. Mitnick and William L. Simon, The Art of Deception: Controlling the Human Element of Security, John Wiley & Sons, Inc., 2003, doi:10.5555/861316.

[14] J.P. Jeretta Horn Nord,Alex Koohang,Kevin Floyd, "IMPACT OF HABITS ON INFORMATION SECURITY POLICY COMPLIANCE," Issues in Information Systems, **21**(3), 217–226, 2020, doi:https://doi.org/10.48009/3_iis_2020_217-226.

[15] A. Longras, T. Pereira, P. Carneiro, P. Pinto, "On the Track of ISO/IEC 27001:2013 Implementation Difficulties in Portuguese Organizations," in 2018 International Conference on Intelligent Systems (IS), IEEE: 886–890, 2018, doi:10.1109/IS.2018.8710558.

[16] K.H. Rose, "A Guide to the Project Management Body of Knowledge (PMBOK® Guide)-Fifth Edition," Project Management Journal, **44**(3), e1–e1, 2013, doi:10.1002/pmj.21345.

[17] R.B. Fuller, "A comprehensive anticipatory design science," **34**, 357–361, 1957.

[18] K.H. Guo, Y. Yuan, N.P. Archer, C.E. Connelly, "Understanding Nonmalicious Security Violations in the Workplace: A Composite Behavior Model," Journal of Management Information Systems, **28**(2), 203–236, 2011, doi:10.2753/MIS0742-1222280208.

[19] K. Huang, K. Pearlson, "For What Technology Can't Fix: Building a Model of Organizational Cybersecurity Culture," 2019, doi:10.24251/HICSS.2019.769.

[20] S. Ouchani, O.A. Mohamed, M. Debbabi, "A Security Risk Assessment Framework for SysML Activity Diagrams," in 2013 IEEE 7th International Conference on Software Security and Reliability, IEEE: 227–236, 2013, doi:10.1109/SERE.2013.11.

[21] L.A. Yves Roudier, "SysML-Sec - A Model Driven Approach for Designing Safe and Secure Systems," in Proceedings of the 3rd International Conference on Model-Driven Engineering and Software Development, SCITEPRESS - Science and and Technology Publications: 655–664, 2015, doi:10.5220/0005402006550664.

[22] F. Lugou, L.W. Li, L. Apvrille, R. Ameur-Boulifa, "SysML Models and Model Transformation for Security," in Proceedings of the 4th International Conference on Model-Driven Engineering and Software Development, SCITEPRESS - Science and and Technology Publications: 331–338, 2016, doi:10.5220/0005748703310338.

[23] A. Dardenne, A. van Lamsweerde, S. Fickas, "Goal-directed requirements acquisition," Science of Computer Programming, **20**(1–2), 3–50, 1993, doi:10.1016/0167-6423(93)90021-G.

[24] J.A. Highsmith, Adaptive Software Development: A Collaborative Approach to Managing Complex Systems, Dorset House Publishing Co., Inc., 2000, doi:10.5555/323922.

[25] W.E. Deming, Out of the Crisis, The MIT Press, 2018, doi:10.7551/mitpress/11457.001.0001.

[26] T. Hollweck, "Robert K. Yin. (2014). Case Study Research Design and Methods (5th ed.). Thousand Oaks, CA: Sage. 282 pages.," The Canadian Journal of Program Evaluation, 2016, doi:10.3138/cjpe.30.1.108.

[27] P. Baxter, S. Jack, "The Qualitative Report The Qualitative Report Qualitative Case Study Methodology: Study Design and Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers Implementation for Novice Researchers," Number 4 Article, **13**(4), 12–13, 2008.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1164-1170 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Performance Evaluation of a Gamified Physical Rehabilitation Balance Platform through System Usability and Intrinsic Motivation Metrics

Rosula Reyes[*,1], Justine Cris Borromeo[1,2], Derrick Sze[1]

[1]Ateneo de Manila University, Department of Electronics, Computer, and Communications Engineering, Quezon City, 917, Philippines

[2]Scuola Superiore Sant'Anna, Pisa PI, 56127, Italy

ARTICLE INFO

Article history:
Received: 09 December, 2020
Accepted: 16 February, 2021
Online: 25 February, 2021

Keywords:
Balance Board
Physical Rehabilitation
Stroke patients
Intrinsic Motivation
System Usability

ABSTRACT

Motivation significantly influences the outcome in the rehabilitation of patients. Several developments have been made to assess and increase patient motivation by addressing factors linked to motivation such as the personality of the patient, professional administering rehabilitation, and the rehabilitation environment. The main objective of the study is to evaluate the reliability of a gamified environment for the rehabilitation of stroke patients by testing its functionalities within standard physical therapy time and intervals. To achieve this, calibration was characterized. Also, user feedback was taken in the form of questionnaires based on the System usability scale (SUS) and Intrinsic Motivation Inventory (IMI). Based on the SUS scale, results show that the game manipulability is good, the game concept and design is satisfactory, and the game comprehensibility is also good based on the qualitative conclusion per SUS score. For the IMI ratings, it was found out that the highest rating was the perceived choice which indicates their voluntary participation in the game. Some improvements can still be added to the game itself to increase the motivation of patients. The balance board manipulability and the recalibration time interval can be further improved for comfort and ease of use by the patients.

1. Introduction

Balance deficiency is one of the common issues for post-stroke and post-injury patients, as well as for individuals who wish to train his/her balance even without deficits (e.g. athletes). Out-patients who had such deficiency will undergo a balance therapy program. In the current setting, a patient will attend therapy sessions about six to ten times. Each session, the patient will perform various balance exercises such as static balance, dynamic balance, and manual perturbation exercises using balance board, wobble board, balance rails, and mirrors extensively assisted by physical therapists. The patient will then perform the exercise, but most of the time, the patient will feel the pain of such exercises as well as boredom due to its repetitive nature. Also, the therapist will record the patient's balance quality according to what is observed, making several corrections to the number of balance sessions as well as the physician's initial observation. This also introduces an inconsistency problem when a different therapist observes the patient and records the score for interpretation. Finally, the patient has to rely on his/her safety on the therapist, since the current tools do not have robust feedback for the patient to observe his/her balance performance. With these, the patient is usually discouraged and unmotivated to attend another therapy session as well as not trying to perform them by himself/herself.

In line with this, one of the determinants that greatly affect the success rate of undergoing physical rehabilitation is the motivation level of the patient. Motivation has been identified to be influenced by a variety of factors such as the personality of the patient, the professional administering rehabilitation, and the rehabilitation environment. Research work integrating the use of the Microsoft Kinector the Wii Fit Board with various games to change the rehabilitation environment has proven to be effective in increasing patient motivation and overall affecting the outcomes on a more positive result [1, 2]. However, improvements on these previous researches can still be made, such as creating a more affordable and open-sourced design of the hardware and incorporating more game elements that engage the patient which will aid in the success rate of the physical rehabilitation.

*Corresponding Author: Rosula Reyes, rsjreyes@ateneo.edu

https://dx.doi.org/10.25046/aj0601131

Balance-on-action-Team (BOAT) implements affordable hardware and open-source software to enable a wider reach in the utilization of gamification in physical rehabilitation. It is an automated balance board medical device with a gamified element for conducting balance exercises for patients undergoing balance rehabilitation therapy. The overall goal is to provide rehabilitation centers, sports clinics, and personal trainers with safer, faster, and more engaging means of conducting balance training and therapy. In this way, balance exercises in the Philippines setting will become faster, more consistent from therapist-to-therapist, and more fun to perform.

The main objective of the study is to evaluate the reliability of the equipment (calibration and functionalities) based on the standard physical therapy time and intervals. Furthermore, conduct surveys to evaluate the game's impact on motivation.

This paper describes the various studies related to rehabilitation and patient motivation in Section 2 of the Review of Related Literature. Section 3 discusses the Methodology which focuses on the hardware and software implementation in Section 3.1 and the testing procedure in Section 3.2, using different metrics. The results are shown and discussed in Section 4 which includes the calibration in Section 4.1, Game Self Evaluation in Section 4.2, Motivation Analysis in Section 4.3, and Game Walkthrough Evaluation in Section 4.4. Lastly, the Conclusion and Recommendation can be found in Section 5.

2. Review of Related Literature

Several studies have been conducted on the rehabilitation of stroke patients using varying platforms, approaches, and frameworks. A similar balancing platform was studied by [3] where the focus is on utilizing the Nintendo Wii Balance Board for rehabilitation. The study developed a *WeHab* system that added visual biofeedback based on the center of pressure location. This system also allows for multiple balanced boards to be used together, customization of activities and difficulty level, and integration with a webcam to capture video footage during the sessions. Similarly, visual biofeedback was also used by [4] using Microsoft Kinect for postural rehabilitation. This allows the system to determine if the patient performed the correct postural exercises through the built-in capabilities of the Kinect in movement and gesture recognition. Another type of balancing platform that can be used is a wobble board which was studied by [5]. Here, the instrumented wobble board is used to create multidirectional perturbations and obtain vibrotactile feedback for training dynamic sitting balance. This research work also uses a microprocessor and an inertial measurement unit but has an added eight vibrating tractors. Aside from these approaches, the rehabilitation of patients can also be monitored using EMG on the lower limbs, such as in the work of [6]. In this case, a balance board was also used but accompanied with obtaining sEMGs signals to analyze how the muscle activity reacts from dynamic leaning caused by balance reactions.

Assessing patient motivation and system usability have been determined to help improve rehabilitation for chronic stroke patients. Instead of balance exercises used in this research, the study by [7] evaluates the feasibility of a new technology-supported task-oriented arm training regime (T-TOAT) for chronic stroke patients. The system is comprised of movement tracking sensors, an exercise board, and a software-based toolkit used for skills training. The patient motivation was assessed on the Health Care Self Determination Questionnaire (HCSDT) based on the self-determination theory. Furthermore, system usability was also assessed and was found to be rated good by the users. After performing various tests, it was discovered that the T-TOAT approach improved arm-hand performance significantly for the duration of the post-training. Similar to this work, system usability and motivation were evaluated through various questionnaires and assessment tools for improving rehabilitation.

Another patient motivation assessment tool applied in this work is the Intrinsic Motivation metrics which was also conducted in a study by [8]. This study aims to characterize motivation based on the information of activities for smart wearable health equipment applications. The participants in the study were tasked to solve a computer-generated puzzle under one of the four random conditions including the long-term feedback (LFB), long term graphical feedback (LGFB), short term feedback (SFB), and non-feedback for the trend of correct answers. Intrinsic motivation was then measured using both a 12-item version of the Intrinsic Motivation Inventory and the time it takes the participant to solve the puzzle. Based on these, it was observed that the most desirable type of feedback is one that improves perceived competence, which can then be adapted to health equipment displaying methods.

The BOAT is a research project that was conceptualized and developed through the collaboration between the Ateneo de Manila University and the Philippine Orthopedic Center. The gamified environment was initially developed by [9] using an inertial measurement unit. The goal was to provide the gamified elements in the form of quests and rewards system for providing an interactive and fun rehabilitation process. The hardware prototype was developed and tested in [10] using intervention techniques, motivation, and game design analysis. This was then continued by [11] through a comparative game design analysis. Several changes have been made for the hardware design, shown in Figure 1, and software game mechanics implementation by [12]. The research work added more customizability, and personalized information from the patients to help the physician perform better assessments. This research focuses on evaluating the gamified rehabilitation platform of [12] through various metrics such as system usability and intrinsic motivation metrics.

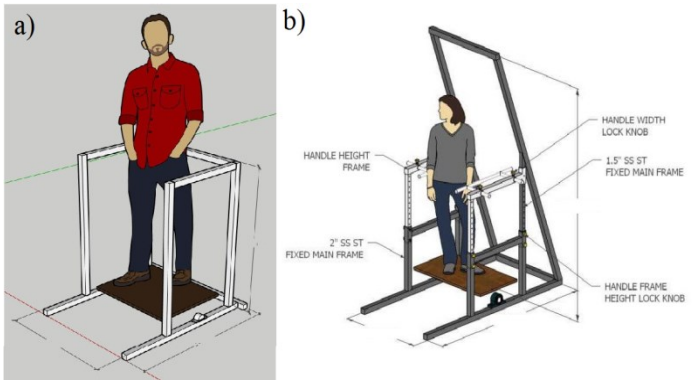


Figure 1: Hardware Changes from the Prototype [6] to the Commercialized [8]

3. Methodology

3.1. Design of the System

For the software, the gamified environment of the balance board lets the user experience a seamlessly integrated, open-world seafaring game integrated with a quest system [12]. The quest system allows the patient to perform physical rehabilitation exercises under the guise of the game. There are three quests or game modes designed to correspond with different physical rehabilitation exercises and specific metrics from the Tinetti Balance Assessment Tool and the Equilibrium Score. Each game mode and its corresponding mechanics are shown in Table 1.

Table 1: Game Modes and Mechanics [12]

Game Mode	Mechanics	Player Input
Collect the Crates	The player must be able to collect as much trash (in the form of crates) as possible within a certain amount of time	Lateral Dynamic Balance Exercise. The player must be able to challenge himself/herself a postural sway and recover from it as much as possible using lateral movements.
Avoid the Bombs	The player must be able to avoid all possible bombs within the boat's path in a certain amount of time	Static Balance Exercise. Players must be able to stand still as much as possible to avoid the bombs.
Follow the Light	The player must be able to follow the light in front of the boat without getting too near or too far from it	Posterior-Anterior Dynamic Balance Exercise. The player must be able to challenge himself/herself a postural sway and recover from it as much as possible using posterior-anterior movements.

The hardware design includes the following components: the frame, balance board, accelerometer, and Gizduino Uno. Most of the hardware components were mounted on the frame which features adjustable mechanisms and can be disassembled. The balance board is commonly used in physical rehabilitation for static and dynamic exercises, especially in the Philippine Orthopedic Center. There are two types of balance boards, the two-directional balance board which was used in this study, and the multi-dimensional balance board or wobble board. Relatively, the bidirectional balance board provides more accurate data from the accelerometer readings since only two directions are measured and the rest are disregarded. Several adjustable components are used to cater to different characteristics and categories of patients. One of these is the stopper which adjusts the maximum tilting angle of the balance board for both sides with increments of 5°. This allows the physician to change the maximum angle depending on the patient's capabilities. Furthermore, this helps in stabilizing the platform at 0° before running the application. Also, the handrails where the patient will hold on to when playing the game can be raised or lowered to compensate for patients with different heights. Each consecutive hole for moving the handrail is about 2 inches apart. Lastly, the friction of the balance board can be altered using the tension control knobs with a range of 1 to 8 for each side. Adjusting the magnitude of the sway of the board will help more balance-challenged patients gradually adapt to the game mechanics.

As a safety precaution, the patient is advised to wear a harness that is attached to an adjustable lanyard, presented in Figure 2. The lanyard is then mounted on the hooks of a metal bar which can be rotated 360° along the vertical axis to allow a smoother transition for the dynamic posterior-anterior to lateral sway and vice versa. Moreover, wearing the harness greatly minimizes the risk of falling from the balance board.



Figure 2: Harness for Lateral (a) and Posterior-Anterior (b) Movement

For the improvements tracking, the calibration of the equipment is accomplished through linear regression using two data points from the sensor readings. From the generated equation in slope-intercept form, the slope and y-intercept will be saved in a .txt file inside the Results folder on the desktop. This will enable the user to see any variations in calibration.

Similarly, the game metrics are recorded in the form of a .txt format in the Results folder. It publishes the patient's name, the date and time when the exercise is completed, the quest played, the type of balance exercise, their in-game score, the Equilibrium Score (ES), the maximum posterior-anterior sway, maximum lateral sway, and game settings information. This information allows the physician to track and evaluate the progress of the patients after engaging in the physical rehabilitation procedure.

Data from each game mode are recorded in the form of the ES, shown in (1), which indicates their quality-of-balance progress as patients play the game. The ES is the probability of fall in which a score of 100 indicates a 0% chance of fall while a score of 0 indicates a 100% chance of fall.

$$ES = \frac{\gamma-(\theta_{max}-\theta_{min})}{\gamma} X\ 100\% \quad (1)$$

where θ_{max} is the maximum postural sway relative to the vertical

θ_{min} is the minimum postural sway relative to the vertical (natural standing angle)

γ is the limits of stability angle for the clinical setting

3.2. Testing the System

To fully test the reliability of the system, several methodologies were adapted based on metrics related to the study, which includes the following:

An SUS (System Usability Scale) is a self-rated scoring system that measures the usability of a certain device or software. It is usually a ten-item test that has open-ended statements with user responses from "Strongly Disagree" to "Strongly Agree". Each item is scored on a 5 (or 7) point scale. To prevent purely positive or purely negative responses, this questionnaire is designed to have alternate positive-worded and negative-worded statements. Thus, the following rules for computing the SUS are the following:

- For every negative statement (odd-numbered items for this test): take the max score (e.g. 7) and subtract it by the raw score (e.g. raw score = 3, SUS score = 7 – 3 = 4)
- For every positive statement (even-numbered items for this test): take the raw score then subtract by 1 (e.g. raw score = 6, SUS score = 6 – 1 = 5)
- Add all of the SUS scores of every item. The maximum score should be:

$$(\text{max points} - 1) * \text{max items} \quad (2)$$

- Depending on the maximum number of items, the overall SUS score should be on a scale of 0 to 100. Therefore, for a 10-item questionnaire with 7 as maximum points, the entire sum should be divided by 0.6 since the maximum score is 60.
- The overall SUS score is then tallied and averaged. Normalization is considered for the average therefore a score of 68 and above is a favorable result, while a score of below 68 is an unfavorable result. Figure 3 shows how the scores will be interpreted qualitatively in the conclusion.

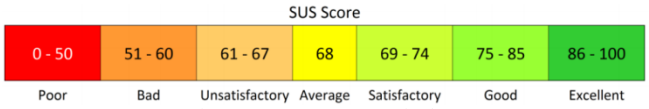


Figure 3: SUS Score Equivalent

The Intrinsic Motivational Inventory (IMI) measures the motivation of the test subjects in playing the game. All 22 questions are modified from the template provided by the Self-Determination Theory Organization but have the same measurement (http://www.selfdeterminationtheory.org). For determining the IMI Score of the test subjects, the following procedures are done.

- The responses for items 2, 9, 11, 14, 19, and 21 are reversed. That is, the item score is equal to 8 minus the response (item score = 8 – response).
- These subscale scores in Table 2 are added together:

Table 2: IMI Subscales

Interest/Enjoyment	Items 1, 5, 8, 10, 14, 17, 20
Perceived Competence	Items 4, 7, 12, 16, 22
Perceived Choice	Items 3, 11, 15, 19, 21
Pressure/Tension	Items 2, 6, 9, 13, 18

- The subscale scores are analyzed as follows:
  - Interest/Enjoyment, Perceived Competence, and Perceived Choice is summed up into the IMI score for the subject's motivation. This sum is high if the test subject is motivated enough to perform the balance exercises through the game
  - Pressure/Tension is the IMI score for the subject's loss of motivation. This score is high if the test subject is not motivated enough to do the balancing exercise even if the game is played.

The Motivation-to-No-Motivation (MNM) ratio is then determined to check how much the test subjects are motivated based on their responses on the IMI. MNM is achieved by applying the equation:

$$\text{MNM} = \frac{\text{IE+PCT+PCH}}{\text{IE+PCT+PCH+PT}} \quad (3)$$

where MNM = Motivation-to-no-Motivation ratio,

IE = IMI subscale score for Interest/Enjoyment,

PCT = IMI subscale score for Perceived Competence.

PCH = IMI subscale score for Perceived Choice, and

PT = IMI subscale for Pressure/Tension.

4. Results and Discussion

The reliability of the prototype was evaluated to assess the accuracy and precision of the components based on the calibration values before and after using the balance board. The test was also conducted with twenty (20) healthy test subjects consisting of 10 males and 10 females within an age range of 20-30 years old. Each subject has no history of stroke and no injury 6 months prior.

4.1. Calibration

Calibration evaluates and adjusts the precision and accuracy of the balance board. In calibrating, the balance board was referenced at +20 and -20 degrees, and sensor values of the MPU 6050 were then recorded on these angles to calculate the slope.

For the reliability of the balance board based on calibration, the slope was recorded before and after the game based on the following parameters: 10min (Game)-15min (Rest), 15min (Game)-20min (Rest), varying tension control, and varying weight of the test subjects.

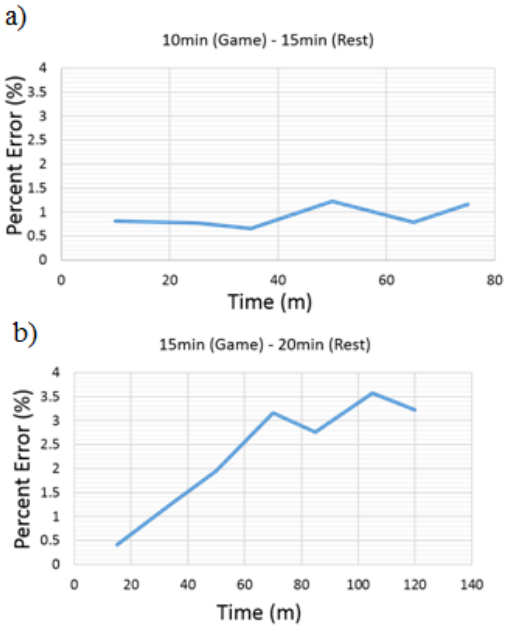


Figure 4: Slope Percentage Error vs. Time Graph: (a) 10min (game) – 15min (rest), (b) 15min (game) – 20min (rest)

Figure 4 shows the percentage error of the balance board slope as time increases. The percentage error was computed using this equation

$$\text{Error (\%)} = \frac{2\,|x-y|}{x+y} * 100 \quad (4)$$

where x is the initial slope while y is the value of the slope after some time. With *3.5%* as the maximum percentage error (4) at around *105 minutes*, it is safe to say the calibration of the prototype is still precise after turning it on for around *1 hour and 30 min.*

The slope of the balance board was also recorded as the tension control adjusts from 1(low friction) to 8(high friction). The correlation between the percentage error of the slope and the tension control is shown in Figure 5. Since the maximum percentage error is around *0.8%*, it shows that the calibration value is still precise even if the tension control is varied.

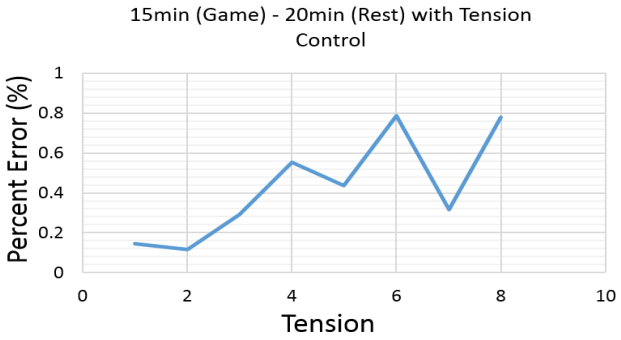


Figure 5: Slope Percentage Error vs Tension

Figure 6 shows the percentage error of the slope before and after doing the game testing as the weight (kg) of the test subject varies. Based on the figure, we can say that there is no correlation between the weight of the person doing the test and the slope percentage error. It also shows that the person's weight doesn't affect the calibration of the balance board with a maximum error of around *1.5%.*

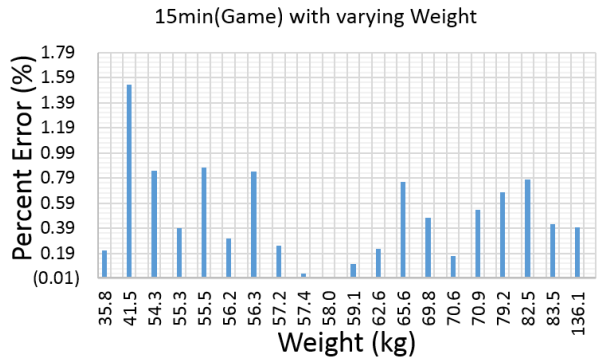


Figure 6: Slope Percentage Error vs Weight

4.2. Game Self Evaluation

Game Self Evaluation is an SUS questionnaire to be filled out by the test subjects for their evaluation of both the hardware and the game. It is divided into four groups: Balance Board Manipulability Measures, Game Manipulability Measures, Game Concept and Design Measures, and Game Comprehensibility Measures.

The Balance Board Manipulability Measures is a self-rated survey on how the balance board is used to them. This is collated and calculated based on SUS. For Balance Board Manipulability Measures, the average SUS for all 20 test subjects is 64.06, which is 5.89% lower than the average SUS score of 68. Overall, the result is unsatisfactory and the balance board manipulability measures can be improved.

The lowest SUS is 35.42 while the highest is 93.75. The highest-rated item is *#1* (I thought the board was too high for me to use) with an average SUS of 5 (0 being the lowest and 6 being the highest). This is because the handle on the sides of the balance board can be adjusted depending on the height of the test subject. The lowest rated item is *#8* (I felt it did not need any much muscle effort in using this platform) with an average SUS of 2.55. This is because the platform was tested for healthy patients while the level of difficulty of the game was intended for post-stroke and post-injury patients. Improvements can be made by adding a motor on the balance board to assist the test subjects when they feel tired or when they executed too much muscle effort.

The Game Manipulability Measures part of the evaluation is based on how playable the game is for the test subjects. The results are also collated and calculated on an SUS basis. The average SUS for all 20 test subjects on Game Manipulability Measures is around *75.94%,* which is *11.67%* higher than the average SUS score of *68*. Based on the results, we can say that game manipulability is good.

The lowest SUS is *47.92* while the highest is *100.* Item #8 (I thought the game was simple and uncomplicated) is the highest rated item with an average SUS rate of 5.65. The lowest rated item is #7 (My legs got tired very easily when playing the game) with an average SUS rate of 3.6.

The Game Concept and Design Measures is a self-rated survey on how enjoyable and how appropriate the game is for balance board rehabilitation. The results are collated and calculated on an SUS basis. The average SUS for all 20 test subjects on Game Concept and Design Measures is around *70.64%,* which is *3.88%* higher than the average SUS score of *68*. Based on the results, we can say that the game's concept and design are satisfactory.

The lowest SUS is *51.28* while the highest is *96.15.* Item #1 (I thought the game was too hard for me) is the highest rated item with an average SUS rate of 5.7. This proves that the game is easy to play and understand by the test subjects. The lowest rated item is #3 (I got bored at this game) with an average SUS rate of 2.75. This might be because the gameplay was a bit easy for healthy test subjects but it might be challenging for post-stroke and post-injury patients.

The Game Comprehensibility Measures is a self-rated survey on how easy and understandable the instructions of the game when playing, which is collated and calculated based on SUS. Based on the results of 20 test subjects, the average SUS is around *80.42,* which is *18.26%* higher than the average SUS of *68*. This shows that the game comprehensibility measures are on a good level based on the qualitative conclusion per SUS score.

The lowest SUS is *63.89* while the highest is *100.* Item #2 (I thought the amount of information to follow this game was enough) is the highest rated item with an average SUS rate of 5.39. The lowest rated item is #11 (The sound was too loud or too soft, or had an annoying burst of sound) with an average SUS rate of 4.17, which is still above the average possible score of 3.

4.3. Motivation Analysis

In this project, all twenty (20) test subjects' responses had been collated and analyzed based on the IMI subscales. Table 3 shows the IMI results while Figure 7 shows a graph of the IMI rate in each IMI subgroup.

Table 3: IMI Results

IMI Subscale	Average	IMI rate (%)	Highest Score	Lowest Score
Interest/Enjoyment	35.25	71.94 %	49	21
Perceived Competence	29.7	84.86 %	35	19
Perceived Choice	30.3	86.57 %	35	18
Pressure/Tension	12.5	35.71 %	22	5

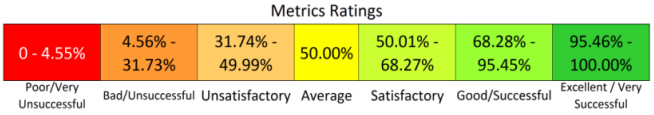


Figure 7: Qualitative Conclusion per Metrics Reading for MNM

In calculating their IMI scores, the average MNM is *0.88* with a minimum MNM of *0.78* and a maximum MNM of *0.95*. Results show that the average motivation of all 20 test subjects is in Good/Successful rating based on qualitative conclusion on metrics rating (Figure 7). This indicates that overall, the test subjects have good motivation in performing the balance exercises.

On all the IMI subscales as shown in Figure 8, the perceived choice is their highest response rate with *86.57%*, which indicates that they rated their voluntary participation and continued participation of playing the game. It is followed by perceived competence with *84.86%*. This means that the test subjects are skilled and they felt competent in playing the game. Interest/enjoyment which is another motivation subscale is also rated high with *71.94%*. The graph also shows that the test subjects didn't feel any pressure/tense and they feel relaxed while playing the game since the Pressure/Tension subscale only has a rate of *35.71%* and is below the average rate of *50%.*

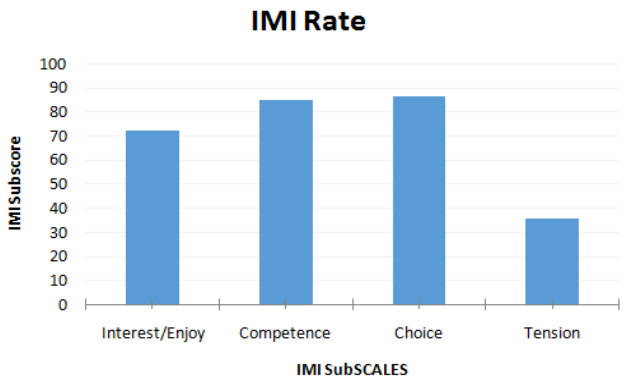


Figure 8: IMI Bar Graph

Game Walkthrough Evaluation is a self-rated survey consisting of short answers on what the test subjects thought of the game. Each item is analyzed and compared with other metrics to find out which part of the game is easy or hard, and what would be the difficulty in using the game.

4.4. Game Walkthrough Evaluation

Figure 9 shows the feedback of the test subjects while playing the game. The figure shows that most of the test subjects enjoyed while experimenting. They considered "Collect the crates" as their favorite part of the game while Follow the light was the most disliked one. Also, "Collect the crates" was the easiest part of the game while "Avoid the bombs" and "Follow the lights" tied to be the hardest part of the game. Additionally, the participants considered "Follow the lights" as the most tiring game, and difficult to control.

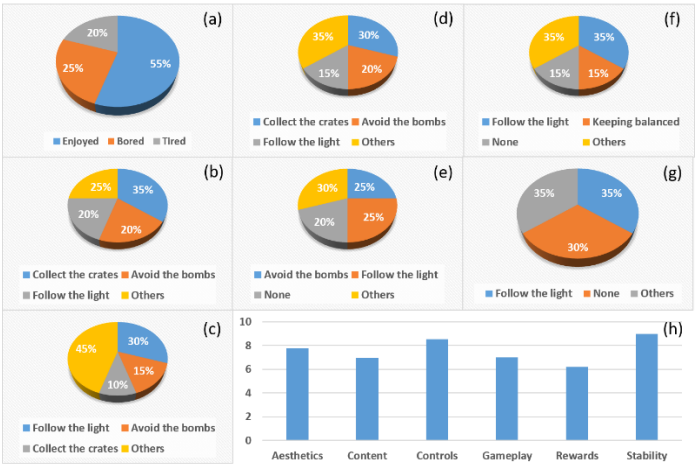


Figure 9: Game Walkthrough Evaluation Results: (a) Feeling after playing the game, (b) Favorite part of the game, (c) Disliked part of the game, (d) Easiest part of the game, (e) the hardest part of the game, (f) Tiring part of the game, (g) Control difficulty in the game, (h) Game overall Score.

The game is scored 7.58 over 10 (shown in Figure 9 (h)), with stability being the strongest feature (9) and rewards being its weakest feature (6.2). The participants also noticed some bugs/errors while playing like delayed movement on the projector screen, squeaking noise during the game, slow pace on the levels, monotonous gameplay, small text on the screen, etc. These need to be improved to increase overall game satisfaction.

5. Conclusion

The gamified environment and balancing platform were successfully tested on twenty individuals in compliance with the standard physical therapy time and procedures. Based on tests with varying tensions and weights, the equipment was observed to require recalibration every 1 hr and 30 minutes. The users also gave feedback based on a rating from the System usability scale (SUS) and IMI (Intrinsic Motivation Inventory) with regards to the different aspects of the game and equipment. The overall evaluation of the game got a score of *7.58* with stability as the strongest feature and rewards as the lowest one.

Further improvements can be made on the overall system such as the need for technical adjustments in the current build such as bugs, glitches, and sounds, adding a "Congratulations" instead of "Game Over" to increase motivation, integrating more natural

special effects (e.g. fish, seagulls, rain, etc.) and also, make the subject immersed in a relaxing-natural coastal area where the endpoint is a docking area/lighthouse. Lastly, testing the system with post-stroke and post-injury patients of different age groups, height, and body mass index (BMI) instead of healthy individuals would yield a more accurate and better reliability evaluation of the platform.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is a product of collaboration with the Philippine Orthopedic Center, funded by the Department of Science and Technology.

References

[1] B. Lange, C. Chang, E. Suma, B. Newman, A. Rizzo, M. Bolas, "Development and Evaluation of Low Cost Game-Based Balance Rehabilitation Tool Using the Microsoft Kinect Sensor," in IEEE Engineering in Medicine and Biology Society, 1831-1834, 2011, doi:10.1109/IEMBS.2011.6090521

[2] A. Reinthal, K. Szirony, C. Clark, J. Swiers, M. Kellicker, S. Linder, "ENGAGE: Guided Activity-Based Gaming in Neurorehabilitation after Stroke: A Pilot Study," Stroke Research and Treatment, 1-10, 2012, doi:10.1155/2012/784232

[3] M. W. Kennedy, J. P. Schmiedeler, C. R. Crowell, M. Villano, A. D. Striegel and J. Kuitse, "Enhanced feedback in balance rehabilitation using the Nintendo Wii Balance Board," 2011 IEEE 13th International Conference on e-Health Networking, Applications and Services, Columbia, MO, 2011, 162-168, doi: 10.1109/HEALTH.2011.6026735.

[4] L. Neri, G. Adorante, G. Brighetti and E. Franciosi, "Postural rehabilitation through Kinect-based biofeedback," 2013 International Conference on Virtual Rehabilitation (ICVR), Philadelphia, PA, 2013, 218-219, doi: 10.1109/ICVR.2013.6662110.

[5] A. D. Williams, A. Kumawat, K. Agarwal, Q. A. Boser, H. Rouhani and A. H. Vette, "An instrumented wobble board for assessing and training dynamic sitting balance," 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, 2017, 2249-2254, doi: 10.1109/SMC.2017.8122955.

[6] K. Song, S. Shin, H. Kim, S. Chung, J. An and C. Lim, "Effect of balance board training on lower limb muscle activity," 2012 5th International Conference on BioMedical Engineering and Informatics, Chongqing, 2012, 515-518, doi: 10.1109/BMEI.2012.6513165.

[7] A. Timmermans, H. Seelen, R. Geers, P. Saini, S. Winter, J. Vrugt, H. Kingma, "Sensor-Based Arm Skill Training in Chronic Stroke Patients: Results on Treatment Outcome, Patient Motivation, and System Usability,"in IEEE Transactions on Neural Systems and Rehabilitation Engineering, **18**(3), 284-292, 2010, doi:10.1109/TNSRE.2010.2047608.

[8] M. Yonei, K. Tanaka, "Toward measurement display content for improving instrinsic motivation in smart health equipment," in 2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech 2019), 206-209, 2019,doi:10.1109/LifeTech.2019.8884067.

[9] H. Lim, M. Tan, R. SJ-Reyes, Aiding Physical Rehabilitation Through Gamification: Development of a Gamified Environment Utilizing an Inertial Measurement Unit mounted on a Balance Board for Physical Therapy, M.S. Thesis, Ateneo de Manila University, 2016.

[10] J. Garcia, M. Rigor, R. SJ-Reyes, Intervention Study of an Enhanced Serious Game: A Comparative Game Design Analysis between Previous and Latest Version, M. S. Thesis, Ateneo de Manila University,2017.

[11] J. Chua, R. SJ-Reyes, A Pre-Clinical Study of Gamification of Balance Exercises with Intervention Techniques and Feedback System using Intervention, Motivation, and Game Design Analyses, M. S. Thesis,Ateneo de Manila University, 2017.

[12] M. Retirado, R. SJ-Reyes, "Development of an Active Balance Training Platform for a Gamified Physical Rehabilitation," in2nd European Conference on Electrical Engineering and Computer Science (EECS),279-289, 2018, doi:10.1109/EECS.2018.00059.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1171-1178 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An algorithm for Peruvian counterfeit Banknote Detection based on Digital Image Processing and SVM

Bryan Huaytalla, Diego Humari, Guillermo Kemper*

Faculty of Engineering, School of Electronic Engineering, Universidad Peruana de Ciencias Aplicadas, Lima 15023, Peru

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 30 January, 2021
Online: 25 February, 2021

Keywords:
Counterfeit
Detection
SVM
Digital Image Processing
Banknote

ABSTRACT

In this work we propose an algorithm for Peruvian counterfeit banknotes detection. Our algorithm operates in banknotes with 50, 100 and 200 soles denominations that were manufactured from 2009 onwards. This algorithm offers an automatic diagnosis based on digital image processing and support vector machines (SVM). Current Peruvian counterfeit detection systems are specially designed to analyze relevant characteristics in dollars and euros. Then, some counterfeiters can fool these systems. We made our detection system robust because we focus on the image acquisition and the segmentation of intaglio marks engraved over the banknotes. After segmentation, we applied embossing and Sobel filters followed by an aperture morphological operation to obtain special characteristics that were then classified by an SVM. We have validated our methodology using real and fake banknotes from a dataset of 240 samples provided by Central Reserve Bank of Peru (BCRP). Our final identification accuracy was 96.5%.

1. Introduction

In order to properly introduce the proposed algorithm, it's been divided into three sections: motivation and incitement, literature review and contribution.

1.1. Motivation and incitement

In Peru, counterfeiting money is a well-known fraud method. In fact, according to the Central Office Against Counterfeiting Money from Peru, the amounts of fraudulent banknotes have reached millionaire quantities (Oficina Central de Lucha contra la Falsificación de Numerario – OCN – Peru: "Principales operativos contra la falsificación" Available: http://www.ocn.gob.pe). As a result, the annual report of the United States Secret Service, an entity with which the Peruvian National Police works together to combat national and foreign banknote counterfeiting, places Peru as the main banknote counterfeiter country worldwide, surpassing other South American countries such as Colombia , which had this title before.

In Peru, this type of practice is especially problematic and constitutes an attack on society, because it makes people to lose trust on currency in circulation which increases money transaction time. Even worse, this distrust in Peruvian money hurts the economic growth and, in the long term, it could also lead to inflation because, as there is more money circulating than the amount the government can handle, people purchasing power increases, which would turn into an increase in prices. Finally, this ends up affecting the economy of any person or any company that receives the fradulent banknotes, especially when those banknotes are the ones with the highest denomination.

In this context, some enterprises offer solutions in the form of counterfeit banknote detectors. However, most of these detectors are based on foreign banknote manufacturing technologies, such as the "Elwic BC 3500 UV / MG next-generation banknote counter" or "PREZZOPAZZO DST-38D banknote detector", which use verification parameters such as magnetism, infrared or banknote thickness, all of which have already been fooled by Peruvian banknotes counterfeiters.

A common method used by scammers is to discolorate low-denomination banknotes (i.e. 10 or 20 soles banknotes), and then overprint them with high-denomination banknote values (i.e. 50, 100 or 200 soles banknote values). These overprints can be of very high detail so they successfully fool the aforementioned verification parameters used by commercial detectors. However these type of techniques are less common on lower denomination banknote (10 and 20 soles) as they would be less cost effective.

Our algorithm focuses on intaglio marks which are distinctive for Peruvian banknotes and whose imitations have not reached such a high level of sophistication.

*Corresponding Author: Guillermo Kemper, Av. Prolongación Primavera 2390, Monterrico, Santiago de Surco, Lima 15023, Peru, email: guillermo.kemper@upc.pe

The focus of the algorithm is on the higher denomination banknotes (50, 100 and 200 soles) since these are the more commonly counterfeited and the more damaging to the victims.

Some other commercial detectors are aimed to serve more as inspection tools for the user, using ultraviolet lights or magnifying glasses as is in the case of "Elwic Professional Banknote Detector IRD-2200" or the "Anizun Magnifying glass mini microscope". These solutions do not provide a clear detection verdict. Rather, they are based on users' prior knowledge of the distinctive marks and details that a legitimate banknote must have. By relying on the human senses, these methods have an intrinsic error. Furthermore, counterfeiters are capable of mimicking several of the most distinctive characteristics of authentic banknotes, such as fluorescent markings or some images that give the illusion of being seen against the light (Arjowiggins Security. "Security features: Watermarks, security fibres and security threads" Available: https://securitypapers.arjowiggins.com). Compared to these detection tools, our method offers a clear and readable verdict on the denomination and legitimacy of the analyzed banknote.

Another common counterfeit technique used in Peruvian banknotes is the removal of security bands from 100 and 200 soles banknotes. These bands are manually removed from a legitimate banknote and placed on a falsified one, then the legitimate but damaged banknote can be exchanged for a new one in banks. This way, counterfeiters have a fake one with a security band that would pass even a meticulous examination. Our algorithm can tackle this problem as it uses banknotes intaglio marks that cannot be removed and placed in falsified ones.

1.2. Literature review

We reviewed various scientific publications related to banknote analysis.

For example in [1] and [2] the authors tackled the automatic banknote denomination recognition problem by using digital image processing techniques. However in our proposed algorithm we not only determine banknote denomination but, in addition, perform a procedure aimed to identify the veracity of the banknote by analyzing intaglio marks with digital image processing techniques.

On the other hand, in [3] and [4] the authors were concerned about banknotes veracity but did not contemplate the use of an enclosure to standardize image light conditions. In comparison, our proposed solution does use an enclosure to ensure uniform illumination across all processed training and validation images.

Publications such as [5] and [6] that are based on digital image processing techniques have some flaws, as they did not consider the use of supervised learning algorithms to solve the classification task. On the contrary, our proposed solution uses a support vector machine (SVM) to tackle the classification problem after a training process performed with labeled images.

Some recent publications include the use of artificial intelligence [7] and deep learning [8] to detect counterfeit currency. However, these algorithms were designed and tested to work on other currencies such as dollars, euros, etc. Its performance on peruvian banknote has not been tested and they don't use the intanglio marks as a classification parameter, which is the most difficult detail to forge.

Likewise, convolution techniques and neural networks are used in [8] which demand a lot of computational load and therefore requires to be implemented in hardware boards such as Jetson TX2. In contrast, we propose a low complexity algorithm that can be easily implemented on a small single-board computer (such as Raspberry Pi 3) with very satisfactory results.

1.3. Contribution and paper organization

Given that intaglio marks (Banco Central de Reserva del Perú, BCRP: "Elementos de seguridad de los billetes de 100 soles", Available: http://www.bcrp.gob.pe) are banknotes difficult-to-falsificate features we decided to design an algorithm that analyzes them (see Figure 1 for a view of some intaglio marks). In order to analyze intaglio marks we capture banknote images using a digital camera and an enclosure to guarantee proper light conditions. These images are then analyzed and enhanced using image processing techniques and a support vector machine (SVM) to determine whether the image corresponds to a legitimate or a fake banknote. Thus, our proposed system offers an automatic diagnostic on the veracity of the banknote analyzed.

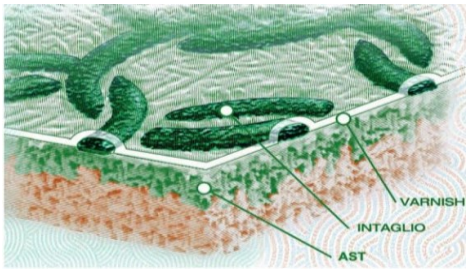


Figure 1: Intaglio marks [9].

Taking all these considerations into account, our methodology was able to perform a 96.5% of detection over the image dataset comprising 240 banknotes provided by the Central Reserve Bank of Peru (BCRP).

The amount of 240 samples was the result of the source of information provided by the Central Reserve Bank of Peru.

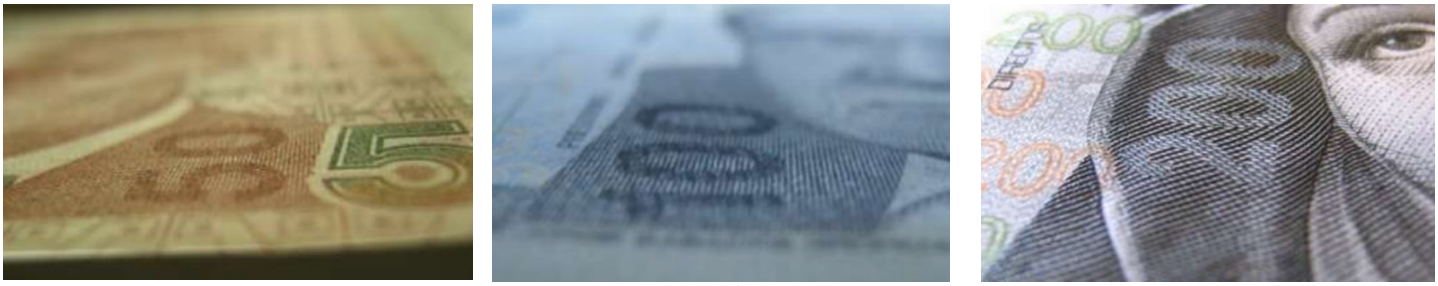
The selection of these samples had the following filters.

- The number of samples for the indicated denominations is reduced to 2 groups.
- The first group is confirmed by banknotes that did not enter into circulation, i.e., that did not deceive people.
- The second group of banknotes belongs to the group that entered circulation, i.e. the user did not detect the counterfeit banknote at the beginning.

For the validation process the second group was selected. In this case, BCRP specialists selected the banknotes that presented the greatest difficulty in detecting counterfeits. The details of the proposed algorithm, the results of the tests and the conclusions will be described in the following sections.

2. Proposed algorithm

For a better understanding, the distinctive characteristics of Peruvian banknotes will be first described and then a discussion of the main blocks of our algorithm will be presented.



(a) (b) (c)

Figure 2: Intaglio marks for a (a) Fifty soles banknote, (b) One-hundred soles banknote, (c) Two-hundred soles banknote.

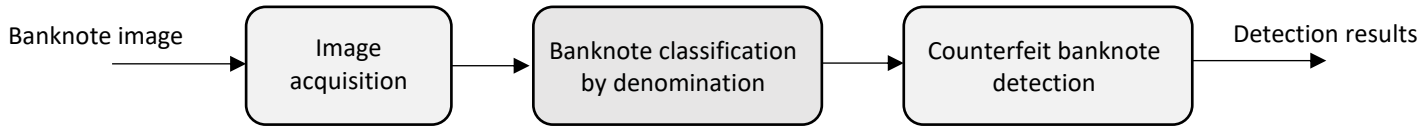
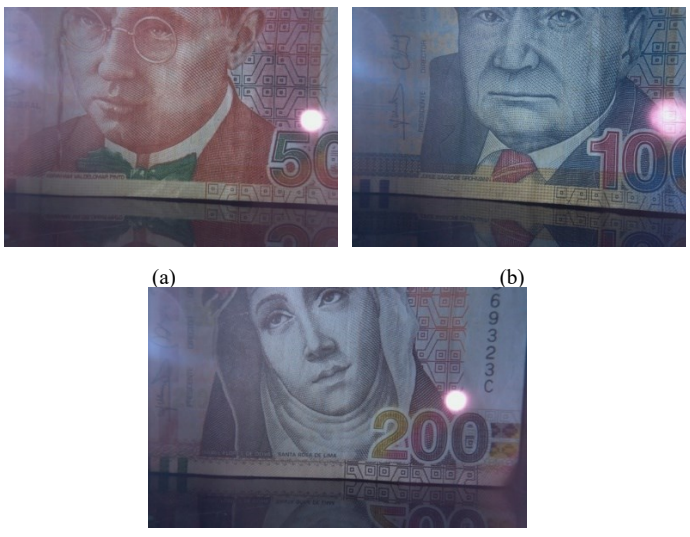


Figure 3: Block diagram of the proposed algorithm

2.1. Banknotes characteristics

We will focus our analysis on the intaglio marks from 50, 100 and 200 soles banknotes (Banco Central de Reserva del Perú, BCRP: "Nuevos billetes de 50, 100 y 200 soles", Available: http://www.bcrp.gob.pe). These marks can be appreciated by the human eye only if banknotes are positioned as in Figure 2 and with good illumination conditions. These marks depend on the banknote denomination, for this reason it is important to first determine the banknote denomination before proceeding with the analysis of intaglio marks.

In Figure 3 we present the block diagram of the proposed algorithm. In the following sections we will describe each block.



(a) (b)

(c)

Figure 4: Sample of acquired images for (a) Fifty soles banknote, (b) One-hundred soles banknote, (c) Two-hundred soles banknote.

2.1. Image acquisicion

We used a closed enclosure to control light conditions, the enclosure was illuminated by a 3W white light LED. To capture the intaglio marks, every picture was taken from the right side of the banknote obverse. The sample banknotes are placed into the enclosure using mechanical guides that help with their correct and quick placement. LED lights were positioned so the area of interest would be distinguished more clearly.

A Raspberry v1 camera was used to take the picture, this camera has a CMOS OmniVision OV56476 with 5 megapixels (2592 x 1944 pixels resolution). The banknotes region of interest has RGB format (true color, 24 bits per pixel). Each primary component is defined as $I_R(x,y)$, $I_G(x,y)$ and $I_B(x,y)$, where (x,y) are the spatial coordinates.

All the computations were made by the Raspberry Pi 3B + single board computer which is compatible with the camera used and can perform all the required processing steps in less than 3 seconds with low power consumption (5 V /2.5 A). The acquired image is set to M=1644 rows and N=2592 columns.

In Figure 4 we showed 50, 100 and 200 soles banknotes taken under the described conditions.

2.2. Banknote classification by denomination

Given the captured images, we defined three regions of interest that were fixed so we can identify relevant characteristics in each banknote. These regions of interest can be expressed as:

Region 1:

$$I1_R(x,y) = I_R(x+400, y+500) \quad (1)$$

$$I1_G(x,y) = I_G(x+400, y+500) \quad (2)$$

$$I1_B(x,y) = I_B(x+400, y+500) \quad (3)$$

where:

$$x = 0,1,\dots,600$$
$$y = 0,1,\dots,500$$

Region 2:

$$I2_R(x,y) = I_R(x+780, y+750) \quad (4)$$

$$I2_G(x,y) = I_G(x+780, y+750) \quad (5)$$

$$I2_B(x,y) = I_B(x+780, y+750) \quad (6)$$

where:

$$x = 0,1,\dots,519$$
$$y = 0,1,\dots,699$$

Region 3:

$$I3_R(x,y) = I_R(x+780, y+1500) \quad (7)$$

$$I3_G(x,y) = I_G(x+780, y+1500) \quad (8)$$

$$I3_B(x,y) = I_B(x+780, y+1500) \quad (9)$$

where:

$$x = 0,1,\dots,519$$
$$y = 0,1,\dots,1024$$

These regions were chosen because they allow us to differentiate between banknotes denomination by their coloration (See Fig. 5 for an example in a 200 soles banknote).



Figure 5: Two-hundred soles banknote with the three regions of interest (red boxes).

We averaged color components in every region of interest as features. To compute these averages, we use the following equations:

$$P1_R = \frac{1}{M1 \times N1} \sum_{x=0}^{M1-1} \sum_{y=0}^{N1-1} I1_R(x,y) \quad (10)$$

$$P1_G = \frac{1}{M1 \times N1} \sum_{x=0}^{M1-1} \sum_{y=0}^{N1-1} I1_G(x,y) \quad (11)$$

$$P1_B = \frac{1}{M1 \times N1} \sum_{x=0}^{M1-1} \sum_{y=0}^{N1-1} I1_B(x,y) \quad (12)$$

where M1 = 601 and N1 = 501

$$P2_R = \frac{1}{M2 \times N2} \sum_{x=0}^{M2-1} \sum_{y=0}^{N2-1} I2_R(x,y) \quad (13)$$

$$P2_G = \frac{1}{M2 \times N2} \sum_{x=0}^{M2-1} \sum_{y=0}^{N2-1} I2_G(x,y) \quad (14)$$

$$P2_B = \frac{1}{M2 \times N2} \sum_{x=0}^{M2-1} \sum_{y=0}^{N2-1} I2_B(x,y) \quad (15)$$

where M2 = 520 and N2 = 650

$$P3_R = \frac{1}{M3 \times N3} \sum_{x=0}^{M3-1} \sum_{y=0}^{N3-1} I3_R(x,y) \quad (16)$$

$$P3_G = \frac{1}{M3 \times N3} \sum_{x=0}^{M3-1} \sum_{y=0}^{N3-1} I3_G(x,y) \quad (17)$$

$$P3_B = \frac{1}{M3 \times N3} \sum_{x=0}^{M3-1} \sum_{y=0}^{N3-1} I3_B(x,y) \quad (18)$$

where M3 = 520 and N3 = 1025

We compound this averages in a feature vector V_{C1} defined as:

$$V_{C1} = [P1_R\ P1_G\ P1_B\ P2_R\ P2_G\ P2_B\ P3_R\ P3_G\ P3_B] \quad (19)$$

The compound feature vector V_{C1} is then classified using a Support Vector Machine (SVM) with a radial basis function kernel [10]. This kernel, which is described as a Gaussian function, was chosen because of its high performance to classify nonlinear data, as it is in our case. SVM was selected as the classifier because of its better performance compared with other techniques like KNN and because of its robustness to classify nonlinear data [11]. We used 70% of the banknotes for training and 30% for testing. We obtained a 99% of accuracy in this task as will be detailed in the results section.

2.3. Counterfeit banknote detection

After banknote classification by denomination, we now proceed to determine if the banknote is false or legitimate. To do so, we first isolate a counterfeit detection region of interest per banknote denomination as follows:

For fifty soles banknotes:

$$I4_R(x,y) = I_R(x+760, y+1460) \quad (20)$$

$$I4_G(x,y) = I_G(x+760, y+1460) \quad (21)$$

$$I4_B(x,y) = I_B(x+760, y+1460) \quad (22)$$

where:

$$x = 0,1,\dots,520$$
$$y = 0,1,\dots,365$$

For hundred soles banknotes:

$$I4_R(x,y) = I_R(x+630, y+465) \quad (23)$$

$$I4_G(x,y) = I_G(x+630, y+465) \quad (24)$$

$$I4_B(x,y) = I_B(x+630, y+465) \quad (25)$$

where:

$$x = 0,1,\dots,295$$
$$y = 0,1,\dots,550$$

For two-hundred soles banknotes:

$$I4_R(x,y) = I_R(x+200, y+545) \quad (26)$$

$$I4_G(x,y) = I_G(x+200, y+545) \quad (27)$$

$$I4_B(x,y) = I_B(x+200, y+545) \quad (28)$$

where:

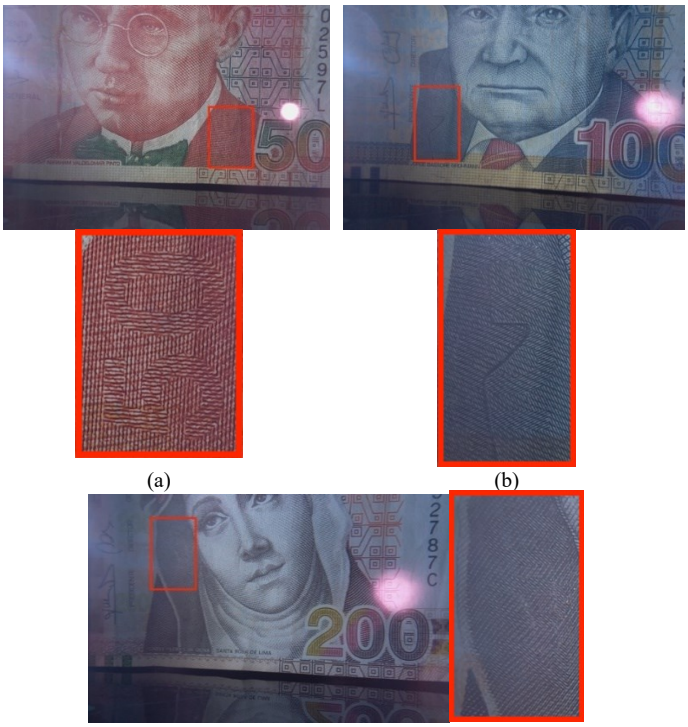
$$x = 0,1,\dots,500$$
$$y = 0,1,\dots,255$$

In Figure 6, we zoom in these regions of interest.

After the selection of the counterfeit detection region of interest, we compute averages in each component using the following equations:

$$P4_R = \frac{1}{M4 \times N4} \sum_{x=0}^{M4-1} \sum_{y=0}^{N4-1} I4_R(x,y) \quad (29)$$

$$P4_G = \frac{1}{M4 \times N4} \sum_{x=0}^{M4-1} \sum_{y=0}^{N4-1} I4_G(x,y) \quad (30)$$



(a) (b)

(c)

Figure 6: Amplified regions of interest for counterfeit banknote detection in the case of a (a) Fifty soles banknote, (b) One-hundred soles banknote, (c) Two-hundred soles banknote.

$$P4_B = \frac{1}{M4 \times N4} \sum_{x=0}^{M4-1} \sum_{y=0}^{N4-1} I4_B(x,y) \quad (31)$$

Every color component is then enhanced using an embossing filter [12] which is designed according to the banknote denomination, as is shown in equations (32), (33) and (34).

$$\mathbf{H1_{50}} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (32)$$

$$\mathbf{H1_{100}} = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad (33)$$

$$\mathbf{H1_{200}} = \begin{bmatrix} -2 & 0 & -2 \\ 0 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix} \quad (34)$$

We then obtain the filtered components $I4'_R(x,y)$, $I4'_G(x,y)$ and $I4'_B(x,y)$ which have highlighted textures as can be seen in Figure 7.

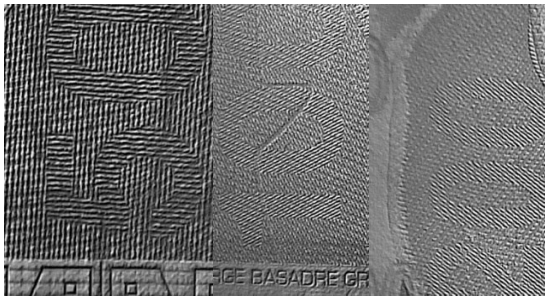


Figure 7: Results of the embossing filter in every region of interest for 50,100 and 200 soles banknotes.

To magnify edges, we applied Sobel filters [13] over the embossing components. In the case of the one-hundred banknote, we used a sobel variation filter due to the diagonal orientation of intaglio marks in this banknote. Final sobel filters are defined in equations (35), (36) and (37).

$$\mathbf{H2_{50}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \times 0.5 \quad (35)$$

$$\mathbf{H2_{100}} = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 0 & -10 \\ 0 & 0 & 2 \end{bmatrix} \times 0.7 \quad (36)$$

$$\mathbf{H2_{200}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & -2 & -1 \end{bmatrix} \times 0.5 \quad (37)$$

After this step, we obtain as results $I4''_R(x,y)$, $I4''_G(x,y)$ and $I4''_B(x,y)$. Sobel filters are designed to sharp edges and variations [14], as we can see in Figure 8.



Figure 8: Sobel filter results for 50, 100 and 200 soles banknotes.

Now, the components are binarized using a threshold that varies with the banknote denomination [15]. Since constant illumination conditions are guaranteed, constant thresholds are used for each denomination in the binarization process. The resulting components are expressed using the following equations:

For fifty soles banknotes:

$$V4_R(x,y) = \begin{cases} 1, & I4''_R(x,y) \geq 40 \\ 0, & I4''_R(x,y) < 40 \end{cases} \quad (38)$$

$$V4_G(x,y) = \begin{cases} 1, & I4''_G(x,y) \geq 40 \\ 0, & I4''_G(x,y) < 40 \end{cases} \quad (39)$$

$$V4_B(x,y) = \begin{cases} 1, & I4''_B(x,y) \geq 40 \\ 0, & I4''_B(x,y) < 40 \end{cases} \quad (40)$$

For one-hundred soles banknotes:

$$V4_R(x,y) = \begin{cases} 1, & I4''_R(x,y) \geq 130 \\ 0, & I4''_R(x,y) < 130 \end{cases} \quad (41)$$

$$V4_G(x,y) = \begin{cases} 1, & I4''_G(x,y) \geq 130 \\ 0, & I4''_G(x,y) < 130 \end{cases} \quad (42)$$

$$V4_B(x,y) = \begin{cases} 1, & I4''_B(x,y) \geq 130 \\ 0, & I4''_B(x,y) < 130 \end{cases} \quad (43)$$

For two-hundred soles banknotes:

$$V4_R(x,y) = \begin{cases} 1, & I4''_R(x,y) \geq 150 \\ 0, & I4''_R(x,y) < 150 \end{cases} \quad (44)$$

$$V4_G(x,y) = \begin{cases} 1, & I4''_G(x,y) \geq 150 \\ 0, & I4''_G(x,y) < 150 \end{cases} \quad (45)$$

$$V4_B(x,y) = \begin{cases} 1, & I4''_B(x,y) \geq 150 \\ 0, & I4''_B(x,y) < 150 \end{cases} \quad (46)$$

The binarization process allowed us to reduce information from grayscale images [16]. Instead of an 8-bit image, now we have a 1-bit image, as seen in Figure 9.



Figure 9: Results from the binarization process

After this, we filtered the results using the aperture morphological operation. We do this to eliminate some noisy elements and enhance visibility of objects of interest (intaglio marks) [17]. Structuring elements are designed depending on the banknote denomination as follows.

$$\mathbf{H3_{50}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (47)$$

$$\mathbf{H3_{100}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (48)$$

$$\mathbf{H3_{200}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad (49)$$

$$\mathbf{I4'''_n} = (\mathbf{V4_n} \ominus \mathbf{H3_n}) \oplus \mathbf{H3_n} = \mathbf{V4_n} \circ \mathbf{H3_n} \quad (50)$$

Thus, we obtain new image components $\mathbf{I4'''_R}$, $\mathbf{I4'''_G}$ and $\mathbf{I4'''_B}$. Figure 10 shows some results.

We then apply connectivity 4 to count the number of objects in each component. The number of objects per component are denominated $K4_R$, $K4_G$ y $K4_B$. We use these values to create a new compound feature vector V_{C2} as follows:

$$V_{C2} = [P4_R \ P4_G \ P4_B \ K4_R \ K4_G \ K4_B] \quad (51)$$

This feature vector is classified using an SVM trained per each banknote denomination (i.e. we used 3 SVMs). Thus, we used independent hyperplanes to perform our classification for our 6-dimensional vector [18]. Again, we used a radial basis function kernel [11] and 70% of the banknotes for training and 30% for testing.

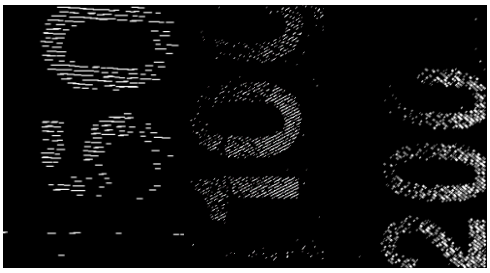


Figure 10: Results after aperture morphological operation using specially designed structural elements.

3. Results

In this section we present our results for the classification and detection tasks using our 240 banknote dataset which are distributed as shown in Table 1. The ground-truth veracity of the banknotes was determined by a specialist of the BCRP.

Result for the banknote classification by denomination task are presented in Table 2. As we used 70% of the data to train the SVM classifier these results correspond to the remaining 30%. To compute our performance metric we divide the number of correct results by total number of samples.

As we can see, the results from this step were almost perfect, we only had one mistake in the classification of a one-hundred soles banknote.

In Table 3 we show the results for the detection process.

Results obtained in the detection process are satisfactory. There was only one mistake on a 50 soles banknote and another one on a 100 soles banknote.

Table 1: Banknotes distribution

Denomination	Veracity	Number of samples
50	Legitimate	32
	Fake	50
100	Legitimate	40
	Fake	49
200	Legitimate	32
	Fake	37

Table 2: Results for the classification process

Denomination	Number of test samples	Correctly classified samples	Performance
50	22	22	100%
100	33	32	97%
200	21	21	100%

Table 3: Results for the detection process.

Denomination	Number of test samples	Correctly detected samples	Percentage of correct detection
50	25	24	96%
100	28	27	96.4%
200	21	21	100%

In Tables 4, 5 and 6, we show correct and incorrect detections of the true and false banknote by denomination. In Table 4, a legitimate 50 soles banknote was incorrectly detected as false. This was caused by a deteriorated banknote that was bent on several places. In Table 5, a legitimate 100 soles banknote was incorrectly detected as false. This was caused by an incorrect insertion of the banknote on the image acquisition zone.

Table 4: Results for the 50 soles banknote detection process

Sample\Result	Legitimate	Fake
Legitimate	14	1
Fake	0	10

Table 5: Results for the 100 soles banknote detection process

Sample\Result	Legitimate	Fake
Legitimate	14	1
Fake	0	12

Table 6: Results for the 200 soles banknote detection process

Sample\Result	Legitimate	Fake
Legitimate	7	0
Fake	0	14

The average performance (according with Table 3) is then 97.47%. We observed that mistakes were generated because of banknote positioning problems (they were not inserted correctly or maybe the banknote was bent so its image could not be acquired correctly).

According to Table 3, the average detection performance of the proposed algorithm is 97.47%. The detection errors were mainly generated by banknote positioning problems, i.e. the banknotes were not inserted correctly or perhaps the banknote was bent, so its image could not be acquired correctly.

These problems could be solved by the use of additional hardware or by training the SVM classifier with images with these imperfections.

Finally, the total effectiveness of our algorithm $P_{(A)}$ (including classification and detection) is computed using the following equation:

$$P_{(A)} = P_{(B)} \times P_{(C/B)} \quad (52)$$

where $P_{(B)}$ is the percentage of correct classifications, $P_{(C/B)}$ is the percentage of correct detections given a correct classification. Table 7 summarizes our final results.

Then, the total effectiveness average is 96.5%, which are results that could improve with a bigger sample. Compared to other studies with different currencies, it reached a higher percentage than Ballado's (95%) [6] and a lower percentage than Han's (100%) [7].

Table 7: Total effectiveness

Denomination	$P_{(B)}$	$P_{(C/B)}$	$P_{(A)}$
50	100%	96%	96%
100	97%	96.4%	93.5%
200	100%	100%	100%

4. Conclusions

The proposed algorithm was able to identify the denomination and veracity of Peruvian banknotes that were manufactured from 2009 onwards using the intanglio marks as a deciding factor, and support vector machines (SVM). The practical operation of this algorithm was validated by specialists from the Central Reserve Bank of Peru (BCRP). The few cases in which the algorithm failed were caused by the deteriorated state of the banknote. It is suggested not to use banknote that is ripped or folded in order to avoid missreadings.

Our automatic algorithm was efficient and can be used by any user who needs to determine the veracity of Peruvian 50, 100 or 200 soles banknotes.

The algorithm was entirely developed in Python and was embedded into a Raspberry single board computer. Additionally, does not consume high computational resources.

As we increase the number of training samples we can improve SVM performance for classification.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank the Research Directorate of the Peruvian University of Applied Sciences (UPC) for technical and economic support and the Central Reserve Bank of Peru (BCRP) for providing the banknotes used in this study.

References

[1] G. Ramasamy and S. Subramanian, "Identification of currency denomination using image processing," International Journal of Advanced Research, Ideas and Innovations in Technology, **5**(2), 2019.

[2] S. Guzmán, Detector de billetes para personas con discapacidad visual, BSc. Thesis, Universidad Tecnológica del Perú, 2012.
[3] D. Alekhya, G. Prabha, G. D. Rao, "Fake currency detection using image processing and other standard methods," International Journal of Research in Computer and Communication Technology, **3**(1), 2014.
[4] J. L. Estela, Análisis comparativo de algoritmos de reconocimiento de imágenes por descriptores de color para la identificación de billetes, BSc. Thesis, Universidad Señor de Sipán, Chiclayo, Peru, 2016.
[5] Abdullah M. Shoeb, Hadeer M. Sayed, Norah F. Saleh, Ehab I. Khafagy, Sameh Neji, "Software system to detect counterfeit Egyptian currency," in International Conference on Control, Instrumentation, Communication and Computational Technologies, 2016, doi:10.1109/ICCICCT.2016.7988010.
[6] A. H. Ballado, J. C. Dela Cruz, G. O. Avendaño, N. M. Echano, J. E. Ella, M.E.M. Medina, B.K.C. Paquiz, "Philippine currency paper bill counterfeit detection through image processing using Canny Edge Technology," in 8th IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2015, doi:10.1109/HNICEM.2015.7393184.
[7] M. Han, J. Kim, "Joint banknote recognition and counterfeit detection using explainable artificial intelligence," Sensors, **19**(16), 3607, 2019, doi.org/10.3390/s19163607.
[8] T. D. Pham, C. Park, D. T. Nguyen, G. Batchuluun, K. R. Park, "Deep learning-based fake-banknote detection for the visually impaired people using visible-light images captured by smartphone cameras," IEEE Access, **8**, 63144 – 63161, 2020, doi:10.1109/ACCESS.2020.2984019.
[9] T. Crane, "Some observations on technological developments in durable substrates," Billeteria: International review on cash management, **5**(9), 2011.
[10] D. Duvenaud. Automatic model construction with Gaussian processes, Ph.D. Thesis, Pembroke College, England, 2014.
[11] C. M. Bishop, Patter recognition and machine learning, Springer, 2006.
[12] S. Umbaugh, Computer imaging: Digital image analysis and processing, CRC Press, 2010.
[13] J. Álvarez, J. Fernández, "Procesamiento de imágenes bajo Windows CE utilizando el procesaror ARMv 4I," ITECKNE Innovación e Investigación en Ingeniería, **8**(2), 2008, doi: 10.15332/iteckne.v8i2.41.
[14] R. Muthukrishnan, M. Radha, "Edge detection techniques for image segmentation," International Journal of Computer Science and Information Technology, **3**(6), 2011, doi:10.5121/ijcsit.
[15] R. Gonzalez, R. Woods. Digital Image Processing, Pearson, 2018.
[16] C. Solomon, T. Breckon, Fundamentals of digital image processing: A practical approach with examples in Matlab, John Wiley & Sons, 2010.
[17] G. Aguilar, Procesamiento digital de imágenes utilizando filtros morfológicos, BSc. Thesis, Escuela Politécnica Nacional, Quito, Ecuador, 1995.
[18] C. C. Chang, C. J. Lin, "LIBSVM: A Library for Support Vector Machines," ACM Transactions on Intelligent Systems and Technology, **2**(3), 2011, doi.org/10.1145/1961189.1961199.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1191-1198 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Novel Approach to Design a Process Design Kit Digital for CMOS 180nm Technology

Thinh Dang Cong[1,2], Toi Le Thanh[1,2], Phuc Ton That Bao[1,2], Trang Hoang[1,2,*]

[1]*Department of Electronics Engineering, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, 72506, Vietnam*

[2]*Vietnam National University, Ho Chi Minh City, 71308, Vietnam*

ARTICLE INFO

Article history:
Received: 08 December, 2020
Accepted: 31 January, 2021
Online: 25 February, 2021

Keywords:
Process Design Kit (PDK)
Standard Cell Library (SCL)
Liberty format
Wire-load Model (WLM)
Ocean language

ABSTRACT

In this paper, a novel approach to design a Process Design Kit Digital for CMOS 180nm process is presented. This work proposes a detailed flow to design a PDK Digital using Ocean language, which is a vital element in the semi-custom design and applied in education purposes in universities in Vietnam. The PDK digital includes Standard Cell Library containing 47 standard cells and Wire-Load Model. The library is designed based on the CMOS 180nm process with a supply voltage of 1.8V.

1 Introduction

CMOS technology has been developing in recent years due to its benefits such as high integration density, low fabrication cost, etc. Many new techniques of circuit design have been given ranging from RFIC [1]–[5] to mm-Wave IC design [6], [7].

In the past, the full custom design took many resources, efforts, and design time for designing a big chip. Meanwhile, the semi-custom design has significant decreased in the design and verification time for chip design by using the Electronic Design Automation (EDA) tools and PDK digital [8]. Especially, the latter provides a considerable decrease in design time, decreasing around 25% to 50%. Moreover, the latter would be applying for many more complex designs than the former [9].

Focusing on the PDK digital, the one consists of Synopsys library and P&R library [10]. While the former is applied for timing, power consumption, and noise analysis, the latter is used for place and route design in a design flow [11]. There are two ways to design an SCL. The first one uses the Liberty NCX tool of Synopsys. The Liberty NCX determines the function of standard cells. Following this, the tool creates the complete simulations by using HSPICE that would make the liberty file for each logic cell in the library [12], [13]. The second way performs characterization for all standard cells manually using the Spectre tool. The process of characterization consists of many simulations where the transition of the input and the capacitance of output load are varied. Timing and other parameters would be extracted, and outputted as a liberty format file.

The license of the Liberty NCX tool is unavailable at universities in developing countries due to its cost. Therefore, designing an SCL using this tool is almost impossible [13]. With regards to the manual characterization, the challenge is that we have to do many simulations to generate timing, power consumption, etc [14]. This way takes much time to design each logic cell. For instance, to design an INV logic cell, we have to do 7x7 simulations and 49x6 calculations to generate all parameters for the liberty file.

In this work, the Ocean language is proposed to solve problems about the licensing and running manually many simulations. The language is an extension of the Skill language used to control Analog simulation. The flow automatically performs parametric simulations of each logic cell and generates output parameters with the format as the liberty file [15]. Regarding the Wire-Load Model (WLM), this paper proposed a method and used accuracy formulas to design a complete model.

The authors presented ideas about designing a PDK Digital for CMOS technology in the ”2019 19th International Symposium on

*Corresponding Author: Trang Hoang, Department of Electronics Engineering, Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Ho Chi Minh City, 72506, Vietnam, Tel: (+84) 988 071 579 & Email: hoangtrang@hcmut.edu.vn

Communications and Information Technologies (ISCIT)" [1] and the following paper is an extension of this work.

The remainder of this paper is organized as follows: Section II shows the SCL design flow from the transistor level. This section also gives the schematic design, Euler rules for layout design, DRC/LVS definition, etc. A detailed liberty format generation flow is given in section III. Section IV shows how to design a wire-load model. Finally, section V presents the conclusion.

2 Standard Cell Library Design

An SCL is made up of combinational logic cells such as INV, AND, FA, sequential logic cells such as DFF, Latch, and physical cells, etc. The design flow is composed of many steps. To begin with the front-end design stage, firstly, the schematic of all standard cells in the library are designed from transistor level. After that, symbol creation and do simulation at corners [16]. Next, in the back-end design stage, to optimize the layout design in term of area and routing, the layout design needs to have some rules as follows:

1. All the input and output pins must be placed on the intersections of the vertical and horizontal routing tracks with the special exception for power and ground pins.
2. All the cells in the library are designed to be multiple of the unit tile which is defined by a library developer. The height of the cells must be equal and be multiple of the unit tile's height.
3. This work uses M1 and M2 for the layout design.
4. Minimize the width of all standard cells in the library.
5. Optimize pins access to prevent the routing congestion in place and route stage.
6. Some special cells such as filler, decap cells can be put in the library to make sure the connection of supply trails and N-well.

The figure below shows the schematic, Euler path for optimization, and layout design of an AOI22 cell in the library. The layout should follow the considerations as above.

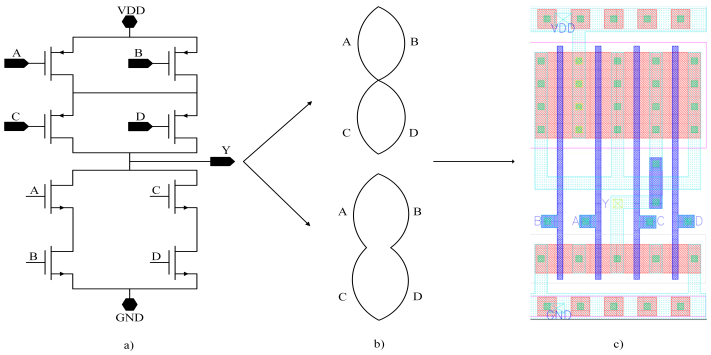


Figure 1: a. Schematic design, b. Euler path, c. Layout design

The Physical Verification process will be performed after the layout design. This process will accomplish sequentially Design Rule Check (DRC) verification and Layout Versus Schematic (LVS) verification.

Firstly, by using the rules from the fabricators, the layout design is checked to ensure no potential violation, which is called DRC verification. The design rules are supplied by the manufacturers to ensure the chip will function properly when fabricated. The design rules include many rules such as the minimal critical dimensions, shape angles, pattern density requirements, etc [17]. The DRC flow chart is shown in Figure 2 and the example design rules are presented in Figure 3.

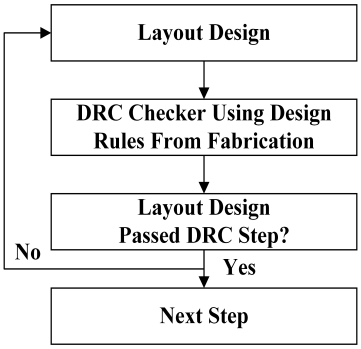


Figure 2: DRC flow chart

For example, the resolution rules for metal 1 and metal 2 layers are 0.23μm and 0.28μm, respectively. With regards to the rules for poly layer, the min.poly width is 0.18μm, min.gate to contact spacing is 0.375μm, and the min.poly overlap of diffusion is 0.22μm. For contact rules, the exact size and min.contact overlap are 0.22μm and 0.12μm, respectively.

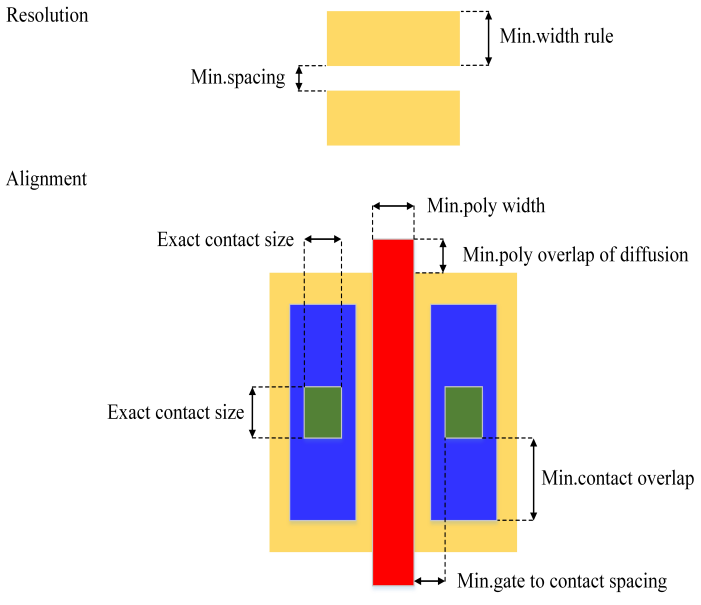


Figure 3: The example rules are used for DRC check

Followed by the LVS verification, which is the more critical verification step. Particularly, the LVS compares any combination of physical or schematic designs.

Depending on the layout design, so different parasitic components can be generated. These parasitic parameters can make the functionality and characteristic of the layout design different from schematic design. For that reason, if the whole parameters in the schematic design are different from the parasitic parameters of the layout design, which makes it impossible to complete the LVS step. Hence, this verification proves the chip after being fabricated will function as the schematic design [17]. When the LVS errors occur, we have to fix the layout design and perform the DRC step again to complete the LVS step without errors. The LVS flow chart is shown in Figure 4.

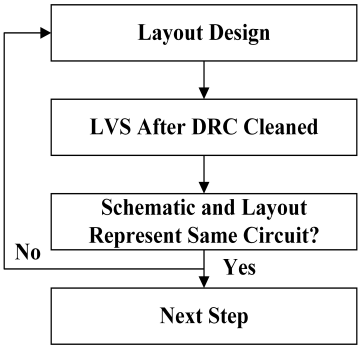


Figure 4: LVS flow chart

The Physical Verification process plays a vital role in the design flow. Therefore, the designers ensure that each cell in the library passed the DRC and LVS without any violations [18]. After the Physical Verification process, the post-layout simulation at corners will be executed before the cell characterization step. Next, Layout Parasitic Extraction (LPE) step is performed. Afterward, the post-layout simulations at corners are performed by using the parasitic parameters extracted. Therefore, the characterization of the standard cells will be calculated [19]. The post-layout simulation becomes possible and correct only after the LVS step is completed, then the extracted parameters are accurate. The post-layout simulation process will be shown in Figure 5.

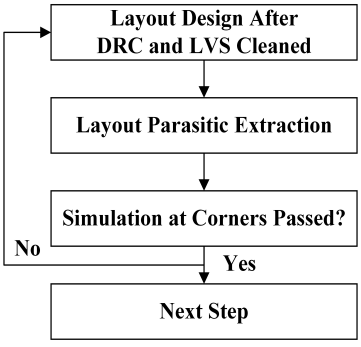


Figure 5: Post Layout Simulation Process Flow chart

Finally, the characterization step is divided into Liberty Format Generation step and Abstract View & GDS step. While the timing models, power models including dynamic and static power are extracted by the former to make the liberty file (*.lib) [20]. The Synopsys internal database format (*.db) may be translated from the liberty file by the Library Compile tool of Synopsys. The latter is used to create the library exchange format (*.lef) file by abstracting from the layout design, and the last step is the generating GDS file [21]. The block diagram in the figure below (Figure 6) illustrates the SCL design flow.

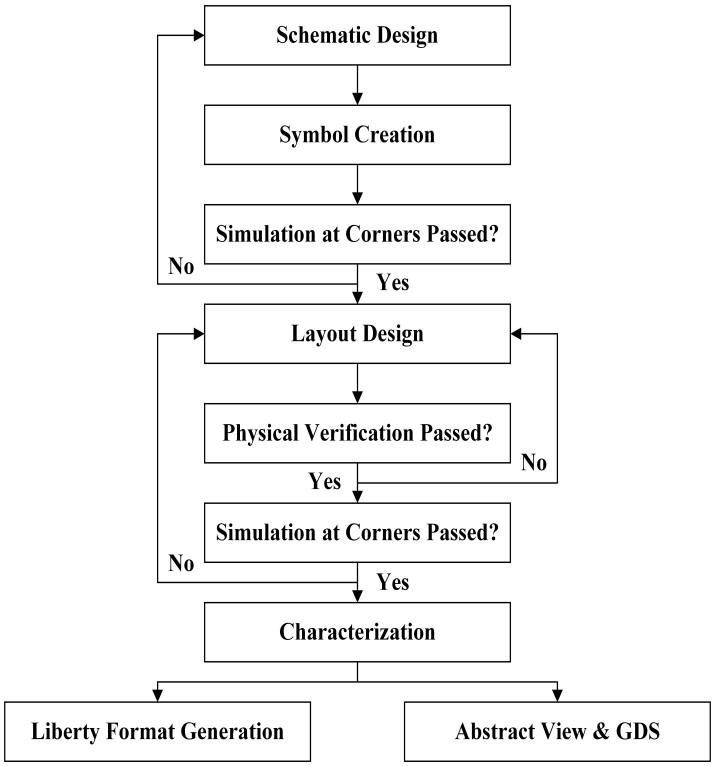


Figure 6: SCL design flow chart

3 Liberty Format Generation

After the post-layout simulation process, the standard cell characterization process is performed. Especially, the liberty format generation process, this process will generate an output file (*.lib) which is one of the most important files in the standard cell library design. The liberty file includes structural information, functional information, timing information, power information, and environmental information of each standard cell.

Firstly, the structural one declares connectivity of a logic cell including cell, bus, and in-out pin descriptions. Secondly, the functional one describes the logical function of each output pin of each cell. Thirdly, timing models are given including cell rise delay, cell fall delay, rise transition, fall transition. Next, power models give the dynamic and static power consumption of each cell. Finally, environmental information describes the manufacturing process, temperature of operation, supply voltage, etc.

In this part, to perform cell characterization and generate output file same as the liberty file's format, all cells in the library will have the same load capacitance C_{load} = 2 fF, 5 fF, 6 fF, 7 fF, 8 fF, 9 fF, 9.5 fF, the input slew of the waveform *slope* = 0.01 ns, 0.02 ns, 0.04 ns, 0.06 ns, 0.08 ns, 0.09 ns, 0.095 ns [22]. The load capacitance and input slew are declared in the Ocean file to create 7x7 = 49 simulation times and extract the timing model and dynamic power

consumption for each cell.

These figures below (Figure7 and Figure 8) show the testbench circuits for the Ocean script run and the results of a NAND gate cell with two inputs.

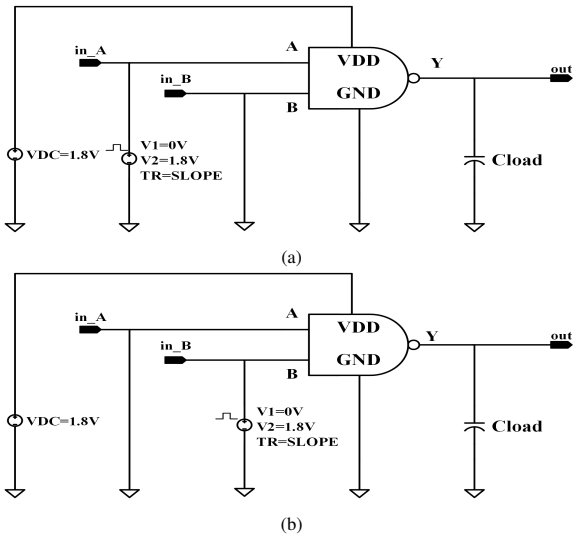


Figure 7: Testbench circuits for characterization

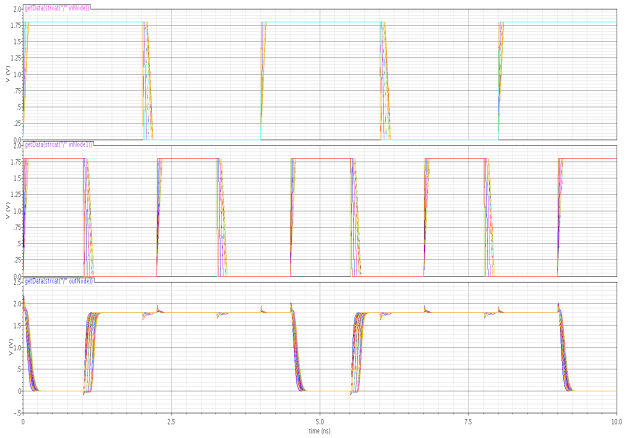


Figure 8: Waveform results

3.1 Input capacitance

In SCL design, one of the most important parameters that we need to calculate is input capacitance. The output capacitance of each cell can be determined but in most cases, the capacitance is specified only for the inputs and not for the outputs, which means that the output pin capacitance is zero in the library design [23].

The following formula illustrates the relationship between input capacitance, current, and supply voltage at an input pin.

$$I = C\frac{dV}{dt} \tag{1}$$

where I is the current at the input pin, V is the supply voltage and it is equal to 1.8V in this paper. By inputting a pulse power to the input pin and calculating the current, the input capacitance value is determined as the formula below:

$$C_{input} = \frac{\int_t^{t+\Delta t} Idt}{\int_t^{t+\Delta t} dV} = \frac{\int_t^{t+\Delta t} Idt}{VDD} = \frac{\int_t^{t+\Delta t} Idt}{1.8} \tag{2}$$

where Δt is the time when the voltage at each input pin changes the logic level.

The two tables below show the capacitance of each input pin of the NAND2 cell. These tables also provide the rise power and fall power which are presented in the next section.

Table 1: PIN A

t_n(ns)	0.01	0.02	0.04	0.06	0.08	0.09	0.095
RISE POWER	0.001098	0.001099	0.001101	0.001098	0.001100	0.001093	0.001097
FALL POWER	0.002029	0.002049	0.002053	0.002040	0.002044	0.002045	0.002044
INPUT CAPACITANCE = 3.079fF							

Table 2: PIN B

t_n(ns)	0.01	0.02	0.04	0.06	0.08	0.09	0.095
RISE POWER	0.001404	0.001406	0.001407	0.001405	0.001412	0.001409	0.001407
FALL POWER	0.001552	0.001534	0.001516	0.001504	0.001491	0.001483	0.001478
INPUT CAPACITANCE = 2.774fF							

3.2 Timing models

All the timing arcs of the cell are declared in the timing model. There are two types of timing models that are linear and non-linear timing. Whereas the former provides less accurate for the sub-micron process, the latter is better. Therefore, most standard cell libraries use the latter. [20]. The definitions of all the timing arcs in the timing model are given in these figures as follows (Figure 9-11):

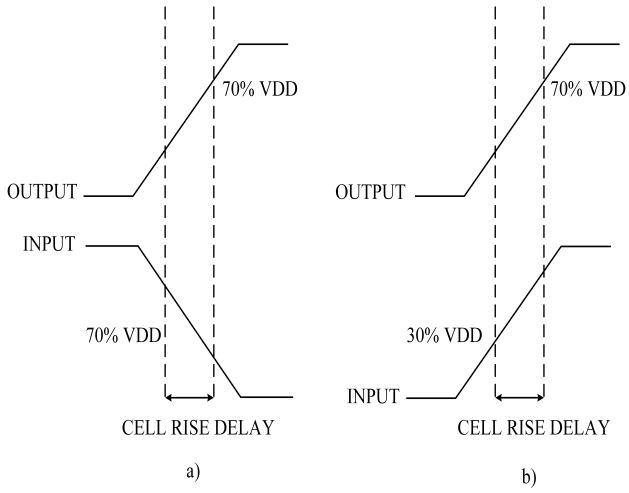


Figure 9: a. Cell rise delay at negative unate case, b. Cell rise delay at positive unate case

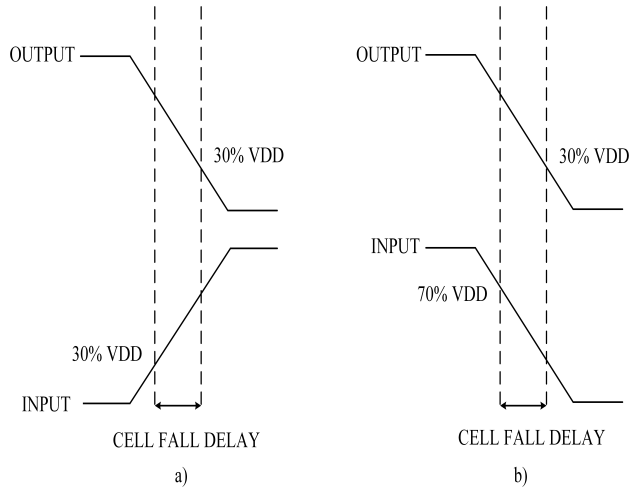


Figure 10: a. Cell fall delay at negative unate case, b. Cell fall delay at positive unate case

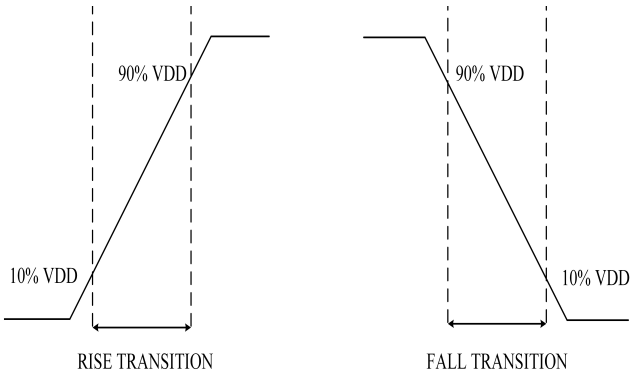


Figure 11: a. Rise transition time, b. Fall transition time

Table 3: Cell Rise Delay

C_{load} \ t_n	0.01	0.02	0.04	0.06	0.08	0.09	0.095
2	0.0284388	0.0459522	0.0481537	0.0503347	0.0524551	0.0545356	0.0555650
5	0.0430937	0.0497589	0.0518424	0.0538036	0.0562515	0.0583046	0.0593685
6	0.0506982	0.0573539	0.0592578	0.0614723	0.0635472	0.0659118	0.0670066
7	0.0587234	0.0650961	0.0671953	0.0692366	0.0712060	0.0733503	0.0744634
8	0.0665493	0.0729249	0.0751495	0.0772798	0.0793351	0.0813871	0.0825202
9	0.0706249	0.0768656	0.0790399	0.0810872	0.0832540	0.0854972	0.0864483
9.5	0.0726732	0.0790870	0.0809977	0.0832278	0.0853940	0.0873195	0.0882354

Table 4: Cell Fall Delay

C_{load} \ t_n	0.01	0.02	0.04	0.06	0.08	0.09	0.095
2	0.0352318	0.0874551	0.0911567	0.0954046	0.0996116	0.103758	0.105814
5	0.0779892	0.0903443	0.0936665	0.0979536	0.1021390	0.105976	0.108099
6	0.0836045	0.0958160	0.0996820	0.1035520	0.1078860	0.112125	0.114218
7	0.0899048	0.1018210	0.1061800	0.1103530	0.1144480	0.118486	0.120481
8	0.0966727	0.1086380	0.1126100	0.1169130	0.1210420	0.125246	0.127239
9	0.1002420	0.1123940	0.1164030	0.1203470	0.1240760	0.128452	0.130556
9.5	0.1019340	0.1139970	0.1181590	0.1222040	0.1261880	0.130108	0.131898

Table 5: Rise Transition

C_{load} \ t_n	0.01	0.02	0.04	0.06	0.08	0.09	0.095
2	0.0285015	0.0525788	0.0565313	0.0590775	0.0617117	0.0643864	0.0657346
5	0.0444646	0.0525918	0.0556488	0.0590963	0.0622286	0.0641798	0.0666422
6	0.0454360	0.0532991	0.0566497	0.0593233	0.0623244	0.0653851	0.0665893
7	0.0483083	0.0564303	0.0588248	0.0616149	0.0644307	0.0672005	0.0683990
8	0.0519840	0.0593587	0.0620207	0.0648213	0.0673532	0.0702575	0.0715643
9	0.0544479	0.0608779	0.0637309	0.0668407	0.0693130	0.0715348	0.0725691
9.5	0.0554892	0.0626609	0.0649285	0.0673040	0.0697900	0.0727158	0.0742639

Table 6: Fall Transition

C_{load} \ t_n	0.01	0.02	0.04	0.06	0.08	0.09	0.095
2	0.0294817	0.0918175	0.0966698	0.102074	0.107603	0.113067	0.115792
5	0.0765303	0.0921913	0.0968935	0.102306	0.107671	0.112408	0.115192
6	0.0763769	0.0914495	0.0964209	0.102027	0.107571	0.113009	0.115709
7	0.0763993	0.0916860	0.0970722	0.102287	0.107523	0.112521	0.115014
8	0.0779230	0.0930608	0.0984027	0.103380	0.108256	0.113244	0.115654
9	0.0789342	0.0936807	0.0983939	0.103534	0.108970	0.114131	0.116615
9.5	0.0800157	0.0946327	0.0993140	0.104174	0.108800	0.113987	0.116831

Table 7: Rise Power Consumption

C_{load} \ t_n	0.01	0.02	0.04	0.06	0.08	0.09	0.095
2	-3.460e-05	-5.033e-05	-5.365e-05	-5.692e-05	-6.020e-05	-6.348e-05	-6.512e-05
5	-4.006e-05	-5.040e-05	-5.376e-05	-5.703e-05	-6.031e-05	-6.351e-05	-6.485e-05
6	-4.021e-05	-5.025e-05	-5.357e-05	-5.691e-05	-6.019e-05	-6.347e-05	-6.510e-05
7	-4.022e-05	-5.011e-05	-5.339e-05	-5.680e-05	-6.010e-05	-6.329e-05	-6.497e-05
8	-4.043e-05	-5.025e-05	-5.353e-05	-5.673e-05	-6.013e-05	-6.327e-05	-6.495e-05
9	-4.050e-05	-5.033e-05	-5.357e-05	-5.686e-05	-6.015e-05	-6.342e-05	-6.508e-05
9.5	-4.051e-05	-5.030e-05	-5.356e-05	-5.685e-05	-6.013e-05	-6.350e-05	-6.515e-05

Table 8: Fall Power Consumption

C_{load} \ t_n	0.01	0.02	0.04	0.06	0.08	0.09	0.095
2	-2.729e-06	-1.502e-05	-1.504e-05	-1.506e-05	-1.506e-05	-1.507e-05	-1.507e-05
5	-1.384e-05	-1.407e-05	-1.404e-05	-1.413e-05	-1.414e-05	-1.424e-05	-1.407e-05
6	-1.276e-05	-1.296e-05	-1.312e-05	-1.305e-05	-1.310e-05	-1.311e-05	-1.316e-05
7	-1.229e-05	-1.246e-05	-1.240e-05	-1.247e-05	-1.251e-05	-1.255e-05	-1.256e-05
8	-1.190e-05	-1.219e-05	-1.222e-05	-1.226e-05	-1.224e-05	-1.222e-05	-1.222e-05
9	-1.209e-05	-1.207e-05	-1.213e-05	-1.204e-05	-1.212e-05	-1.219e-05	-1.223e-05
9.5	-1.211e-05	-1.209e-05	-1.214e-05	-1.219e-05	-1.224e-05	-1.229e-05	-1.230e-05

In this work, threshold values for timing arc calculation are 30%, 70% for cell rise/fall delay calculation and 10%, 90% for rise/fall transition calculation. These values are declared in the header section of the liberty file [24]. These four tables below show the timing arc results of the NAND2 cell after the characterizing step.

3.3 *Power models*

There are two types of power in the power model that are dynamic power and static power consumption. For dynamic power, this one

includes rise power and fall power. While the former is calculated in the time when output changes from 0 to VDD. The latter is determined when output goes from VDD to 0. The formula below shows the dynamic power calculation.

$$VDD^{*}integ(i(\text{"}/V0/PLUS\text{"}), t, t+\Delta t)/(1e-9) \quad (3)$$

These two tables (Table 7 and Table 8) above give the dynamic power consumption of the NAND2 cell after the characterizing step.

In contrast, the static power consumption occurs when the cell is in inactive condition. There are many leakage currents caused the cell has this power consumption such as sub-threshold current which is the main component of the currents, tunneling current through the gate oxide [25]. Static power consumption can be shown as the following formula:

$$LEAKAGE\ POWER = \sum I_{LEAKAGE}.VDD \quad (4)$$

With regards to calculate the static power, firstly, a testbench circuit is designed with all inputs are pulled down as Figure 12a and calculate the power. Next, a testbench circuit is designed with all inputs are pulled up as Figure 12b and then calculate the power. Finally, taking the bigger power number to represent the static power of a cell.

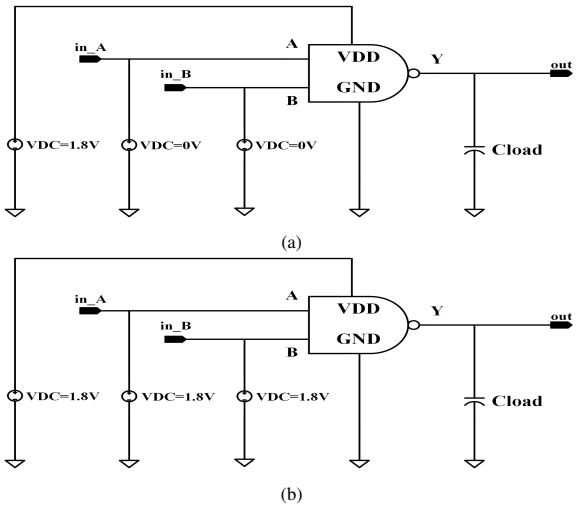


Figure 12: Testbench circuits for static power calculation, (a) Testbench circuit for static power calculation in pull down condition, (b) Testbench circuit for static power calculation in pull up condition

4 Wire-Load Model Design

The wire delays must be estimated during the timing analysis before placement and routing. Using a wire-load model is the simplest method of estimation. The wire-load model gets a rough value of the total wire capacitance based on the size of the chip and the fanout of the net [26]. The wire-load model is specified in the Liberty file.

In this work, the wire capacitance is computed by using the results of T. Sakurai and K. Tamaru's paper [27]. The accuracy of these formulas is practically sufficient for a wide range of wire thickness, wire width, and inter-wire spacing. When two or three nets are placed on bulk silicon, the total capacitance of one net includes the "coupling" capacitance between lines C_{side} and "ground" capacitance between the net and the ground C_{plate}. Below are the detailed steps to determine the wire capacitance of a net.

First, the coupling capacitance C_{plate} between line and bulk silicon is given as formula:

$$C_{plate} = E_{ox}(1.15(\frac{W_{int}}{T_{fox}}) + 2.8(\frac{T_{int}}{T_{fox}})^{0.222}) \quad (5)$$

Next, the coupling capacitance C_{side} between lines is shown as:

$$C_{side} = 2E_{ox}(0.03\frac{W_{int}}{T_{fox}} + 0.83\frac{W_{int}}{T_{fox}} - 0.07\frac{W_{int}}{T_{fox}}^{0.222})(\frac{W_{sp}}{T_{fox}})^{-1.34} \quad (6)$$

Where E_{ox} is a dielectronic constant of SiO_2. It is given as:

$$E_{ox} = E_0K_r = 3.9x8.85xE^{-03} = 0.035(\frac{fF}{\mu m}) \quad (7)$$

Final, the total capacitance or wire capacitance is determined as:

$$C_{total} = C_{plate} + MFC.C_{side}\ (\frac{fF}{\mu m}) \quad (8)$$

Where Miller Coupling Factor (MFC) is equal to 1.5 in this work.

Regarding the resistance of a interconnect net, this value is computed by the formula as below:

$$R_{total} = \rho\frac{L}{S} = \rho\frac{L}{W_{int}T_{int}}\ (\frac{\Omega}{\mu m}) \quad (9)$$

The figure below defines all variables used in the formulas above.

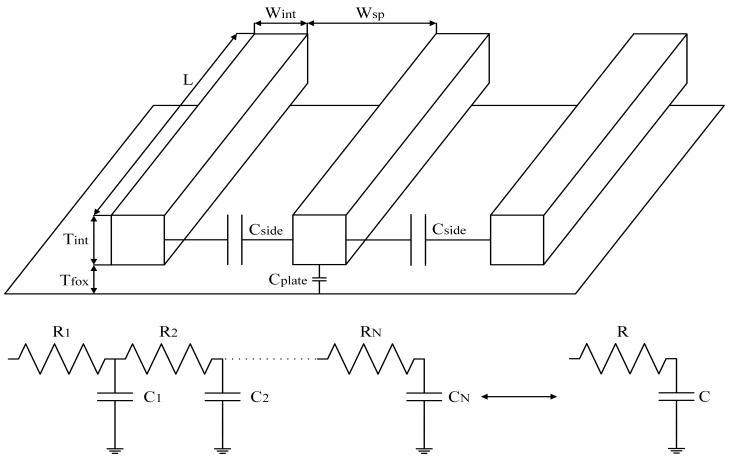


Figure 13: Variables definition used to compute the wire capacitance

This work is performed based on the CMOS 180nm process which has six metal layers. The following table provides the values of all variables defined in the formulas above.

Table 9: Values Of Variables For Wire Capacitance Calculation

Metal Layer	W_{INT} (μm)	W_{SP} (μm)	T_{INT} (μm)	T_{FOX} (μm)	C_{side} ($\frac{fF}{\mu m}$)	C_{plate} ($\frac{fF}{\mu m}$)
M1	0.23	0.23	0.53	0.003981	0.019	1.456
M2	0.28	0.28	0.53	0.003981	0.015	1.737
M3	0.28	0.28	0.53	0.003981	0.015	1.737
M4	0.28	0.28	0.53	0.003981	0.015	1.737
M5	0.28	0.28	0.53	0.003981	0.015	1.737
M6	1.50	1.50	2.34	0.003981	0.007	8.665

The table below shows the results of wire capacitance and resistance with the unit length of 1μm. In addition, the delay of each metal layer is computed by the following formula:

$$t_{delay} = 4^{*}RC \tag{10}$$

Table 10: Values Of Wire Capacitance, Resistance, Delay

Metal Layer	Length (μm)	Capacitance (fF)	Resistance ($ohms$)	Delay (ps)
M1	1	1.484	0.164	0.097
M2	1	1.759	0.135	0.095
M3	1	1.759	0.135	0.095
M4	1	1.759	0.135	0.095
M5	1	1.759	0.135	0.095
M6	1	8.675	0.006	0.020

Based on the results above, the resistance and capacitance of a net will be calculated as the following formulas:

$$R = length.R_{unitlength} \tag{11}$$

$$C = length.C_{unitlength} \tag{12}$$

The tool in the timing analysis stage, especially the synthesis stage will calculate the capacitance and resistance of a interconnect net based on the particular *length* value.

5 Conclusion

In this paper, a novel approach to design a Process Design Kit (PDK) Digital for CMOS 180nm process is presented. For the Standard Cell Library design, by using the Ocean language, this work solved the problems about the licensing of the Liberty NCX tool which is unavailable in universities in developing countries, especially Vietnam, and running manually many simulations when the tool is not applied. Regarding the Wire-Load Model, this paper proposed a method and used accuracy formulas to design a complete model.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment This research is funded by Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, under grant number BK-SDH-2021-2080912.

References

[1] C. THINH Dang, T. TOI Le, H. TRANG, "Technology Education Challenges and Solution to Design a Process Design Kit for Digital CMOS Technology in Vietnam," in 2019 19th International Symposium on Communications and Information Technologies (ISCIT), 381–385, 2019, doi: 10.1109/ISCIT.2019.8905155.

[2] X. He, X. Zhu, L. Duan, Y. Sun, C. Ma, "A 14-mW PLL-Less Receiver in 0.18- μm CMOS for Chinese Electronic Toll Collection Standard," IEEE Transactions on Circuits and Systems II: Express Briefs, **61**(10), 763–767, 2014, doi:10.1109/TCSII.2014.2345303.

[3] J. Sun, C. C. Boon, X. Zhu, X. Yi, K. Devrishi, F. Meng, "A Low-Power Low-Phase-Noise VCO With Self-Adjusted Active Resistor," IEEE Microwave and Wireless Components Letters, **26**(3), 201–203, 2016, doi: 10.1109/LMWC.2016.2521167.

[4] H. Liu, C. C. Boon, X. He, X. Zhu, X. Yi, L. Kong, M. C. Heimlich, "A Wideband Analog-Controlled Variable-Gain Amplifier With dB-Linear Characteristic for High-Frequency Applications," IEEE Transactions on Microwave Theory and Techniques, **64**(2), 533–540, 2016, doi:10.1109/TMTT.2015.2513403.

[5] H. Liu, X. Zhu, C. Boon, X. He, "Cell-Based Variable-Gain Amplifiers With Accurate dB-Linear Characteristic in 0.18 μm CMOS Technology," Solid-State Circuits, IEEE Journal of, **50**, 586–596, 2015, doi:10.1109/JSSC.2014.2368132.

[6] X. Tong, Y. Yang, Y. Zhong, X. Zhu, J. Lin, E. Dutkiewicz, "Design of an On-Chip Highly Sensitive Misalignment Sensor in Silicon Technology," IEEE Sensors Journal, **17**(5), 1211–1212, 2017, doi:10.1109/JSEN.2016.2638438.

[7] S. Chakraborty, Y. Yang, X. Zhu, O. Sevimli, Q. Xue, K. Esselle, M. Heimlich, "A Broadside-Coupled Meander-Line Resonator in 0.13- μm SiGe Technology for Millimeter-Wave Application," IEEE Electron Device Letters, **37**(3), 329–332, 2016, doi:10.1109/LED.2016.2520960.

[8] M. A. Capilayan, R. Minguez, J. A. Hora, "Analysis on HSpice performance trade-off versus simulation time," in 2015 International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), 1–6, 2015, doi: 10.1109/HNICEM.2015.7393176.

[9] A. Sadhu, P. Bhattacharjee, "Methodology of Standard Cell Library Design in .LIB Format," Journal of VLSI Design Tool & Technology, 30–38, 2014.

[10] J. Hu, J. Wang, "Low power design of a full adder standard cell," in 2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), 1–4, 2011, doi:10.1109/MWSCAS.2011.6026457.

[11] S. Srinivasan, M. Hogan, "Physics to Tapeout: The Challenge of Scaling Reliability Verification," in 2019 IEEE International Reliability Physics Symposium (IRPS), 1–5, 2019, doi:10.1109/IRPS.2019.8720440.

[12] Y. Kuo, L. J. Arana, L. Seva, C. Marchese, L. Tozzi, "Educational design kit for synopsys tools with a set of characterized standard cell library," in 2018 IEEE 9th Latin American Symposium on Circuits Systems (LASCAS), 1–4, 2018, doi:10.1109/LASCAS.2018.8399907.

[13] M. S. I. bin Hussin, Y. W. Lim, N. A. Kamsani, S. J. Hashim, F. Z. Rokhani, "Development of automated standard cell library characterization (ASCLIC) for nanometer system-on-chip design," in 2017 IEEE 15th Student Conference on Research and Development (SCOReD), 93–97, 2017, doi:10.1109/SCORED.2017.8305414.

[14] S. Jain, M. Alioto, "A closed-form energy model for VLSI circuits under wide voltage scaling," in 2016 IEEE International Conference on Electronics, Circuits and Systems (ICECS), 548–551, 2016, doi:10.1109/ICECS.2016.7841260.

[15] C. K. Jha, S. N. Ved, I. Anand, J. Mekie, "Energy and Error Analysis Framework for Approximate Computing in Mobile Applications," IEEE Transactions on Circuits and Systems II: Express Briefs, **67**(2), 385–389, 2020, doi: 10.1109/TCSII.2019.2910137.

[16] M. Li, C.-I. Ieong, M.-K. Law, P.-I. Mak, M. Vai, R. Martins, "Subthreshold standard cell library design for ultra-low power biomedical applications," Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, **2013**, 1454–1457, 2013, doi:10.1109/EMBC.2013.6609785.

[17] M. U. Khan, Y. Xing, Y. Ye, W. Bogaerts, "Photonic Integrated Circuit Design in a Foundry+Fabless Ecosystem," IEEE Journal of Selected Topics in Quantum Electronics, **25**(5), 1–14, 2019, doi:10.1109/JSTQE.2019.2918949.

[18] M. Lee, S. Cho, N. Lee, J. Kim, "Design for High Reliability of CMOS IC With Tolerance on Total Ionizing Dose Effect," IEEE Transactions on Device and Materials Reliability, **20**(2), 459–467, 2020, doi:10.1109/TDMR.2020.2994390.

[19] B. Contreras, G. Ducoudray, R. Palomera, C. Bernal, "Automated Parameter Extraction and SPICE Model Modification For Gate Enclosed MOSFETs Simulation," in 2019 16th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), 189–192, 2019, doi:10.1109/SMACD.2019.8795289.

[20] J. Bhasker, R. Chadha, Static timing analysis for nanometer designs: A practical approach, 2009, doi:10.1007/978-0-387-93820-2.

[21] T. J. Chang, Z. Yao, B. P. Rand, D. Wentzlaff, "Organic-Flow: An Open-Source Organic Standard Cell Library and Process Development Kit," in 2020 Design, Automation Test in Europe Conference Exhibition (DATE), 49–54, 2020, doi:10.23919/DATE48585.2020.9116540.

[22] B. Gupta, W. R. Davis, "Characterization of Fast, Accurate Leakage Power Models for IEEE P2416," in 20th International Symposium on Quality Electronic Design (ISQED), 39–44, 2019, doi:10.1109/ISQED.2019.8697565.

[23] M. Olivieri, A. Mastrandrea, "Logic Drivers: A Propagation Delay Modeling Paradigm for Statistical Simulation of Standard Cell Designs," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, **22**(6), 1429–1440, 2014, doi:10.1109/TVLSI.2013.2269838.

[24] K. Vaidyanathan, D. H. Morris, U. E. Avci, H. Liu, T. Karnik, H. Wang, I. A. Young, "Improving Energy Efficiency of Low-Voltage Logic by Technology-Driven Design," IEEE Journal on Exploratory Solid-State Computational Devices and Circuits, **4**(1), 10–18, 2018, doi:10.1109/JXCDC.2018.2812242.

[25] N. Karagiorgos, S. Nikolaidis, "Application of the power contributors method for the AOI22 CMOS cell," in 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), 1–4, 2019, doi:10.1109/MOCAST.2019.8742028.

[26] M. R. F. Abingosa, C. Receno, J. Imperial, J. A. Hora, "Interconnect modeling of global metals for 40nm node," in 2015 International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), 1–6, 2015, doi:10.1109/HNICEM.2015.7393172.

[27] T. Sakurai, K. Tamaru, "Simple formulas for two- and three-dimensional capacitances," IEEE Transactions on Electron Devices, **30**(2), 183–185, 1983, doi:10.1109/T-ED.1983.21093.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1237-1243 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Creativity in Prototypes Design and Sustainability - The case of Social Organizations

Clara Silveira[*,1], Leonilde Reis[2], Vitor Santos[3], Henrique S. Mamede[4]

[1]*School of Technology and Management, Polytechnic Institute of Guarda, Guarda, 6300-559, Portugal*

[2]*School of Business Administration, Polytechnic Institute of Setúbal, Setúbal, 2914-503, Portugal*

[3]*NOVA Information Management School (NOVA IMS), Lisboa, 1070-312, Portugal*

[4]*INESC TEC, Universidade Aberta, Lisboa, 1269-001, Portugal*

ARTICLE INFO

Article history:
Received: 07 October, 2020
Accepted: 04 December, 2020
Online: 16 December, 2020

Keywords:
Creativity
Design
Sustainability
Social Organization

ABSTRACT

The role of creativity techniques in the design of prototypes is of particular interest given its potential for innovation. At same time, despite the efforts of decades in terms of policies and programs of action, humanity has not yet come close to global sustainability. Sustainability design must involve society and creatively employ all available knowledge sources for creating sustainable software. This paper proposes a prototype design approach rooted in employing creativity techniques, while being guided by the dimensions and principles of the Karlskrona Manifesto. This approach is applied to the development of a multidisciplinary aggregator for the optimization of social services. As a result. guidelines for the use of creativity in requirements engineering will be presented, as well as on how to include sustainability issues, namely the Sustainable Development Goals and the five dimensions of sustainability in the design of prototypes.

1. Introduction

This work is an extended version of the paper [1] originally presented in 15th Iberian Conference on Information Systems and Technologies (CISTI 2020).

The software industry has become extremely competitive. Often, a choice is made between multiple software systems that strive to meet the needs of users in the same application domain. Therefore, in order to be sustainable and grow in the market, a software product needs to differentiate itself from other similar products and captivate users with new and useful features. Given such context, the use of creativity techniques is a determining factor for those who develop software. In Grady Booch's opinion [2], the nature of the systems we build continues to change and as they cross collectively in our lives, we must take into account not only the technical elements of software development, but also human needs. The use of creativity techniques allows requirements engineers to deal with new problems and domains, unique contexts, new applications and a combination of existing methods and techniques [3], [4]. However, research on how the application of creativity occurs has been neglected. Various creativity techniques, such as Brainstorming, scenario generation and simulations, and ad hoc design can be used to support and promote creativity in the requirements engineering process [3].

The Karlskrona Manifesto [5] and other authors [6]-[11] gave visibility and increased awareness of the importance of sustainability in software and Information Systems (IS). Specifically, the Karlskrona Manifesto for Sustainability Design (KMSD) serves as a guide for designing and developing sustainable software systems. Thus, the study of sustainability in the area of software development and IS is of extreme importance in promoting good sustainability practices and the integration of the 17 Sustainable Development Goals (SDGs) defined by the United Nations [12].

This study aims to describe the Creative process and techniques to be used in the design of integrative prototypes for a Social Organization, taking into account sustainability issues and the SDGs.

The paper is organized into five sections. After the introduction, the fundamentals and motivation for using creative techniques in the design of software prototypes are presented, as well as the principles of the Karlskrona Manifesto. Section three

[*]Corresponding Author: Clara Silveira, Polytechnic of Guarda, mclara@ipg.pt

describes the proposed approach. Next, section four presents a case of applying the approach to a social organisation as well as the proposal for a prototype in relation to the specificity of the problem. Finally, in the last section, the conclusions are presented, as well as the perspectives of future works, showing the spaces still open in the field of work development.

2. Creativity and Sustainability Design

2.1. Creativity in Requirements Engineering

The field of software engineering, especially requirements engineering, is extremely creative. However, professionals and researchers seem to have different perceptions about creativity, making knowledge transfer difficult [13]. The types of knowledge that we find in the literature that impact creativity from the perspective of software engineers are classified as [14]: previous knowledge, analogies and dynamic knowledge.

Prior knowledge ensures useful information about previous projects for the early stages of ideas generation. This background information provides a starting point or motivation to start with ideas for solving the problem. Analogies help software engineers create a link between two objects in the same or another domain and create new ideas. Dynamic knowledge refers to the knowledge embedded in the tools and available at any time.

We also emphasize that knowledge itself is creative [15] and, therefore, should be encouraged to develop. Results from various experiments show that creativity during the design phase can be improved with the help of analogies [16]. Robert Sternberg [17] defines creativity as the ability to produce work that is simultaneously new (i.e., original and unexpected) and appropriate (i.e., useful and adaptable).

It is also considered that the design of the prototype must contemplate aspects of creativity, including also the concern with the level of comprehensiveness of the information in face of the organizational reality. One concern focuses on the excessive coverage underlying duplication of information [18]. Such a situation can occur when there are duplicates in the individual sources.

In this sense, prototypes must be created to stimulate the creativity of the actors. It may be useful to think about the possible causes of prototype coverage [18], when compared to the target population. Therefore, it is considered essential that a prototype is developed based on research, knowledge, analogies, creative techniques, sustainability factors and, ultimately, validated to answer if it works as specified [19].

Thus, given the specificity of the organization used in this study (of a social nature), it is considered that the purpose of a prototype [19], is to create a semi-realistic representation of something in order to be able to interact and be tested with end users. Its main function is to identify problems and opportunities for improvement long before implementation begins. As it is an agile and efficient method, prototyping will help to define and include more assertiveness in problem solving.

Within the scope of the works related to the thematic area and considering the objective of the paper, practical works developed in organizational context are cited, ([1], [3], [4]).

2.2. Sustainability Design for Software Development

The concept of sustainability is often extrapolated to software engineering with the KMSD [5]. The KMSD principles cover several aspects of sustainability, serving as a point of reference and guide during the software design process [20].

The KMSD manifesto appeared at the Third International Workshop on Requirements Engineering for Sustainable Systems (RE4SuSy), held in Karlskrona, Sweden. It is considered a guide to design and develop more sustainable software systems. The KMSD includes nine principles ([5], [21]):

- **P1.Sustainability is systemic**: Sustainability is never an isolated feature. Systematic thinking should be the main focus to include interdisciplinarity.
- **P2.Sustainability has multiple dimensions**: Advocates the inclusion of the five dimensions (Table 1) in our analysis to be able to understand the nature of sustainability in any situation.
- **P3.Sustainability transcends multiple disciplines**: Sustainability exceeds several disciplines, working on challenges from several perspectives.
- **P4.Sustainability is a concern independent of the purpose of the system**: Regardless of the system's development context, sustainability must always be included.
- **P5.Sustainability applies to both a system and its wider contexts**: Concerned with considering at least two spheres in the design of the system: the sustainability of the system itself and how it may affect the sustainability of the context of the entire system.
- **P6.System visibility is a necessary precondition and enabler for sustainability design**: Concerned with sharing the vision of the system and its context in different perspectives and with the various levels of abstraction, allowing an informed choice. In this principle, it is important to cite the application of the *SharedClearVision* pattern [22].
- **P7.Sustainability requires action on multiple levels**: Consists of researching measures for sustainability that are considered the most effective way to intervene in comparison with other alternative actions.
- **P8.It is possible to meet the needs of future generations without sacrificing the prosperity of the current generation**: Sustainability innovation can play a key role in defining present and future needs. Thus, it is possible to make options that benefit both the present and the future.
- **P9.Sustainability requires long-term thinking**: Assess, in the long term, the possible impacts, costs and benefits of decisions made.

The manifesto presents general and abstract principles as well as values of sustainability. The KMSD Manifesto also includes the dimensions (five) shown in Table 1.

Table 1: Karlskrona Manifesto - Sustainability Dimensions [5], [21]

Dimension	Focus
Human	It focuses on the well-being of the individual.
Social	Refers to social issues like groups of people and organizations in a community. Social

	responsibility, justice, job, and democracy are essential elements.
Economic	Economic sustainability aims at maintaining capital and added value.
Environ-mental	The main mission of environmental sustainability is to improve the well-being of the person, protecting natural resources and ecosystems.
Technical	Concerned with the longevity of information, systems, their evolution, maintenance and innovation.

The KMSD principles and commitments provide an overview of the various dimensions of sustainability and their relationships. In this way, we consider that they show the impact that concerns with sustainability can have in the design of an information system. We added a new vector, combining creativity with the various dimensions of sustainability, allowing to preserve the environment and achieve new technological innovations in social organizations.

3. Creative Approach in Prototype Design

In this section, we describe the creative approach used in prototyping within the framework of a more comprehensive methodology for requirements engineering. Among the various creativity techniques that can be used to support and promote creativity in the design of prototypes, the following were selected by the creativity expert [2], namely: Brainstorming, Brutethinking, Reversal and Ideabox. The selection of these techniques took into account the author's accumulated experience in conducting workshops.

Brainstorming is an idea generation technique, which consists of a meeting designed to encourage the full release of mental activity without restrictions. This technique works so well because, among many other reasons, "ideas call ideas". Other people's ideas are sometimes starting points for our best ideas [3].

The Brutethinking technique [3] is characterized by being a simple process, which develops in four stages: choosing a random word; choose things/elements associated with the randomly chosen word; force a link between the word and the problem and between associations and the problem; relate the ideas obtained and analyze them.

The Reversal technique originates from the transformations that generate ideas. In some cases it is better to think first about the negative and then invert the negative, indicate "the worst" and then revert to "the best" [3]. One of the ways is double inversion: initially the reversal identifies the ways to make the situation worse by realizing what can aggravate the problem, instead of improving it. Then, you make an inversion again, identifying the paths that can lead to the improvement of the situation.

IdeaBox is a tool that allows you to combine the parameters (characteristics, factors, variables or aspects) of a challenge (it can be a prototype) into new ideas.

In the creative approach, it is proposed to integrate multiple existing paradigms, in which the creative contribution results from the integration and combination of different approaches that can interrelate, in the search for the integration, in a single proposal, of different forms of thought about a phenomenon [23]. On the other hand, still in the context of the creative process, it is essential to mention the importance of the characteristics of the creative product, the different views on the models of creative processes, the role of the field of application, the personal characteristics found in highly creative people and, finally, the socio-organizational aspects involved [24].

And we have to note that [23] the multiple views of creativity are complementary and not mutually exclusive. This suggests that future models of creativity can integrate different sources and diverse individual perspectives. The creative approach proposed here is included in a more comprehensive methodology [3] and focuses on the step "Understanding user requirements and sustainability factors", which aims to apply Brainstorming, Brutethinking and reversal creative techniques to: identify the requirements of the user in a process of seeking a creative and sustainable solution; interaction with stakeholders to identify their needs and find out what needs to be built.

In this approach, existing applications and prototyping are used as a source for collecting and validating requirements. Sustainability principles and commitments are incorporated [5]. The preliminary list of requirements is drawn up, as well as the design of exploratory prototypes. In the analysis of prototypes it is proposed to use the Ideabox creativity technique.

4. Prototype Design in Social Organization

The Parish Social Center of the Parish of São Sebastião is a Social organization, founded in 1998 that develops strong intervention in the São Domingos Community (SDC). In his daily intervention integrates food distribution programs, monthly food aid (food products given by individuals and companies), distribution of clothing and hygiene products [25].

4.1. SDC Characterization

The organization under study, previously presented, has been the subject of work in the thematic area, contributing to the optimization of established practices [7]. In this sense, a brief characterization of the SDC (Figure 1) is presented to frame the problem.



Figure 1: SDC Characterization [7]

From the meetings between the Project Manager and the SDC Manager and also with the participation of the Software Architect, the organization's current situation was made. The development of the project also had the online collaboration of the Requirements Engineer element.

Figure 1 shows the various aspects that the prototype must include in Housing, Skills, Health, Training, Volunteering for the objective of Social Integration of the Citizen.

In view of the problem under study and the specificity of the social organization under analysis, the team used the Brutethinking technique, in three iterations, to identify the problem, the main objective, the primary causes, the actors of the system and their objectives, the functional and non-functional requirements, creative and non-creative requirements. The Reversal technique was used to prioritize requirements. The team also used the techniques of debate and Brainstorming, in order to evaluate and review artifacts and solutions found. In order to aggregate the information underlying the organization's activity, Figure 1 was conceived in order to highlight the themes involved in a transversal logic of support to the citizen inserted in a community, which, by its nature, has the collaboration of several official entities.

The development of the prototype for SDC was based on a multiple perspectives in order to describe each individual in the various components and taking into account the collaboration of several partners.The valence of volunteering has greater prominence (Figure 2) given the importance of collaboration / interaction of people in the community.



Figure 2: Volunteering characterization

Next, the other objectives and components of Figure 1 are characterized [26]:

- Social Center - support to activities –
  - Objective: aggregating skills and knowledge, so that the Social Center can offer high quality and integrated services. The objective is to provide a platform to the various actors, enabling them to access certain SDC data, depending on their access profile;
  - Characterization: the platform will allow: managing citizen profile, volunteer members, health professionals; nursing treatments; medication, physical activities; sharing goods and resources; share vehicle and register availability on the platform for volunteering.
- Official Entities – Partnerships
  - Objective: to list the entities with which they will establish partnerships, namely: parish council, health center, Public Security Police, Foreign Service and Borders. Collaboration with official entities will enable more effective responses to SDC people;
  - Characterization: the entity and the manager are identified, the mission, location, preferred contacts, partnership/protocol, association with other stakeholders are described.
- Experience - knowledge kills -
  - Objective: an SDC member's life experience (or that of their caregiver), can be valuable information to enhance cooperation with the SDC. Each member's level of literacy is also important so that their skills can be developed, culminating in the sharing of knowledge;
  - Characterization: the things a person did throughout their life; the things they like to do and would like to do; the activities they would like to participate in: these can be assessed from one's life experience.
- Resource/car sharing
  - Objective: to allow the sharing/donation of consumer goods or resources (equipment) and car sharing. Reducing carbon emissions and fuel consumption is promoted;
  - Characterization: application with resources and goods and availability; places and times to car sharing;
- Health and well-being
  - Objective: the aim is to thoroughly characterize the health status of community members, to help with specific and necessary care;
  - Characterization: in the various health valences, the importance of aggregating by areas is emphasized: health situation, rehabilitation/physiotherapy, motricity/biomechanics, medical and nursing care, taking medication.

The detailed characterization of the information previously exposed allows to present a multidisciplinary and aggregating vision in order to enhance the provision of services by the Social Center, corresponding to the expectations of its users.

4.2. Prototypes

The Social Center, as an entity that aggregates all information and considering the specificity of its mission and its social context, presents needs for collaboration with various entities.

The exploratory prototypes that were presented were conceived using the creative approach, namely the following techniques: Brainstorming, Brutethinking, Reversal and Ideabox. The following principles of the KMSD (described in section 2.2) have also been incorporated: P1; P2; P3; P4; P7 and P9.

Figure 3 shows the exploratory prototype for the main aspects of the Social Center. The prototype multidisciplinary enhances the aggregation of information. The project aims to define an access profile considering data protection concerns within the National Data Protection Commission, an entity that regulates citizens' access to and data protection in Portugal. Figure 3 presents as one of its components Partnerships representing the various entities

involved in the process, namely the Foreigners and Borders Service in order to support the organization of the process of inhabitants residing in the SDC district. The partners' name is also relevant in this context is social security since there is a regulatory and legal framework in Portugal that enhances the analysis of the economic and professional situation of citizens and families in order to enhance the creation of subsidies, namely the Social Income of Insertion. In this context, the prototype is of particular interest since it aggregates the information in various valences for each citizen and his family.

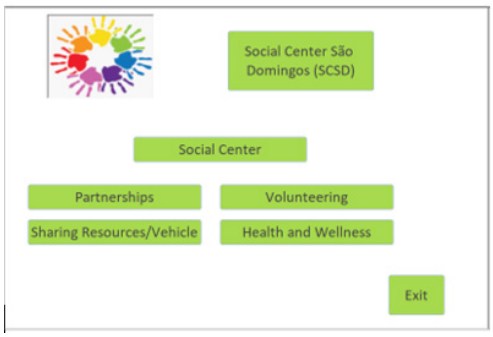


Figure 3: Main menu prototype

The Social Center São Domingos includes a low-level, low-economic community, mainly made up of elderly people with low economic levels, as they are retired people from the manufacturing environment. However, SDC also includes families with adolescents and children with economic needs. Thus, the creation of a prototype of this nature can enhance the centralization of community information in order to allow the delineation of support strategies. Figures 4 present prototypes for Citizen Profile.



Figure 4: Prototype Citizen Profile

The Citizen Profile characterization aims to collect and keep up to date a set of information referring to each citizen and his family in order to aggregate the various valences. Thus, it is intended a comprehensive characterization with regard to personal data in order to support the various sharing of information in order to create better living conditions. The medication and Health Problems includes the valences of improving the health care of each person, but also knowing the medication necessary to each person in the sense that in case of need the Social Center São Domingos provide its acquisition. Another of the concerns of the Social Center São Domingos is the mobility and conviviality of the elderly in which the Life Experience is an appeal to the sharing of knowledge promoting initiatives that allow to enhance the enormous knowledge of the elderly with other elderly and younger age groups through the organization of workshops and other events. Another aspect of the project focuses on recording in detail the Academic and Professional Training in which the practical experience of each one is emphasized by fostering this sharing through events. This need and also willingness to collaborate and explore the accumulated knowledge has allowed the create of availability for volunteering since it allows the taking of older people from home, enhances collaboration among community members increases the feeling of being useful and actively contributing in their community. Sometimes this collaboration is carried out in exchange for perishable and non-perishable goods increasing also of this were the resources of the family.

The regard to Volunteering - Actions to be promoted - it includes the concerns of involving SDC citizens in concrete voluntary actions. Particularly about the elderly, the objective is to enhance the coexistence and sharing of knowledge. The aspect willingness to receive volunteer support is carried out in actions such as: small reparations in housing, receiving non-perishable goods or other needs that are identified. In view of the sustainability concerns that are intended to be fostered, it is considered that in this context and in view of the specificity of the SDC Sharing – Resources/Vehicles has a view to promoting sustainability by appealing to the implementation of SDGs practices, in particular with regard to the re-use of resources and optimization of car sharing. Regarding Activities - Health and Well-being is intended to promote mobility and active life, as well as to have assertive monitoring particularly with regard to the senior population.

4.3. Discussion

The characterization of the SDC is of particular interest in the sense of having the opportunity, within the scope of a partnership, to study a community in order to test the applicability of the creative approach, contributing to the design of the prototype to meet the needs of the organization, although innovative and appropriate.

The prototype underlies the feedback from the team that includes the SDC manager. The creativity techniques Brutethinking, Reversal and Brainstorming were considered adequate for the design. The application of creativity techniques and the incorporation of concerns about sustainability allowed the inclusion of new services: car sharing; resource sharing; take medicine; volunteering; prototype integrating several valences and entities. This corroborates with the authors, in which the involvement and decision of the interested parties in the software project is an integral part of the constructions of software for effective use.

It is considered that the proposed prototype is an aggregator of information based on a multidisciplinary approach allowing the collaboration of several entities taking into account the specificity of intervention of each one so that the Social Center provides a more complete and integrating service.

In terms of sustainability concerns, it is noteworthy, for example, that volunteer actions allow contributing to SDGs as shown in Figure 5.



Figure 5: SDGs in the context of volunteering (adapted from [12])

The main contributions to the SDGs are highlighted [12]:

- SDG 4: The prototype under development aims to enhance the participation of young people from the Community of São Domingos to the extent that by getting involved in the various activities they acquire relational and technical skills in the context of their collaboration, even promoting their participation as entrepreneurs;
- SDG 5: The aim of this area is to promote women's empowerment by advocating a set of initiatives. A set of strategies have already been delineated, including short courses on the use of ICT and a financial education course in order to raise women's awareness of the importance of assertive management suggesting in this context the use of ICT to control expenditure;
- SDG 10: The Community of São Domingos has as its strategy inclusive policies given the multicultural provenance of its inhabitants. In this sense, it organizes social events serving the various cultures and religions existing in the neighbourhood.
- SDG 11: The Community collaborate in promoting efforts to protect the world's cultural and natural heritage organizes events;
- SDG 12: Aiming to promote sustainability policies, the Community of São Domingos promotes the distribution of non-perishable goods: clothing, specific health equipment such as wheelchairs and articulated beds and furniture. The Community not only aims to contribute to the well-being of its inhabitants, but also to promote the re-use and recycling of goods;
- SGD 14: Given its geographical location, the Community of São Domingos organizes awareness-raising actions about pollution mainly regarding the Sado River.

It should be noted that the SDGs cannot be analyzed in isolation, since the improvement of one SDGs will have a positive and/or negative impact(s) on another SDGs [27]. Integration adds valuable knowledge that can help society meet the SDGs [28].

5. Conclusions

This work presents a proposed approach to the design of prototypes using creativity techniques. It includes a study of a social nature, which contains a set of assumptions that constitute a challenge. Throughout the development of the work, conditions were created for the application and validation of techniques, namely in the field of creativity in IS. In this sense, Brainstorming, Brutethinking, Reversal and Ideabox techniques were applied. Another technique applied in the work was collaboration, where software engineers shared their ideas and thoughts with others. They build on each other's ideas and provide constructive feedback. In this way, it was possible to fill the knowledge gap of each one, creating a group. On the other hand, the principles of the Karslkrona Manifesto have already been applied in several areas [11], [27]-[31].

In this research was possible to apply the theoretical references mentioned to study a real problem, proposing an integrative and multidisciplinary solution in order to enhance the activity of the Social Center in the São Domingos Community in order to provide better services to its community.

As discussed in section 4.3, the results obtained show that creativity techniques can play an important role in transforming of existing Information Systems, bringing innovative and unexpected solutions, which without removing functionality or performance from Information Systems can make them more sustainable.

Therefore, in this context, the understanding that there is an important role to be played by creativity techniques in promoting SDGs and contributing to the development of more sustainable IS is reinforced.

The process was applied in another context [26] and a significant contribution was perceived in the field of sustainability, in the areas: social, individual, economic, environmental and technical [7].

As future work prospects, the case of the chosen Social Organization will be replicated in other communities/institutions in order to contribute to their sustainability. Also, the development of an application to make available to the Social Center is recommended, underlying concerns of a graphic nature, screen design, usability and accessibility. In this sense, solutions are being studied that address the specificities and constraints, particularly financial constraints implicit to social organizations. Thus, the choice of the platform for the development and hosting of information, licenses and maintenance are concerns about the importance of keeping the solution in full operation over time. For the validation of the application we´ll use inspection techniques, usability tests and accessibility.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] C. Silveira, L. Reis, V. Santos, H. Mamede, "Criatividade no Design de Protótipos - O caso das Organizações Sociais," in CISTI'2020 - 15th Iberian Conference on Information Systems and Technologies, Sevilha, Espanha, 2020, doi:10.23919/CISTI49556.2020.9140870.

[2] G. Booch, "The future of software engineering," in Proc. 37th International Conference on Software Engineering (ICSE 15), 2015.

[3] V. Santos, Criatividade em Sistemas de Informação, Lisboa: FCA, 2018.

[4] V. Santos, L. Amaral, H. Mamede, R. Gonçalves, "Creativity in the Information Systems Planning Process," in Handbook of Research on

Innovations in Information Retrieval, Analysis, and Management, IGI-Global, 2015.

[5] C. Becker, R. Chitchyan, L. Duboc, S. Easterbrook, B. Penzenstadler, N. Seyff, C. Venters, "Sustainability Design and Software: The Karlskrona Manifesto," in Proc. 37th International Conference on Software Engineering (ICSE 15), 2015.

[6] Software Sustainability Institute, "Manifesto," 2019. [Online]. Available: www.software.ac.uk/about/manifesto.

[7] L. Reis, C. Silveira, "Sustentabilidade Multidimensional em Sistemas de Informação," in XXX Jornadas Luso-Espanholas de Gestão Científica, Bragança, 2020.

[8] L. Castro, V. Santos, "A perceptional study on the role of development agents and global leaders in leveraging technology for sustainable development," in Multi Conference on Computer Science and Information Systems, 169-176, 2019.

[9] J. Meneses, "A sustentabilidade em 2019," 2019. [Online]. Available: www.bcsdportugal.org/noticias/a-sustentabilidade-em-2019. [Accessed 25 outubro 2019].

[10] S. Hettrick, Research Software Sustainability: Report on a Knowledge Exchange Workshop, The Software Sustainability Institute, 2016.

[11] C. Venters, R. Capilla, S. Betz, B. Penzenstadler, T. Crick, S. Crouch, E. Nakagawa, C. Becker, C. Carrillo, "Software sustainability: Research and practice from a software architecture viewpoint" Journal of Systems and Software, **138**, 174-188, 2018. https://doi.org/10.1016/j.jss.2017.12.026

[12] United Nations Development Programme, "Sustainable Development Goals," 2015. [Online]. Available: www.undp.org/content/undp/en/home/sustainable-development-goals.html. [Accessed 02 12 2019].

[13] R. Mohanani, P. Ram, A. Lasisi, P. Ralph, B. Turhan, "Perceptions of Creativity in Software Engineering Research and Practice," in 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 210-217, Vienna, 2017, doi:10.1109/SEAA.2017.21.

[14] R. Hegde, G. Walia, "How to Enhance the Creativity of Software Developers: A Systematic Literature Review," in International Conference on Software Engineering and Knowledge Engineering, SEKE 2014, 2014.

[15] T. Davenport, L. Prusak, Conhecimento Empresarial, Rio: Campus, 1998.

[16] N. Bonnardel, "Creativity in Design Activities: The Role of Analogies in a Constrained Cognitive Environment," Creativity and Cognition conference, pp. 158-165, 1999.

[17] R. J. Sternberg, Handbook of creativity, Cambridge University Press, 1999.

[18] A. Black, "The IDI prototype spine's creation and coverage," Statistics New Zealand Working Paper No 16–03, 2016.

[19] G. Silvestri, "Prototipação e a criação de produtos que resolvem problemas," 18 09 2017. [Online]. Available: https://uxdesign.blog.br/o-objetivo-de-um-protótipo-é-criar-uma-representação-semi-realística-de-algo-que-seja-possível-e0445d85ca93. [Accessed 18 02 2020].

[20] B. Penzenstadler, S. Oyedeji, "Experiences from Applying the Karlskrona Manifesto Principles for Sustainability in Software System Design," Proceedings of the 8th International Workshop on Requirements Engineering for Sustainable Systems, 2020.

[21] B. Ovelheiro, C. Silveira, L. Reis, "Sustainability Design Applied to the Digital Signature of Documents," in In Carvalho, L. C., Reis, L., Prata, A., & Pereira, R. (Ed.), Handbook of Research on Multidisciplinary Approaches to Entrepreneurship, Innovation, and ICTs, IGI Global, 2021, 349-374, doi:10.4018/978-1-7998-4099-2.ch016.

[22] S. Adolph, P. Bramble, Patterns for Effective Use Cases, Addison-WesleyPearson Education, 2003.

[23] R. Stenberg, "The development of creativity as a decision-making process," in Creativity and development V.J.-S. K.., New York, Oxford University Press, 2003, 91-138.

[24] L. Nguyen, G. Shanks, "A framework for understanding creativity in requirements engineering," Information and Software Technology, **51**, 655–662, 2009. https://doi.org/10.1016/j.infsof.2008.09.002

[25] CSPSS, "Centro Social Paroquial de São Sebastião," Documento interno, Setúbal, 2019.

[26] C. Silveira, L. Reis, "Sustainability in Information and Communication Technologies," in Handbook of Research on Multidisciplinary Approaches to Entrepreneurship, Innovation, and ICTs, In Carvalho, L. C., Reis, L., Prata, A., & Pereira, R. (Ed.), IGI Global, 2021, 375-396, doi:10.4018/978-1-7998-4099-2.ch017.

[27] L. Reis, C. Silveira, L. Carvalho, C. Mata, R. Madeira, "Design de Protótipo para Pessoa sem-abrigo - Contributos de Sustentabilidade," in IV International Forum on Management, Universidade Aberta. Lisboa, 2020.

[28] P. Shrivastava, M. Smith, K. O'Brien, L. Zsolnai, "Transforming Sustainability Science to Generate Positive Social and Environmental Change Globally," One Earth, **2**(4), 329-340, 2020. https://doi.org/10.1016/j.oneear.2020.04.010

[29] S. Oyedeji, A. Seffah, B. Penzenstadler, "A Catalogue Supporting Software Sustainability Design," Sustainability, **10**, 2296, 2018. https://doi.org/10.3390/su10072296

[30] B. Penzenstadler, S. Betz, C. Venters, R. Chitchyan, J. Porras, N. Seyff, L. Duboc, C. Becker, "Everything is Interrelated: Teaching Software Engineering for Sustainability," in IEEE/ACM 40th International Conference on Software Engineering: Software Engineering Education and Training, 2018.

[31] P. Lago, "Architecture Design Decision Maps for Software Sustainability," in IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1244-1249 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Computer Vision for Industrial Robot in Planar Bin Picking Application

Le Duc Hanh*, Huynh Buu Tu

Mechanical Department, Ho Chi Minh City University of Technology, Vietnam National University, Ho Chi Minh city, 70000, Vietnam

ARTICLE INFO

Article history:
Received: 05 August, 2020
Accepted: 03 December, 2020
Online: 16 December, 2020

Keywords:
Bin picking
Computer vision
Industrial Robot
3D Point cloud
Watershed

ABSTRACT

This research presents an effectively autonomous method that can save time and increases productivity for an assembly line in industry by using a 6DOF manipulator and computer vision. The objects are flat, aluminum type and randomly stacked in a box. Firstly, 2D color image processing is performed to label the object, then using 3D pose estimating algorithm, the surface normal, angle and position of an object are calculated. To reduce the burden time of 3D pose estimation, a voxel grid filter is implemented to reduce the number of points for the 3D cloud of the objects. As known the 3D image object was obtained by camera in bin often involves both heavy noise and edge distortions, so to prepare for the assembly a manipulator will pick and place it at sub position then a 2D camera is used to estimate the pose of the object correctly. Through implement experiment, the system proved that it is stable and have good precision. Installing time and maintaining is fast and not complicated. It is applicable in production line where mass product is produced. It is also the good foundation for other deep researches.

1. Introduction

For many reasons, such as people, productivity, management and working environment; the robots gradually replace the people to carry out the tasks flexibly, accurately and safely. The tasks such as assembling and disassembling a machine part, arranging large quantities of goods requiring a lot of effort. In addition to exposure to hazardous environments, the risks of machinery such as billet cutting machines, corrugated iron stamping, turning, heavy chemical containers, grippers and leveling of embryos when welding can be dangerous to humans. Therefore, the participant of intelligent robot is indispensable because The robot can recognize the surrounding environment and objects correctly by the sensors as well as very high performance. Currently, there are many researches in the world about automatically supplying embryos in the factory or especially the task of picking up objects occluded each other inside a bin and placing it in the assembly area or welding. This task is called "Bin Picking". Casado and Weiss proposed method to apply algebraic relations between the invariants of a 3D model and those of its 2D image under general projective projection for estimating object 3D poses [1-2]. However, because of the occlusion, pose estimation using 2D image matching only works when their poses are limited to a few cases especially in automatic random bin-picking systems. Ban and Xu et proposed a laser vision system consists of using a projector and a fixed camera to project vertical and horizontal lines, thereby using the camera to determine the X, Y, Z coordinate and direction of objects to pick [3-4]. Chang recommended attaching a laser and projector on the hand of the mobile device to reconstruct the 3D image of the workspace and objects [5]. However, the system [3-5] has limited speed because it requires waiting times for projector operating and also of process complex images. Recently, the work has been done to use the stereo vision system to view the work area and to calculate the position and normal vectors of the gripping objects or using both stereo vision, projector and CAD models to determine the position and coordinates of the object to be picked up in a box [6-9]. However, CAD models of objects in actual production are often not provided. Recently, the structured-light vision system and Point cloud are also recommended as Schraft proposed an algorithm to calculate the position and coordinate of the object based on the structure light system and a sample point cloud [10]. After obtaining 3D working area point cloud the system will compare with the sample point cloud to determine the location of an object. The system has the disadvantage that the model points must be saved at first as a data system so the algorithm will take a lot of time to compare with this sample data point cloud. More over the database will be huged and complicated. Bousmalis [11] and Mahler [12] use the learning method to identify the object to be picked, but the algorithm is

*Corresponding Author: Le Duc Hanh, Email: ldhanh@hcmut.edu.vn

limited because the object is randomly stacked, so the objects can overlap each other [11-12]. When the algorithm performs learning, it will also learn the black overlap region, so the amount of data learned is very large and when performing the pick-up, the robot will have a lot of difficulties due to these black shadows data. Over past decade, deep machine learning has also been applied to object recognition such as Tuan which uses large data collection and neural network to identify and take rectangular objects in containers [13]. Chen re-modeled the 3D object to be picked up, simulating and training the position and gripping position of the object to be picked [14]. However, the technique of [13-14] need to be retrained the database if using other objects and training data is very large.

This research is different from discussed above. A Kinect sensor, in addition to returning the location data of the pixel, also produces the color of each location, so combining the location and color data will improve the algorithm accuracy. Firstly, all the point cloud information consists of pixel location and its RGB image are converted to 2D image for labeled data and then converted again to point cloud data. Secondly, the 3D pose of object is roughly estimated. The data is then transmitted to a Manipulator to perform a grasping task. The object will be placed on the V-groove and the system uses a 2D camera to process the image to refine the object's pose.

2. Segmentation

2.1. 2D image processing

Watershed is an algorithm used in image segmentation [15]. Algorithms are popular in fields such as biomedical and medical processing, computer vision. Firstly, due to the high density of 3D data acquired from Kinect, the location of point cloud inside the registered range are kept by using pass-through filter algorithm. By using filer algorithm, the number of points in cloud reduced from 307200 points to about 30000 points. After filtering, all the point cloud information consisting of pixel location and its RGB image are converted to 2D image. The canny and the distance image transform algorithm are applied to find the contour of the object. After that the watershed algorithm is applied to label object. The labeled objects contained pixel locations and labeled colors are then converted back to 3D point cloud data as shown in Figure 1.

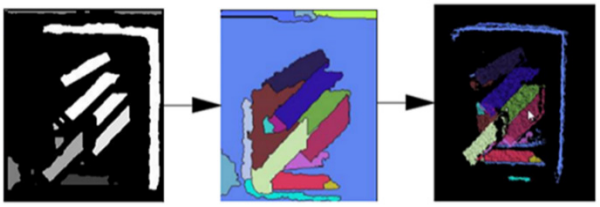


Figure 1: Watershed algorithm and colors assigning to Point cloud.

After objects are labeled, the data remains as unorganized objects data and will be segmented into each planar surface which represents for object cloud called cluster.

2.2. 3D image processing

When using 3D devices, the collected data often appears to be hovering and not in any plane. These points usually appear near the interface between surfaces, or near the edges of an object and. These points need eliminating, so the outlier filter is applied by considering the number of neighbor points K and the radius R of each point. If any point in point cloud whose number of neighbors in sphere of radius R is not greater than a threshold number, that point will be removed from point cloud. After that normal vector and the cluster segmentation are calculated. Normal vector is the most important point feature that distinguishes between different 3D cloud. the normal vector is the normative vector of a plane estimated from a given set of points. The problem of determining the normal to a point on the surface is approximated by the problem of estimating the normal of a plane tangent to the surface. To determine a plane equation in space, it is necessary to determine an x point in that plane and its normal vector $\vec{n}$. The point x is determined by focusing on nearby points.

$$x = \overline{p} = \frac{1}{k}\sum_{i=1}^{k} p_i \tag{1}$$

As discussed in [16], the solution for $\vec{n}$ is given by finding the eigenvalues and eigenvectors of covariance matrix $C \in R^3$ of P^k with:

$$C = \frac{1}{k}\sum_{i=1}^{k} \xi_i (p_i - \overline{p})(p_i - \overline{p})^T \tag{2}$$

The eigenvectors $\vec{v}_j$ form an orthogonal frame, corresponding to the principal components of P_i^k

$$C\vec{v}_j = \lambda_j \vec{v}_j, \quad j \in \{0,1,2\} \tag{3}$$

If $0 < \lambda_0 < \lambda_1 < \lambda_2$ the eigenvector $\vec{v}_0$ corresponding to the smallest eigenvalue λ_0 . Therefore the approximation of $\vec{n}$ is $\vec{n} = \{n_x, n_y, n_z\}$ or $-\vec{n}$. The calculation of the normal vector at the center will have errors, so to increase the reliability for the estimated vector, the maximum allowed error between the normal vector at the center and the normal vector at the adjacent point satisfy the following conditions:

$$\theta_{\overline{xi}} = ar\cos(\vec{n}_x \vec{n}_i) \tag{4}$$

Through conducting experiment, the $\theta_{\overline{xi}} = 5\ ^\circ$ is found to be appropriate.

Although the Kinect was placed perpendicular to the bin, the barrel edge point cloud still existed and connected to the real object in the bin formed an unexpected object in bin. When calculating the normal vector, the program will consider the barrel edge and the object are one object then the normal vector calculation and the position of the real object are not correct. Therefore, based on the normal vector, it is possible to separate the sides of the barrel with the idea that the angle between the two normal vectors on the side and the bottom edge will be larger than threshold value α. From experiment, the angle $\alpha = 50^0$ is appropriate. The results are shown in Figures 2 and 3.

The Euclidean clustering algorithm basically used to segment cluster and based on their neighbors [17]. Therefore, to estimate the normal vector to recognize that they belong to one plane, the Clustering procedure are through the following steps:

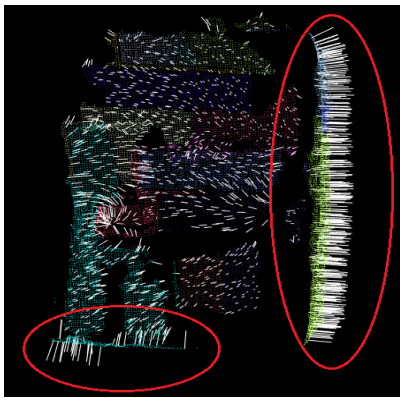


Figure 2: Normal vector of points

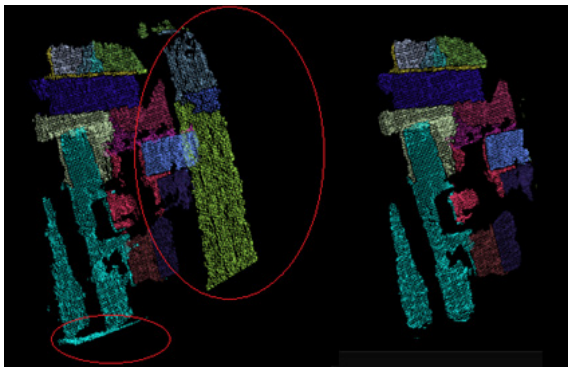


Figure 3: Image before and after removing barrel edge.

- Normal calculation of point cloud.
- Apply Kd-tree algorithm by define radius searching to find nearby points.
- Create list of cluster C and list of points to handle L.
- For every point p_i in P, do the following step:
- Add p_i in to L
- Search the list of neighbor n of p_i within a sphere of radius r using kd-tree algorithm.
- For every neighbor n_i, check if point exists in L. If it exists we skip to the next neighbor, otherwise we apply the following check.
- Calculate the angle between 2 point p_i and its neighbor n_i

$$\theta_{ij} = a\cos(\vec{p}_l, \vec{n}_l) \tag{5}$$

- Determine the color difference of 3 channels R, G, B of 2 points:

$$\left|R_p - R_n\right| + \left|G_p - G_n\right| + \left|B_p - B_n\right| < \text{Threshold} \tag{6}$$

- if the point p_i and its neighbor n_i satisfies (5) and (6) then add n_i to L, otherwise skip to the next neighbor.
- After checking all the points in L, put L into C and reset L.
- If the number of points in the cluster are in the range (Min_clust, Max_clust), it will save the cluster. Among the cluster that meet the conditions of color, normal, size, the cluster will be selected if it has the highest height as shown in Figure 4.

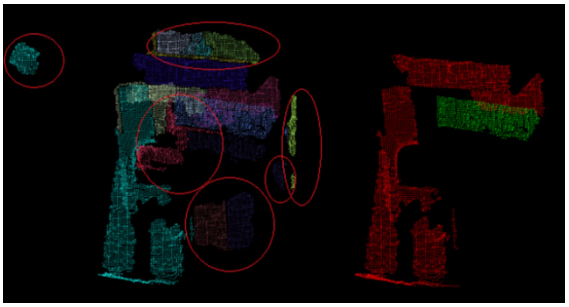


Figure 4: Image before and after extract cluster.

In Figure 4, the area circled to the left is the cluster having a smaller number of points than the Min_clust. Therefore, they were removed. Finally, the result is shown on the right with the green cluster which has the highest height. The robot will use this cluster to calculate the position and direction.

3. Manipulator and camera calibration

Because it is hard to determine the position of the camera coordinate that corresponds to the coordinate of the manipulator; therefore, using the method of measuring the distance from the manipulator to the camera will not give accurate results. Therefore, this research uses intermediate positions with known coordinates in both camera coordinates and manipulator to find transfer matrices. Suppose we have any point M in 3D real space, the coordinates of this point M in the camera coordinate system is denoted M_C and in the robot coordinate system is M_R, then we have the formula to transfer coordinates from the camera coordinate system to the robot coordinate system as follows:

$$M_R = RM_C + T \tag{7}$$

Where R is a rotation matrix and T is the translation matrix containing the translational vector from the robot origin to camera origin. If $M_R = \begin{bmatrix} X & Y & Z \end{bmatrix}^T$ and $M_C = \begin{bmatrix} x & y & z \end{bmatrix}^T$ then (7) becomes:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T$$

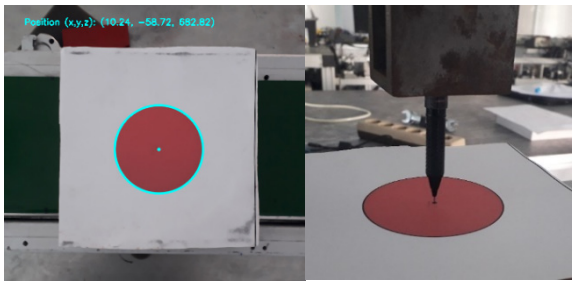
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \tag{8}$$

From (8), there are 12 variables to find, so it is necessary to have 12 equations corresponding to 12 variables, each point coordinates can only provide 3 equations so we need to take 4 points to be able to find two matrices R and T. After getting the necessary points, the system is solved as linear equation

$$A = BX \Leftrightarrow X = B\text{-}1A \tag{9}$$

With Kinect cameras, both 2D and depth images have a resolution of 640x480 and each pixel in 2D images corresponds to a point obtained from the depth camera. Kinect's true coordinate

system is defined to coincide with the color camera's coordinate system, so when the pixel coordinates of a point in an RGB image are obtained, the x, y, and z values of that point can be retrieved in the camera coordinates. So, the method to get coordinates is as follows: Use the Hough Circle algorithm to determine the center of a circle. From the pixel coordinates of the center of the circle in the color image, the center of the circle in the camera coordinate system is retrieved as shown in Figure 5a. Then, use the Teach-pendant to control the manipulator's head to the center of the circle to get the coordinates of this point in the robotic coordinate system. The work head needs to be small enough to increase accuracy Figure 5b. The data table of this process and error checking are shown in Table 1 and Table 2



(a) (b)

Figure 5: (a) Coordinate from camera, (b) coordinate from robot

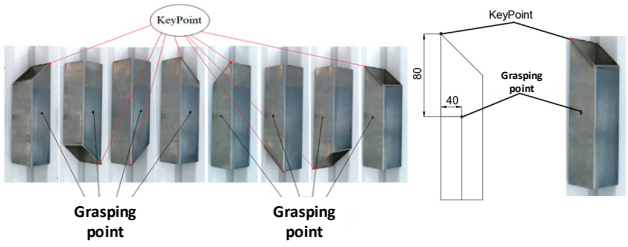
Table 1: Coordinates from camera and robot

Point	Camera Coordinate (mm)	Robot Coordinate (mm)
1	-11.238; -95.1482; 798	298.218; -339.686; 0.0484
2	64.4799; -101.508; 680	289.673;-419.271; 120.996
3	-35.7193; -88.948; 746	307.759 -314.414 50.322
4	22.9576; 28.523; 741	427.611; -374.943; 51.3479

Table 2: Calibration error checking table

Camera coordinate (mm)	Robot coordinate (mm)	Absolute error (mm)		
		X	Y	Z
48.2097; 15.5754; 790	414.531; -399.18; 1.72738	0.08498	1.2113	1.3723
-5.20496; 33.4604; 792	432.75; -344.522; 1.66852	0.7737	0.7121	2.4602
-29.8731; -62.924; 677	330.144; -319.86; 119.984	1.9182	2.0250	0.3202
14.6488; -16.0439; 743	378.403;-367.813; 51.472	2.4365	1.1975	0.2991
Mean error (mm)		1.3033	1.2865	1.1130

After identifying separated object, object is grasped and placed on the V-groove and there are eight case of placing as shown in Figure 6a. Since it is placed on a notched groove, the position of the object just changes along the axis of the groove. To determine the position O for grasping, it is only necessary to determine the coordinates of one point called the key point on the object as shown in Figure 6b and then using the dimension of the object to calculate the O position.



(a) (b)

Figure 6: (a) eight cases placing on V-groove, (b) Key point and grasping point.

To determine the key point, Firstly, the template matching algorithm are performed then the Hough Line detection applied to find the line equation. Then based on line equation, the intersection point coordinate of each two lines are calculated. By combining each case result of template matching and calculated points, the appropriate coordinate of the key point is determined.

4. Experiment

Working space for experiment includes one RGB-D camera fixed above the working space, a monocular camera 2D attached at the end effector of Manipulator, the 6DOF Nachi Manipulator and pneumatic suction cup with maximum 1 kg payload and the V-groove for fine tuning location of object as shown in Figure 7. The 3D Camera is the Microsoft Kinect v1 with a resolution 640x480 combined with 2D camera Logitech. Both of them are parallel fixed with table surface having rectangular box objects. The specification of computer is: Laptop, RAM 16GB, Core I7, CPU 3.1GHZ. Six objects, representing for six random cases, are aluminium type and randomly stacked inside the box. They will be picked up and put at registered position.

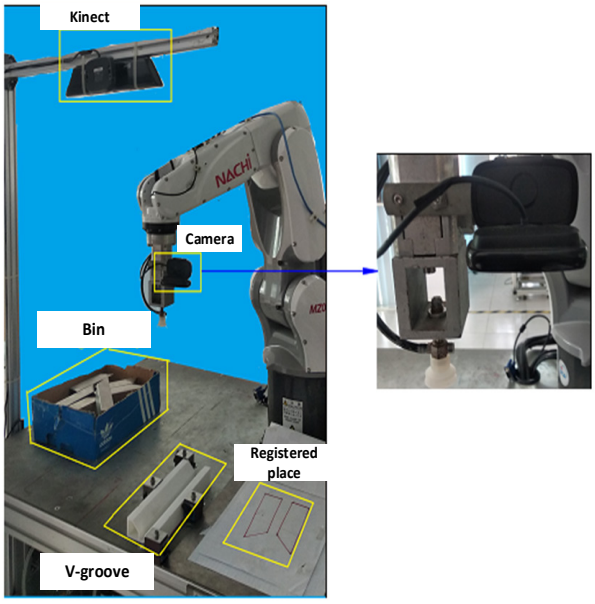


Figure 7: Experiment apparatus

By using 2D and 3D image processing discussed above, the positions of the objects are calculated as shown in Fig.8. The cluster having the highest height, green colour, is chosen to be grasped

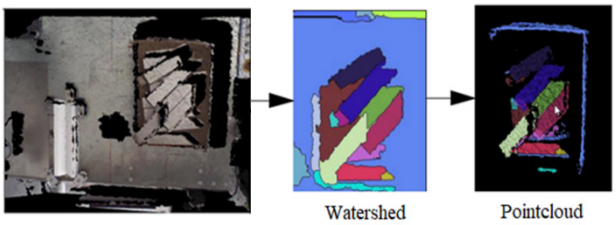


Figure 8: Image processing procedure

After having the coordinates and directions of object in the bin, the manipulator will perform the grasping task and then place it on the V-groove for the fine-tuning step Figure 9. After placing the object in the V-groove, the robot will use the second monocular camera to take the image to find the location and side of object to be picked Figure 10. Then the robot will put it to the registered position as shown in Figure 11.

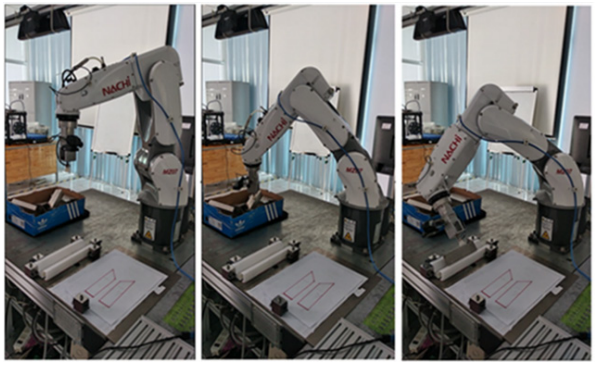


Figure 9: Process of grasping the object in the bin and putting it on V-groove

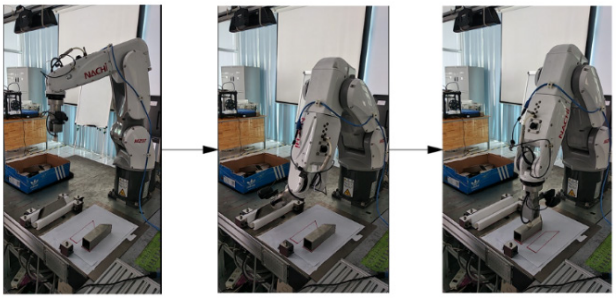


Figure 10: Process of grasping object on V-groove and putting it at registered

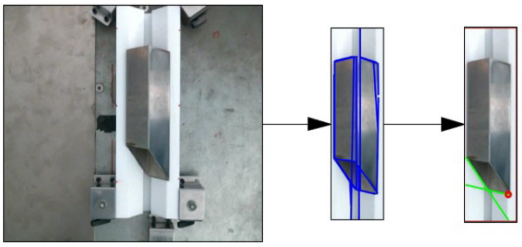


Figure. 11: Finding the key point of the object

After finish grasping task, the angular error, position error of the objects and the desired position are compared as shown in Figure 12. The real position of objects is obtained by reading the coordinate value on teach-pendant when moving robot manually. From that position the error of angle and position are calculated. The data are taken many times to calculate the mean error of the system as shown in following Table 3. The processing time is also calculated Figure 13

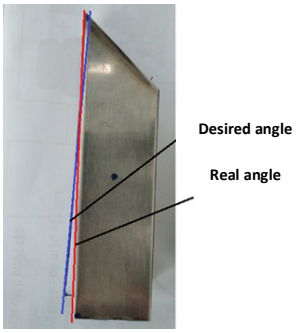


Figure 12: Desired position and real position

Table 3: Error position and angle calculation

No	Position Error (mm)			Angle error (degree)
	X	Y	Z	
1	1.00	1.73	0.08	1.1243
2	1.25	1.62	0.11	1.6973
3	1.01	1.52	0.19	1.0981
4	1.18	1.91	0.36	1.8352
5	0.50	1.15	0.34	1.9559
6	1.61	1.27	0.52	1.0600
Mean Error	1.0917	1.533	0.2667	1.4618

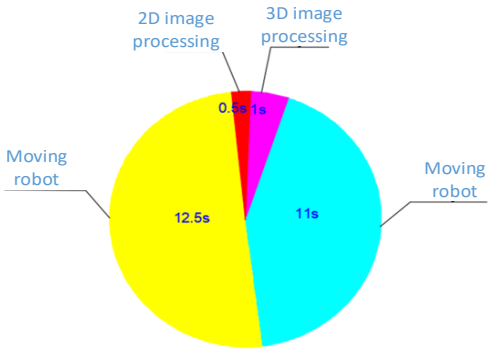


Figure 13: Processing time of system

From Table 3 it can be seen that the error of angle and position compared to the desired position is small and acceptable. The time to move the robot takes up most of the time by 23 seconds. Time to process images and data transmission is small about 1.5 seconds. The total time of implementation is 25 seconds, which is less than 30s. The image processing time is just about 6% of total time. To improve speed of system, the speed of robot could be increased to 100%. But for safety reasons it should only be done at 50% speed.

5. Conclusion

The paper proposed an effective method for the bin-picking problem. The image processing time is small and the accuracy is high. The method consists of two steps: the first step is to use the

3D camera and to combine 2D and 3D image processing for object segmentation and the second step is to use 2D camera for fining the grasping task. The effectiveness of the proposed approach was confirmed by a series of experimental results. The proposed method is also easy for installation and maintenance and applicable for assembly task in real industry. The comparing performance of proposed system and existing ones which are discussed in the introduction section are summarized in below Table 4.

Table 4: Comparing performance of Systems

	Proposed System	Existing system
Complexity	Not complicated	Very complicated
Time executing	Fast	Time consuming
Quality	Good and stable	Good and stable
Quantity	Mass product	Mass product
Setup	Easy	Complicated

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was supported by Saigon Hi-Tech Park Training Center and Vietnam National University, HoChiMinh City

References

[1] F. Casado, Y.L. Lapido, D.P. Losada, A.S. Alonso, " Pose estimation and object tracking using 2D image," In 27th International Conference on Flexible Automation and Intelligent Manufacturing, Modena, Italia, 63-71, 27-30 June, 2017, doi: 10.1016/j.promfg.2017.07.134.

[2] I. Wesis, M. Ray, ""Model-based recognition of 3D objects from single images," IEEE Transactions on Pattern Analysis and Machine Intelligence, **23**(2), 116-128, 2001, doi: 10.1109/34.908963.

[3] K. Ban, F. Warashina, I. Kanno, H. Kumiya, " Industrial intelligent robot," FANUC Tech. Rev., **16**(2), 29-34, 2003.

[4] J. Xu, " Rapid 3D surface profile measurement of industrial parts using two-level structured light patterns," Optics and Lasers in Engineering, **49**(7), 907-914, 2011, doi: 10.1016/j.optlaseng.2011.02.010.

[5] W.C. Chang, C.H. Wu, " Eye-in-hand vision-based robotic bin-picking with active laser project," International Journal of Advanced Manufacturing Technology, 85, 2873–2885, 2016, doi: 10.1007/s00170-015-8120-0

[6] K. Rahardja, A. Kosaka, "Vision-based binpicking: recognition and localization of multiple complex objects using simple visual cues," In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 100-107, 1996, doi: 10.1109/IROS.1996.569005.

[7] M. Berger, G. Bachler, S. Scherer, " Vision guided bin picking and mounting in a flexible assembly cell," In Proeeding of the 13th International Conference on Industrial & Engineering Applications of Artificial Intelligence and Expert Systems, 109-118, 2000, doi: doi.org/10.1007/3-540-45049-1_14

[8] M.Y. Liu, O. Tuzel, A. Veeraraghavan, Y. Taguchi, T.K. Marks, R. Chellappa, "Fast object localization and pose estimation in heavy clutter for robotic bin picking," The International Journal of Robotics Research, **31**(8), 951–973, 2012, doi: 10.1177/0278364911436018

[9] A. Adan, S.V. Andrés, P. Merchán, R. Heradio, "Direction Kernels: Using a simplified 3D model representation for grasping," Machine Vision and Applications, **24**(1), 351-370. 2013, doi: 10.1007/s00138-011-0351-y

[10] R. D. Schraft, T. Ledermann, "Intelligent picking of chaotically stored objects," Assembly Automation, **23**(1), 38-42, 2003, doi: 10.1108/01445150310460079

[11] K. Bousmalis, A. Irpan, P. Wohlhart, Y. Bai, M. Kelcey, M. Kalakrishnan, L. Downs, J. Ibarz, P. Pastor, K. Konolige, S. Levine, V. Vanhoucke, "Using simulation and domain adaptation to improve efficiency of deep robotic grasping," In Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 4243 – 4250, 2018, Australia, doi: 10.1109/ICRA.2018.8460875.

[12] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J.A. Ojea, K. Goldberg," Dex-net 2.0: deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics," In Robotics: Science and Systems, 245-251, 2017, doi: 10.15607/RSS.2017.XIII.058

[13] L. Tuan-Tang, L. and Chyi-Yeu, "Bin-Picking for Planar Objects Based on a Deep Learning Network: A Case Study of USB Packs," Sensors, **19**(16), pp. 3602-3618, 2019, doi: 10.3390/s19163602

[14] J. Chen, T. Fujinami, E. Li, "Deep Bin Picking with Reinforcement Learning," In Proceedings of the 35th International Conference on Machine Learning, 1-8, 2018, Sweden.

[15] S. K. Anton, V.S. Ilia, "An Overview of Watershed Algorithm Implementations in Open Source Librarie," Journal of Imaging, **4**(10), 123-138, 2018 doi: 10.3390/jimaging4100123.

[16] L.D. Hanh, L.M. Duc, "Planar Object Recognition For Bin Picking Application," In Proceedings of 5th NAFOSTED Conference on Information and Computer Science (NICS), 213-218, 2018, doi: 10.1109/NICS.2018.8606884.

[17] Conditional Euclidean Clustering. Available online: http://pointclouds.org/documentation/tutorials/conditional_euclidean_clustering.php(accessed on 12 January 2020).

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1250-1257 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Crystallinity and Hardness Enhancement of Polypropylene using Atmospheric Pressure Plasma Discharge Treatment

Oscar Xosocotla[1,2,*], Horacio Martinez[1], Bernardo Campillo[2]

[1]*Laboratorio de Plasmas Atmosféricos y Espectroscopia Raman, Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, 62220, México*

[2]*Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México, 05000, México*

ARTICLE INFO

Article history:
Received: 10 October, 2020
Accepted: 06 December, 2020
Online: 16 December, 2020

Keywords:
Polypropylene
Crystallinity
Hardness
Atmospheric pressure plasma

ABSTRACT

Atmospheric pressure plasma was used to treat polypropylene (PP) surfaces. Optical emission spectroscopy (OES) was used to determine the chemical species formed in the plasma as well as the electron temperature and density of the plasma. The flux of species (O and OH) produced during the plasma treatment interact with the polymer surface creating polar groups on the PP surface, which were evaluated by X-ray, Raman, and Attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) analyses. Moreover, the thermal properties of PP were investigated by thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) analysis. The plasma treatment increased the crystallite size and microhardness. This enhanced effect was produced from high concentrations of O and OH that induced functional polar groups containing C–O, C=O, and C=O–OH bonds appended to the PP surface. The polar groups produced by the air plasma can be attributed primarily to the oxygen radicals in the air plasma hitting the PP surface and to heat-induced oxidation rather than the incorporation of oxygen radicals or UV-induced oxidation from the plasma. The increase in hardness is attributed to the introduction of carbonyl and hydroxyl groups, cross-linking, annealing effects, and chemical etching.

1. Introduction

Atmospheric pressure plasma discharge treatment has become important because of its relatively simple process, low cost, easy operation, and applicability to vacuum-incompatible materials and processes compared with the low-pressure plasma process [1]. Typically, two categories of plasmas exist based on their gas temperatures: thermal and nonthermal (cold) plasmas [2, 3, 4]. In thermal plasmas, electron temperature is very close to those of heavy particles (both charged and neutral), and therefore, it cannot be used for the surface modification of heat-sensitive materials. However, nonthermal plasma electrons have a temperature that is 2–3 orders of magnitude higher than the temperature corresponding to the random motion of the background gas molecules. As a result of the very low heat capacity and density of electrons, plasma temperature is almost the same as room temperature. Subsequently, the ions remain rather cold, and a nonthermal plasma does not produce any thermal damage to the treated surfaces [2-4]. These characteristics are particularly advantageous for the treatment of heat-sensitive and fragile materials, such as polymers [5, 6].

Surface modification can be applied either directly or indirectly. Direct surface modification methods are used to modify an inert surface with free radicals for specific applications, whereas indirect surface modification methods are based on polymer grafting [5, 6]. The efficiency of plasma treatment methods depends on surface type, the gas phase of the plasma, and treatment conditions such as discharge type, pressure, frequency, discharge voltage and current, and treatment time. To modify surface properties without any change in the bulk material, it is essential to use successful surface modification methods such as atmospheric pressure plasma [5, 6].

Surface modification of polypropylene (PP) has been extensively studied to improve its properties, such as wettability (surface free energy), hydrophilicity, biocompatibility, and adhesion, using different atmospheric pressure plasma reactors [5,

*Corresponding Author: Oscar Xosocotla, Email: hm@icf.unam.mx

7-11]. However, only two studies have shown PP hardness improvement: the first study [12] involved beta irradiation at several doses, showing substantial changes in PP structure and mechanical and micromechanical properties, and the second study [13] utilized atmospheric pressure plasma treatment by diffuse coplanar surface barrier discharge using ambient air as the working gas. They observed no significant changes in the mechanical properties of the films after plasma treatment. Therefore, the purpose of this study was to characterize and test the impact of the commercial plasma device on the PP surface. Therefore, optical emission spectroscopy (OES) was used to investigate chemical species that cause plasma surface modification, electron temperature (T_e) and density (n_e) were used to characterize the plasma, differential scanning calorimetry (DSC) was used to determine the melting points of materials, thermogravimetric analysis (TGA) was used to evaluate the degradation of materials subjected to plasma treatment, X-ray diffraction (XRD) was employed to analyze crystallite orientation and size, Raman spectroscopy and Attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectroscopy were performed to investigate the conformations of polymer chains, and atomic force microscopy (AFM) with a micro-durometer was used to study the morphology of the surfaces and determine the microhardness of PP.

2. Experimental

2.1. Plasma Source

A schematic illustration of the APC 500 plasma torch source (Diener Electronics, Germany) used for surface modification is shown in Figure 1. The plasma source [14] was operated at a voltage of 20 kV, discharge current of 25 mA, frequency of 40 kHz, and gas pressure of 1500 torr. The plasma gliding discharge was generated by two electrodes. The discharge extended from the electrode region because of air flow, and the active plasma covered an area of ~10 cm^2. The samples were placed in the most homogenous region, which was at the center of the gliding plasma. OES was used for optical characterization of the plasma. For OES measurements, an Ocean Optics spectrometer (model HR2000CG-UV-NIR) was used in the spectral range 200–1100 nm, and the data were obtained with a step of 0.35 nm and an optical resolution of 1.0 nm (FWHM).

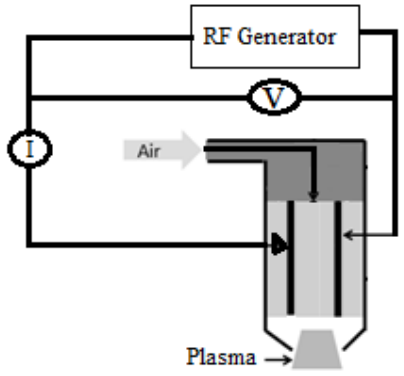


Figure 1: Diagram of the radio frequency atmospheric plasma generator

2.2. Materials

A commercially available isotactic (Lyondell Basell) PP cylinder with a density of 0.91 grams per cubic centimeter (g cc^{-1}) was cut into 1.5 × 1.5 × 0.5 cm^3 pieces. One side of each PP sample was polished to achieve uniform roughness using silicon carbide sandpaper with a grain size of 1,500. The samples were then washed in an ultrasonic bath for 20 min. To remove the residual liquid, the samples were placed in an oven at 30 °C for 30 min.

2.3. X Ray analysis

The XRD patterns of PP samples were recorded by a Rigaku D MAX-2000 diffractometer using Cu K_α (λ = 0.154 nm) radiation with 2θ in the range 10°–20° using a step size of 0.06° and step time of 10 s. The diffraction patterns of untreated and treated PP samples consisted of well-defined peaks. The high-intensity peaks were chosen to estimate the crystalline size and lattice strain using Scherrer's formula and Williamson–Hall (W–H) model [15, 16].

2.4. Raman analysis

Raman spectra were obtained using a SENTERRA II spectrometer (Bruker). The characteristic peaks of the crystalline and amorphous phases in the spectra of treated and untreated PP samples were then evaluated. The Raman spectrometer was equipped with a 785-nm laser and was operated at a power of 100 mW. The integration time was 1,000 ms.

2.5. DSC and TGA

DSC and TGA were performed to analyze the thermal properties of the samples, which were placed in aluminum specimen holders. To correct for their thermal histories, the treated PP samples were heated from 10 °C to 250 °C at a rate of 10 °C/min in a nitrogen atmosphere.

2.6. ATR FTIR analysis

The chemical compositions of the PP surfaces were analyzed before and after plasma treatment using a Spectrum 400 ATR-FTIR spectrometer (PerkinElmer). The vibrational modes appearing in the ATR-FTIR spectra were consistent with those reported in the literature [17-20].

2.7. Roughness

A NX10 atomic force microscope (Park) was used to quantitatively analyze the surface roughness of the PP samples due to plasma treatment [21-23].

2.8. Microhardness

A micro-durometer (Matsuzawa) was used to measure PP microhardness under a 300 g load. The penetration time was set to 20 s, and five indentations were made on each PP surface to obtain an average value [12, 24].

3. Results and analysis

3.1. Plasma characterization

The efficiency of plasma treatment depends on the surface type and treatment conditions. Therefore, the plasma discharge

was produced by applying a voltage (V) of 10 kV, and the electrodes were separated by a distance (d) of 2 cm. The resulting electric field strength (E) was equal to V/d. E was calculated to be 5.0×10^3 V cm^{-1}, and E/P was equal to 6.58 V cm^{-1} torr^{-1}.

The electron drift velocity (v_d) and E/N, where N is the number density, were obtained from the literature [25], considering the experimental conditions used. We obtained a v_d of 3.2×10^6 cm s^{-1} and E/N of 0.19 Td (1 $Td = 10^{-17}$V cm^2). Thus, N is 2.63×10^{19} cm^{-3}.

The electron density (n_e) was estimated from the measured current density (J) [26]:

$$n_e = \frac{J}{e v_d}. \quad (1)$$

The electron density was found to be $(1.95 \pm 0.24) \times 10^{12}$ cm^{-3}, and the order of magnitude was consistent with reported values [27, 28]. The electron density, n_e, was assumed to be equal to the ion density n_i in the plasma bulk [29].

3.2. Optical emission spectroscopy (OES)

OES was performed to characterize the plasma discharge. The OES spectrum obtained from the discharge is shown in Figure 2. It is well known that atomic and molecular transitions emit photons with specific wavelengths and energies that correspond to the difference between the upper and lower energy levels involved in the transitions. The OH($A^2\Sigma^+ - X^2\Pi_i$) molecular emission near 309 nm indicated that the plasma was not in equilibrium because of strong transitions and trace amounts of water vapor in the air. It is often difficult to prevent water from affecting the plasma. The peak at 337 nm was due to a $C^3\Pi_u \rightarrow B^3\Pi_g$ transition in N_2 and was the most intense peak in the region of 330–530 nm. Moreover, the first negative system, N_2^+ ($B^2\Sigma_u^+ \rightarrow X^2\Sigma_g^+$) generating a 0–0 band between 388.5 and 391.5 nm was observed together with two low-intensity peaks at 777.19 and 844.71 nm, which were attributed to O ($^5S^0 - {}^5P$) and O ($^3S^0 - {}^3P$) transitions, respectively.

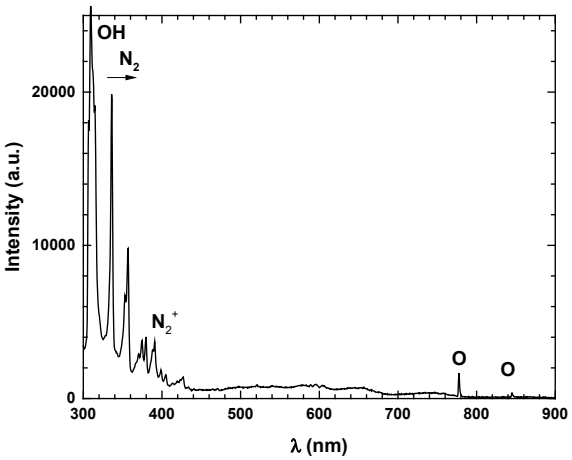


Figure 2: Optical emission spectroscopy (OES) spectrum.

OES is often used to determine the energy of electrons in the gas phase [30-33]. Thus, it can be applied to plasmas that are not in equilibrium at atmospheric pressure, which is often the case during the analysis of complex mixtures. The optical characterization of the discharge was based on the ratios of the O atomic line intensities in the emission spectrum. The electron energy was found to be (0.36 ± 0.04) eV, which was consistent with the results of previous atmospheric plasma studies [34-36]. The electron temperature (Te) can be several orders of magnitude higher than the temperature of the ions (Ti) or the neutral gas (Tg). Therefore, the heavy particles stay close to ambient temperature (27^0C) [29, 37, 38].

The species observed by (OES) in the plasma are formed due to electrons are accelerated by the electric field and can induce electron impact excitation, ionization and dissociation of the primary gas phase (humid-air containing 79% N_2/20% O_2/ 1%) reactions, these are:

$$e + N_2 \rightarrow N_2^* + e, \quad (2)$$
$$e + O_2 \rightarrow O_2^* + e, \quad (3)$$
$$e + N_2 \rightarrow N_2^+ + e + e, \quad (4)$$
$$e + O_2 \rightarrow O_2^+ + e + e, \quad (5)$$
$$e + N_2 \rightarrow N + N + e, \quad (6)$$

The formation in OH• radicals and H species during the plasma treatment may be attributed to the fragmentation/decomposition of H_2O molecules:

$$e + H2O \rightarrow H + OH + e \quad (7)$$

or

$$e + O_2 \rightarrow 2O + e \quad (8)$$
$$O + H_2O \rightarrow OH + OH \quad (9)$$

whereas the formation of atomic oxygen may be due to the dissociation of oxygen molecules by reaction (8).

3.3. X-Ray Diffraction analysis

The XRD patterns of the treated and untreated PP samples are shown in Figure 3. The peak characteristics of the α phase were observed at (2θ) 14.02° (110), 16.8° (040), and 18.44° (130), and between 20.80^0 and 22.40^0 for (111), (131), and (041). The same peaks appeared in the patterns of the treated PP samples, indicating that the plasma had no effect on the crystalline structure of PP. However, the peak intensity of the untreated sample at 16.8° differs from those of the treated samples.

The XRD patterns were used to calculate the crystallite size and orientation of planes in each sample.

The size of crystallites (D) in each PP sample based on the (110) and (040) planes was calculated using the Scherrer equation [16]:

$$D = \frac{0.9\lambda}{\beta \cos\theta}, \quad (10)$$

where λ is the wavelength, β is the full width at half maximum (FWHM), and θ is the maximum angle of dispersion.

The crystallite size (D) was compared with the values obtained from the Williamson-Hall (W–H) model, where the total peak broadening is represented by the sum of the contributions from crystallite size (D) and strain (ε) in the material. Assuming that the strain present in the material is uniform and considering the isotropic nature of the crystal, the W–H equation [15, 16] is given by

$$\beta \cos(\theta) = \frac{0.9\lambda}{D} + 4\varepsilon \sin(\theta). \quad (11)$$

The estimated values of the strain and crystallite size are given in Table 1.

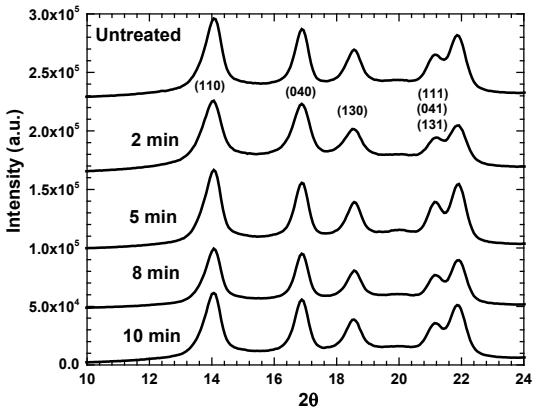


Figure 3: XRD patterns of PP samples subjected to plasma treatment for various periods.

Table 1: Geometric parameters for PP as a function of treatment duration.

Treatment duration	Scherrer´method	Williamson-Hall method	
min	D (nm)	D (nm)	ε (10^{-7})
0	11.7	10.7	4.4
2	10.5	8.0	9.9
5	12.4	10.5	5.6
8	12.9	11.9	3.4
10	12.5	11.8	2.8

We observed that increasing the treatment duration increases the crystal size, except for 2 min of treatment, which may be due to the rearrangement of the orientation parameters of the crystal planes due to the introduction of polar groups on the PP surface by the plasma treatment. In addition, the strain at 2 min of plasma treatment increases almost 100% and then decreases as the treatment duration increases. That behavior may be attributed to dislocations that increase the volume of grain boundaries (increase in the crystal size) due to the plasma treatment and incorporation of the polar groups on the surface of PP.

To evaluate the orientation parameters of the (040) crystal plane, A_{110} (040), and (110) crystal plane, C_{040} (110), we used Zipper's equation, which takes into account the intensities of the peaks observed [39-41]:

$$A_{110}(040) = \frac{I_{110}}{I_{110}+I_{111}+I_{131+041}}, \quad (12)$$

$$C_{040}(110) = \frac{I_{040}}{I_{110}+I_{040}+I_{130}}, \quad (13)$$

where I_{110}, I_{111}, I_{104}, I_{040}, I_{130}, and $I_{131+041}$ correspond to the peak intensity of the respective crystal plane.

Figure 4a shows the orientation indices of A_{110} (040) and C_{040} (110) crystal planes as a function of treatment duration. The orientation indices appear to increase as the duration of plasma treatment increases due to the surface functionalization by the plasma treatment. Changes in crystallite sizes in the (110) and (040) planes during plasma treatment are illustrated in Figure 4b. Crystallites of the samples became smaller during the first 2 min of treatment. The crystallite size increased as the duration of plasma treatment increased. It can be explained because the plasma treatment acts as a nucleating agent on the PP surface, which promotes the formation of perfect crystals [12, 39-41].

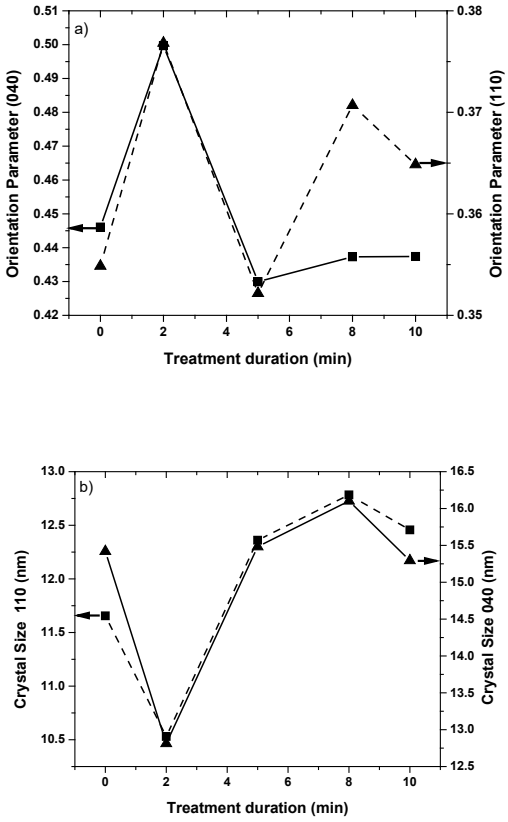


Figure 4: Crystallographic parameters of the treated PP samples. 4.a) Orientation indices of the crystal planes and 4.b) crystallite size.

3.4. Raman analysis

Raman spectra from 809 to 841 cm^{-1} are shown in Figure 5. The peak attributable to the crystalline phases in the PP samples was observed near 809 cm^{-1}, and the peak near 841 cm^{-1} was ascribed to the phases that contained helical chains with isomeric defects [17, 25, 42, 43].

The intensities of the observed peaks (I_{809} and I_{841}) were used to evaluate dispersion changes that were not related to crystallinity. The sum of the intensities ($\bar{I} = I_{809} + I_{841}$) was

independent of crystallinity within the range of measurement. The crystallinity of the treated PP samples was estimated using equations 14 and 15. The proportions of phases with defects and crystalline are denoted by x^b and x^c, respectively [40, 41]:

$$x^b = \frac{I_{841}}{\bar{I}}, \tag{14}$$

$$x^c = \frac{I_{808}}{\bar{I}}. \tag{15}$$

Variations in the proportions of phases with defects (x^b) and crystalline phases (x^c) in the samples are shown in Figure 6. Plasma treatment for 10 min increased the crystallinity of the PP samples, although the x^c values of samples treated for 2, 5, or 8 min slightly decreased. Variations in the proportions of phases with defects indicated that phase defects were associated with crystallinity. Regular chains in the samples subjected to plasma treatment for the greatest amounts of time reorganized to generate helices and local order. The amorphous phases contained irregular atactic chains that could not be reorganized into helices because of the random orientations of their side groups. Those changes are correlated with the increase in crystallite size and orientation parameters due to the plasma treatment and the functionalization of the PP surface [12, 21, 22, 25, 39-43].

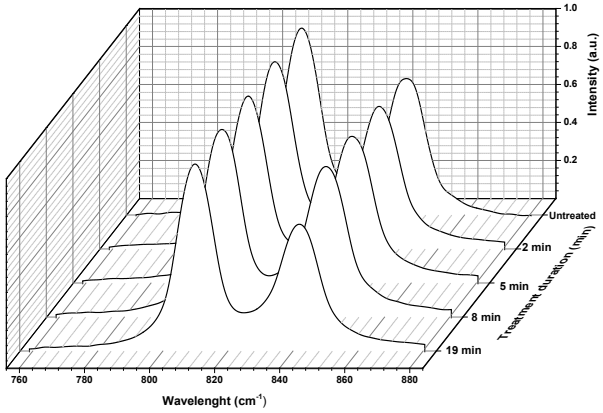


Figure 5: Raman spectra as a function of treatment duration.

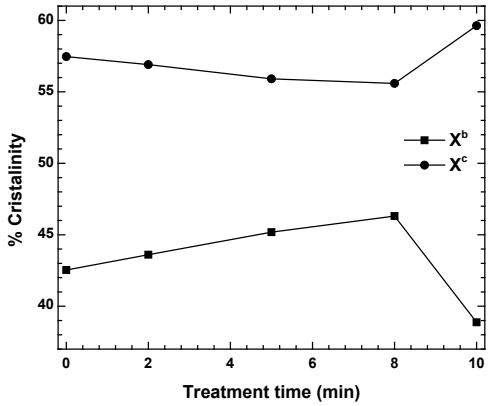


Figure 6: % crystallinity as a function of treated time.

3.5. DSC and TGA analysis

Changes in the thermal properties of the treated PP samples were Changes in the thermal properties of the treated PP samples were analyzed via DSC and TGA. Degradation is shown as a function of temperature in Figure 7. Plasma treatment did not affect mass loss under a nitrogen atmosphere because the percentages of mass lost by the untreated and treated samples did not differ. The first derivative of each mass loss curve indicated the largest change in mass. The highest rates of change in the weight loss curves of PP samples treated for 2 or 5 min occurred at 404 °C. In the weight loss curves of the untreated samples and samples treated for 8 or 10 min, the highest rates of change were observed at 416 °C (Figure 8).

Structural and compositional transformations in a polymer due to plasma treatment can be analyzed by evaluating the changes in its properties at different temperatures [39]. The DSC thermograms of the treated and untreated samples are shown in Figure 9. The untreated sample and samples treated for 8 or 10 min melted at 165 °C, whereas the samples treated for 2 or 5 min melted at 167 °C. Therefore, the melting points of the treated and untreated samples did not differ significantly [24].

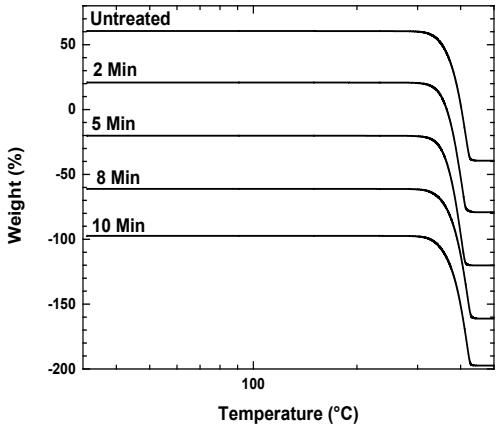


Figure 7: TGA thermograms of the treated and untreated PP samples, the curves have been shifted vertically for better comparison.

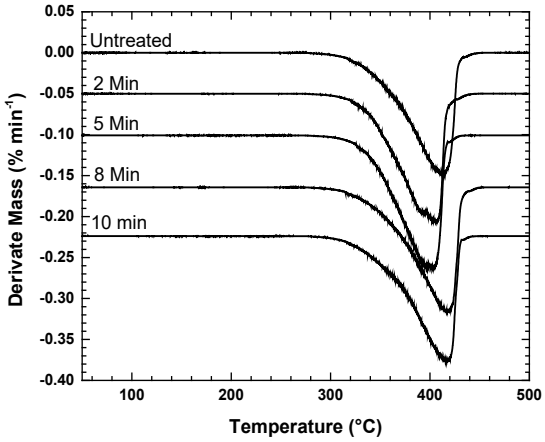


Figure 8: Weight loss curves of the treated and untreated PP samples, the curves have been shifted vertically for better comparison.

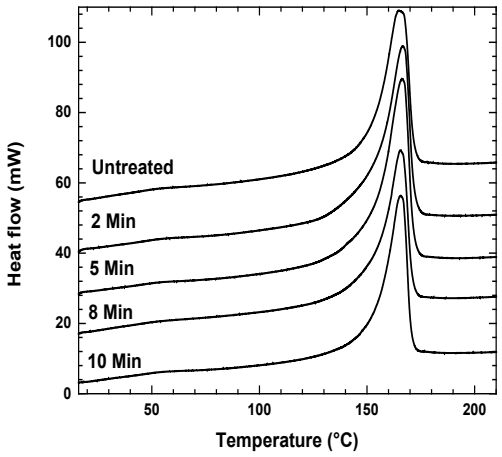


Figure 9: DSC thermograms of the treated and untreated PP samples, the curves have been shifted vertically for better comparison.

3.6. ATR-FTIR analysis

The FTIR spectra of the untreated and treated PP samples are shown in Figure 10. Chemical differences in the surface properties of the plasma-treated PP samples can be observed in Figure 10, indicating changes in their surface chemical compositions. The spectra confirmed the introduction of polar groups into all treated samples. Vibrational bands attributed to carbonyl (C=O) groups appeared from 1,500 to 1,700 cm^{-1} [18-21]. Moreover, the carbonyl (C=O) band was observed at 1,720 cm^{-1} in the spectra of the samples treated for 5 and 10 min but did not appear in the spectra of samples treated for 2 or 8 min. It means that the C content decreases, and the O content increases on the surface because of the plasma treatment. Additionally, the peak around 3400 cm^{-1} corresponding to the hydroxyl (–OH) group was produced by the plasma treatment. The relative intensities of all absorption peaks observed by FTIR were enhanced, indicating the increases in the carbonyl and hydroxyl groups on the surface [44]. In the present treatment, large amounts of O and OH polar groups were appended to the surface of PP because of the plasma treatment, improving the crystallinity and microhardness of PP.

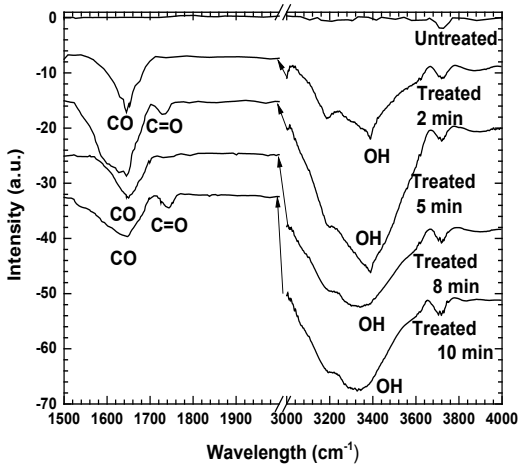


Figure 10: ATR-FTIR spectra as a function of treatment duration.

3.7. Roughness

The AFM images of the untreated and treated samples at 5 min are shown in Figures 11 and 12, respectively. The morphologies of the samples treated by plasmas were smoother than those of the untreated sample. Those figures were used to study the changes in the morphology of the PP samples, and the RMS roughness found for the untreated sample was 0.554 ± 0.064 pm, whereas for the samples treated at 2, 5, 8, and 10 min, the surface roughness values were 0.153 ± 0.031, 0.136 ± 0.013, 0.111 ± 0.024, and 0.090 ± 0.008 pm, respectively. The sample treated at 8 min shows the lowest value of RMS roughness. Those morphological variations may be due to the removal of the top few monolayers of the PP samples produced by the bombardment of plasma particles on the PP surface [13, 18], reducing the surface roughness and improving the microhardness of PP. Although their surfaces were smoothed, the ATR-FTIR spectra of the treated samples indicated that the introduction of polar groups was possible [13, 18-20, 22]

3.8. Microhardness

The microhardness values of the treated and untreated PP samples are plotted in Figure 13. The hardness values of the treated samples were higher than those of the untreated sample, and the hardness of the sample treated for 10 min increased by nearly 56%. An increase in hardness is attributed to annealing effects and chemical etching produced by the plasma treatment [45]. Also, the hardness increase may be related to either the hardening of interlamellar regions due to the polar group penetration or the hardening of the crystals themselves caused by the surface attachment of polar groups, as observed by Raman and ATR-FTIR. Plausible changes on the PP surface caused by the plasma treatment include cross-linking and oxidation. Cross-linking is due to the formation of tertiary carbon radicals. Tertiary hydrogen abstraction generates carbon radicals, which react with radicals in the nearest molecular chain. Oxidation primarily affects secondary carbon atoms. The introduction of oxygen-bearing groups, including –C–OH, –CO–OH, and –C=O moieties, occurs when the bonds between hydrogen and secondary carbon atoms on the PP surface are broken during the plasma treatment [13, 22-23].

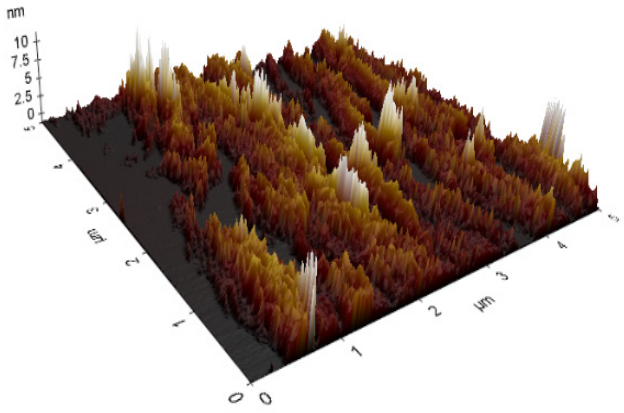


Figure 11 Roughness profiles of the untreated PP sample.

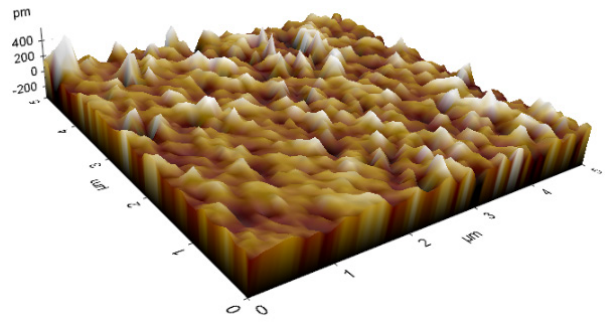


Figure 12: Roughness profiles of PP treated for 5 min.

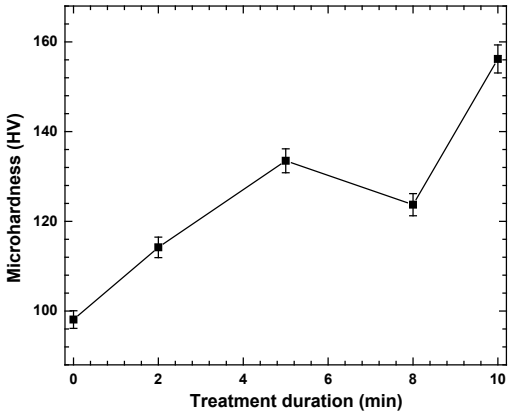


Figure 13: Microhardness values of the samples before and after plasma treatment

4. Conclusions

A commercial plasma device was used for the surface modification of PP. Using OES, we determined the chemical species formed in the plasma. High concentrations of OH and O radicals were observed in the plasma, and they led to the formation of functional polar groups, C–O, C=O, and C–O–OH bonds, on the PP surface. The plasma treatment increased the crystallinity size and microhardness. This enhanced effect is attributed to the formation of polar groups produced by the air plasma due to the oxygen radicals in the air plasma hitting the PP surface. The formation of polar groups in the PP surface was corroborated by Raman and ATR-FTIR spectroscopies. DSC and TGA indicated that plasma treatment had no effect on the thermal properties of PP. AFM analysis showed that plasma treatment for extended periods reduced the surface roughness of PP. Finally, the increase in PP hardness is attributed to the introduction of carbonyl and hydroxyl groups by the plasma treatment, cross-linking, annealing effects, and chemical etching produced by the plasma treatment.

Acknowledgment

The authors thank E. Vazquez, H. H. Hinojosa, F. Castillo, and O. Flores for their technical assistance. This work was partially supported by DGAPA project IN-105519 and CONACyT projects 225991 and 268644.

References

[1] E.E. Kunhardt, "Generation of large-volume, atmospheric-pressure, nonequilibrium plasmas", IEEE Trans. Plasma Sci. **28**, 189-200 (2000), https://doi.org/10.1109/27.842901

[2] P.J. Bruggeman, F. Iza and R. Brandenburg, "Foundations of atmospheric pressure nonequilibrium plasmas", Plasma Sources Sci. Technol. **26**, 123002, 2017, https://doi.org/10.1088/1361-6595/aa97af.

[3] A. Fridman, A. Chirokov and A. Gutsol, "Non-thermal atmospheric pressure Discharges", J. Phys. D: Appl. Phys. **38,** R1–R24, 2005, doi:10.1088/0022-3727/38/2/R01

[4] A. Schutze, J. Y. Jeong, S. E. Babayan, J. Park, G. S. Selwyn and R. F. Hicks, "The atmospheric-pressure plasma jet: a review and comparison to other plasma sources", IEEE Transactions on Plasma Science, **26**(6), 1685-1694, 1998, doi: 10.1109/27.747887.

[5] S.K. Nemani, R. K. Annavarapu, B. Mohammadian, A. Raiyan, J. Heil, Md. A. Haque, A. Abdelaal, and H. Sojoudi, "Surface Modification of Polymers: Methods and Applications", Adv. Mater. Interfaces **5**, 1801247, 2018, doi:10.1002/admi.201801247

[6] V. Palma, M. Cortese, S. Renda, C. Ruocco, M. Martino and E. Meloni, "A Review about the Recent Advances in Selected NonThermal Plasma Assisted Solid–Gas Phase Chemical Processes", Nanomaterials **10**, 1596, 2020, doi:10.3390/nano10081596

[7] M. Kehrera, A. Rottensteiner, W. Hartl, J. Duchoslav, S. Thomas, D. Stifter, "Cold atmospheric pressure plasma treatment for adhesion improvement on polypropylene surfaces", Surface & Coatings Technology in press (2020), https://doi.org/10.1016/j.surfcoat.2020.126389

[8] H.B. Baniya, R. P. Guragain, B. Baniya, and D. P. Subedi, "Cold Atmospheric Pressure Plasma Jet for the Improvement of Wettability of Polypropylene", International Journal of Polymer Science, **2020**, 3860259, 2020, https://doi.org/10.1155/2020/3860259;

[9] R. Janík, M. Kohutiar, M. Pajtášová, A. Dubec, J. Pagáčová and J. Šulcová, "The impact of DCSBD plasma discharge on polypropylene", IOP Conf. Series: Materials Science and Engineering **776,** 012090, 2020, 10.1088/175799X/776/1/012090

[10] F. Darvish, N. M. Sarkari, M. Khani, E. Eslami, B. Shokri, M. Mohseni, M. Ebrahimi, M. Alizadeh, C. Fu Dee, "Direct plasma treatment approach based on non-thermal gliding arc for surface modification of biaxially-oriented polypropylene with post-exposure hydrophilicity improvement and minus aging effects", Applied Surface Science **509**, 144815, 2020, https://doi.org/10.1016/j.apsusc.2019.144815

[11] Y. Akishev, G. Aponin, A. Petryakov and N. Trushkin, "On the composition of reactive species in air plasma jets and their influence on the adhesion of polyurethane foam to low-pressure polyethylene", J. Phys. D: Appl. Phys. **51** 274006, 2018, https://doi.org/10.1088/1361-6463/aac5fd

[12] D. Manas, M. Hribova, M. Manas, M. Ovsik, M. Stanek, D. Samek, "The effect of beta irradiation on morphology and micro hardness of polypropylene thin layers", Thin Solid Films **530**, 49–52, 2013, http://dx.doi.org/10.1016/j.tsf.2012.09.051

[13] A. Oravcova, I. Hudec, "The Influence of Atmospheric Pressure Plasma Treatment on Surface Properties of Polypropylene Films", Acta Chimica Slovaca, **3**(2), 57 - 62, 2010, http://www.acs.chtf.stuba.sk/?id=7&paper=68

[14] http://plasma.de/en/plasmasystems/plasma-apc500.html, accessed on 28.9.2015

[15] W. Burton, O. Kenneth, R. Thomas, I.Y. Chan, "On the estimation of average crystallite size of zeolites from the Scherrer equation: A critical evaluation of its application to zeolites with one-dimensional pore systems", Microporous Mesoporous Mater. **117**, 75-90, 2009, doi:10.1016/j.micromeso.2008.06.010

[16] G.K. Williamson and W. H. Hall, "X-Ray line broadening from filed aluminium and wolfram", Acta Metallurgica **1**, 22-31, 1953, https://doi.org/10.1016/0001-6160(53)90006-6

[17] E. Andreassen and J. Karger-Kocsis, "Infrared and Raman spectroscopy of polypropylene", in: Karger-Kocsis J. (Eds.), Polypropylene. Polymer Science and Technology Series, **2**, 320-328, 1999, https://doi.org/10.1007/978-94-011-4421-6_46

[18] M.R. Monzó Pérez. "Investigación de la influencia del tipo de gas utilizado en el plasma atmosférico, en la modificación superficial del polipropileno" (tesis de doctorado). Universidad Politécnica de Valencia. Valencia, España, 2015.

[19] M.J. Shenton and G. C. Stevens. "Surface modification of polymer surfaces: atmospheric plasma versus vacuum plasma treatments". Journal of Physics

D: Applied Physics, **34**, 2761-2768, 2001, doi:10.1088/0022-3727/34/18/308

[20] N.V. Bhat and D. J. Upadhyay, "Plasma-Induced Surface Modification and Adhesion Enhancement of Polypropylene Surface", Journal of Applied Polymer Science, **86**, 925-936, 2002. https://doi.org/10.1002/app.11024

[21] H.H. Šourková, G. Primc, and P. Špatenka, "Surface Functionalization of Polyethylene Granules by Treatment with Low-Pressure Air Plasma", Journal Materials, **11**, 885, 2018. 10.3390/ma11060885

[22] S. D. Lee, M. Sarmadi, F. Denes, and J. L. Shohet, "Surface Modification of Polypropylene Under Argon and Oxygen-RF-Plasma Conditions", Plasmas and Polymers, **2**, 177-198, 1997. https://link.springer.com/article/10.1007/BF0276613

[23] M. Kehrer, J. Duchoslav, A. Hinterreiter, A. Mehic, T. Stehrer, D. Stifter, "Surface functionalization of polypropylene using a cold atmospheric pressure plasma jet with gas water mixtures", Surface and Coatings Technology, **384**, 125170, 2020, https://doi.org/10.1016/j.surfcoat.2019.125170.

[24] A. García, M. Domínguez-Díaz, R. Paniagua, J. A. Hernández-Péreza, H. Martínez. "Microhardness modification of ultrahigh-molecular-weight polyethylene by oxygen plasma: Effect of the polymer crosslinking". Nuclear Inst. and Methods in Physics Research B, **445**, 8-12, 2019, https://doi.org/10.1016/j.nimb.2019.02.020

[25] Y. Hiejima, K. Takeda, and K-h Nitta, "Investigation of the Molecular Mechanisms of Melting and Crystallization of Isotactic Polypropylene by in Situ Raman Spectroscopy", Macromolecules, **50**, 5867−5876, 2017, https://doi.org/10.1021/acs.macromol.7b00229.

[26] J. Dutton, "A survey of electron swarm data", Journal of Physical and Chemical Reference Data **4,** 577, 1975, https://doi.org/10.1063/1.555525

[27] K.K. Trusov, "Dynamics of multichannel and quasi-homogeneous sliding discharge formation in rare gases", Journal of Physics D: Applied Physics, **40**(3), 786–794, 2007, doi:10.1088/0022-3727/40/3/015

[28] A. Yu Nikiforov, Ch Leys, M. A. Gonzalez and J. L. Walsh, "Electron density measurement in atmospheric pressure plasma jet: Stark broadening of hydrogenated and non-hydrogenated lines", Plasma Sources Sci. Technol, **24**(3), 034001, 2015, https://doi.org/10.1088/0963-0252/24/3/034001

[29] K. Yoshida, K. Nitta, H. Ohmi, K. Yasutake, and H. Kakiuchi, "Plasma parameters in very high frequency helium and argon plasmas at atmospheric pressure", J. Appl. Phys**.** **128**, 133303, 2020, https://doi.org/10.1063/5.0010195

[30] C. Laux, T. Spence, C. Kruger and R. Zare, "Optical diagnostics of atmospheric pressure air plasmas", Plasma Sources Sci. Technol*,* **12**, 125–138, 2003, https://iopscience.iop.org/article/10.1088/0963-0252/12/2/301

[31] S. Moon and W. Choe, "A comparative study of rotational temperatures using diatomic OH, O_2 and N_2^+ molecular spectra emitted from atmospheric plasmas", Spectrochim. Acta B, **58**, 249–257, 2003, https://doi.org/10.1016/S0584-8547(02)00259-8

[32] S. Astashkevich, M. Kaning, E. Kaning, N. Kokina, B. Lavrov, A. Ohl and J. Ropcke, "Radiative characteristics of 3p Σ, Π; 3d Π^-, Δ^- states of H_2 and determination of gas temperature of low pressure hydrogen containing plasmas", J. Quant. Spectrosc. Radiat. Transfer **56**, 725–751, 1996, https://doi.org/10.1016/S0022-4073(96)00103-3

[33] P.J. Bruggeman, N. Sadeghi, D. C. Schram and V. Linss, "Gas temperature determination from rotational lines in non-equilibrium plasmas: a review", Plasma Sources Science and Technology, **23**(2), 023001, 2014, https://doi.org/10.1088/0963-0252/23/2/023001

[34] P.J. Bruggeman, F. Iza and R. Brandenburg, "Foundations of atmospheric pressure non-equlibrium plasmas, Plasma Sources Sci. Technol. **26**(12), 123002, 2017, https://doi-org.biblioteca.ibt.unam.mx:8080/10.1088/1361-6595/aa97af

[35] H. B. Baniya, R. Shrestha, R. P. Guragain, M. B. Kshetri, B. P. Pandey, and D. P. Subedi, " Cold Atmospheric Pressure Plasma Jet for the Improvement of Wettability of Polypropylene", International Journal of Polymer Science, **2020**, 3860259, 2020, https://doi.org/10.1155/2020/3860259

[36] A. Czernichowski, H. Nassar H, A. Ranaivosoloarimanana A.A. Fridman, M. Simeek, K. Musiol, E. Pawelec and L. Dittrichova, "Spectral and Electrical Diagnostics of Gliding Arc" Acta Phys. Pol. A89, 595, 1996, doi:10.12693/APhysPolA.89.595

[37] K. Ollegott, P. Wirth, C. Oberste-Beulmann, P. Awakowicz, and M. Muhler, "Fundamental Properties and Applications of Dielectric Barrier Discharges in Plasma-Catalytic Processes at Atmospheric Pressure", Chem. Ing. Tech, **92**(10), 1542–155, 20208, doi:10.1002/cite.202000075;

[38] M. Keidar, A. Shashurin, O. Volotskova, M.A. Stepp, P. Srinivasan, A. Sandler, and B. Trink, "Cold atmospheric plasma in cancer therapy", Physics of Plasmas **20**, 057101, 2013, https://doi.org/10.1063/1.4801516;

[39] S.H. Jafari, A. Kalati-vahid, H. A. Khonakdar, A. Asadinezhad, U. Wagenknecht, D. Jehnichen. "Crystallization and melting behavior of nanoclay-containing polypropylene/poly (trimethylene terephthalate) blends", Express Polymer Letters, **6**(2), 148–158, 2012, 10.3144/expresspolymlett.2012.16

[40] M.R. Abadchi, A. Jalali-Arani, "Crystallization and melting behavior of polypropylene(PP) in (vulcanized nanoscale polybutadiene rubberpowder/PP) polymer-nanocomposites", Thermochimica Acta, **617**, 120-128, 2015, https://doi.org/10.1016/j.tca.2015.08.027

[41] M.M. Favaro, M. C. Branciforti, R. E. S. Bretas, "A X-ray Study of β-Phase and Molecular Orientation in Nucleated and Non-Nucleated Injection Molded Polypropylene Resins", Materials Research, **12**(4), 455-464, 2009, https://doi.org/10.1590/S1516-14392009000400014

[42] A.S. Mielsen, D. N.Batchelder and R. Pyrz, "Estimation of crystallinity os isotactic polypropylene using Raman spectroscopy", Polymer **43**(9), 2671-2676, 2002, https://doi.org/10.1016/S0032-3861(02)00053-8

[43] T. Kida, Y. Hiejima, K.-H. Nitta, "Molecular orientation behavior of isotactic polypropylene under uniaxial stretching by rheo-Raman spectroscopy", eXPRESS Polymer Letters, **10**(8), 701–709, 2016, 10.3144/expresspolymlett.2016.63

[44] N. L. Singh, A. Qureshi, N. Shah et al., "Surface modification of polyethylene terephthalate by plasma treatment," Radiation Measurements, vol. **40**(2-6), 746–749, 2005, https://doi.org/10.1016/j.radmeas.2005.01.014

[45] F.J. Baltá Calleja, Microhardness relating to crystalline polymers. In: Kaush H.H., Zachman H.G. (Eds), "Characterization of Polymers" in the Solid State I: Part A: NMR and Other Spectroscopic Methods Part B: Mechanical Methods. Advances in Polymer Science, **66**, 117-148, 1985, https://doi.org/10.1007/3-540-13779-3_19

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1258-1266 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

LEACH Based Protocols: A Survey

Nour Najeeb Abdalkareem Qubbaj*, Anas Abu Taleb, Walid Salameh

Computer Science Department, Princess Sumaya University of Technology, Amman, 11941, Jordan

ARTICLE INFO

Article history:
Received: 10 October, 2020
Accepted: 06 December, 2020
Online: 16 December, 2020

Keywords:
Fuzzy Logic
Energy efficiency
Hierarchical routing protocols
LEACH
CHs (clusters Head)
WSN

ABSTRACT

Advances in the world of communications and information technology, as well as the urgent necessity to monitor particular areas and regions, have led to a considerable and influential development in the world of wireless sensor nodes. As they are small, low-cost multi-purpose nodes with limited energy and capabilities. The most important points that deserve research and solution is the problem of energy conservation and increasing the lifetime of the network as a whole. Whereas, many routing protocols have been proposed in the wireless sensor network, some of which are based on homogenous networks while others are not, some are location-based, some are based on hierarchical orientation, some are based on the hybrid approach and many more. Despite the long time that has passed since the discovery and activation of the Low Energy Adaptive Clustering Hierarchy (LEACH) protocol, it and its descendants still have the lion's share of attention, research, and development. In this research, we will be presenting the most prominent routing protocols while pointing to some pros and cons of each protocol.

1. Introduction

"This paper is an extension of work originally presented in "11th International Conference on Information and Communication Systems (ICICS). IEEE, 2020" [1].

Wireless sensor networks (WSN) can be defined as collection of thousands or hundreds of thousands of small self-configuring sensor nodes that bond together and collaborate to aggregate information and send it to a central station. The purpose simply is to get rid of the human aspect [2]. Usually, (these nodes being distributed randomly by aircrafts to serve military interests, detecting natural disasters like fires and earthquakes, also it plays a huge role in some peaceful applications such as detection of a specific disease in a specific crop, health care, electronic sales, and others [2]. On the other hand, routing is defined as the process of exactly determining the path from the main source to the final destination, and it must be fast and reliable as well.

As a reminder, these nodes are thrown from aircraft randomly and to places that are difficult for humans to reach, so they must remain functional as long as possible. Unfortunately, the only energy source for them is the battery and it is known that the battery depletes quickly and the node dies at that time. Therefore, the most important challenge is to keep the nodes alive as much as possible and distribute the total energy wisely, and not lose their energy, especially in the early stages. Researchers have provided a massive number of routing protocols and classify them on several levels, considerations, and parameters [1].

The rest of our research is split into the following sections: in section II the problems of routing are presented. In section III. Most important WSN routing methods. In section IV. Furthermore, Low energy adaptive clustering hierarchy (LEACH), the before the last section V. Different routing protocols based on (LEACH) protocol, the last VI. Conclusion and future work subsequently reference.

2. The Problems of Routing

In WSN, routing protocols can be considered to be dependent on the application, the data to be collected, and the good use of limited energy. A good routing protocol is a protocol that has several advantages, including simplicity, ability to know its energy level, ability to expand and adapt under the limited capabilities of the nodes, to name a few: power, memory, computation capabilities, and bandwidth allocated to the network according to [3]. The main purpose of the network design is to

*Corresponding Author: Nour Najeeb Abdalkareem Qubbaj, Email: nou20188107@std.psut.edu.jo

deliver all the information from the deployed nodes to the base station (BS) with the lowest possible power. This design faces several obstacles, including [1,3]:

2.1. Deployment of nodes

Deployment of nodes relies on the type of application that massively affects the efficiency of the routing protocol. Where the deployment of the sensor nodes may be random (self-organizing), by plane, for example, or be in a specific and thoughtful place. The node may be fixed, not moving, and it may be mobile, or all of them are stationary and the main station is mobile.

Location and movement play an important bit part in energy efficiency.

2.2. Conservation of energy

Data and information are transmitted by radio through the network, the power of the radio is inversely proportional to the square of the distance traveled, and the power consumption will be affected in the event of obstacles. Therefore, it is better to use the multi-hop approach rather than one-hop. It is worth noting that the one-hop orientation is suitable for a network with a small number of nodes and close to each other, which confirms that energy is affected by the paths that you take.

2.3. Tolerance of Faults

Error tolerance can be described as the ability of the network to continue its operation in the presence of obstacles. Such as the nodes death. For example, the routing protocol must quickly find an alternate route to cancel an insufficiency of information to the main (Base) station.

2.4. WSN Scalability

The number of nodes in the target area to be monitored will exceed hundreds or thousands. Thus the routing protocol must be able to accommodate this large density of nodes.

2.5. Constraints on hardware

All components of the nodes including, the power unit (battery), the processing unit (memory and the microprocessor unit), the sensor unit (sensor), and the radio communication unit, must be smaller in size, less expensive, and energy-consuming.

2.6. The surrounding environment

The node must be prepared for the difficult conditions in which it will monitor.

2.7. Transmission Media

From the name (WSN), we find that the mean of communication is wireless communication, as it is vulnerable to the rate of error and greatly influences the correct operation of the network.

2.8. Models for data delivery

Delivering the data to the main station is related to the type of application, it may be continuous, where all the data is sent to the station continuously and periodically. While if the connection is related to an event, the node sends its data when a specific event occurs. The communication may also be inquiry-oriented, so that when the station needs to inquire about a specific piece of information from the nodes, and the connection may be a hybrid connection consisting of two or more of the mentioned methods. Finally, it's worth pointing that the routing protocol being influenced by the type of that connection significantly.

2.9. Information aggregation (fusion)

The cluster head (CH) collects data from more than one node, so there is a possibility of duplicating data, especially adjacent nodes, so CH does this data fusion process to minimize the volume of the transmitted data to the main station.

2.10. Capacity of node

The capacity of a node this means that it is possible to assign a specific sensor node to one of the functions(tasks) (relaying, sensing, aggregation) according to the application, instead of giving the node all the functions(tasks) together to avoid consuming excess energy.

3. Most Important WSN Routing Methods

Route pathways may be defined in one of three forms i.e. reactive, proactive, or mixed, the reactive paths calculate the paths when the actual necessity of them. While the proactive route precedes the occurrence of the real event and stores the possible paths in a table for each node, it is heavy-duty due to the huge number of network nodes. The last type is the combination of the characteristics of both proactive and reactive see Fig 1.

Based on the design of network flow, the routing protocols are classified into flat-routing, hierarchical and location-based routing. In flat-based routing, all nodes execute the same task such that data is transferred from each sensor node with a strong data repetition [3].

Hierarchical Protocols, the clusters are formed in this type to raise energy efficiency and accommodate a large number of nodes and cover a larger area of the network with better efficiency by applying two main steps. The first one is determining the head of the cluster. The second is the path from the head of the cluster to the main station. [3].

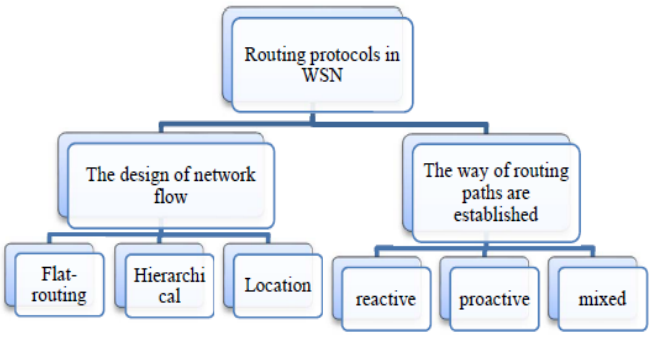


Figure 1: Routing protocols kinds [1].

Location-based protocols, some applications may require the

location of the node to calculate the distance between the nodes. There are several ways to calculate this, including (Global Positioning System) GPS or other ways [3]. See Fig 1.

The most rudimentary way to transfer data from the nodes to the base station is that each node does this by itself without an intermediary, which is known as Direct Transmission (DT), as it is considered one of the flat and simple protocols. And another example of it too is Minimum Transmission Energy (MTE). Whereas, the nodes that are farther from the base station and the nodes with a large volume of data require more energy than others. Therefore, the hierarchical protocols are distinguished from the flat and simple protocols, such as the ones mentioned.

The working principle of the clustering methods of WSN, is based on dividing the entire network into groups called the clusters, and in each cluster, there is a header that gathers data from the rest of the nodes, merges them, reduces their size, and sends them at once to the Base Station (BS) and the role of this head is exchanged with the rest of the nodes in the clustering according to the followed protocol [4]. Cluster formation has several characteristic and advantages, so it excels over flat protocols, including [5]:

- Easily handle the complexity of the network: The network is composed of nodes that act as the head of the cluster and that communicate directly with the main station on behalf of all the nodes of the cluster, and other nodes that collect data from the surrounding environment and send it to the head of the cluster that does not have another load, thus the master station receives data from a small number of nodes rather than all of the cluster members and thus facilitates control of the network as a whole see Figure 2.
- Aggregation / Fusion of information: Where the CH integrates and collects information from the rest of the members and reduces its size before sending it to the base station, the base station receives data from a small number of nodes rather than all of the cluster members.

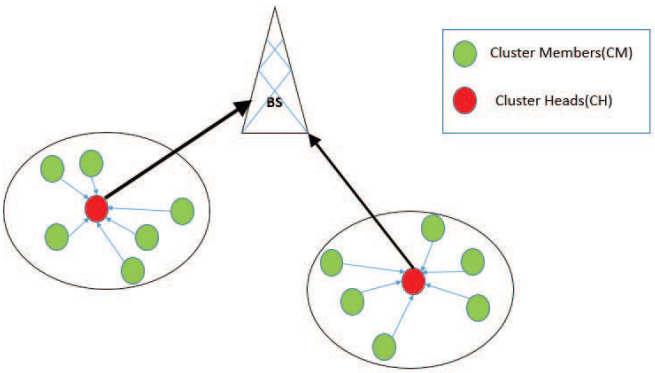


Figure 2: Creation of clusters [6].

- The life of the network is stronger and longer: By rotating the head of the cluster in every round to distribute an equal distribution of energy among the network members.
- Reducing collisions and traffic: In a WSN with a flat communication model, all the nodes must send their information to the main station at the same time. In the clustering organized, the communication with the main station is limited only to the head of the clusters.

4. Low Energy Adaptive Clustering Hierarchy (LEACH)

It is the parent protocol for hierarchical routing based on Clustering proposed by W.B. Heinzelman [7]. As has been reported, the clustering protocols have proven effective in distributing energy consumption among members and increasing the life of the network as a whole.

LEACH works on the principle of dividing the network into random groups (clusters) and then choosing a node from each group to be the head node to represent its group while connecting to BS. The role of the cluster head rotates to keep the energy consumption distributed among the network members. It's worth pointing that the head nodes have the heading role as an additional role to their original tasks (tasks of the non-cluster head nodes). It is also a kind of one-step approach (one-hop routing) to the BS. The LEACH principle is divided into several rounds, and each round contains two basic stages: One is the Setup stage, and the other is the Steady stage as shown in Figure 3.

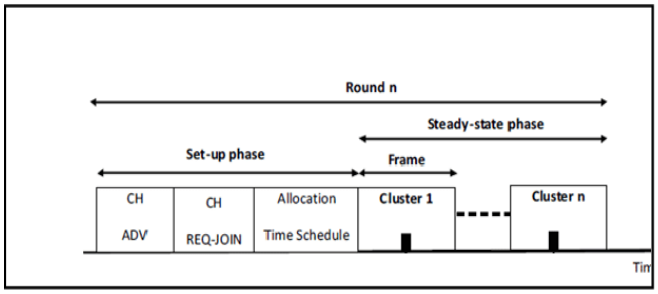


Figure 3: Operation of LEACH [8].

The node must initially pick a random number (n). If this number (n) is less than the threshold specified in the first equation T(n), then the node is selected as the cluster header (CH).

$$T(n) = \begin{cases} \dfrac{p}{1 - p.\left(r \bmod \frac{1}{p}\right)} & if\ n \epsilon\ G \\ 0 & otherwise \end{cases} \quad (1)$$

while p is the possibility of a cluster head (CH), G is a list of nodes that will never be chosen as cluster-head (CH) nodes until the 1 / p level.

After selecting the heads of the cluster according to the number (n), each cluster head will broadcast its status to all other nodes via CDMA (Code Division Multiple Access). After receiving broadcast messages from CHs, each normal node uses the Received Signal Strength (RSS) to determine which CHs are closest to it. Then the CH creates TDMA (Time Division Multiple Access) to divide a schedule for members of the same group (cluster) so that the normal node is activated and connected to CH at the time allotted to it according to the TDMA schedule. It will

be inactive at other times to preserve its energy, which means that CH deals with all nodes from the same cluster and is the only one who deals with BS from cluster members [9]. Note that the period of the Steady stage is longer than the Setup stage.

The significant drawbacks of LEACH are the following, according to [1,2,10]:

- Since LEACH (one-hop routing) to the BS is never suitable for a large network of sensors.
- Not suitable for time-compulsory applications due to delays.
- Unacceptable distribution for cluster heads: as have been stated earlier, a random selection of CH that is not based on clear criteria such as the distance between CH and BS, causes an imbalance in the power consumption of the network and this is considered a significant drawback.
- Unacceptable selection of cluster head: The randomness in choosing the head of the cluster so that the primordial energy of each node is not taken before it is considered as the head of the cluster.
- Note that the types of data delivery to the base station may be continuous, event-related, directed to an inquiry, or a mixture between them. The LEACH protocol is more appropriate with continuous delivery only.
- Different sizes for each cluster: The randomness of choosing the cluster causes the existence of large-numbered clusters and other small-number without specific study due to randomness only.

Several routing protocols have been proposed to solve some of these problems faced by the LEACH protocol, the most prominent of which will be presented in section Five.

5. Different Routing Protocols Based On (LEACH) Protocol

5.1. Low Energy Adaptive Clustering Hierarchy Centralized (Leach-C)

LEACH-C is one of the LEACH versions, which is a centralized protocol. It is divided into several rounds as the LEACH protocol and each round into two basic stages. This protocol assumes that each node can communicate directly with the BS and can calculate its energy and current position.

Unlike LEACH, LEACH-C take into consideration the distance from the node to the base station. The base station determines who is the CH in this round based on its energy and the distance between the node and the base station.

Where in the setup stage all nodes send their energy and location to the base station, the station calculates the average value of the energy values, and the node that contains energy higher than the average value is elected as CH for this round. In the Steady stage, the non-head nodes send their information to the head of their cluster, and the head of the cluster, in turn, collects the data and sends it back to the base station [5].

5.2. LEACH-Balanced (LEACH-B)

The LEACH-B protocol overcomes the original LEACH protocol by taking into account the remaining energy of the node and also strives to maintain the optimum number of cluster heads per round of the protocol. The optimum value for the number of cluster heads varies from 3 to 5% based on what was mentioned in [11]. Thus reduce energy consumption, it is a decentralized protocol, LEACH-B has improved energy performance compared to the parent LEACH protocol see Figure 4.

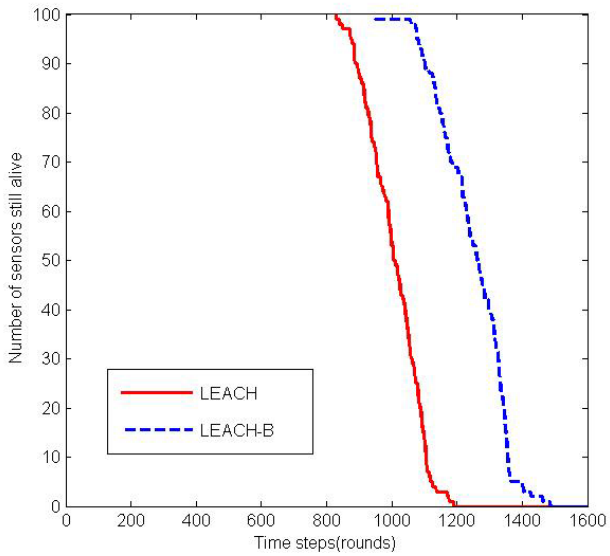


Figure 4: System life of LEACH and LEACH-B [11].

5.3. Fixed Number of Clusters LEACH (LEACH- F)

This protocol is a centralized approach, like the LEACH-C protocol. clusters are set, and rotation is just for the CH inside the same cluster. The steady-state is identical to the original LEACH since the number of clusters has been determined in advance, the energy wanted for re-clustering is saved. It is not elastic in the case of extension, elimination, or death of a sensor node [12].

5.4. Advanced Low Energy Adaptive Clustering Hierarchy (A-LEACH)

As a result of the overload on CH, it is likely to die early before the other nodes, A-LEACH prolongs the duration of stability (the time before the death of the first node). It worth mentioning mention that A-LEACH is a heterogeneous protocol.

Its working principle is as follows: Every sensor knows how to continue each round utilizing the synchronized clock. Let's assume that the total number of nodes is (n), and (m) is the number of nodes that contain more power than others (also called (CGA) nodes that chosen as gateways or CH), so the number of nodes other than nodes can be found using equation 2, see Fig 5.

$$head = (1 - m)\,.\,(n) \qquad (2)$$

A-LEACH is defined as a Distributed Algorithm because the selection of clusters is independent of the base station. CAG nodes can start to transmit data even when all the non-cluster head nodes collapse [12].

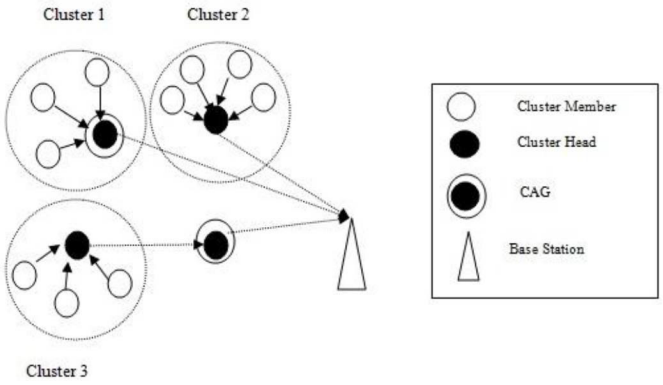


Figure 5: LEACH -A system [1].

5.5. Energy Low Energy Adaptive Clustering Hierarchy (LEACH-E)

E-LEACH is a variant of the LEACH protocol that takes into account the residual energy of a node selected as a cluster head. So the node with the maximum and most appropriate energy is chosen. In other words, E-LEACH is concerned with the number of CHs where the large number of cluster heads consumes higher energy while the small number of cluster heads will cover a greater distance. If the nodes are scattered away from the BS, it will consume more energy. The most appropriate number of cluster heads should be determined, their location and energy. The minimal spanning tree of cluster heads is used in the E-LEACH see Figure 6 [1,2].

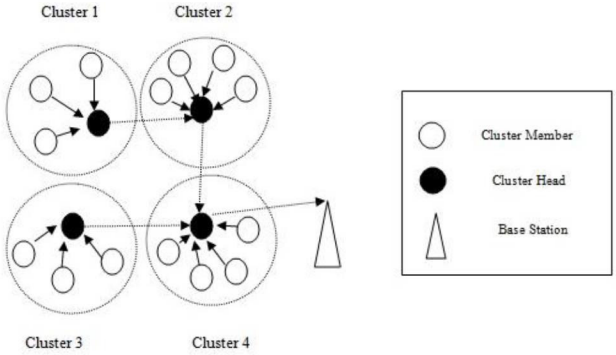


Figure 6: LEACH -E system [1].

5.6. Vice Cluster Head LEACH (V- LEACH)

V-LEACH is another version of the LEACH protocol. It is known that the load on the cluster head is greater because it aggregates data from the rest of the nodes and performs simple operations to reduce its size and transfer it to the main station.

If the CH energy is reduced, it will die thus, the whole cluster is useless because it does not communicate its information to the main station. V-LEACH suggested that the presence of a deputy head of the cluster in the event of his death. The vice-CH is responsible for transferring the information to the main station thus, the whole cluster remains connected with the main station, but in case of death of the vice-CH, there is no solution in this protocol see Figure 7 [10].

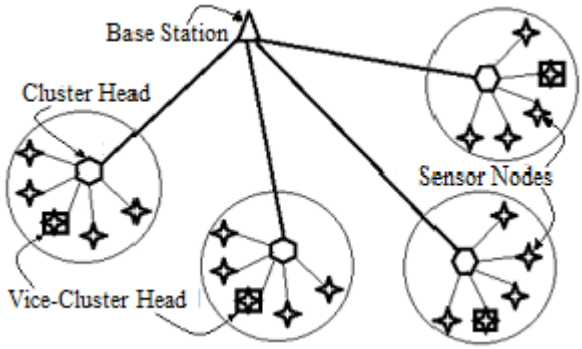


Figure 7: LEACH -V system [13].

5.7. I-LEACH

A new algorithm, I-LEACH, has been proposed, which takes into account the following: the geographic location of the CH, the number of neighbors, and the residual energy to improve network performance and lengthen its life. So the node with the most residual energy and closest to the BS is chosen. This protocol overcame a set of drawbacks of the original LEACH protocol some of them: Different size for each cluster, Unacceptable distribution for cluster heads, and Unacceptable selection of cluster head.

I-LEACH has been evaluated and proven to be effective in average energy consumption, the number of data packets received, and the flexibility to change the location of the BS in the network. Where it reduced energy consumption by approximately 62%, increased the percentage of data packets received by approximately 56%, and increased the life of the network each by 65% at least [14].

I-LEACH, has some of the following assumptions:

- All nodes have the same primary energy, are not rechargeable, and in the same random position when the nodes are thrown from the airplane. They remain in place and never move.
- All nodes have their identifier and can know their exact location and the remaining energy in its battery.
- It is known that each node contains a processing unit and a memory unit, and it can also define the sensed and the aggregated data.
- A node can sense different types of data.
- Each node can communicate directly with the BS and has the power required for this.

5.8. Cell Low Energy Adaptive Clustering Hierarchy (Cell-LEACH)

Cell-LEACH is another version of the LEACH, as in the original LEACH protocol, the entire network is split into groups (clusters). Here too, the network is divided into groups where each group is divided into 7sub-groups called cells. Each group has a cell head whose task is collecting the information from the members of the cell and communicates directly with the head of the cluster. The cell is randomly assigned and then chosen based on its remaining energy to be cell head see Figure 8 [15]. Cell-LEACH has proven its ability to lengthen the network life without

reducing efficiency as can be seen in Figure 9.

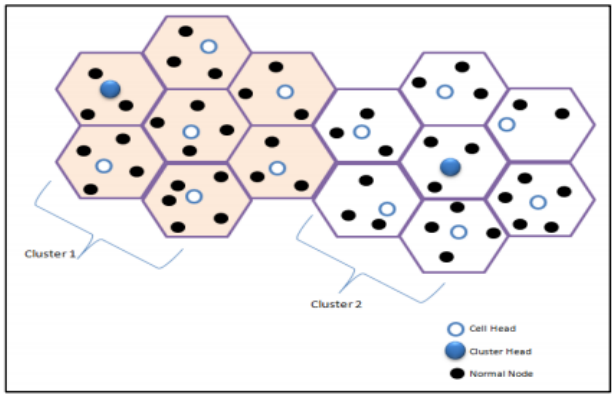


Figure 8: LEACH -Cell system [16]

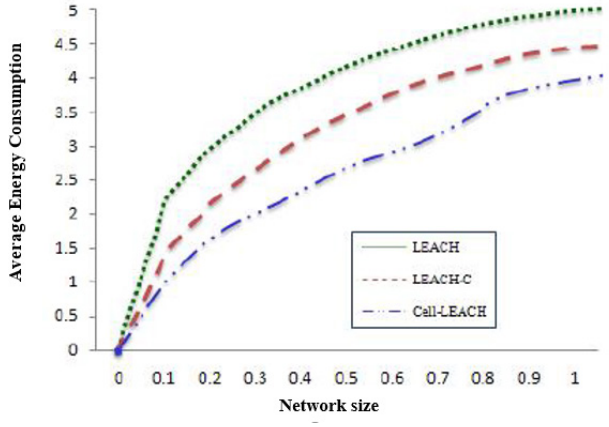


Figure 9: Average power consumed with network size [15].

5.9. Multi-Hop LEACH

5.9.1. TL-LEACH (Two levels LEACH)

Two levels LEACH) was suggested in [17]. Instead of aggregating and transmitting data in one hop, the data in CH being aggregated at a certain level and transferred to a closer CH level to BS thus, reducing the transmission power, which enhances energy efficiency, and extending the network life. To illustrate the idea further, we have Fig 10 that shows that there are 3 CHs at the top level is closest to BS and also 5 CHs on the second level closest to the normal nodes sensor, and the other 10 SN (Sensor Node) normal nodes at the final level. It has been shown that TL-LEACH increases network life by nearly 30 percent and also increases the percentage of packets received by 20 percent [17].

5.9.2. Multi-Hop LEACH:

Due to the disadvantage in " LEACH (one-hop routing) to the BS", and as mentioned earlier, the task of the cluster head is to transfer the collected data from the non-head nodes to the BS. This implies that if the distance is large, then the energy of the cluster head runs out quickly. So a midway cluster head has been considered to reduce distance, thus reducing energy consumption this is known as multi-hop communication [6].

SN	**Sensor Nodes**
CH_{ij}	**Cluster Head to second level**
CH_i	**Cluster Head to first level**
BS	**Base Station**

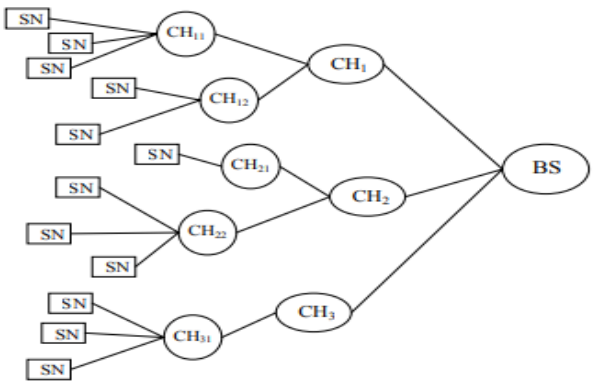


Figure 10: The TL-LEACH network format after the setup phase [17].

The original LEACH protocol does not take into account the primitive energy of the head of the cluster, CH may lose its energy more quickly than others and die, and this cluster loses its connection with the main station. Therefore, it has been proposed to amend the basis of CH selection to take into account energy.

A term was added to the original formula for the original LEACH protocol as in Equation 3 [6].

$$T(n) = \begin{cases} \dfrac{p}{1 - p.\left(r \bmod \frac{1}{p}\right)} \cdot \dfrac{E_{rem}}{E_{totnet}} & if\ n \epsilon\ G \\ 0 & otherwise \end{cases} \quad (3)$$

while p is the possibility of a cluster head (CH), G is a list of nodes that will never be chosen as cluster-head (CH) nodes until the 1 / p level. E_{rem} is the remaining energy level of the node, E_{totnet} is the total energy network. This happens in the setup phase.

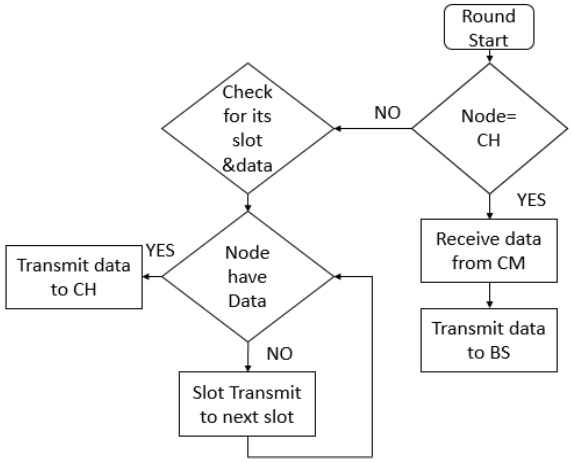


Fig 11: Freshly revised TDMA description [6].

In several protocols associated with an event, the most appropriate thing to do is to take advantage of the free time of the nodes. For example, when a node does not have information to send to the main station, an allocation technique activates to allocate the time assigned to the free node in the TDMA schedule to another node that has information to send. Thus, reducing energy and time consumption as well, as can be seen in Fig 11 [6].

5.10. Clustering Process Based on Fuzzy Logic

Fuzzy logic simulates human decision-making and experience to judge things and deals with decisions better than probabilistic approaches. Also, it is more suitable for real-time applications [18]. Due to the suspicion and doubt in the WSN environment, some protocols use Fuzzy Logic technology, as it simulates human logic less than others, which include computational complexities and the need for memory. Fuzzy Logic protocols used to choose CH more effectively and flexibly due to its flexibility and error tolerance choose CH more. The Fuzzy Logic model requires four steps to accomplish, namely:

- Fuzzification: In it, the inputs are inserted with clear values and are distorted into a fuzzy group, and this happens in the fuzzifier.
- Rule evaluation (Fuzzy Rule base): This is storing Rule IF-THEN.
- Fuzzy Inference Engine: This engine handles both input and law IF-THEN to produce a fuzzy group.
- Defuzzification: Converts the fuzzy group to clear values.

These steps are shown in Figure 12. Most of the Fuzzification methods use Mamdani's Method Fuzzy Inference technique to select the CH.

This block is called fuzzy logic controller (FLC). Not too long ago, the authors began leveraging Fuzzy Logic (FL) and incorporating it into WSN to lengthen the life of the network and improve the method for selecting the appropriate CH. Some of the popular Clustering Process Based on Fuzzy Logic will be featured in this section:

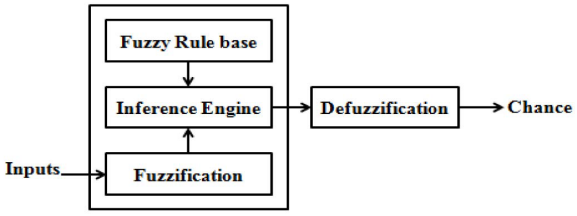


Figure 12: The Fuzzy Inference Process structure diagram.

5.10.1 Cluster Head Election using Fuzzy Logic (CHEF):

In this type, CH is chosen by entering two inputs: power and local distance. (The local distance is the total distance between the temporary CH in that round and the cluster members (normal nodes) within a predetermined fixed competition radius) [19]. So that the likelihood of the closest and more energy nodes to be selected as CH is higher than others, As the experiments of [18] show that CHEF is 22.7% more effective than LEACH.

However, some of the disadvantages of this type are that not all sensor nodes have a GPS device and sometimes are not able to know their geographical location [18]. Besides, this type is not suitable for network sized about 200m × 200m or more [19,20]. clusters closer to BS, run out of energy faster because traffic becomes congested the closer to BS, so an unequal clustering has been suggested to balance energy consumption [19].

5.10.2 Energy Efficient Unequal Clustering (EEUC):

This type divides all nodes into clusters that are not equal in number so that the clusters close to the BS contain fewer nodes. The distant clusters contain a greater number because the closer we get to the base station, the congested the traffic becomes, causing the closest nodes to consume more energy from the farthest [21]. As in the LEACH, the choice of CH is random here is random as well, a random number between 0 and 1 assigned for each node. The number indicates whether the node is going to be a CH or not [19]. Experiments proved that EEUC is more efficient than LEACH see Figure 13.

5.10.3. Energy-Aware Unequal Clustering with Fuzzy (EAUCF):

Battery level goes down to the CH [19]. In other words, when the residual energy of the temporary CH and the distance to the BS is large then the competition radius is large and vice versa [21]. Let us say that each temporary CH will forward its remaining energy and also if it is within the competition radius of this CH, another temporary CH will be compared between their energy, and the least energy will be out of competition and not CH [20]. In the experiments. It has been proven that EAUCF is more efficient than LEACH, CHEF, and EEUC see Figure 13. While this type does not pay attention to the degree of the node, it means the density of the internal cluster, as the density of communication inside the cluster costs energy and reduces the life of the wireless network as a whole [20].

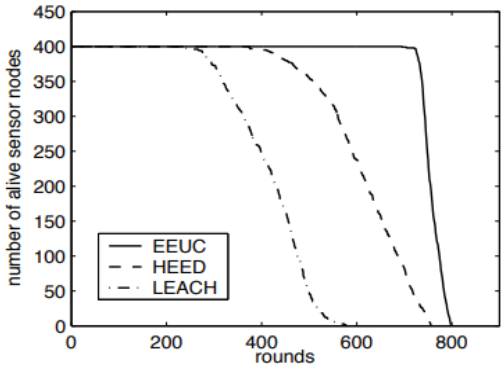


Figure 13: Over time, the number of live sensor nodes [22].

5.10.4. Fuzzy Cluster Head Selection Algorithm based on LEACH:

The authors of [16] noticed that when more than one entry is taken into account to determine one output (appropriate CH selection), Fuzzy Logic, is more accurate and appropriate than others. This type requires three fuzzy input variables. The first variable expresses the number of surviving neighbors, the second

variable expresses residual (remaining) energy, the last variable expresses the distance to BS. The one variable of fuzzy output is named a chance [20] as shown in Figure 15. The CH is selected with the most living neighbors, the highest residual energy, and the least distant to BS [20].

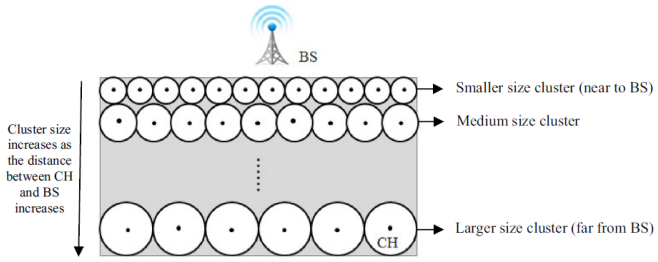


Figure 14: Unequal clustering architecture at WSN [22]].

Table 1: At round 250, the total remains energy for each algorithm [19]

Algorithm	Total Remaining Energy (J)
LEACH	27.14
CHEF	43.66
EEUC	44.78
EAUCF	46.41

Just like the LEACH, in each cycle, every node picks a random number between 0 and 1, and if the number was chosen is less than the threshold T, then it becomes within the Concurrent CH. Then each Concurrent CH node calculates its chance from the output of the Fuzzy and shares it with the rest of the Concurrent CH who are within its competition radius. Consequently, each Concurrent node CH has a list of all CH node and its chance, where the highest is elected as CH node, and the rest within the competition radius is a normal node and the elected CH node joins. Also, as usual, the elected CH node is required to create a TDMA schedule and to publish it to the rest of the cluster. The normal nodes send their information to the selected CH node in the time allotted to them and turn off their transmitter device at other times to conserve energy [20]. Compared to LEACH, this form improves network life by 58.38 percent see Table 1.

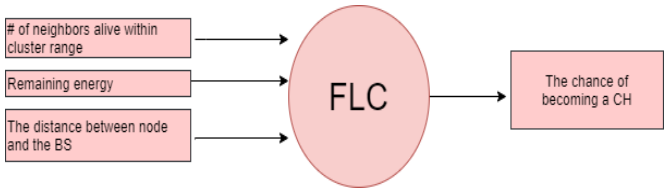


Figure 15: The Fuzzy-Logic system which controls the clustering

5.10.5. A Fuzzy-Logic based Energy-Efficient Clustering algorithm (FLEEC):

The authors of [21] noticed that when more than one entry is taken into account to determine one output (appropriate CH selection), Fuzzy Logic, is more accurate and appropriate than others. It should be noted that the process of calculating the distance between the node and the BS is done based on the Received Signal Strength (RSS).

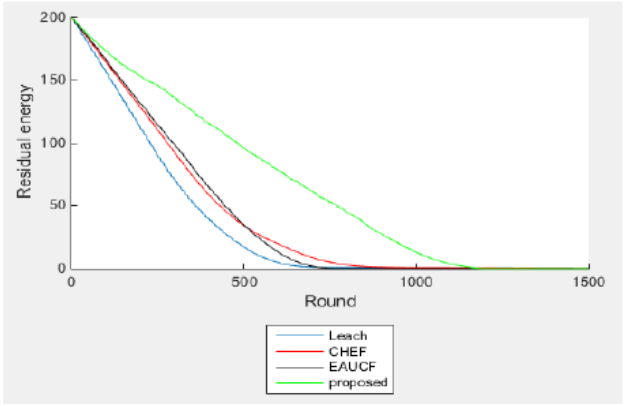


Figure 16: The full remaining energy of every cycle [20].

If clusters of the same size are created without taking into account the distance to the BS, then a "hot spot problem" begins to arise. As was evident from Figure 14, it requires more energy for the cluster closer to the BS. To create uneven clusters, the first fuzzy logic controller (FLC) is constructed, which takes two inputs (The distance to the base station and Node Density), and one output (cluster Radius) see Figure 15. Then the head of potential clusters measures the total distance of the neighbors that are in their radius [21]. The total distance to the CH and Residual energy are inputs to the second FLC, and the output is the probability of each node to be CH. The node that has the highest likelihood to become CH is when the node has the largest residual energy and more neighbors.

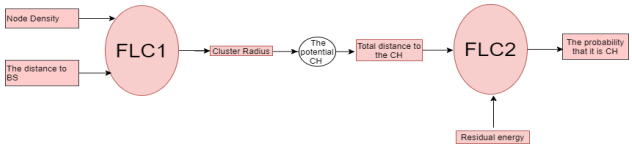


Figure 17: The Fuzzy-Logic system which controls the clustering in (FLEEC)

FLLEC is composed of three major stages [21]:

- Determination of cluster size process: First of all, the BS node broadcasts a welcome message to every node in the region, and then each node would like to calculate its distance from BS via Received Signals Strength (RSS). Each node sends to the BS node an ACK message, which contains its ID, residual energy, geographic distance from the BS node. Then the BS computes the density distribution from the information it obtained from every node. Here comes the role of the first FLC by calculating the corresponding cluster radius for each sensor node, and this information is saved in the memory of each node.
- Determination of likelihood to be a cluster head process: After the cluster radius determined successfully, all nodes would like to broadcast a message of exploring the neighbors around it, and each would like it to receive that message followed an ACK message. After that, the nodes calculate the total distance of their neighbors according to RSS. Here comes the role of the second FLC to calculate the likelihood to be CHs depending on the fuzzy rules reserved in the fuzzy rules base.

- Determination of CH process: Each Node broadcasts the HELLO message to all of its neighbors within the competition radius that it has both CH and its ID. During this information exchange process, the CH that has the greatest probability is the one who becomes CH and the rest is a normal node. CH broadcasts an advertisement message (ADV). If the normal node receives more than the ADV, then it determines the most suitable one that will belong to it, then each node sends a request message to the most appropriate CH.

Comparing FLEEC to LEACH (the main clustering algorithm) and CHEF (the main fuzzy-logic-based algorithm) algorithms. As can be seen from Figure 18, the results showed the superiority of FLEEC over both LEACH and CHEF, especially in the issue of energy consumption.

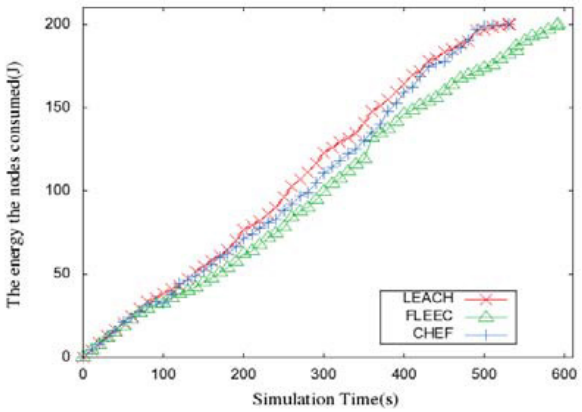


Figure 18: The energy consumed by the three main algorithms is comparable [21].

6. Conclusion and future work

Due to the emergence of nanotechnology and the persistence of multiple computing fields, the field of WSN and the optimal management of node energy remains a fertile field for research and experiment. This study covered two main topics, presenting and discussing some of the most paramount routing protocols and describing their key types and issues, and going through the most prominent hierarchical protocols and their development process.

We concluded from this research that the clustering design is one of the most energy-saving designs. Where we noticed that Multi-Hop LEACH saves energy because it reduces the distance over the CH and also excels and the fuzzy logic strategy based on clustering is superior to other clustering strategies because it is in line with human logic and its ability to convert more from one input to one output, and as it was presented in the research, it was saving more energy than any other algorithms.

In future work, machine learning with the wireless sensor network can be studied to further rationalize energy consumption and merge data before sending it to BS.

References

[1] N. Qubbaj, A.A. Taleb, W. Salameh, "Review on LEACH Protocol," 2020 11th International Conference on Information and Communication Systems, ICICS 2020, 414–419, 2020, doi:10.1109/ICICS49469.2020.239516.

[2] K. Sohraby, D. Minoli, T. Znati, Basic Wireless Sensor Technology, 2007, doi:10.1002/9780470112762.ch3.

[3] P. Kumar, M.P. Singh, U.. Triar, "A review of routing protocols in wireless sensor networks," 2008 International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2008, **1**(4), 1–14, 2012, doi:10.1109/WiCom.2008.946.

[4] A. HNINI, A. EZZATI, M. FIHRI, A. HAJAMI, "Time Slots Investment of a dead node in LEACH protocol on Wireless Sensor Network," International Journal of Advanced Computer Science and Applications, **4**(2), 8–12, 2014, doi:10.14569/specialissue.2014.040202.

[5] A. Kumar, R.. Saini, E. Brijbhusahn, "Comparison and Review of LEACH ," (June), 2016.

[6] C.T. Sony, C.P. Sangeetha, C.D. Suriyakala, "Multi-hop LEACH protocol with modified cluster head selection and TDMA schedule for wireless sensor networks," in Global Conference on Communication Technologies, GCCT 2015, 2015, doi:10.1109/GCCT.2015.7342720.

[7] J.S. Leu, T.H. Chiang, M.C. Yu, K.W. Su, "Energy efficient clustering scheme for prolonging the lifetime of wireless sensor network with isolated nodes," IEEE Communications Letters, **19**(2), 259–262, 2015, doi:10.1109/LCOMM.2014.2379715.

[8] N. Nejah, "Propagation Approach for Clustering in Wireless Sensor Network," (March), 2018.

[9] S. Nasr, M. Quwaider, "LEACH Protocol Enhancement for Increasing WSN Lifetime," 2020 11th International Conference on Information and Communication Systems, ICICS 2020, 102–107, 2020, doi:10.1109/ICICS49469.2020.239542.

[10] S. Varshney, R. Kumar, "Variants of LEACH Routing Protocol in WSN: A Comparative Analysis," 199–204, 2018.

[11] T. Mu, M. Tang, "LEACH-B: An improved LEACH protocol for wireless sensor network," in 2010 6th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2010, 2010, doi:10.1109/WICOM.2010.5601113.

[12] H. Dhawan, S. Waraich, "A Comparative Study on LEACH Routing Protocol and its Variants in Wireless Sensor Networks: A Survey," International Journal of Computer Applications, **117**(12), 27–33, 2015, doi:10.5120/20608-3231.

[13] P. Maurya, A. Kaur, "A Survey on Descendants of LEACH Protocol," International Journal of Information Engineering and Electronic Business, **8**(2), 46–58, 2016, doi:10.5815/ijieeb.2016.02.06.

[14] Z. Beiranvand, A. Patooghy, M. Fazeli, "I-LEACH: An efficient routing algorithm to improve performance & to reduce energy consumption in wireless sensor networks," in IKT 2013 - 2013 5th Conference on Information and Knowledge Technology, 2013, doi:10.1109/IKT.2013.6620030.

[15] A. Yektaparast, F.-H. Nabavi, A. Sarmast, "An improvement on LEACH protocol (EZ-LEACH)," Conference Proceedings of 2015 2nd International Conference on Knowledge-Based Engineering and Innovation, KBEI 2015, 956–960, 2016, doi:10.1109/KBEI.2015.7436173.

[16] M. Al-Shalabi, M. Anbar, T.C. Wan, A. Khasawneh, "Variants of the low-energy adaptive clustering hierarchy protocol: Survey, issues and challenges," Electronics (Switzerland), **7**(8), 1–28, 2018, doi:10.3390/electronics7080136.

[17] V. Loscrì, G. Morabito, S. Marano, "A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH)," IEEE Vehicular Technology Conference, **3**, 1809–1813, 2005, doi:10.1109/VETECF.2005.1558418.

[18] P. Nayak, A. Devulapalli, "A Fuzzy Logic-Based Clustering Algorithm for WSN to Extend the Network Lifetime," IEEE Sensors Journal, 2016, doi:10.1109/JSEN.2015.2472970.

[19] H. Bagci, A. Yazici, "An energy aware fuzzy unequal clustering algorithm for wireless sensor networks," in 2010 IEEE World Congress on Computational Intelligence, WCCI 2010, 2010, doi:10.1109/FUZZY.2010.5584580.

[20] W. Abidi, T. Ezzedine, "Fuzzy Cluster Head Election algorithm based on LEACH protocol for wireless sensor networks," 2017 13th International Wireless Communications and Mobile Computing Conference, IWCMC 2017, 993–997, 2017, doi:10.1109/IWCMC.2017.7986421.

[21] Q. Wang, D. Lin, P. Yang, Z. Zhang, "A Fuzzy-logic Based Energy-efficient Clustering Algorithm for the Wireless Sensor Networks," 2018 26th International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2018, 424–428, 2018, doi:10.23919/SOFTCOM.2018.8555848.

[22] C. Li, M. Ye, G. Chen, J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks," 2nd IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, MASS 2005, **2005**, 597–604, 2005, doi:10.1109/MAHSS.2005.1542849.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1304-1311 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Robust Adaptive Feedforward Sliding Mode Current Controller for Fast-Scale Dynamics of Switching Multicellular Power Converter

Rihab Hamdi*,[1], Amel Hadri Hamida[1], Ouafae Bennis[2], Fatima Babaa[3]

[1]*LMSE Laboratory, Electrical Engineering Department, Biskra University, Biskra, 07000, Algeria*

[2]*PRISME Institute, University of Orléans, 21 rue Loigny La Bataille, Chartres Orléans, 28000, France*

[3]*LEC Laboratory, Electrical Engineering Department, Constantine1 University, Constantine, 49042, Algeria*

ARTICLE INFO

Article history:
Received: 07 December, 2020
Accepted: 09 February, 2021
Online: 28 February, 2021

Keywords:
Power Converter
Sliding Mode Control
High Switching Frequency
Multicellular Converters
Hysteresis Modulation

ABSTRACT

Higher efficiency and lower losses are widely considered as the best metrics to optimize, in a high-power converter performance context. To provide a solution to the ever-increase of high switching frequencies challenges, we must explore soft-switching technologies to resolve interface issues and reduce the switching losses. This manuscript describes a comparative analysis between the fixed-bandwidth (FBW) and the variable-bandwidth (VBW) of the hysteresis modulation (HM) based on the conventional sliding mode (CSM) strategy. The two adopted techniques are applied to a bidirectional multichannel DC-DC asynchronous Buck converter. The cells are parallel-connected and operating in continuous conduction mode (CCM). The objective is to have a system that is more stable, more efficient and able to cope with variations in input voltage, load and desired output voltage. That requires, first, to attenuate the non-linearity phenomenon of the conventional sliding mode by placing a hysteresis modulation. Then, after applying this technique, we confronted the dilemma of the variable switching frequency. Our hypothesis was to incorporate a variable bandwidth of the hysteresis modulation. The results obtained under parametric variation clearly show the areas where significant differences have been found between the two approaches. Likewise, they both share several key features. Simulation studies in the MATLAB® / Simulink™ environment are performed to analyze system performance and assess its robustness and stability.

1. Introduction

DC-DC converters, with many different topologies, represent a hotspot in modern technology and have revolutionized switching control. Although the characteristics of power electronics, in terms of nominal voltage and current values, continue to improve, the range of applications continues to expand in several areas including energy storage, automation, transport, high power industrial drives and transmission / distribution of electrical energy [1], [2]. Whereas the needs for higher efficiency, reliability, modularity in high electronics power, multicellular converters are on the rise. Studies [3–6] based on a parallel multiphase topology have shown that the latter provides better performance, lower losses, reduced cost, and fast dynamic responses. Parallel-connected multi-cell DC-DC converters have an important role in several applications such as inverters, powertrain systems and equipment of the power factor correction and they have countless advantages, the major one is to enhance performances such as energy efficiency and high capacity [7–11]. One approach dedicated to small converters is to run them at a higher switching frequency, this will reduce the size of reactive components (inductors and capacitors) and it contributes to the improvement of the output voltage's quality of the converter. High losses in the IGBT switches remains the major limitation of the high frequencies, which reduces efficiency. Significant information and in-depth investigations in the same framework can be found in these studies [12–18].

Higher efficiency and more stringent control features of power electronics become very attractive. Nonlinear Sliding mode controllers applied to parallel multi-cell converters combine control concepts with multiple sliding surfaces and integral

*Corresponding Author Rihab Hamdi, Email: rihab.hamdi2012@gmail.com

variable structure. They are easy to design, robust and shows good transient and stable performances. Furthermore, they restrict the switching frequency variation, mitigate the undesired transient response and achieve a good compromise between transient and stationary performances. Hence, there are two ways of control involved for multicellular converters. The first consists in controlling the multi-commutations with phase shifting from one another, one therefore speaks of interleaving. The second is to apply the same control signal simultaneously to the multi-switchings, it is therefore synchronicity. The controllers developed for paralleled multiphase converters take into consideration most of the characteristics mentioned above [19]. To date, several researchers have reported that these control schemes have many desirable features: easy-designed because each sliding surface is independently controlled, smooth transient responses tested under parametric variations, eliminate the problem of chattering, the impact of high switching frequencies is managed by the integrators. All these promising results have been experimentally proven on a closed loop system. Previous studies proposed to use the extended linearization method for the design of the sliding mode controller, and then improving robustness using the combined controllers. Thus, they developed a procedure to choose the high-pass filter parameters of the sliding mode-controlled converters. In addition, they presented derivatives of sliding mode control, including the direct /indirect sliding mode, the integral sliding mode, the proportional sliding mode and the high order sliding mode [20–28]. As far, sliding mode control has been widely studied and has proven to be a future solution for generations of power converters, yet, it suffers from the problem of high frequency oscillation (chattering) due to the practical limitations of system components and the problem of switching frequency variation by applying the traditional sliding mode.

This leads us to propose an implementation of the sliding mode controller by hysteresis function. The implementation is easy to reproduce and does not require additional auxiliary circuits or complex calculations. The introduction of a hysteresis band as the limit conditions of the sliding mode, delays the period of states exchange and gives a sign to control the switching frequency of the system, and therefore, the trajectory of the system variates in close proximity to the excesses of the band. So, the practical problem of chattering will be solved. Particularly, boundary layer-based control schemes allow multicell converters to operate with a fixed switching frequency [29–35]. This strategy will be fully covered in the rest of our work. The remainder of the paper is organized as demonstrated: In Section II, the adopted system is described and modeled. Section III reviews the control mechanism based on the sliding mode controller. Section IV is devoted to analyze the simulation results and to discuss the effectiveness of the control strategies. The conclusion is reported in Section V.

2. Paralleled Non-Isolated Multi-Cell Converter

The studied three-cell DC-DC buck converter is shown in Figure 1. It consists of a constant DC input voltage source V_{in} connected to three similar modules sharing a resistive load R and a filter capacitor C. Each module is based on an implemented controllable power MOSFETs S, the antiparallel diode D that authorizes the embedded switch to exhibit a bidirectional current conduction property and a filter inductor L. Note that the parallel switching buck converter operates in continuous conduction mode (CCM). The three cells are identical, in other words, $L_1=L_2=L_3$.

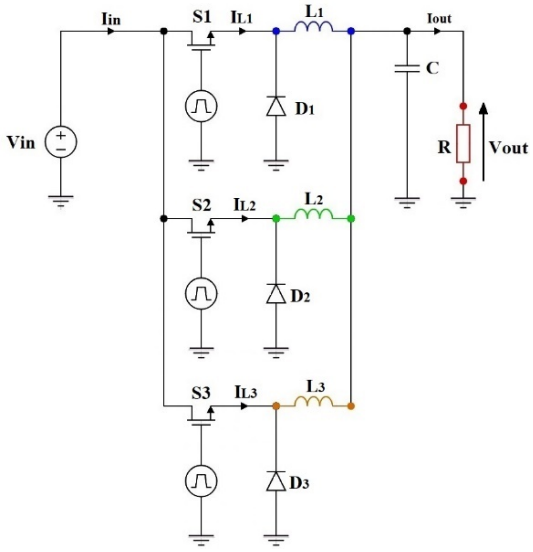


Figure 1: Three-channel Buck converter parallel-associated

The converter can be represented by a system of equations based on a model with mean values, then the system dynamics of the studied converter can be expressed as:

$$L\frac{dI_{Lj}}{dt} = V_{in}D - V_{out} \tag{1}$$

$$C\frac{dv_{out}}{dt} = I_L + \frac{V_{out}}{R} \tag{2}$$

Where I_{Lj} is the current flowing through inductance $j = 1,2,3$, I_L the current at the output of the converter, V_{in} is the input voltage and V_{out} is the output voltage of our converter and V_{ref} is the desired output voltage. The state of space can be expressed as:

$$\begin{aligned} x_1 &= V_{ref} - V_{out} \\ x_2 = \dot{x}_1 &= -\frac{dV_{out}}{dt} = \frac{1}{C}\left(\frac{V_{out}}{R} - \int \frac{uV_{in} - V_{out}}{L} dt\right) \end{aligned} \tag{3}$$

This model can be used directly to simulate the converter, in MATLAB / Simulink type environment.

3. Hysteresis Modulation-Based Sliding Mode Controller

We propose two different control schemes, developed in the continuous domain, to stabilize parallel multi-phase standalone converter and compare them. The proposed sliding mode control (SMC) strategy based on hysteresis modulation (HM) automatically compensates for the parametric variations of the converter and the variation of the line impedance, such that it allows the three-cell paralleled converter to share the load and does not require the modules to be interconnected. We illustrate the fixed-bandwidth of the hysteresis modulation technique-based sliding mode control; however, this approach has accentuated the problem of variable switching frequencies that requires a high-bandwidth current sensor. At this point in our work, we are interested in an adaptive and feedforward current control solution to overcome this problem and adapt a variable-bandwidth of the

hysteresis modulation to mitigate the nonlinearity phenomenon in conventional sliding mode control to fix the switching frequency.

Describing the mathematical model of the system in question, operating in CCM, and especially its states of space necessary for the design of the sliding mode voltage control (SMV), the reformed expression of the state of space is:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{LC} & -\frac{1}{RC} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{V_{in}}{LC} \end{bmatrix} u + \begin{bmatrix} 0 \\ \frac{V_{ref}}{LC} \end{bmatrix} \quad (4)$$

Employing the sliding mode voltage control (SMVC) and taking into account the state trajectory that includes the control parameters x_1 and x_2, we can determine the switching function u.

$$S = \alpha x_1 + x_2 = Jx \quad (5)$$

$$\begin{cases} J = [\alpha, 1] \\ x = [x_1, x_2]^T \end{cases} \quad (6)$$

That forms the control law:

$$u = \begin{cases} 1 = 'ON' & When\ S > k \\ 0 = 'OFF' & When\ S < k \end{cases} \quad (7)$$

The convergence conditions are the criteria that allow the dynamic errors of the system to converge towards the sliding surface. They are ensured by a judicious choice of the function of LYAPUNOV, which guarantees the attraction of variables to regulate to their references, to build the discontinuous control size and meet the stability criterion:

$$\lim_{S \to 0} S.\dot{S} < 0 \quad (8)$$

Thus, the existence condition for the adopted control strategy is:

$$\dot{S} = \begin{cases} J\dot{x} < 0 & for\ 0 < S < \xi \\ J\dot{x} > 0 & for\ -\xi < S < 0 \end{cases} \quad (9)$$

where ξ is a positive quantity, infinitely small and arbitrarily chosen.

Substituting (3) and (5), we find:

$$S = \frac{1}{R}(V_{ref} - V_{out}) - i_C \quad (10)$$

The sliding coefficient is set as (11) and thus the above equation is valid:

$$\alpha = \frac{1}{RC} \quad (11)$$

We can therefore deduce the conditions of existence:

$$\lambda_1 = (\alpha C - \frac{1}{R})x_2 - \frac{1}{L}x_1 + \frac{V_{ref} - V_i}{L} < 0 \quad (12)$$

$$\lambda_2 = (\alpha C - \frac{1}{R})x_2 - \frac{1}{L}x_1 + \frac{V_{ref}}{L} > 0$$

The bandwidth of the FBW regulator based sliding mode controller has been selected according to the following equation:

$$B_{Fixed-Width} = \frac{(V_{in0} - V_{out})}{2.L.fw} \cdot \left(\frac{V_{out}}{V_{in0}} \right) \quad (13)$$

$$V_{in0} = \frac{3 \cdot V_{in} \cdot \sqrt{6}}{\pi} \quad (14)$$

Hence, we focus on the adaptive feedforward current control technique in continuous conduction mode (CCM). So that, the expression of the current variation ΔI_L, as a function of the input voltage V_{in}, the output voltage V_{out}, the switching frequency fw, and the value of the inductance L, can be shown as:

$$\Delta I_L = \frac{(V_{in} - V_{out})}{L.fw} \cdot \left(\frac{V_{out}}{V_{in}} \right) \quad (15)$$

Still with the aim of ensuring the operation at fixed frequency of the proposed hysteresis modulator, a requirement is imposed, and that the hysteresis bandwidth must satisfy it, is:

$$B_{Variable-Width} = \frac{\Delta I_L}{2} = \frac{(V_{in} - V_{out})}{2.L.fw} \cdot \left(\frac{V_{out}}{V_{in}} \right) \quad (16)$$

The dynamic performance of the two schemes needs to be demonstrated under conditions when some changes are imposed and a feedforward transient occurs. Simulations are performed for various cases of the circuit parameter variations: load changes, input voltage changes and reference voltage changes.

4. Simulation Results and Robustness Assessment

In this part, and to properly study the behavior and performance evaluation of the closed-loop converter and the various control techniques synthesized in the previous section, particularly when modifying the operating conditions, we will simulate the behavior of a 3-cell chopper connected to an R load. A robustness test was carried out to analyze the sensitivity of the strategies and the correctors implemented with regard to possible variations in the parameters of the model. For the parameter's identification of the studied converter, we determine the nominal values of the inductance L and the capacitance C using the conditions below:

- The current through the inductance, must be within a reasonable interval, for all the conditions of charge, because our converter is operating in continuous conduction mode.
- The maximum ripple of the output voltage should not exceed a small percentage, usually 5% of the output voltage V_{out}.

Taking these conditions into account, the LC filter must obey the conditions of the relationships below, where fw corresponds to the switching frequency and β is the duty cycle.

$$L \geq \frac{V_{in}}{\Delta I_{out,\max} \cdot f_w} \beta(1 - \beta) \quad (17)$$

$$C \geq \frac{(1-\beta)}{8L.f_w^{\,2}}\left(\frac{V_{out}}{\Delta V_{out}}\right) \quad (18)$$

Table 1 shows the specifications that we adopted of the studied buck converter simulated in the MATLAB® environment.

Table 1: Specifications of The Paralleled Three-Channel Buck Converter

Description	Parameters	Nominal Value
Input voltage	V_{in}	*24 V*
Capacitance	*C*	*60 µF*
Inductance	*L*	*50 mH*
Resistance	*R*	*12 Ω*
Switching frequency	*fw*	*400 kHz*
Desired output voltage	V_{ref}	*12 V*

Figure 2 duplicates the variation of the average switching frequency as a function of the input voltage variation. The VBW HM-based SMC adaptive feedforward control reduces switching frequency variation compared to the FBW HM-based SMC control. This result highlights the major contribution of our work, thanks to the developed strategy, we achieve a significant reduction in the variation of the switching frequency.

Figure 3 and Figure 4 show the output voltage response for the two control strategies used. it seems clear that the system operating with VBW HM-based SMC achieves the best compromise passing from the transient regime to the stationary regime. Thus, its transition response is smoother, more stable, without overshoot and with less chatter, on the other hand, the response of our system under the control of the FBW HM-based SMC is faster but it represents more variation in amplitude of the oscillations, with an overshoot.

It appears clearly from Figure 5 and Figure 6, during an increase/decrease in resistance (50% variation) every 0.2s, that the closed-loop output voltage response reacting with the FBW HM-based SM exhibit an undesired transient drop that lasts a few seconds and disturbs the smoothness of the output voltage followed abruptly by a steady-state. While, in the other case, the system compensates for a transient drop in a soft way. Through the output current simulation results, we can observe that the VBW HM-based SM goes perfectly with the load changes to even almost assure that there is no transient regime. Figure 7 and Figure 8 point-out the output voltage response and the output load current response for the two controllers when the system undergoes a change in the input supply (V_{in} varied from 24V to 50V) at 0.2s. For the two cases, the chattering increase, however, the output voltage and the output load current vary around the reference values. The zoom in clearly shows an increase in the amplitude of the oscillations accompanied by an upward shift of the output voltage in steady-state regime when the input voltage increases, it comes down to the variation of the switching frequency.

Figure 9, Figure 10, Figure 11 and Figure 12 present the output voltage response and the output load current response for the two controllers when the system undergoes a change in the output reference (V_{ref} varied from 12V to 20V). Switching from one value to another is almost similar, there is no big difference except that the VBW HM-based SM is more precise. The carried-out simulations show extremely encouraging outcomes as far as reference tracking effectiveness and robustness, a quick estimation of the control law permits a quicker dismissal of the unsure load impact. They demonstrate the propriety of sliding mode control for such sort of framework.

The results in terms of stability assessment and performance analysis under system parameters variation are recapitulated in Table 2.

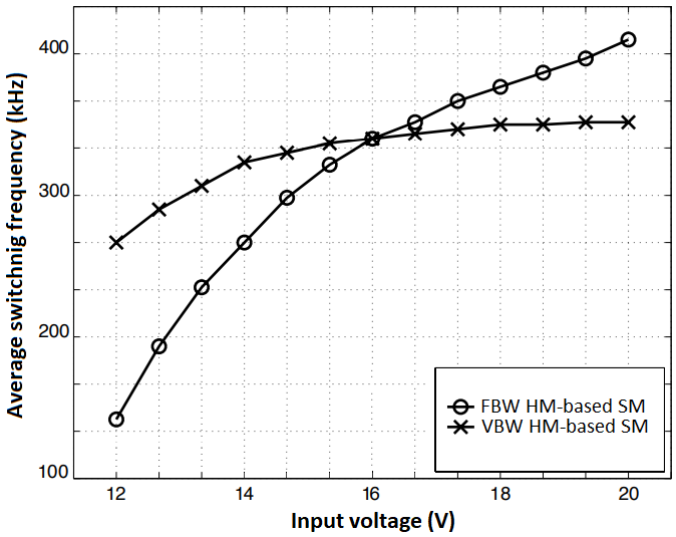


Figure 2: The variation of the average switching frequency as a function of the input voltage variation.

Table 2: Comparative Performances of FBW HM-SMC with VBW HM-SMC

	FBW HM-based SMC	VBW HM-based SMC
Load variation		
Time of compensation	*0.01 s*	*0.015 s*
The drop width	*0.22 V*	*0.13 V*
Amplitude of oscillations	*±0.15 V*	*±0.1 V*
Response time	*0.02 s*	*0.03 s*

Input variation		
Amplitude of oscillations		
For $t < 0.2\ s$	*±0.14 V*	*±0.12 V*
For $t > 0.2\ s$	*+0.3 V* *-0.15 V*	*±0.12 V*
Reference variation		
Time of transition	*0.015 s*	*0.03 s*
Amplitude of oscillations	*±0.14 V*	*±0.12 V*

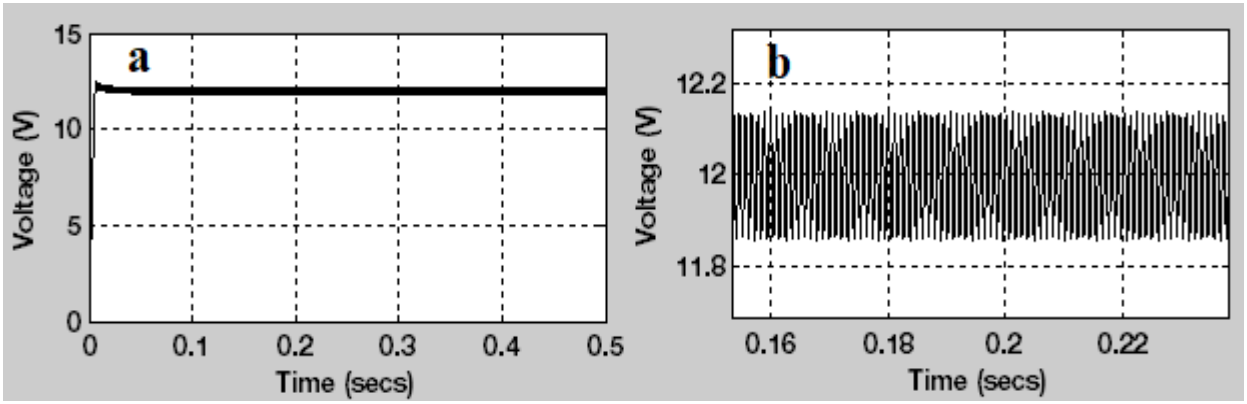


Figure 3: (a) The Output voltage response for the FBW HM-based SMC. (b) Zoom in

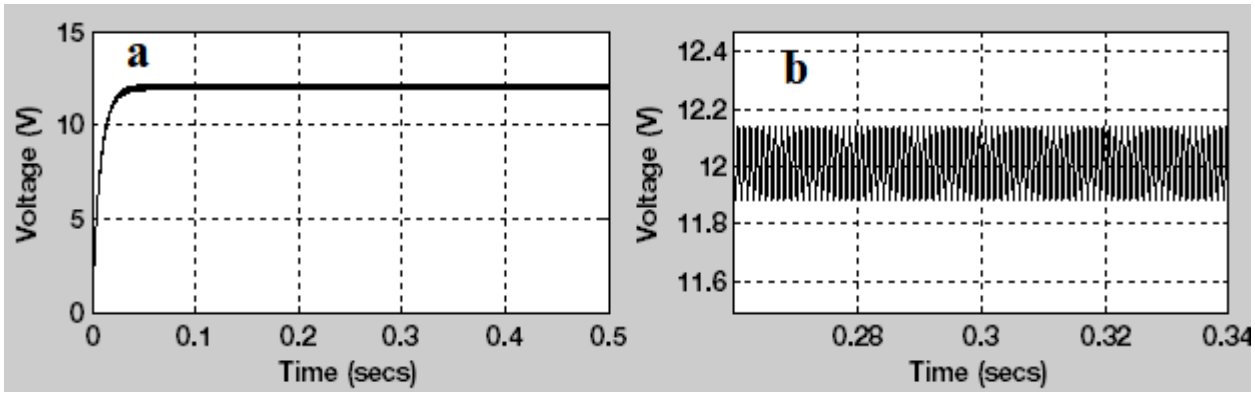


Figure 4: (a) The Output voltage response for the VBW HM-based SMC. (b) Zoom in

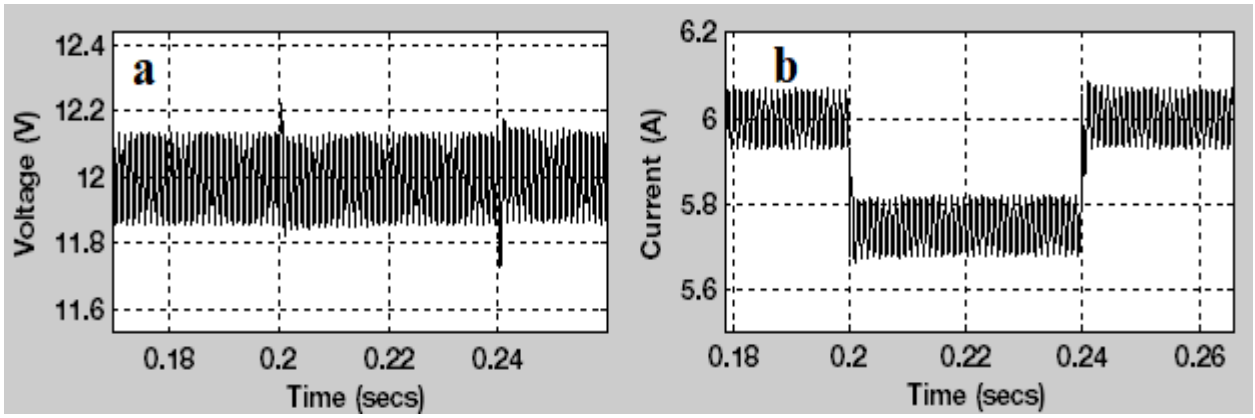


Figure 5: (a) The output voltage response. (b) The output load current response.
Both for the FBW HM-based SMC when the system undergoes a slight change in the load resistance.

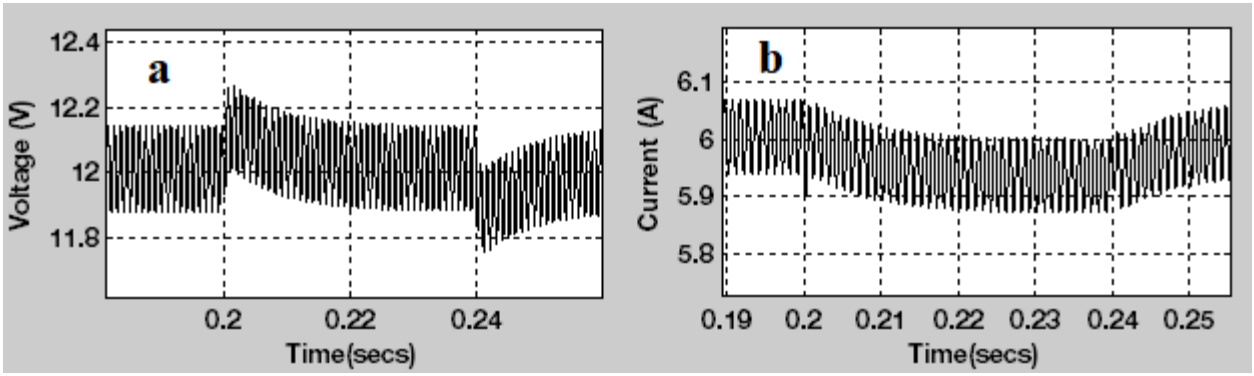


Figure 6: (a) The output voltage response. (b) The output load current response.
Both for the VBW HM-based SMC when the system undergoes a slight change in the load resistance.

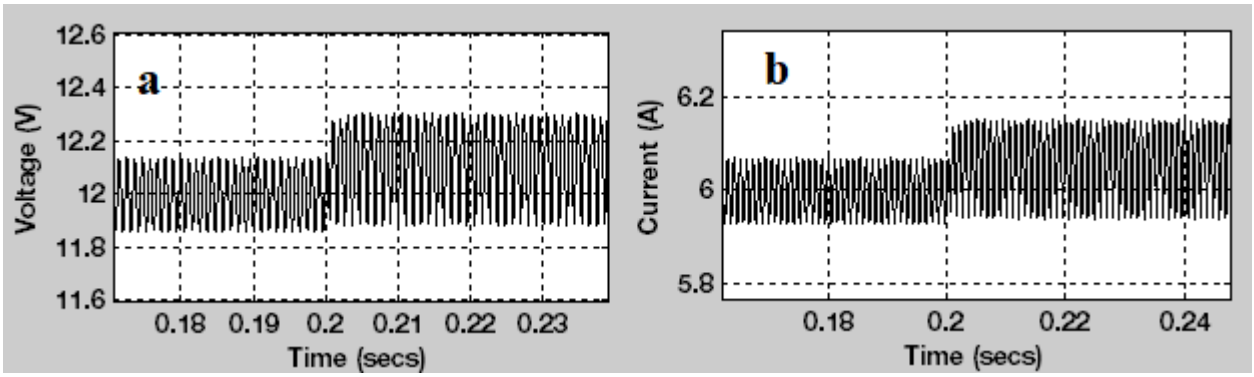


Figure 7: (a) The output voltage response. (b) The output load current response.
Both for the FBW HM-based SMC when the system undergoes a change in the input supply.

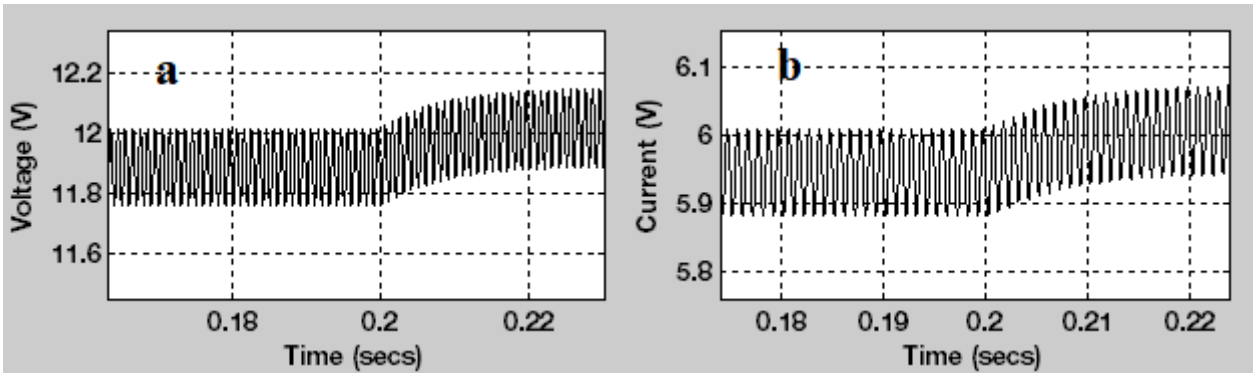


Figure 8: (a) The output voltage response. (b) The output load current response.
Both for the VBW HM-based SMC when the system undergoes a change in the input supply.

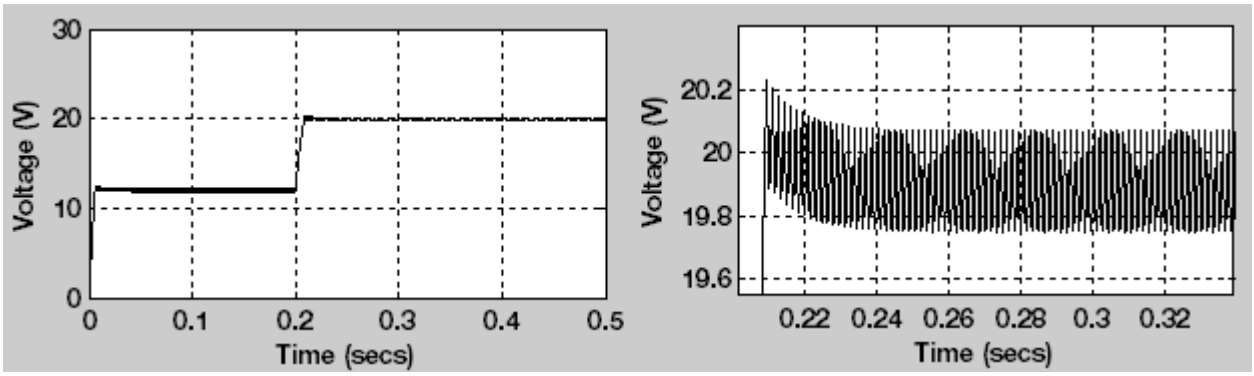


Figure 9: (a) The output voltage response for the FBW HM-based SMC when the system undergoes a change in the output reference.
(b) Zoom in.

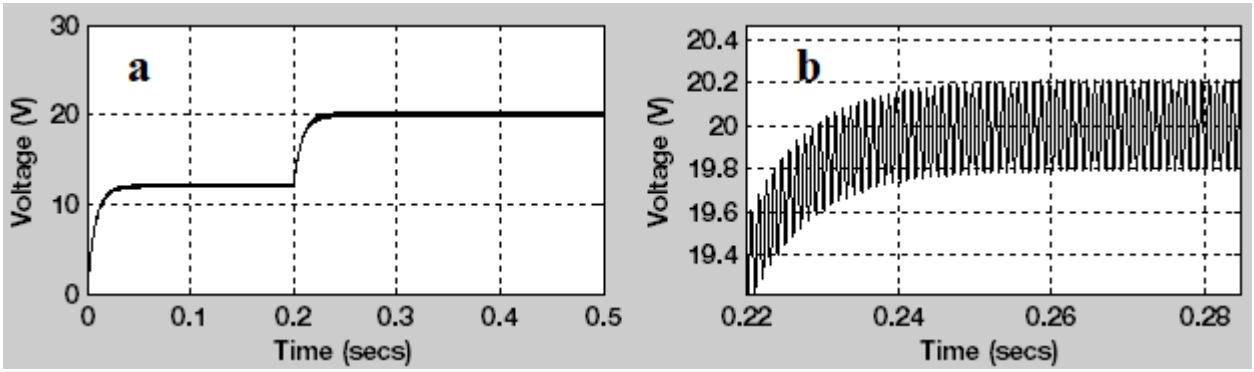


Figure 10: (a) The output voltage response for the VBW HM-based SMC when the system undergoes a change in the output reference. (b) Zoom in.

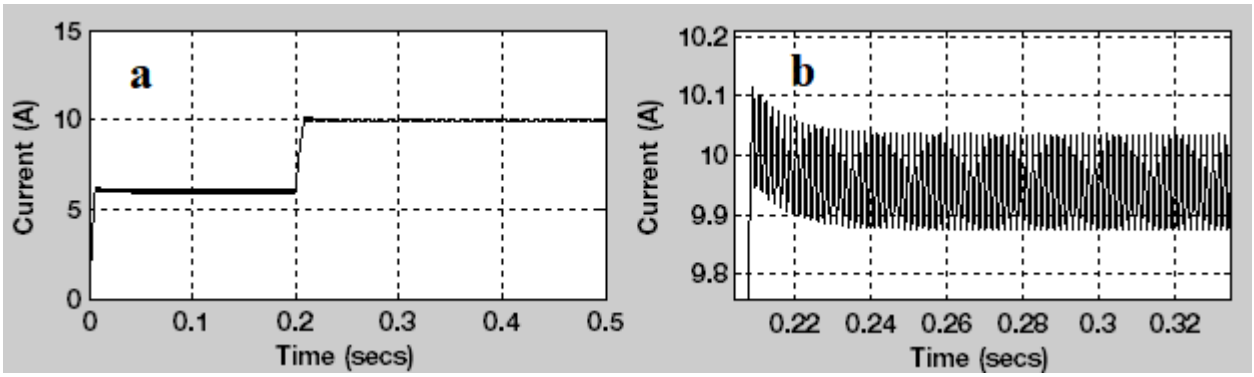


Figure 11: (a) The output load current response for the FBW HM-based SMC when the system undergoes a change in the output reference. (b) Zoom in.

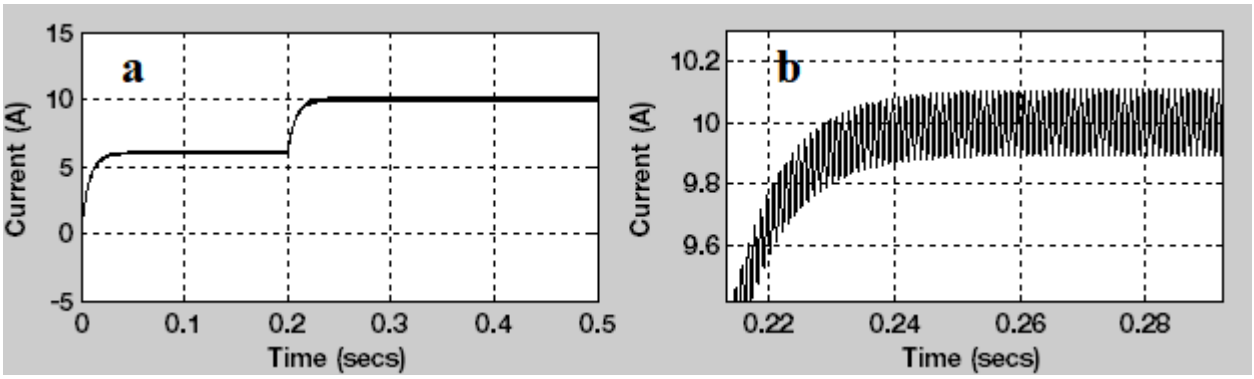


Figure 12: (a) The output load current response for the VBW HM-based SMC when the system undergoes a change in output reference. (b) Zoom in

5. Conclusion

The variable-bandwidth hysteresis modulation is favored because it ensures the chattering suppression caused by the switching of the MOSFETs and provides high efficiency since it offers a greatly desired output voltage tracking. The VBW HM-based SMC is faster to the point that the response time at 5% is mostly short. His appropriate dynamic accuracy is characterized by zero overshoot during the transient regime for the output voltage response, contrariwise to the FBW HM-based SMC that represents 2.5% overshoot. The output voltage regulation is improved from 2% (V_{in}=24V) to 0.15% (V_{in}=24V) through the VBW HM-based SMC. Add to that, the frequency variation is reduced from 30% to 15% with the developed strategy. The various control techniques provide similar performances, they have improved the dynamic conduct and they can adapt to include voltage and load varieties. The use of variable bandwidth hysteresis modulation and fixing of the switching frequency leads to large scale dynamics.

References

[1] R. Hamdi, A.H. Hamida, O. Bennis, F. Babaa, "HM-Based SMVC with Adaptive Feedforward Controller Applied to DC-DC Converter," International Conference on Electrical and Information Technologies (ICEIT), IEEE: 1–6, 2020, doi:10.1109/ICEIT48248.2020.9113220.

[2] W. Li, He X, "Review of nonisolated high-step-up DC/DC converters in photovoltaic grid-connected applications," IEEE Trans Ind Electron, **58**, 1239–1250, 2011. doi:10.1109/TIE.2010.2049715.

[3] W. Do, K. Eguchi, A. Shibata, "An analytical approach for parallel switched capacitor converter," Energy Reports, **6**, 338–342, 2020, doi:10.1016/j.egyr.2020.11.233.

[4] M. Srinivasan, A. Kwasinski, "Control analysis of parallel DC-DC converters in a DC microgrid with constant power loads," International Journal of Electrical Power and Energy Systems, **122**, 106-207, 2020,

doi:10.1016/j.ijepes.2020.106207.
[5] H. Li, X. Jiang, Y. Zou, C. Liu, "A time-domain stability analysis method for paralleled LLC resonant converter system based on Floquet theory," Microelectronics Reliability, **114**, 113-849, 2020, doi:10.1016/j.microrel.2020.113849.
[6] MSB. Ranjana, N. Sreeramulareddy, RKP. Kumar, "A novel nonisolated switched inductor floating output DC-DC multilevel boost converter for fuelcell applications," Conference on Electrical, Electronics and Computer Science, 2014, doi:10.1109/SCEECS.2014.6804492.
[7] M. Sagar, B. Ranjana, "A Novel Non-Isolated Switched Inductor Floating Output DC-DC Multilevel Boost Converter For Fuelcell Applications", 2014, doi:10.1109/SCEECS.2014.6804492.
[8] S. Ben Said, K. Ben Saad, M. Benrejeb, "ScienceDirect HIL simulation approach for a multicellular converter controlled by sliding mode," International Journal of Hydrogen Energy, 1–7, 2017, doi:10.1016/j.ijhydene.2017.01.198.
[9] P. Djondiné, "Overview of Control Techniques for Multicellular Converter," Journal of Engineering Sciences, **5(**1), 10–14, 2018, doi:10.21272/jes.2018.5(1).e3.
[10] P. Taylor, M. Defoort, M. Djemai, T. Floquet, W. Perruquetti, "Robust finite time observer design for multicellular converters," International Journal of Systems Science Robust finite time observer design for multicellular converters", 37–41, 2011, doi:10.1080/00207721.2010.543494.
[11] F. Engelkemeir, A. Gattozzi, G. Hallock, R. Hebner, "Electrical Power and Energy Systems An improved topology for high power softswitched power converters," Electrical Power and Energy Systems, **104**, 575–582, 2019, doi:10.1016/j.ijepes.2018.07.049.
[12] S. Ogasawara, H. Akagi, A. Nabae, "A novel PWM scheme of voltage source inverters based on space vector theory," Arch Für Elektrotechnik, **74**, 33-41, doi:10.1007/BF01573229.
[13] N. Patin, "Power Electronics Applied to Industrial Systems and Transports," 2016, doi:10.1016/C2015-0-04476-8.
[14] O. Hegazy, J. Van Mierlo, P. Lataire, "Analysis , Modeling , and Implementation of a Multidevice Interleaved DC / DC Converter for Fuel Cell Hybrid Electric Vehicles," **27**(11), 4445–4458, 2012.
[15] N.H. Baharudin, TMNT. Mansur, FA. Hamid, R. Ali, MI. Misrun, "Performance Analysis of DC-DC Buck Converter for Renewable Energy Application," J. Phys. Conf. Ser., vol. 1019, Institute of Physics Publishing, doi:10.1088/1742-6596/1019/1/012020.
[16] A. Kolli, A. Gaillard, A. De Bernardinis, O. Bethoux, D. Hissel, "A review on DC / DC converter architectures for power fuel cell applications," Energy Conversion And Management, **105**, 716–730, 2015, doi:10.1016/j.enconman.2015.07.060.
[17] D. Boroyevich, H. Vanlandingham, W.T. Baumann, "Nonlinear Analysis and Control of Standalone , Parallel DC-DC , and Parallel Multi-Phase PWM Converters," 2001.
[18] M. Andrade, V. Costa, "DC-DC Buck Converter with Reduced Impact," Procedia Technol, 17, 791-8, 2014, doi:10.1016/j.protcy.2014.10.209.
[19] K.V.R. Swathi, GVN. Kumar, "Design of intelligent controller for reduction of chattering phenomenon in robotic arm : A rapid prototyping," R. Comput Electr Eng,1-15, 2017, doi:10.1016/j.compeleceng.2017.12.010.
[20] M. Alfayyoumi, AH. Nayfeh, D. Borojevic, "Modeling and analysis of switching-mode DC-DC regulators," Int J Bifurcat Chao, **10**, 373-90, 2000, doi:10.1142/S0218127400000244.
[21] H. Al-Baidhani and al., "Sliding-Mode Voltage Control of Dynamic Power Supply for CCM," IEEE International Symposium on Circuits and Systems (ISCAS), 1-5, 2019, doi:10.1109/ISCAS.2019.8702628.
[22] Y. Wu, Y. Huangfu, R. Ma, A. Ravey, D. Chrenko, "A strong robust DC-DC converter of all-digital high-order sliding mode control for fuel cell power applications," Journal of Power Sources, **413**, 222–232, 2019, doi:10.1016/j.jpowsour.2018.12.049.
[23] S. Das, M. Salim Qureshi, P. Swarnkar, "Design of integral sliding mode control for DC-DC converters," in Materials Today: Proceedings, Elsevier Ltd: 4290–4298, 2018, doi:10.1016/j.matpr.2017.11.694.
[24] S.K. Pandey, S.L. Patil, D. Ginoya, U.M. Chaskar, S.B. Phadke, "Robust control of mismatched buck DC–DC converters by PWM-based sliding mode control schemes," Control Engineering Practice, **84**, 183–193, 2019, doi:10.1016/j.conengprac.2018.11.010.
[25] Y. Gao, J. Liu, G. Sun, M. Liu, L. Wu, "Systems & Control Letters Fault deviation estimation and integral sliding mode control design for Lipschitz nonlinear systems," Systems & Control Letters, **123**, 8–15, 2019, doi:10.1016/j.sysconle.2018.08.006.
[26] SK. Pandey, SL. Patil, D. Ginoya, UM. Chaskar, SB. Phadke, "Control Engineering Practice Robust control of mismatched buck DC – DC converters by PWM-based sliding mode control schemes," Control Engineering Practice, **84**, 183–193, 2019, doi:10.1016/j.conengprac.2018.11.010.
[27] BB. Naik, AJ. Mehta, "Sliding mode controller with modi fi ed sliding function for DC-DC Buck Converter," ISA Transactions, 2017, doi:10.1016/j.isatra.2017.05.009.
[28] A. Fezzani, S. Drid, A. Makouf, "Commande Robuste par Mode Glissant d ' Ordre Supérieur de la Machine Synchrone à Aimant Permanent," 247–251, 2010. http://eprints.univ-batna2.dz/1230/.
[29] M.S. Shaker, AA. Kraidi, "Robust observer-based DC-DC converter control u," Journal of King Saud University - Engineering Sciences, 2017, doi:10.1016/j.jksues.2017.08.002.
[30] KVR. Swathi, GVN. Kumar, "Design of intelligent controller for reduction of chattering phenomenon in robotic arm : A rapid prototyping R," Computers and Electrical Engineering, 1–15, 2017, doi:10.1016/j.compeleceng.2017.12.010.
[31] M. Phattanasak, W. Kaewmanee, J. Martin, S. Pierfederici, B. Davat, "Interleaved Double Dual Boost Converter for Renewable Energy System," **932**, 904–909, 2014, doi:10.4028/www.scientific.net/AMR.931-932.904.
[32] R. Ramos, D. Biel, E. Fossas, R. Grin, "Control Engineering Practice Sliding mode controlled multiphase buck converter with interleaving and current equalization," **21**, 737–746, 2013, doi:10.1016/j.conengprac.2012.09.005.
[33] A. El Aroudi, R. Giral, J. Calvente, "Sliding-Mode Control of DC-DC Switching Converters," IFAC, 2011, doi:10.3182/20110828-6-IT1002.00557.
[34] A. Mehta, B. Naik, "Sliding Mode Controller with PI-Type Sliding Function for DC–DC Buck Converter," In: Sliding Mode Controllers for Power Electronic Converters. Lecture Notes in Electrical Engineering, **534**, Springer, doi:10.1007/978-981-13-3152- 7_3.
[35] F. Blaabjerg. "Control of Power Electronic Converters and Systems," Elsevier 2018, ISBN:978-0-12-816136-4, doi:10.1016/c2017-0-04756-0.

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1312-1318 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Evaluation of Personal Solar UV Exposure in a Group of Italian Dockworkers and Fishermen, and Assessment of Changes in Sun Protection Behaviours After a Sun-Safety Training

Alberto Modenese[1,*], Fabio Bisegna[2], Massimo Borra[3], Giulia Bravo[4], Chiara Burattini[2], Anna Grasso[1], Luca Gugliermetti[2], Francesca Larese Filon[5], Andrea Militello[3], Francesco Pio Ruggieri[1], Fabriziomaria Gobba[1]

[1]*Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, 41125, Italy*

[2]*Department of Astronautical, Electrical and Energy Engineering, University of Rome "Sapienza", Rome, 00184, Italy*

[3]*Department of Occupational and Environmental Medicine, Epidemiology and Hygiene, Italian Workers' Compensation Authority (INAIL-DiMEILA), Monte PorzioCatone (Rome), 00078, Italy*

[4]*Department of Medicine, University of Udine, Udine, 33100, Italy*

[5]*Unit of Occupational Medicine, University of Trieste, Trieste, 34128, Italy*

ARTICLE INFO

Article history:
Received: 23 December, 2020
Accepted: 30 January, 2021
Online: 28 February, 2021

Keywords:
Solar ultraviolet radiation
Occupational exposure
Maritime workers

ABSTRACT

Solar ultraviolet radiation (UVR) is considered a relevant health risk for the workers of the maritime and port sectors, but scant data are available on actual exposure measured using personal dosimeters. Moreover, in outdoor workers sun protection habits are usually poor, while some promising data suggest that sun-safety campaigns can be effective in increasing self-protection at work. Accordingly, our aim was to conduct an assessment of solar UVR exposure in dockworkers and fishermen using personal dosimeters, and to evaluate the use of sun protection measures at work after a sun-safety training. We performed two different UVR measurements campaigns in spring-summer 2018, investigating 7 fishermen and 14 dockworkers. Electronic dosimeters have been placed on the workers for at least a half work-day. Only at the port it was also possible to register the environmental UVR exposure with a specrto-radiometer, while for fishermen we estimated the corresponding environmental exposure using an algorithm. Our results demonstrate a high erythemal UVR dose received by the workers, with an individual exposure up to 542 J/m2 for fishermen in spring and up to 1975 J/m2 for dockworkers in summer. This data indicates an excessive occupational risk, needing more effective prevention. Accordingly, we offered a sun-safety training to the workers. Before the training, protective behaviour of the workers was rather poor: about the 50% never used the hat, the 40% never wore sunglasses and none of the workers referred to apply sunscreens at work. After the training, fishermen reported a relevant improvement in the use of individual UV protections, as hat (+9.6%), sunglasses (+28.5%) and clothes (+5%), even if the use of sunscreens at work was not increased.

1. Introduction

Solar ultraviolet radiation (UVR) exposure represents an important occupational risk in the maritime and port sectors: this paper is an extension of work originally presented in the 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe) [1], integrated with data we published elsewhere [2], further elaborated and with additional inclusion of new results.

The possible adverse health effects occurring in workers exposed to solar UVR include both acute and long-term ones and mainly involve the skin and the eyes. The consequences may be severe, as in the case of cancers: solar UVR is considered the most frequent occupational carcinogenic exposure [3], for its ability to

*Corresponding Author: Alberto Modenese, Email: alberto.modenese@unimore.it

induce in particular non-melanoma skin cancers (NMSC) [4], but also some forms of melanoma skin cancers may be related to occupational UV exposure [5], as well as rare forms of eye tumors, as the squamous cell carcinomas of the cornea and conjunctiva and ocular melanomas [6]. Considering the skin, other frequent adverse health conditions associated with cumulative solar UV exposure, and representing precancerous lesions, are photo-aging and actinic keratosis [7]. Regarding the eyes, it should be noted that usually this organ is much more protected from UV rays compared to the skin, for anatomical reasons and for the actions of various involuntary reflexes (pupillary reflex, squinting, winking) [8]. Nevertheless, UV eye exposure may be relevant in particular conditions, as in the case of UV rays reflected from surfaces like polished metal, water, snow, white sand and marbles, etc [8]. Fishermen and dockworkers, often close to the water, may be at particular risk for relevant UV eye exposure. In addition to the eye tumors, there are various ocular diseases recognizing solar UVR as a risk factor: among all, cataract [6], currently the leading cause of visual impairment worldwide [9]; pterygium, a hyperplasia of the bulbar conjunctiva highly correlated with increased levels of UV radiation [6]; and possibly also age-related macular degeneration [6], a chronic degenerative retinal disease, the second leading cause of blindness in the world [9]. For all these three eye pathologies, positive significant associations with solar radiation exposure and increased risks for outdoor workers of being diagnosed with the diseases have been well documented [10–12].

Despite the recognized health risk, there are scant studies investigating solar UVR exposure levels in the maritime and port sector [2,13–15]. Fishermen spend most of their working time outdoor on the boats, close to the water surface, reflecting a relevant amount of UVR [8]; but also dockworkers, in particular those working at the quay, are almost always outside and quite often close to reflecting surfaces, such as water, but also lucid metallic surfaces on the quay/ships. According to these considerations, it is certainly of interest a detailed assessment of solar UVR exposure of these groups of workers. Furthermore, the extant scientific literature shows a relevant under-estimation of the risk by outdoor workers, with indications of poor protective habits and behaviors with respect to a relevant occupational risk, which is solar UVR exposure [16–22].There is also a growing evidence that sun-safety interventions are effective in increasing outdoor workers' sun-protection habits [21]; nevertheless, there is still a scarcity of these interventions focused on outdoor workers, and in particular on those of the maritime and port sectors.

Our objective is to report the results of two different solar UVR measurements campaigns, respectively performed in a group of Italian dockworkers and fishermen, and a subjective evaluation of sun-exposure in a sub-group of the workers. Moreover, we wanted to test the possible improvement of the use of sun protections one year after our measurement's campaigns, including specific sun-safety trainings.

2. Materials and Methods

2.1. The solar UV radiation measurements campaigns

We registered personal exposure to solar UVR with electronic dosimeters during two different campaigns involving two groups of outdoor workers: fishermen (FM) and dockworkers (DW). Both the groups were voluntary recruited respectively from a fishing and a port company active in the area of the north-east of Italy, on the Adriatic sea (coordinates: 43° 9' N, 12°7' E for FM; 45° 4' N, 13° 5' E for DW). Considering the unstable whether conditions and the job activities mainly in the morning, we monitored FM for two consecutive spring days (15–16 May 2018), while for DW we registered personal solar UVR exposure during a summer day, with almost sunny wheather (4 July 2018). The personal electronic dosimeter used were 10 devices of the Gigahertz-Optik X2000 and X2012 series, while for a reference of the environmental erythemal UV dose in the monitored days we consulted the free online database of the Troposhpheric Emission Monitoring Internet Service (TEMIS) of the European Space Agency, as described in detail in our previous works [1,2]. Briefly, the TEMIS website provides daily erythemal UV dose data for various places of the world; the data provided are derived by means of radiative transfer calculation codes from satellite measurements, considering local meteorological variables; such data were used to calculate the local environmental radiant exposure referred to the horizontal surface plane. The closest location to the dosimetry campaigns for which effective irradiance data have been made available is Venice, in Italy, that was selected as a proxy, being less than 200 km north-east compared to the workplace of the fishermen, and approximately 160 km south-west considering the port of the dockworkers. Erythemal UVR dose in Venice was 2.6 kJ/m^2 on 15th May, 3.6 kJ/m^2 on 16th May and 4.7 kJ/m^2 on 4th July. For dockworkers, we used this data only as comparison, as we registered the local erythemal radiant exposure on the horizontal surface with a Gigahertz-Optik BTS2048-UV-S spectro-radiometer placed at the quay. For fishermen it was not possible to use the spectro-radiometer on the boats, and we applied a formula [2] to reconstruct the percentage of ambient UVR received by the workers. For all the workers, the dosimeters were mainly placed on the upper back, that was the position we found most comfortable in order to not interfere with the job activities. In some cases we were able to put the devices on the chest of the workers, and only for a fisherman we had the possibility to attach the meter on the back side of his cap to simulate nape exposure (Fig. 1).



Figure 1: Placement of the UV electronic dosimeters at the workplace, on the fishermen (FM) and dockworkers (DW). From left to right, top to bottom: (1) a FM with a dosimeter on the back and one on the cap to simulatenape exposure; (2) a FM with a back dosimeter working in a covered area of the boat;(3) two FM with back dosimeters working in an uncovered area of the boat; (4) a DW working as port coordinator with a chest dosimeter; (5) a DW working as longshoreman with a

dosimeter on the upper back; (6)spectroradiometer for ambient solar UV dose measurements placed at the quay of the port.

The measurements were organized within a preventive campaign for the evaluation of the occupational risks in various occupational activities according to the Italian national occupational health and safety legislation, and were also aimed to the development of more adequate information and training of the workers. All the ethic principles considered in the Helsinki declaration were followed.

2.2. The sun exposure habits and protective behaviors investigation and the sun safety intervention

Only for the fishermen, between the two consecutive days of the measurements campaign we were able to collect a self-administered questionnaire, investigating personal solar UVR exposure habits and protecting behaviors during work and leisure activities, previously described [18]. Briefly, the 22 items questionnaire has two sections, one for work and one for leisure exposure; questions investigate the type of outdoor activity and the protections adopted to reduce solar UVR exposure. With the purposes of evaluating possible improvements of sun protection habits and behaviours at work after a sun safety training, we analyzed the results related to four items investigating the frequency of adoption of individual protections at work: use of UV protective hat, sunglasses with UV filtering lenses, UV protective clothes and sunscreens protections. Workers were asked to answer mainly on a 5 point Likert scale (from 0, meaning "never adopted the exposure habit/protective behavior", to 5 "always adopted the exposure habit/protective behavior"). The questionnaire was collected before a four-hours sun-safety training, explaining to the workers the characteristics of the solar UVR risk, the methods to evaluate the risk, the possible adverse health effects associated to solar UVR exposure and the importance of prevention during outdoor work practices, including the use of individual protections. The training was performed by the same experts involved in the measurements campaigns, for both DW and FM, but, as previously mentioned, the questionnaire was collected only for FM. The group of FM who participated in the training was larger compared to the workers we individually investigated with our dosimeters during the measurements campaigns. We decided to administer the questionnaire before the training, i.e. before the workers received any information on solar UVR risk that could be able to influence their answers. Then, one year after the measurements campaigns and the sun-safety training (May 2019), we contacted via phone the fishermen who participated in the sun-safety training and we adminstered the same questionnaire, in order to evaluate a possible improvement of the protective habits and behaviors of the workers, to be possibly attributed to the intervention we performed.

3. Results

3.1. Results of the solar UV radiation measurements campaigns

We measured individual solar UVR exposure of 7 fishermen and 14 dockworkers, all males (Table 1). FM worked on three fishing boats with different characteristics with respect to the availability of protections against solar radiation: one medium size boat (B1), only partially covered, but with various tasks of the work activity (mussel fishing) performed in shielded areas, one small boat (B2), with almost no coverage from UVR, for activities (sea snails and cuttlefish fishing) performed in direct sunlight and another medium size boat (B3), partially covered, with only parts of the activities (trawling fishing) in direct sun. Considering DW, 10 of them worked as longshoremen (LSM) almost always close to the quay and to the water, while 4 subjects worked as coordinators of the port traffic (traffic coordinator, TC), with some activities performed inside a small office (Table 1).

Table 1: Results of individual measurements of occupational solar UVR exposure for 7 fishermen and 14 dockworkers from the North-East of Italy.

Work place / job task		Placement of the dosimeter	Working period / length of the period measured (minutes)*	Personal UV erythemal dose (J/m^2)*	Ambient UV effective radiant exposure (H_{eff}) (J/m^2)* ±	Personal erythemal vs. ambient UVR exposure (%)* ±
B1	FM1	Back	morning / 397	213	1710	12.5
	FM2	Back		79		4.6
	FM3	Back		71		8.8
B2	FM4	Back (& Nape)	morning / 180	542 (380)	830	65.3 (45.8)
	FM5	Back		288		34.7
B3	FM6	Back	morning / 300	151	1560	9.7
	FM7	Back (& chest)		129 (98)		8.3 (7.8)
Quay	DW1-LSM	Back	full-day / 432	1975	3339	59
	DW2-LSM	Back	full-day / 417	1067	3276	32
	DW3-LSM	Back	morning / 144	288	857	33
	DW4-LSM	Back	morning / 191	402	1322	30
	DW5-LSM	Chest	morning / 243	257	1596	16
	DW6-LSM	Back	morning / 251	854	1965	43
	DW7-LSM	Back	morning & early afternoon / 260	551	2675	20
	DW8-LSM	Back	afternoon / 246	622	2191	28
	DW9-LSM	Back	afternoon / 226	458	2001	22
	DW10-LSM	Back	afternoon / 225	413	1981	20
Center of the port (partially indoor)	DW11-TC	Chest	morning / 177	48	1550	3
	DW12-TC	Chest	Morning / 193	139	1759	7
	DW13-TC	Chest	Afternoon / 210	89	1876	4
	DW14-TC	Chest	Afternoon / 248	44	2045	2

Legend: *B1= boat 1, middle-size, partially covered; B2= boat 2, small size, no coverage; B3= boat 3, partially covered; FM= fisherman; DW= dockworker; LSM= longshoreman; TC= traffic coordinator. *= for fishermen the worst exposure scenario occurred in one of the two days with measurements is reported in the Table; ± = for fishermen, ambient exposure was estimated according to a specific formula [2]*

Regarding the length of the individual measurements period, unfortunately for almost all the workers it was not possible to monitor an usual full working day, as FM started their work-shifts few hours before sunrise, so that measuring UVR in that period was meaningless, and they finished between 1:00-2:00 p.m.; considering DW, for the large majority of them the activities were organized on separate morning and afternoon work-shifts, and only two DW worked both in the morning and in the afternoon the day we performed our measurements.

Considering the worst exposure scenario for the seven fishermen during the two days monitored, we found personal UVR exposure at the back ranging between 71 J/m^2 registered for FM3 working on the B1 and 542 for FM4 on the B2. The estimated proportion of environmental erythemal UVR dose received on the back of the FM resulted between a minimum of 4.6% for FM2 on the B1 and a maximum of 65.3% for FM4 on the B2 (Table 1).

Considering DW, we collected personal solar UVR exposure data at the back for LSM while at the chest for the TC, as they had quite often to sit inside a small office in the middle of the port area. The results of the measurements show a higher exposure in LSM compared to TC: the highest exposure of 1975 J/m^2 was collected, not surprisingly, for one of the two LSM (DW1) who worked for a full day (approximately 7 hours and a half). The individual exposure of the other LSM ranged between 257 J/m^2 collected at the chest of DW5 (i.e. the oly LSM with a chest dosimeter, who was monitored for more than four hous in the morning) and 1067 J/m^2 for the other LSM who was monitored for approximately seven hours (DW2). For the other LSM with dosimeters on the back and monitored only in the morning or afternoon, exposures resulted between 288 J/m^2 measured in the morning in about two hours and a half (DW3) and 854 J/m^2 registered again in the morning but in more than four hours (DW6). Percentages of ambient exposure received by the LSM varied between the 16 and the 59%. Considering TC, their individual solar UVR exposures at the chest varied between 44 J/m^2 measured in about four hours in the afternoon (DW14) and 139 J/m^2 registered in approximately three hours in the morning (DW12), with percentages of individual versus ambient exposure between the 2 and the 7% (Table1).

3.2. Results of the investigation of sun exposure habits and protective behaviors among fishermen

Twenty-one fishermen, all males, participated in the sun-safety training we proposed to the fishing company where we performed the individual solar UVR measurements campaign. We administered our questionnaire investigating sun exposure habits and behaviours before the starting of the training, and all of them filled-in the questionnaire. The 47.7% of the FM reported to never use an UVR protective hat at work. The 38.1% of them never used UV filtering sunglasses on the boats. None of the workers reported to use. always, often or even sometimes, sunscreens at work. Only for the clothes we collected some positive responses: the 40% of the sample wore often protective clothes (Table 2). After one year from the sun safety training we contacted all the 21 fishermen via phone, asking them the same questions related to the frequency of adoption of solar UVR individual protections at work.

Table 2: Results of the subjective investigation on the frequency of adoption of solar UVR individual protections at work among 21 fishermen

Individual protecions against solar UVR				
Frequency of adoption on a 5-point Likert scale	**Fishermen adopting the individual protection: % (n)**			
	Protective hat	UV filtering sunglasses	Protective clothes*	Sunscreens
ALWAYS	14.3 (3)	14.3 (3)	15 (3)	0
OFTEN	19 (4)	9.5 (2)	40 (8)	0
SOMETIMES	19 (4)	14.3 (3)	25 (5)	0
SELDOM	0	23.8 (5)	15 (3)	19 (4)
NEVER	47.7 (10)	38.1 (8)	5 (1)	81 (17)

**1 answer was missing*

In order to better appreciate the differences between the answers given before and after the training, we grouped together as positive responses the answers "always", "often" and "sometimes" and we analyzed the percentage of FM who reported to adopt UV protections at the baseline and one year after the sun-safety campaign (Fig. 2). After the sun-safety interventions there was an improvement of the 9.6% for the use of UV protective hat at work, of the 28.5% for the use of UV filtering sunglasses (with a significant difference at the non-parametric McNemar statistical test for paired nominal data, p=0.031) and of the 5% for the wearing of UV protective clothes. No improvements have been found for the use of sunscreens at work by the fishermen (Fig. 2).

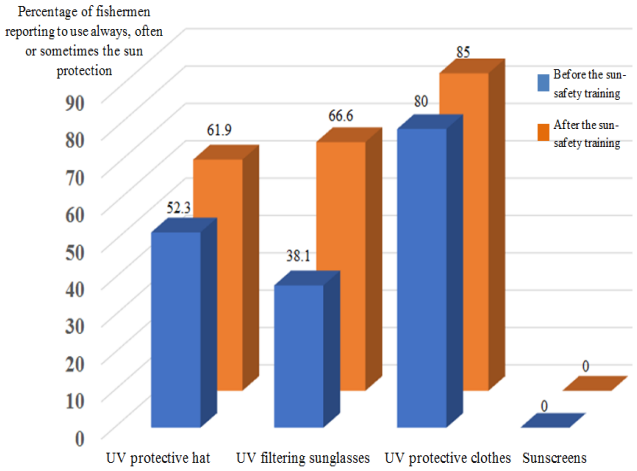


Figure 2: Use of individual solar UVR protections at work before and one-year after the sun-safety training in a group of 21 fishermen.

4. Discussion

Our results on personal solar UVR exposure measured in a group of dockworkers and fishermen of North-East Italy show remarkable high UV exposure levels. For FM we collected measures in late spring, while for DW in summer: not surprisingly, the exposure levels we detected resulted higher in the case of DW. Nevertheless, even if the two days of the measurements campaign for the FM were partially clouded spring days, we registered, in less than five hours on average, an exposure above the equivalent of a Standard Erythemal Dose (SED, i.e. an UVR personal

exposure able to induce an erythema in an individual, adjusted for skin pigmentation, and corresponding approximately to 100 J/m^2) in all but two FM, with a maximum exposure of more than 5 SEDs/day. This indicates a very high risk of acute skin burns, and, considering the photochemical accumulation of UV-induced damages, may also represent an increased risk for long-term adverse health effects, in case of exposures prolonged over years. Considering the specific work tasks, the FM were involved in three different types of fishing performed on three different boats. We collected the highest exposure levels on the smallest boat, with almost no coverage from the solar UV rays, dedicated to an activity (sea snails and cuttlefish fishing) performed almost always in direct sunlight.

Considering now the DW, also in this case we measured different solar UVR exposure levels depending on the specific type of work activity performed. We registered higher solar UVR individual levels for longshoremen compared to traffic coordinators. The personal UV doses collected for LSM, in the 90% of the cases with dosimeters placed on their upper back, varied between 2.6 SED/half-day and 20 SED/day. Also, for the LSM investigated with a chest dosimeter we registered a quite high exposure of 5.5 SED/half-day. Considering dockworkers employed as TC, we had to place the dosimeters on their chests, as they frequently had to be seated at a desk inside a small cabin office, therefore performing also partially indoor activities. For TC we measured exposures between 0.4 and 1.4 SED/half-day.

Discussing now the relations between individual and environmental solar UVR skin exposure, we found quite high variability both in the results related to FM and in those collected for DW. This variability is not surprising and reported also in other studies of personal solar UVR eposure measurements in outdoor workers [14,23–28]. It can be related to the different working postures adopted by the workers and to other factors, as the sun-angle on the horizon, the distance from reflecting surfaces (e.g. water) and the presence of shading structures: accordingly, the exposure of different body districts can change during the day in relation to the ways the activity is performed and to the environment. Percentages of individual vs. environemtal UVR exposure varied between the 4.6 and the 65.3% for FM, 20.6% on average including only data retrieved for dosimeters placed on the backs. For longshormen, the average percentage of ambient exposure received at the workers' backs was of 31.9%, with a maximum of 59 and a minimum of the 20%, dropping to the 16% considering the LSM with a chest dosimeter. It should be noted a lower variability of the results related to the solar radiation exposure of dockworkers employed as port traffic coordinators compared to fishermen and longshoremen. A possible reason is that the activity of the TC is less dynamic compared to that of the other groups: for these workers we found a percentage of personal vs. ambient exposure between the 2 and the 7%, almost stable and quite low, considering that the dosimeters in these cases were placed on the chest and that some working tasks of the workers were performed in an indoor area.

As regards to the results of the investigation on solar UVR individual exposure habits and behaviours before and after a specific sun-safety training, first of all it should be noted that, to the best of our knowledge, this is the first reporting of a sun safety initiative specifically addressing fishermen. As reported also by similar studies performed in other groups of outdoor workers [16–20], our results indicate poor protective habits and behaviors of fishermen, in particular before the sun-safety training. After our intervention, we observed a slight improvement related to the adoption of three specific individual protections: the UV protective hats, sunglasses and clothes, with an increased percentage of fishermen who reported to use, at least sometimes, these protections at work. This is still far from an adequate perception of the risk, as even after our sun-safety campaign approximately the 35-40% of the workers never or only seldom reported to use hats and sunglasses at work, while more adequate habits have been reported for protective clothes. Unfortunately, our intervention apparently failed in raising the awareness of the group of fishermen on the importance of applying sunscreens at work: this is a relevant point, and we should reflect on what can be the issues in achieving the goal of a more widespread use of sunscreens at work. Some possible problems may be related to the quite high costs of the sunscreens, in particular considering that they have to be abundantly applied and re-applied during the work-day, and to the fact that, at least in Italy, these costs are usually not covered by the companies. Nevertheless, it should also be noted that sunscreens can't be considered personal protective equipment as UV protective hats, sunglasses and clothes, but they are an additional protective measure to be adopted in case the other personal protections can't be considered sufficient to limit the exposure.

Our study also has some limitations: first of all, the sample size, in particular in the case of the fishermen, where we collected solar UVR exposure data of only seven workers: nevertheless, having followed the workers for two days, and in some cases with two dosimeters per worker, we retrieved a total number of nineteen point measurements, i.e. even more measurements compared to the campaign performed at the port, where we investigated fourteen dockworkers in only one day. Also, for the subjective investigation of solar UVR exposure habits and behaviours the sample of twenty-one subjects can be considered of a quite small size: in this case, a relevant point, possibly reducing the limitation, is that we were able to re-contact all the investigated workers after one year, and accordingly our analysis of the changes in protective behaviours after our sun-safety training resulted more valid.

A further limitation is related to the estimate of the environmental UVR doses for the measurement campaign in the fishing company. Unfortunately, we performed ambient measurements with an UVR spectro-radiometer only at the harbour, where it was possible to safely place the instrumentation in a stable site. This was not possible on the small boats of the fishermen, so that we had to retrieve environmental data from an online database, and estimate the corresponding dose in the period worked by the fishermen. Moreover, in the database we selected there was no availability of ambient data referred to the place where we actually performed the measurements (43° 9' N, 12°7' E), so that we considered available data for the city of Venice (about 200 km in the northern-east direction, at a slightly higher latitude and similar altitude).

Finally, we want to mention also some possible limitations related to the placement of the dosimeters on the workers. Unfortunately, we were not able to "a priori" select the most appropriate body districts where to fix the dosimeters according to the specific working tasks of the workers with respect to solar UVR

exposure. On the contrary, as this was an on-field measurements campaign during real working situations, we had to place the dosimeters in positions selected as the best compromise between an appropriate evaluation of the personal exposure and the need of not interfering with the usual job activities of the workers. Accordingly, we decided to place the dosimeters mainly at the upper back for fishermen and dockworkers employed as longshoremen, while at the chest for the dockworkers performing as traffic coordinators of the port area. In the few cases we had for the same workers both the data from the back dosimeter and the data from other body districts (chest, and in one case also the nape), the upper back resulted the district with the highest exposure levels.

5. Conclusions

In conclusion, our data are, to the best of our knowledge, the first Italian data demonstrating with on field measurements an intense solar UVR dose received by fishermen and dockworkers in the spring and summer seasons. According to the photo-chemical effect of the UV rays absorbed in the skin, these data indicate an excessive occupational risk, possibily resulting after years of work in an increased occurrence of skin pre-cancerous and cancerous lesions. Therefore, there is an urgent need of sun-safety campaigns for outdoor workers, in particular for those of the maritime and port sectors, as we found no previous reporting of such campaigns for these workers in scientific literature. We offered a sun-safety training to the workers we investigated with dosimeters for their personal solar UVR exposure. Before the training, workers reported very poor sun exposure habits and behaviours. After our training, fishermen reported an improvement in the use of individual UV protections, as hat, sunglasses and clothes. Unfortunately, no positive improvements have been found with respect to the use of sunscreens at work in these workers.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

For the kind collaborations permitting us to conduct this research, we are very grateful to the dockworkers company, to the Port Authority and to the Prevention Department – Occupational Medicine section of the city of Trieste. We are very grateful also to the fishermen company “Casa del Pescatore” of Cattolica, to the Port Authorities and the Coasts Guards of Cattolica and Rimini, to the Ruggieri's family and to the occupational medicine service “Studio Manini”.

References

[1] A. Modenese, F. Bisegna, M. Borra, C. Burattini, L. Gugliermetti, F.L. Filon, A. Militello, P. Toffanin, F. Gobba, “Occupational Exposure to Solar UV Radiation in a Group of Dock-workers in North-East Italy,” Proceedings - 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe, EEEIC / I and CPS Europe 2020, 2020, doi:10.1109/EEEIC/ICPSEurope49358.2020.9160703.

[2] A. Modenese, F.P. Ruggieri, F. Bisegna, M. Borra, C. Burattini, E. Della Vecchia, C. Grandi, A. Grasso, L. Gugliermetti, M. Manini, A. Militello, F. Gobba, “Occupational exposure to solar UV radiation of a group of fishermen working in the Italian north adriatic sea,” International Journal of Environmental Research and Public Health, **16**(16), 1–12, 2019, doi:10.3390/ijerph16163001.

[3] T. Kauppinen, J. Toikkanen, A. Savela, D. Pedersen, R. Young, W. Ahrens, P. Boffetta, M. Kogevinas, J. Hansen, H. Kromhout, V. de La Orden-Rivera, J. Maqueda Blasco, D. Mirabelli, B. Pannett, N. Plato, R. Vincent, “Occupational exposure to carcinogens in the European Union,” Occupational and Environmental Medicine, **57**(1), 10–18, 2000, doi:10.1136/oem.57.1.10.

[4] T. Loney, M.S. Paulo, A. Modenese, F. Gobba, T. Tenkate, D.C. Whiteman, A.C. Green, S.M. John, “Global evidence on occupational sun exposure and keratinocyte cancers: a systematic review,” British Journal of Dermatology, 1–11, 2020, doi:10.1111/bjd.19152.

[5] B.K. Armstrong, A.E. Cust, “Sun exposure and skin cancer, and the puzzle of cutaneous melanoma: A perspective on Fears et al. Mathematical models of age and ultraviolet effects on the incidence of skin cancer among whites in the United States. American Journal of Epidemiology 1977” Cancer Epidemiology, **48**, 147–156, 2017, doi:10.1016/j.canep.2017.04.004.

[6] J.C.S. Yam, A.K.H. Kwok, “Ultraviolet light and ocular diseases,” International Ophthalmology, **34**(2), 383–400, 2014, doi:10.1007/s10792-013-9791-x.

[7] K. Grandahl, J. Olsen, K.B.E. Friis, O.S. Mortensen, K.S. Ibler, “Photoaging and actinic keratosis in Danish outdoor and indoor workers,” Photodermatology Photoimmunology and Photomedicine, **35**(4), 201–207, 2019, doi:10.1111/phpp.12451.

[8] G. Ziegelberger, “Icnirp statement-protection of workers against ultraviolet radiation,” Health Physics, **99**(1), 66–87, 2010, doi:10.1097/HP.0b013e3181d85908.

[9] R.R.A. Bourne, J.B. Jonas, A.M. Bron, M.V. Cicinelli, A. Das, S.R. Flaxman, D.S. Friedman, J.E. Keeffe, J.H. Kempen, J. Leasher, H. Limburg, K. Naidoo, K. Pesudovs, T. Peto, J. Saadine, A.J. Silvester, N. Tahhan, H.R. Taylor, R. Varma, T.Y. Wong, S. Resnikoff, “Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe in 2015: Magnitude, temporal trends and projections,” British Journal of Ophthalmology, **102**(5), 575–585, 2018, doi:10.1136/bjophthalmol-2017-311258.

[10] A. Modenese, F. Gobba, “Cataract frequency and subtypes involved in workers assessed for their solar radiation exposure: a systematic review,” Acta Ophthalmologica, **96**(8), 2018, doi:10.1111/aos.13734.

[11] A. Modenese, F. Gobba, “Occupational exposure to solar radiation at different latitudes and pterygium: A systematic review of the last 10 years of scientific literature,” International Journal of Environmental Research and Public Health, **15**(1), 2018, doi:10.3390/ijerph15010037.

[12] A. Modenese, F. Gobba, “Macular degeneration and occupational risk factors: a systematic review,” International Archives of Occupational and Environmental Health, **92**(1), 2019, doi:10.1007/s00420-018-1355-y.

[13] U. Feister, G. Meyer, U. Kirst, “Solar UV exposure of seafarers along subtropical and tropical shipping Routes,” Photochemistry and Photobiology, **89**(6), 1497–1506, 2013, doi:10.1111/php.12144.

[14] K. Grandahl, P. Eriksen, K.S. Ibler, J.P. Bonde, O.S. Mortensen, “Measurements of Solar Ultraviolet Radiation Exposure at Work and at Leisure in Danish Workers,” Photochemistry and Photobiology, **94**(4), 807–814, 2018, doi:10.1111/php.12920.

[15] M. Oldenburg, B. Kuechmeister, U. Ohnemus, X. Baur, I. Moll, “Actinic keratosis among seafarers,” Archives of Dermatological Research, **305**(9), 787–796, 2013, doi:10.1007/s00403-013-1384-z.

[16] K. Grandahl, K.S. Ibler, G.H. Laier, O.S. Mortensen, “Skin cancer risk perception and sun protection behavior at work, at leisure, and on sun holidays: A survey for Danish outdoor and indoor workers 11 Medical and Health Sciences 1117 Public Health and Health Services,” Environmental Health and Preventive Medicine, **23**(1), 1–11, 2018, doi:10.1186/s12199-018-0736-x.

[17] F. Larese Filon, M. Buric, C. Fluehler, “UV exposure, preventive habits, risk perception, and occupation in NMSC patients: A case-control study in Trieste (NE Italy),” Photodermatology Photoimmunology and Photomedicine, **35**(1), 24–30, 2019, doi:10.1111/phpp.12417.

[18] A. Modenese, T. Loney, F.P. Ruggieri, L. Tornese, F. Gobba, “Sun protection habits and behaviors of a group of outdoor workers and students from the agricultural and construction sectors in north-Italy,” Medicina Del Lavoro, **111**(2), 116–125, 2020, doi:10.23749/mdl.v111i2.8929.

[19] M.M. Nkogatse, M.C. Ramotsehoa, F.C. Eloff, C.Y. Wright, “Solar Ultraviolet Radiation Exposure and Sun Protection Behaviors and Knowledge Among a High-Risk and Overlooked Group of Outdoor Workers in South Africa,” Photochemistry and Photobiology, **95**(1), 439–445, 2019, doi:10.1111/php.13008.

[20] C.E. Peters, M.W. Koehoorn, P.A. Demers, A.M. Nicol, S. Kalia, “Outdoor Workers’ Use of Sun Protection at Work and Leisure,” Safety and Health at Work, **7**(3), 208–212, 2016, doi:10.1016/j.shaw.2016.01.006.

[21] D. Reinau, M. Weiss, C.R. Meier, T.L. Diepgen, C. Surber, "Outdoor workers' sun-related knowledge, attitudes and protective behaviours: A systematic review of cross-sectional and interventional studies," British Journal of Dermatology, **168**(5), 928–940, 2013, doi:10.1111/bjd.12160.

[22] S. Schneider, K. Diehl, L. Schilling, M. Spengler, R. Greinert, T. Görig, "Occupational UV exposure and sun-protective behaviour in German outdoor workers: Results of a nationwide study," Journal of Occupational and Environmental Medicine, **60**(11), 961–967, 2018, doi:10.1097/JOM.0000000000001397.

[23] M. Boniol, A. Koechlin, M. Boniol, F. Valentini, M.C. Chignol, J.F. Doré, J.L. Bulliard, A. Milon, D. Vernez, "Occupational UV exposure in French outdoor workers," Journal of Occupational and Environmental Medicine, **57**(3), 315–320, 2015, doi:10.1097/JOM.0000000000000354.

[24] P. Gies, J. Wright, "Measured Solar Ultraviolet Radiation Exposures of Outdoor Workers in Queensland in the Building and Construction Industry¶," Photochemistry and Photobiology, **78**(4), 342, 2003, doi:10.1562/0031-8655(2003)078<0342:msureo>2.0.co;2.

[25] V. Hammond, A.I. Reeder, A. Gray, "Patterns of real-time occupational ultraviolet radiation exposure among a sample of outdoor workers in New Zealand," Public Health, **123**(2), 182–187, 2009, doi:10.1016/j.puhe.2008.12.007.

[26] M. Antoine, S. Pierre-Edouard, B. Jean-Luc, V. David, "Effective exposure to solar UV in building workers: Influence of local and individual factors," Journal of Exposure Science and Environmental Epidemiology, **17**(1), 58–68, 2007, doi:10.1038/sj.jes.7500521.

[27] M.A. Serrano, J. Cañada, J.C. Moreno, "Solar UV exposure in construction workers in Valencia, Spain," Journal of Exposure Science and Environmental Epidemiology, **23**(5), 525–530, 2013, doi:10.1038/jes.2012.58.

[28] M. Wittlich, S.M. John, G.S. Tiplica, C.M. Sălăvăstru, A.I. Butacu, A. Modenese, V. Paolucci, G. D'Hauw, F. Gobba, P. Sartorelli, J. Macan, J. Kovačić, K. Grandahl, H. Moldovan, "Personal solar ultraviolet radiation dosimetry in an occupational setting across Europe," Journal of the European Academy of Dermatology and Venereology, **34**(8), 1835–1841, 2020, doi:10.1111/jdv.16303.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1319-1327 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Combining ICT Technologies To Serve Societal Challenges

Helen Leligou[*,1], Despina Anastasopoulos[1], Anita Montagna[2], Vassilis Solachidis[3], Nicholas Vretos[3]

[1]*INTRASOFT International S.A., Research & Innovation Development Unit, Luxembourg, L1253, Luxembourg*

[2]*Centro Studi Pluriversum, 53100, Italy*

[3]*Information Technologies Institute, Center for Research and Technology Hellas, Thessaloniki, 60361, Greece*

ARTICLE INFO

Article history:
Received: 27 November, 2020
Accepted: 09 February, 2021
Online: 28 February, 2021

Keywords:
Artificial Intelligence
Data Analytics
Conversational Interfaces

ABSTRACT

European counties continue to receive an increasing number of migrants and refugees from an also increasing number of both European and non-European countries. This results in a huge societal challenge which is societal inclusion of people speaking different languages and of diverse backgrounds. Key for their inclusion is job finding which comes with hurdles like the language, the difficulty in assessing and certifying their skills and many more. In this paper, we present the architecture of a novel platform that aspires to provide migrants with a) assistance in discovering and assessing their hard and soft skills by employing Artificial Intelligence technologies, b) recommendations of appropriate job sectors and positions based on their profile, c) recommendations about training that would allow them to find jobs in the country/region they are located and d) practical information regarding the integration process. Furthermore, the proposed platform aims at assisting host authorities, non-governmental organizations and companies in detecting the needs of the target populations (migrants and refugees) through data analytics and supports them in reaching them. Apart from the technical architecture, we provide the results from the initial testing of the platform in real-life pilots in two countries.

1. Introduction

The current trend of migration for economic reasons combined with large flows of refugees ("people fleeing their countries due to armed conflicts or persecution") challenges our societies. A key to societal inclusion of migrants and societal coherence is their integration in the labour market. Nevertheless, the elaborate human resource recruitment processes established in our economies make it difficult and complicated not only for the migrants but also for host organisations, according to the claims gathered through interviews with field experts. The main barriers (also discussed in [1]) are multiple: a) lack of professional qualifications and certified skills and competences; b) language skills are impeding fluent communication and this is even more intense for refugees; c) the scale of the phenomenon which is large and increasing: over a million refugees in Europe since 2015, 70% of them with ages between 18 and 64, i.e. in working age and low-skilled people percentages greater than the average of the hosting countries' [2]. Refugees and migrants are a heterogeneous group as far as education and other issues are concerned and thus, special vulnerabilities have also an impact on individuals' labour market integration chances.

In the host societies, there are organisations (public, private or Non-Governmental Organisations) that try to assist migrants and refugees in finding their way to the labour market. However, two factors limit the positive impact of these efforts to a certain extent: a) the multiplicity of entities operate usually with no coordination among them and b) the fact that the potential employers have to follow stricter regulations when hiring migrants and also have to afford the cost for screening their skills and profiles at the absence of any certificate and/or proof. At the same time, tailored training is considered of paramount importance to increase the probability of employment of refugees and migrants according to the Committee for Labour Migration [3]. Soft skills are also important and should not be overlooked. The new recommendation on Key Competences for Lifelong Learning [4] has enhanced the importance of key competences and soft skills, which are considered as equally important to hard skills when looking for a job. The evaluation of the soft skills is also difficult to be

[*]Corresponding Author: Helen Leligou, Email: leligou@gmail.com

performed imposing further barriers in the integration of migrants in the labour market.

The exploitation of Information and Communication Technologies (ICT) is considered an appropriate mean to communicate with the migrants and refugees as they usually have access to the internet usually through mobile devices. ICT technologies could also contribute to personalization of the information presented to the users. However, up to now, ICT has been used mainly for information provisioning which is valuable but yet not enough. As a result, the integration of migrants and refugees is still a challenge for all involved actors i.e. migrants, refugees, host authorities, NGOs, private and public organisations and society alike.

In this article, we present a novel digital platform that aspires to support not only migrants and refugees but also the host authorities and organisations towards the integration of the first group in the labour market. The proposed platform employs state-of-the-art technologies like artificial intelligence and data analytics to lower the cost of skills evaluation, facilitate the operation of organisations and significantly shorten the path of migrants to the labour market. In its design, we have combined Artificial Intelligence, data analytics, native mobile applications and web technologies with social science expertise and skill assessment techniques to offer a novel holistic platform to the migrants. The proposed platform is not a "bundling" of functionalities like the ones offered by multiple other platforms; instead, it offers novel functionalities like the skill assessment functionality based on machine learning algorithms trained on data from migrants' populations. The proposed platform has been prototyped and used by migrants/refugees in real life.

The rest of this article is organized as follows: in section 2, we present existing digital platforms tackling the migrants' integration challenge and discuss their strengths and weaknesses. In section 3 we describe the proposed platform, and its technical architecture is presented in section 4. Section 5 presents the results from the evaluation in real life and section 6 concludes the article. A preliminary version of the presented approach was initially published in [1] without technical details and any evaluation results.

2. Existing Platforms

Up to now, few tens of platforms primarily targeting the needs of the migrants/refugees have been deployed exploiting the fact that the majority of them has access to the internet through their mobile phones. The following table is based on the table that appears in [1] and reflects the current situation. The fact that such platforms proliferate, show that they cover actual needs and thus the involved actors try to satisfy them. The table includes the information considered necessary for their comparison.

Examining the table, it is evident that there is no platform adopting a holistic approach. Most of these platforms present to the user information but the supported functionalities of search and the limited languages reduce their actual value. Just one of these platforms supports skill assessment and this targets a limited set of job sectors. This is mainly due to the language barrier which is difficult to overcome with any textual-based assessment. As explained in [5], there are methodologies to assess skills and competencies, but they are mainly twined with the education processes which is not a representative case for the migrants arriving in a new country. The use of neural networks has been proposed in [6] in the context of assessing skill of surgical doctors. With respect to digital tools, they usually evaluate performance in a normative way after the user has completed an exercise. In many cases, this is not possible without the subjective scoring of an external evaluator.

To sum up, guidance through the path from arrival to labour market integration is provided in a rather segmented way by existing platforms. Having a unique platform offering all these services through a single point and having a digital companion guiding the user through it would be of high value.

Table 1: Online platforms for migrants

Platform	Organisation	Target	Type of platform	Services	Language
https://mygrants.it/	Mygrants - social enterprise - Bologna	Migrants	online training; information platform	Information on asylum and rights, social challenges, entrepreneurship; formal training opportunities; non-formal training	English French Italian
https://www.arbeitsagentur.de/fuer-menschen-aus-dem-ausland	German Ministry of Work	Migrants	information platform; online training (language); recognition of qualifications; job vacancies; e-services for inclusion	Information and support on financial, social insurance, German education, starting a business, finding a job in Germany.	German English Arabic
https://jobskills.arbetsformedlingen.se/	Swedish Public Employment Service	Migrants	job matching service	Creation of a profile, career guidance services; job matching services	English Swedish Arabic Farsi Somali Amharic

Platform	Organisation	Target	Type of platform	Services	Language
http://myskills.de/	Berufliche Kompetenzen Erkennen	Migrants	assessment platform	Online recognition of skills for people with no vocational qualification (8 professional profiles assessed)	German English Turkish Arabic Russian Persian
https://ec.europa.eu/migrantskills/#/; https://ec.europa.eu/social/main.jsp?catId=1412&langId=en	European Commission	Third country nationals	self-assessment platform	Online self-assessment tool to create a profile of skills	EU languages Arabic Farsi Pashto Sorani Somali Tigrinya Turkish
https://www.make-it-in-germany.com/en/living-in-germany/integration/advisory-service/	German Government	International qualified professionals; prospective international students; employers; Institutions	information platform	Informational content on jobs, study and training, Visa and living in Germany for those who want to move to Germany	English German Spanish French
https://www.universitaperrifugiati.it/en/default.aspx	Uninettuno telematic University - Rome	Migrants and refugees	training platform; recognition of qualifications	Recognition of study titles previously earned in their countries of origin to access the University; recognition of vocational skills	English French Italian Arabic
http://www.moocs4inclusion.org/index.php	Project funded by European Commission	Institutions with focus on migrants and refugees	training platform	Catalogue of free digital learning initiatives for migrants	English
https://www.refugee.info/	Signpost project Mercy Corps and International Rescue Committee	Migrants	information platform	Informational content for migrants, geo-mapping of services, FAQ, contact with NGOs	Italian Persian Urdu Arabic English French
https://www.infomigrants.net/en/	ANSA, Deutsche Welle, French media companies	Migrants	information platform	Information content and news for migrants	French Arabic Persian Pashto English
https://migrationdataportal.org/?i=stock_abs_&t=2017	IOM	Citizens; organisations	information and data	Data about migration	English
https://openmigration.org/en/	Italian Coalition for Civil Liberties (CILD)	Citizens; organisation	information and data	Data about migration	Italian English
http://migration.iom.int/europe?type=arrivals	IOM Migration	Citizens; organisation	information and data	Data about migration	English
https://iamamigrant.org/	IOM	Migrants and citizens	information platform	Information content and news for migrants	Arabic French Farsi Pashto

Platform	Organisation	Target	Type of platform	Services	Language
					English
https://euraxess.ec.europa.eu/jobs/science4refugees	European Commission	Refugees researchers and research groups with vacancies	matching platform for science vacancies	Information and support services on working and living in Europe; research refugee friendly internships, part-time and full-time jobs, access to an European Research Community	English
https://www.migrantsorganise.org/	a consortium of partners	Migrants and activists	online community		English

3. The Proposed Solution

We have designed a novel platform which has multiple target groups: a) refugees which are seeking for asylum, b) migrants who are seeking a better life, c) vulnerable migrants (e.g. suffering diseases), d) host authorities which are trying to improve their offerings and facilitate the integration of migrants and refugees, e) non-governmental organisations and f) organisations that consider migrants as receivers or target groups of their services.

For the users falling under the first three aforementioned categories, we aspire to assist them in preparing effective job applications and in finding appropriate positions to apply. For local people, finding a job consists primarily of three steps: as shown in Figure 1 below: 1) inspect available job positions which is usually done through a web-based job position portal or a professional social network or (more rarely) through a Human Recruitment agent; 2) filter the jobs and select those of interest to apply. Language is not a barrier for locals and the creation of a CV or an application is usually not an issue. 3) go through (one or more) interviews or follow the recruitment process of the specific organization where again, in principle,communication and specific aid is not required, although ways to maximise the potential of getting the job exist and are gaining momentum, as shown in Figure 1. For migrants and refugees, each of these steps represents a challenge: accessing the job announcement, understanding them (due to language barrier) and selecting some of them are of great difficulty. Often the CV format is not familiar to them, and usually their qualifications are not certified since they are gained via informal and nonformal experiences.

Figure 1 shows the process of labour integration without the proposed platform on the left (marked as a) and with the functionalities offered by the proposed platform on the right. As shown in the right (figure 1b), the 1st step is completely changed and automated by the proposed platform:

1. to prompt the selection of job position and relieve the migrants from searching in relevant portals which they may not be aware of, the platform prompts them to fill in their skills and profile.
2. Additionally, the platform prompts them to answer dedicated questionnaires which enable the evaluation of soft and hard skills. Artificial intelligence technology is used to evaluate the skills through these questionnaires as it will be elaborated in the next section.
3. As many refugees and migrants do not possess certification of their qualifications, the platform supports them in entering other types of proofs (e.g. images or videos). The proposed platform offers to them guidance through a wizard (step-by-step approach) and also provide videos to help them create their portfolio.

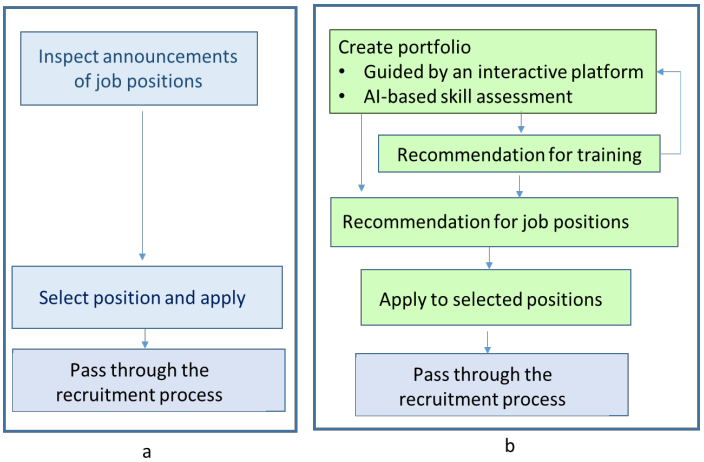


Figure 1: The path to the labour market now (shown in (a)) and with the proposed intervention (shown in (b), on the right)

Once the profile and skills are registered, the user is provided with

- Personalized recommendations about the job categories he/she is appropriate for, based on the definition of skills relevant to each occupation or more specifically to the ESCO code, i.e. to the code of the occupation category in the multilingual classification of European Skills, Competences, Qualifications and Occupations [7].
- Recommendation of additional training modules which would significantly improve their potential for getting a job. For example, assuming a user has a set of certain skills and according to these skills he matches job category A. If a skill is also needed for this job category, then a training that leads to the improvement of this skill is suggested to the migrant. This way, the migrant/refugee becomes aware of opportunities to maximise their potential for being hired.
- Once this is done, the migrant/refugee is ready to get from the proposed platform a list of job positions related with the job categories he is appropriate and then (if the wishes) to apply (e.g. send his CV created in the previous step).

To facilitate the migrants/refugees life easier, a conversational interface (a chatbot) is integrated in the platform. The proposed chatbot and the whole platform support multiple languages as it is crucial for the migrants and refugees to easily find the information they are looking for in their native language and understand the results of their activities.

Turning out attention to the rest target user groups mentioned at the beginning of this section and to offer them valuable assistance, we have designed the platform so as to represent a single-point-of-entry for them with regards to information about refugees/migrants flows, demographics of the migrants' population per country, statistics about their needs and skills as well as the means they find more useful for finding a job. For these groups, once they have set a concrete goal, (e.g. assist vulnerable migrants in finding jobs), they need as accurate information as possible and as fast as possible because the resources that can be devoted are scarce. The proposed platform essentially aggregates information and presents it to the users (indicating of course the source of the information) in an easy to understand way by employing data analytics.

Table 2: The offerings of the platform for the two target groups

Offerings targeting migrants
1. multilingual environment (English, Greek, French, Arabic, Spanish and Farsi) 2. Information about employment 3. Information about social integration 4. Skill assessment using a) Questionnaires and b) Serious games 5. Online training including a) Text, b)Videos and c) Self assessment questionnaires 6. Relocation recommendations based on the migrant's social needs 7. Recommendation of appropriate job categories (using ESCO classification) 8. Recommendation of appropriate job vacancies 9. Recommendation of trainings 10. Chatbot that is trained on the employment/social information and provides answers to migrants questions in many languages
Mobile app for foreign language learning and assessment
Tools for Host authorities, Companies and NGOs
1. Graphs and statistics about the skills of the migrants 2. Recommendations for the host authorities
Statistics regarding migration and employment in Europe

Additionally, to facilitate the task of organizations working with the migrants, the platform supports the organization of users (migrants) in groups. This way, if an organization wants to summarize the skills of a group they are managing/handling, the organization can ask the migrants to use the group identifier and fill in the skill assessment questionnaires. Then the organization can observe the statistics of the skills of this specific group and look for appropriate job positions and relevant trainings. This was considered a very valuable tool during the interviews we organized with such organisations.

To sum up, the major difference between the proposed method and the other in the literature is that no other platform adopts a holistic approach. The offerings of the proposed platform are summarised in the Table 2.

4. Technical Architecture

The technical architecture of the proposed platform is shown in Figure 2, where the distinction between the front-end components, the backend components and the repositories is clearly shown. The platform runs in a cloud environment. The front end is organized in two sets of components: a) those serving the migrants/refugees (shown on the left in green) and b) those serving Host authorities, Companies and NGOs (HCN) (shown on the right in orange). Both sets include a) a home page which is the landing page of the platform and is different per user groups once the user is logged in, b) an account page- here the user creates his/her profile inserting different information depending on the user type they belong and c) the "about" page where information for the platform, the services it provides and the way it handles the collected data is provided. In the account page, the host authorities, companies and NGOs insert their name, the services they offer and the geographical area where they operate (for example, health services in Athens). In the same page, if the user has entered as a migrant, the information collected includes personal information.

The interface for the migrants/refugees further includes: a) the "e-portfolio" component which guides step-by-step the migrant to fill in his/her skills and competences in a very user friendly way; b) the "chatbot" which responds to questions relevant to legal, social and health – it supports relevant questions in six different languages; c) the skill-assessment (SA) questionnaires which include multiple choice questions and questions requiring the association of elements (visual or textual); d) the recommender User Interface (UI) that offers recommendations for trainings and job vacancies and e) the social/career interface which offers general information about relevant issues.

The interface for the HCN group includes (apart from the three components already presented) a) the data analytics component through which the HCN finds information about the migrants' flows and needs as already described above, b) the group analytics, which presents to the HCN statistics about the group they manage and they have the rights to access (e.g. on the results from the skill assessment or their needs) and c) the social/career interface where they can upload pieces of information they consider of value to the migrants/refugees.

The back-end part is where intelligence resides. First of all, the component "neural network algorithms for skills assessment" implements artificial intelligence to evaluate a mix of thirty five hard and soft skills (like numeracy and communication) based on the serious games the migrants played. This component uses a model which has been trained with data that come from the local society. To improve the accuracy of the evaluation, we are using the platform to collect data from migrants and train the model with them to provide more accurate evaluation results.

The chatbot backend component is responsible for the processing of the information available to it (the migrants' question) by providing the appropriate answer. The proposed chat-bot [8] is a Question-Answer oriented agent, capable of replying to migrant administrative service-related queries and it incorporates two different systems, a daily conversation agent

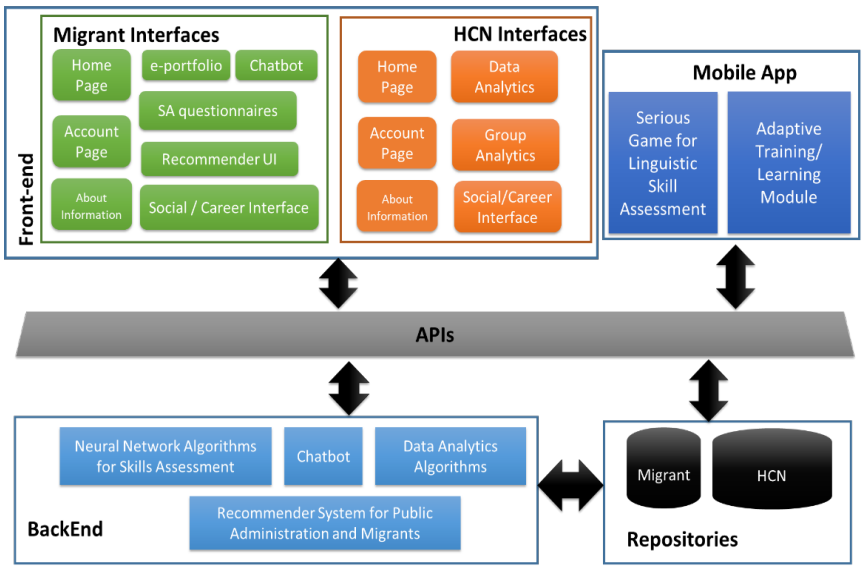


Figure 2: The technical architecture of the presented platform

based on the model released by Huggingface and an answer retrieval module. The latter initially searches for valid in regard to the user query documents, splits them into semantic paragraphs and then applies a BERT [9] fine-tuned model to finally reply to the user.

The methodology of initially splitting the document into paragraphs depending on the structure of the text, rather than directly forwarding the whole document to the model has shown improved results. Towards extending the chatbot's capabilities, aim is to further take advantage of paragraph embeddings during answer prediction, incorporating them online into the models. A challenge faced by this specific chatbot application is the support of multiple languages and more specifically the support of non-European languages like Arabic and Farsi which are spoken by significant percentages of the migrants and refugees arriving in European countries. As presented in [10] the most common countries of origin to Europe (Mediterranean) for 2019 is Afghanistan, Syria, Morocco and Algeria hence Farsi, Arabic as well as European languages (France Spanish and English) cover the vast majority of the migrant spoken languages.

The Data Analytics algorithms is the component that retrieves data from the repositories so as to present to the HCN information about migrants' flows and needs, as well as information about the performance and characteristics of each group they manage. Thus, it is connected to the data analytics and group analytics component of the front-end. More specifically, this module can display not only the skills of individuals but also the average skills of the entire group. As a result, each HCN is able to identify in an easy and user-friendly way the strengths and weaknesses of each group of migrants in order to offer to them tailored trainings based on their skills and competences.

Finally, the recommender system is responsible for the following functions:

a) it retrieves information about the migrants' social needs (which are declared by himself) and suggests relocation options. In order to identify the most suitable city for relocation, the recommender uses Quality of Life tables of European cities and correlates them with the user needs.

b) it retrieves information regarding the skills of the migrant/refugee and recommends appropriate ESCO job categories and job vacancies. More specifically, the recommender system includes 40 occupation categories, that are matched to all the second level of ESCO-classification, ensuring representativeness of all job categories that exist in the labour market. For every job category we set up thresholds that specific skills must pass in order to assign this job category to the specific user. For every job category the algorithm checks if the skill vector of the user (35 skills) passes the thresholds and if the result is positive then it returns the respective job category.

Additionally, the proposed platform suggests specific job vacancies to the users based on their recommended ESCO categories and location (city). The job vacancies are retrieved from job vacancies web sites and then each job title is assigned to one or more ESCO codes using the ESCO Web Services. The vacancy is considered suitable for the user if it is assigned to ESCO categories that have been recommended to the user. The recommender system allows the user to search for vacancies from a long list of European cities. However, the platform offers a personalized list of cities based the social needs matching (the top cities are the most suitable) in order to assist further to his relocation.

c) it recommends to the migrant trainings that would significantly improve his/her potential of being hired. The platform is equipped with training sessions designed for each of the 35 skills. Each session includes training material (text, video presentations) as well as testing and self-assessment activities.

d) it presents to host authorities evaluations of eight indicators (social, health and other defined by European initiatives) for different European cities. This way a mayor of city A could easily compare city A with other cities with respect of the eight considered aspects, check in which aspect(s) city A is lagging behind the rest of the cities and decide to proceed to implementing

measures to improve city A. For example, if city A has low evaluations in health services, the mayor may trigger appropriate actions and measures to improve health services in city A. With respect to the relocation options, it is worth providing an example: the platform prompts the migrant to declare their priorities with respect to their needs and takes them into consideration when the platform suggests to the migrant to relocate. This implies, that if a migrant has declared that health services are of premium importance, the platform will not suggest him to relocate to a city with low evaluations in health services.

To support this rich functionality, we have deployed a relational database which stores the data relevant to migrants and refugees (in one part) and relevant to HCN in the other as shown in Figure 2. To enable the communication between the repository and the frontend and backend components, a set of Application Programming Interfaces (API) has been developed.

Apart from the platform which runs through any browser and can also run as a web-based mobile app, we have considered the case where a native mobile application can be integrated with the platform. We assume a mobile application that allows the user to assess its reading and speaking competence in English or other EU languages. Having a native mobile application in place allows the evaluation of pronunciation apart from understanding. Although this application could include all the front-end interfaces of the platform, we have opted not to do so and we have kept it as different app to showcase that any relevant mobile app developed by a 3rd party, could be integrated with our platform and use the open API we have developed to include the results in the portfolio (upon specific minor modifications). Once the user interacts with the application and completes the assessment, the score (the level A1/A2, or other according to European qualification levels) is stored as a hard skill in the repository. It is stressed that the platform supports the evaluation and storage of a set of hard skills, i.e. skills that are easy to quantify like proficiency in a language and a set of soft skills like communication and leadership which are often also called interpersonal skills. This information is directly integrated in the user's portfolio (and CV) and is considered when they seek for job recommendations. Similarly, if the recommender provides to the user a recommendation to improve his English-speaking competence, a mobile app developed for this purpose will be proposed. This app can be the one described above or any other app providing foreign speaking competence. Any training opportunity can be registered in the platform and suggested to migrants, if it fits their needs. The result from the training can be integrated in the e-portfolio and can be used during job recommendations.

In the following Figure 3, indicative screenshots of the platform are shown. In Figure 3a, the page where the user selects his type (migrant/refugee or HCN) so that the platform provides relevant information. In Figure 3b, the landing page of the migrants is shown while in Figure 3c the web page where the portfolio is created is depicted. Finally, in Figure 3d, the landing page of the HCN user is shown.

5. Evaluation in Real Life

The presented platform has been used in real life by migrants and refugees in two European countries receiving large flows: Spain and Greece. The piloting was organized in the framework of H2020-NADINE project. For the evaluation purposes, three different sub-groups of the first user group was considered: a) asylum seekers, b) recognized refugees and c) vulnerable migrants.



Figure 3: Screenshots from the platform: a) selection of type of user, b) the screen offered to the migrant, c) the page where the migrant creates their portfolio, d) the web page offered to the HCN type of user

They were invited in the premises of NGOs in these countries in small groups. They first watched a demonstration of the platform by a NADINE person which showed the different functionalities and then they were asked to use the platform and evaluate it through online questionnaires to have an objective and harmonized evaluation. The users spent two days in the premises of the NGOs. They also had the option to comment on the platform and the different pages and results they were presented with. The assessment targeted each functionality of the platform separately. The collected results and comments from the two-days tests and the 96 users were used to improve the platform. Further details on the evaluation activities, demographics of the users and details about the evaluation of each functionality can be found in [11].

The results from the real-life evaluation per functionality are shown in Figure 4. The users were asked to rate the functionalities in a 5-level Likert scale. The results showed that for e-portfolio and registration page, the "excellent" and "good" levels were scored by the majority of users. This is considered satisfactory as the proposed platform was used in Farsi and Arabic in the pilot tests. For the rest three functionalities, the results are slightly lower which comes as no surprise because any artificial intelligence-based component requires training. Up to now the AI-based components of the platform have been further trained based on limited datasets. We anticipate that as the platform is more and more used, we will use the data we collect to further train the AI-based components so that these provide more valuable results to the migrants and thus enhance their satisfaction. Namely, for the chat bot, 68% of users evaluated this as "average" and above. It is worth stressing that the migrants "talk" to the chatbot in Farsi and Arabic. Actually, for all functionalities (apart from the registration page), more than 66% of the users evaluate it as average and above. This shows that the platform, although it is currently in a prototype form, succeeds in satisfying the users. We anticipate that this first real time piloting results will help us improve the platform so as to beeter satisfy their needs. This also contributes to the sustainability of the platform. The "language assessment" functionality was tested through a mobile app which injected the results in the presented platform. This way the potential of integrating 3rd party applications was confirmed.

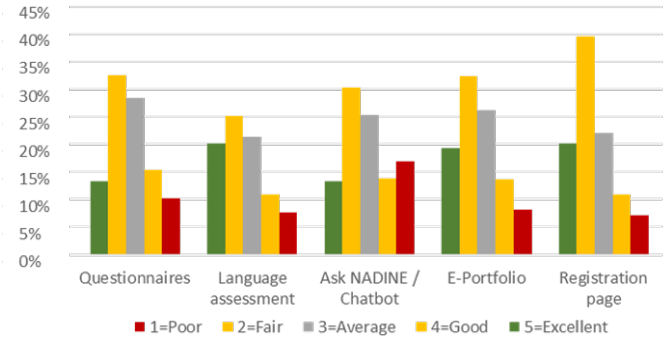


Figure 4: Statistics of the assessment results per component

For each functionality included in the figure, a set of 10 indices were evaluated including: "easy to use", "enjoyable activity", "meet my expectations", "comprehensive/clear", "time spending", "cover my personal needs", "easy to navigate", "overall aesthetics" along with some additional ones which are linked to the specifics of each functionality. These results together with the comments received through free text questions are now used to improve the platform. A common observation was that higher levels of satisfaction were indicated by vulnerable migrants compared to the rest two categories of recognized migrants and asylum seekers. In the language assessment, the percentages of results do not sum up to 100% as a percentage of the pilot users did not test and thus did not rate the language mobile app.

Moreover, and in order to evaluate the participants' overall experience from the platform, more specific aspects were also examined: Likelihood of recommending NADINE platform, User friendliness of NADINE platform, on which level the platform covers the participants' personal needs / requirements, How comprehensive and clear it is, Easiness to use and understand NADINE platform, Easiness to navigate through NADINE platform, Effectiveness and usefulness. All results regarding the evaluation of the participants' overall experience from NADINE platform are thoroughly demonstrated in the Figure 5. The result is very satisfactory, as the strong majority reports good experience for the platform which is currently in prototype level. Further details of the evaluation can be found in [11].

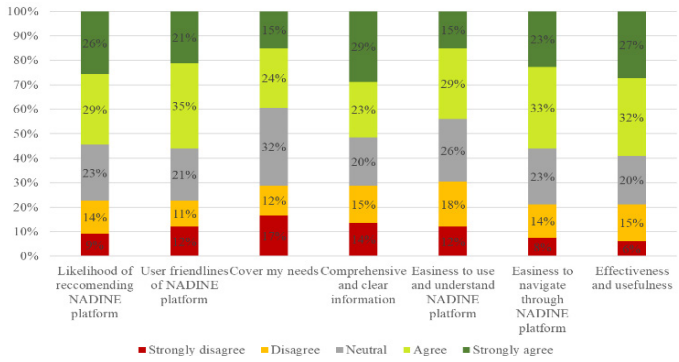


Figure 5: Statistics of the assessment results per component

Turning our attention to the host authorities, NGOs and private organisations, in this phase, ten (10) interviews were conducted. The main outcomes of the interviews were: a) finding information in a single point saves a lot of time for them as they have a clear picture of the situation in significantly shorter time; b) for NGOs working with migrants, the data analytics presented by the platform regarding the information relevant to the groups of migrants they work with was considered an extremely valuable tool; c) for host authorities, the most valuable functionality was the ability to see the sector (social, health, other) that they lack behind and also the ability to communicate with the migrants; for organizations working towards job finding for migrants, having in place a component that evaluates the skills in an accurate way for migrants and refugees is very valuable. They all consider the platform a valuable tool for their everyday job which creates the expectation that the number of migrants and of HCN users engaged with the platform will continuously increase. HCN actors considered that the deployment of a platform like the presented one requires funding from host authorities (regional or national) or EU-wide initiatives.

6. Conclusions

The societal inclusion of migrants and refugees passes through their integration in the labour market. A platform that supports all

their way to the labour market has been proposed in this paper. Its novelty lies not only in the fact that it does support them along the whole path but also on the fact that it exploits mature ICT technologies like artificial intelligence and data analytics to offer unprecedented experience to the users. Utilising artificial intelligence, a set of soft and hard skill are assessed through the platform. Through an appropriately designed user interface, the users are guided to create their e-portfolio. Based on this information, appropriate job sectors and concrete job positions are suggested to the users. The host organizations in turn enjoy mainly data analytics which offer them appropriate information facilitating their decision making. The offerings of the proposed platform prototype have been positively evaluated in real life in two countries receiving large numbers of refugees and migrants.

Acknowledgment

The work presented in this document was funded through H2020-NADINE project. This project has received funding from the European Union's Horizon 2020 Programme under Grant Agreement No. 822601.

References

[1] H.C. Leligou, D. Anastosopoulos, A. Montagna, V. Solachidis, N. Vretos, "Embracing novel ICT technologies to support the journey from camp to job," in 2020 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2020, Institute of Electrical and Electronics Engineers Inc., 2020, doi:10.1109/ICMEW46912.2020.9106035.

[2] H.U. of E. Reallabor Asyl, The labour market integration of refugees ' white paper – A focus on Europe, 2017.

[3] International Labor Organization, "How to Facilitate the Recognition of Skills of Migrant Workers Guide for Employment Services Providers Second Edition," 99, 2020.

[4] Council Recommendation of 22 May 2018 on key competences for lifelong learning Text with EEA relevance., 2018/C 189/01, 2018.

[5] L. Falahati, L. Paim, M. Ismail, S.A. Haron, J. Masud, "Assessment of university students financial management skills and educational needs," African Journal of Business Management, **5**(15), 6085–6091, 2011, doi:10.5897/AJBM10.1583.

[6] A. Jin, S. Yeung, J. Jopling, J. Krause, D. Azagury, A. Milstein, L. Fei-Fei, "Tool Detection and Operative Skill Assessment in Surgical Videos Using Region-Based Convolutional Neural Networks," 2018. arXiv preprint arXiv:1802.08774, 2018.

[7] https://ec.europa.eu/esco/portal/escopedia/Occupation

[8] A. Lelis, N. Vretos, P. Daras, "Nadine-bot: An open domain migrant integration administrative agent," in 2020 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2020, Institute of Electrical and Electronics Engineers Inc., 2020, doi:10.1109/ICMEW46912.2020.9106024.

[9] C. Alberti, K. Lee, M. Collins, "A BERT Baseline for the Natural Questions," arXiv preprint arXiv:1901.08634, 2019.

[10] Refugee and Migrant Arrivals to Europe - Jan to Dec 2019, Feb. 2021. - https://data2.unhcr.org/en/documents/details/74670

[11] https://nadine-project.eu/deliverables/

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1328-1333 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Analysis of qCON and qNOX Anesthesia Indices and EEG Spectral Energy during Natural Sleep Stages

Joana Cañellas*, Anaïs Espinoso, Juan Felipe Ortega, Umberto Melia, Carmen González, Erik Weber Jensen

Quantium Medical, Research and Development Department, 08302 Mataró, Spain

ARTICLE INFO

Article history:
Received: 25 December, 2020
Accepted: 14 February, 2021
Online: 28 February, 2021

Keywords:
sleep, consciousness
nociception, anaesthesia
monitor, REM, NREM

ABSTRACT

The objective of this research is to study the behaviour of the anaesthesia monitor Conox during natural sleep to open the gate for this devices to assess subjects during this stage. The values of qCON and qNOX indices and EEG frequency bands are analysed during night sleep of 10 volunteers when they lose consciousness, in order to determine if they can be used for monitoring sleep. The possibility of using these indexes to differentiate between NREM/REM cycles of night sleep is studied. A reduction in the hypnotic index was observed while the nociception index stayed significantly higher. Statistical differences where found for qCON between sleep cycles, allowing this index to detect REM intervals and possibly opening the gate to use depth of anaesthesia devices to monitor sleep.

1 Introduction

This paper is an extension of work originally reported in 11^{th} Conference of the ESGCO [1].

Sleep disorders such as insomnia, abnormal movements or behavior during night, and inability to control the sleep hours, can cause severe health problems such as depression or attention deficit hyperactivity disorder. Hence, sleep is widely studied in order to develop methods to detect these disorders by its monitoring. Some relevant laboratory tests for studying sleep disorders consists of an overnight polysomnography, measuring how fast the patients fall asleep and their attention for being awake as well as actigraphy [2]. Nowadays, there are commercial medical devices designed to monitor the level of consciousness during general anesthesia, allowing the anesthesiologists to individually tailor the administration of anesthetics. These devices monitor and process the electroencephalography (EEG) that is recorded by non-invasively locating electrodes on the forehead of the patient. The frequency and the amplitude of the EEG evolve because of changes in cerebral activity between wakefulness and sleep stages, and these variations can also be used to differentiate between the rapid eye movement (REM) and non-rapid eye movement (NREM) states of sleep. The energy of the EEG frequency bands Delta, Theta, Alpha, Beta and Gamma (from lower to higher frequencies) also varies between natural sleep stages according to literature [3], being the REM stage characterized by a decrease in the energy of low frequencies such as Delta band, and an increase in high frequencies as Beta and Gamma bands. Previous studies [4, 5] have proposed to use EEG monitoring devices to measure also the natural sleep further than using them only for anesthesia assessment. In this work we analyzed the behavior of the anesthesia monitor Conox (Fresenius Kabi, Bad Homburg, Germany), during natural sleep. Sleep disorders such as insomnia, abnormal movements or behavior during night, and inability to control the sleep hours, can cause severe health problems such as depression or attention deficit hyperactivity disorder. Hence, sleep is widely studied in order to develop methods to detect these disorders by its monitoring. Some relevant laboratory tests for studying sleep disorders consists of an overnight polysomnography, measuring how fast the patients fall asleep and their attention for being awake as well as actigraphy [2].

Nowadays, there are commercial medical devices designed to monitor the level of consciousness during general anesthesia, allowing the anesthesiologists to individually tailor the administration of anesthetics. These devices monitor and process the electroencephalography (EEG) that is recorded by non-invasively locating electrodes on the forehead of the patient. The frequency and the amplitude of the EEG evolve because of changes in cerebral activity between wakefulness and sleep stages, and these variations can also be used to differentiate between the rapid eye movement (REM) and non-rapid eye movement (NREM) states of sleep. The energy of the EEG frequency bands Delta, Theta, Alpha, Beta and Gamma (from lower to higher frequencies) also varies between natural sleep stages according to literature [3], being the REM stage characterized by a decrease in the energy of low frequencies such as Delta

*Corresponding Author: Joana Cañellas, Av. Ernest Lluch 32, Office 3.17, Mataró 08302, +34 619534413 | Email: jc@quantiummedical.com

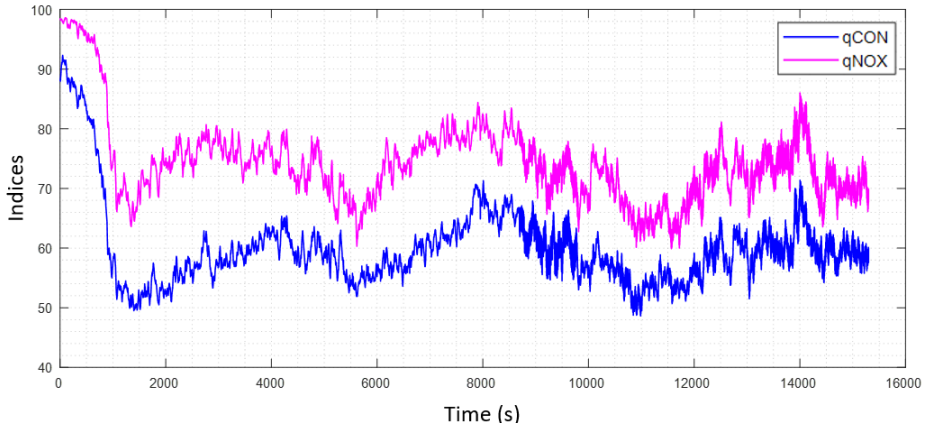


Figure 1: Mean qCON and qNOX for all volunteers from 15 minutes before loss of consciousness (qCON=60) until 4 hours after.

band, and an increase in high frequencies as Beta and Gamma bands. Previous studies [4, 5] have proposed to use EEG monitoring devices to measure also the natural sleep further than using them only for anesthesia assessment. In this work we analyzed the behavior of the anesthesia monitor Conox (Fresenius Kabi, Bad Homburg, Germany), during natural sleep.

2 Methods

In order to obtain the evolution of both anesthesia indexes during natural sleep, 10 volunteers were accepted for this study and agreed to be connected to a Conox device during one night sleep. Signal registration started a few moments before the subject falling asleep and lasted for at least six hours. As an indicator for unconsciousness, the same threshold as in clinical practice during a surgery was used, being a qCON value of 60. A Mann-Whitney U test and also a paired Wilcoxon test were performed in order to analyze if there was significant difference between the two indices. Instants of both indices falling below different thresholds (90, 80, 70 and 60) were also computed in order to know the time decay of each index when the subjects are losing consciousness.

From the EEG signal recorded we analyzed the spectrum in five frequency bands: Delta (0.5 - 4 Hz), Theta (4.5 - 8 Hz), Alpha (8.5 – 13.5 Hz), Beta 1 (14 - 20 Hz) and Beta 2 (20.5 - 30 Hz). Beta energy has been divided in two separate ranges for a better differentiation of the behavior of each frequency since the whole Beta range is very wide. The energy was calculated applying the Fast Fourier Transform (FFT) on 2-seconds EEG windows with overlap of 1 second. The FFT was normalized by the energy contained between the 0.5 and 30 Hz range. A 60 samples moving average smoothing filter was applied for a better visualization of the results, and the mean value of each frequency band was calculated during awake and asleep stages. It was considered that patients were awake when the qCON value was above 60, and asleep otherwise. To determine if there were significant differences between the values of the frequency bands between sleep and awake stages, a Wilcoxon signed rank test was used, at significance level of 0.05.

Taking into consideration the variation of energy in high and low frequencies, to distinguish between REM and NREM stages, we proposed an adaptive threshold for Delta and Beta 2 energy based on a preliminary data analysis which indicated that these frequency bands are the most affected by the transition between sleep stages. Delta threshold is calculated as the 80 % of the difference between maximum and mean energy values during the last 3600 seconds of recording. Beta 2 threshold is calculated as the 80 % of the difference between mean and minimum energy values during the last 3600 seconds of recording. We considered that a REM interval was detected when Delta energy was below the threshold and Beta 2 energy was above it, short intervals being less than 5 minutes were not considered.

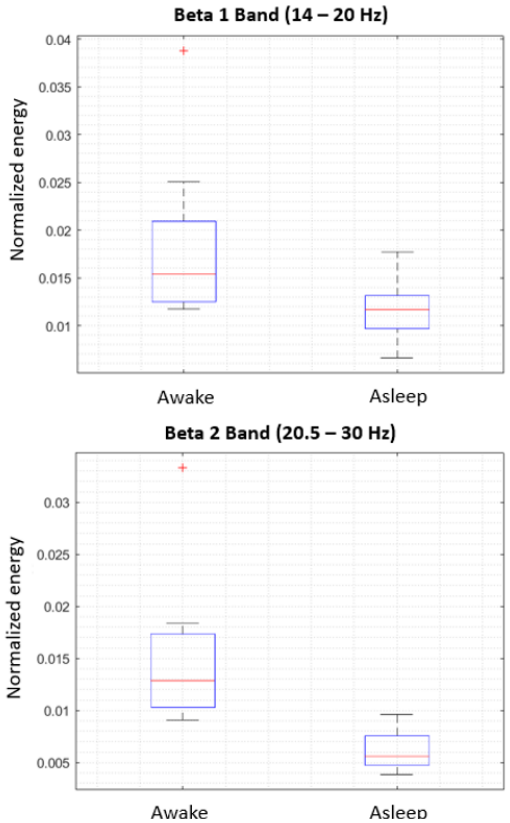


Figure 2: Boxplots of Beta 1 (above) and Beta 2 (below) bands during awake (qCON ≥ 60) and asleep (qCON < 60) stages. On each box, the central mark is the median, the edges of the box are the 25^{th} and 75^{th} percentiles. The outliers are marked with a +.

We studied the values of qCON and qNOX during the REM and NREM stages previously identified, and we performed a Wilcoxon signed rank test, at significance level of 0.05, for comparison of the indices.

Table 1: Column 1 shows the mean value ± standard deviation of the qCON for each volunteer when below 60 and column 2 shows the corresponding mean value of qNOX when qCON is below 60. The last column shows the difference between the mean values of qCON and qNOX. The last row indicates the mean and the standard deviation for all volunteers, where qNOX has significantly higher values as compared with qCON. (*p-value = 0.002)

ID	Mean value qCON<60	Mean value qNOX (when qCON<60)	qNOX-qCON
1	45.46 ± 6.11	62.36 ± 10.07	16.90 ± 11.78
2	50.02 ± 5.71	61.72 ± 7.34	11.69 ± 9.30
3	50.73 ± 5.87	61.93 ± 10.71	11.20 ± 12.22
4	45.09 ± 6.92	57.03 ± 12.64	11.94 ± 14.41
5	50.78 ± 4.83	65.51 ± 8.32	14.74 ± 9.62
6	50.77 ± 4.68	64.97 ± 7.20	14.19 ± 8.59
7	48.99 ± 4.71	58.28 ± 8.65	9.30 ± 9.85
8	52.06 ± 4.19	72.69 ± 6.91	20.63 ± 8.08
9	56.09 ± 2.52	72.62 ± 5.47	16.53 ± 6.02
10	55.45 ± 2.84	77.89 ± 5.83	22.44 ± 6.48
Mean all volunteers	50.54 ± 3.58*	65.50 ± 6.81*	14.96 ± 4.23

Table 2: Difference of time decay between qCON and qNOX to reach thresholds during the first loss of consciousness (N.A: not available due to qNOX always above the threshold).

	Time decay difference between qNOX and qCON for different thresholds in minutes			
ID	**qCON and qNOX ≤ 90**	**qCON and qNOX ≤ 80**	**qCON and qNOX ≤ 70**	**qCON and qNOX ≤ 60**
1	12.30	14.93	11.57	2.35
2	5.78	4.70	3.65	4.12
3	7.30	N.A.	7.35	2.85
4	6.20	9.93	7.80	N.A.
5	3.25	4.02	2.02	7.35
6	0.30	1.38	2.72	4.55
7	N.A.	9.51	5.03	0.80
8	N.A.	N.A.	10.65	N.A.
9	N.A.	5.60	4.83	N.A.
10	14.85	11.15	8.72	N.A.
Mean all volunteers	7.14 ± 5.01	7.65 ± 4.46	6.43 ± 3.29	3.67 ± 2.24

3 Results

In Figure 1, average indices values across volunteers are shown and it can be observed that nociception (qNOX index) stays higher than consciousness (qCON index) during sleep. The differences between the two indices were statistically significant (p-value < 0.05, Mann-Whitney U test). In Table 1 mean values of both indexes are compared when qCON is below 60 for every subject. Results confirm that qNOX was higher for all patients than qCON (p-value < 0.05, Wilcoxon paired test). Table 2 includes the differences of time decay between indexes, showing that qNOX takes more time to fall under all thresholds set than qCON, and in some case never decreases under those values.

The spectrum analysis of the EEG showed that the energy in the low frequency bands Delta, Theta and Alpha present no significant differences between awake and asleep stages (p-values > 0.05, Wilcoxon signed rank test). On the contrary, the energy from Beta 1 and Beta 2 did have significant differences and showed on average higher values while the volunteers were awake (Table 3). The difference in energy value of these bands for both stages can also be observed in Figure 2.

In Figure 3 we can observe the transition between NREM and REM intervals after applying the adaptive thresholds designed. It can be observed that these transitions occurred with certain periodicity during all recording.

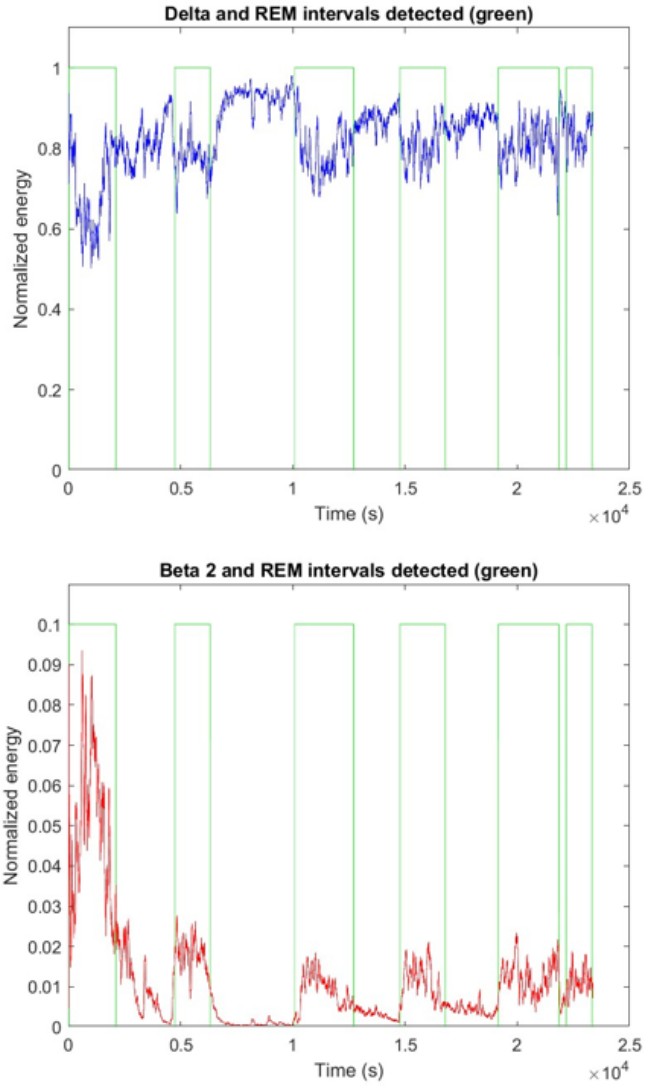


Figure 3: Delta (above) and Beta 2 (below) energy from one subject. The transition between REM and NREM, which corresponds with the moment where the value of Delta and Beta 2 pass their respective thresholds, is represented with the green line.

Table 3: Mean ± standard deviation values of the five selected EEG frequency bands for the periods where the subjects are awake and when they are asleep. P-value is obtained from a Wilcoxon signed rank test, at significance level of 0.05.

	Mean value Awake (qCon ≥ 60)	Mean value Asleep (qCon < 60)	p-value
Delta (0.5 - 4 Hz)	0.856 ± 0.137	0.866 ± 0.122	0.625
Theta (4.5 - 8 Hz)	0.068 ± 0.071	0.069 ± 0.064	0.920
Alfa (8.5 - 13.5 Hz)	0.042 ± 0.054	0.047 ± 0.065	0.695
Beta 1 (14 – 20 Hz)	0.019 ± 0.023	0.012 ± 0.016	0.019
Beta 2 (20.5 – 30 Hz)	0.015 ± 0.020	0.006 ± 0.008	0.002

Table 4: Mean values of qCON during REM and NREM stages for each volunteer and Wilcoxon signed rank test between both sleep stages revealing statistically significant differences ($p < 0.05$).

qCON			
ID	**Mean value NREM**	**Mean value REM**	**p-value**
1	48.91 ± 12.46	63.50 ± 10.23	
2	55.25 ± 10.39	68.96 ± 13.35	
3	57.19 ± 11.64	56.15 ± 15.73	
4	68.17 ± 21.49	60.45 ± 18.32	
5	53.84 ± 10.67	65.87 ± 7.47	
6	56.91 ± 9.01	67.26 ± 15.32	
7	51.49 ± 8.21	53.64 ± 12.95	
8	55.74 ± 9.38	74.84 ± 10.49	
9	72.23 ± 12.20	69.33 ± 8.64	
10	58.44 ± 6.95	68.08 ± 8.40	
Mean all volunteers	57.82 ± 11.24	64.81 ± 12.09	0.049

Table 5: Mean values of qNOX during REM and NREM stages for each volunteer and Wilcoxon signed rank test, no significant differences were obtained ($p > 0.05$).

qNOX			
ID	**Mean value NREM**	**Mean value REM**	**p-value**
1	65.94 ± 13.97	81.59 ± 13.49	
2	66.66 ± 10.77	77.21 ± 13.56	
3	75.08 ± 14.86	64.23 ± 14.50	
4	81.99 ± 19.66	71.14 ± 18.54	
5	70.77 ± 9.99	72.61 ± 11.74	
6	72.21 ± 11.26	77.65 ± 12.64	
7	64.13 ± 10.63	60.57 ± 12.77	
8	76.34 ± 8.30	79.67 ± 13.16	
9	88.38 ± 9.50	82.16 ± 10.70	
10	81.35 ± 6.72	84.28 ± 9.65	
Mean all volunteers	74.29 ± 11.57	75.11 ± 13.07	1.00

The results regarding the qCON and the qNOX values during the REM and NREM stages, previously delimited with the thresholds, showed significant differences in qCON values between stages (p-value < 0.05, Wilcoxon signed rank test), which was not observed in qNOX values, as shown in Table 4 and Table 5 respectively. In the boxplot of Figure 4 there is a representation of the differences in the qCON values for REM and NREM.

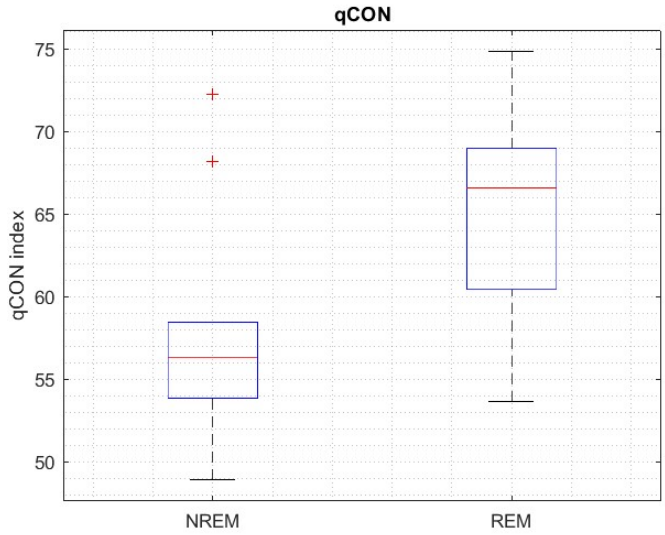


Figure 4: Boxplot containing the values of the qCON index during the sections of NREM and REM of night sleep. On each box, the central mark is the median, the edges of the box are the 25^{th} and 75^{th} percentiles. The outliers are marked with a +.

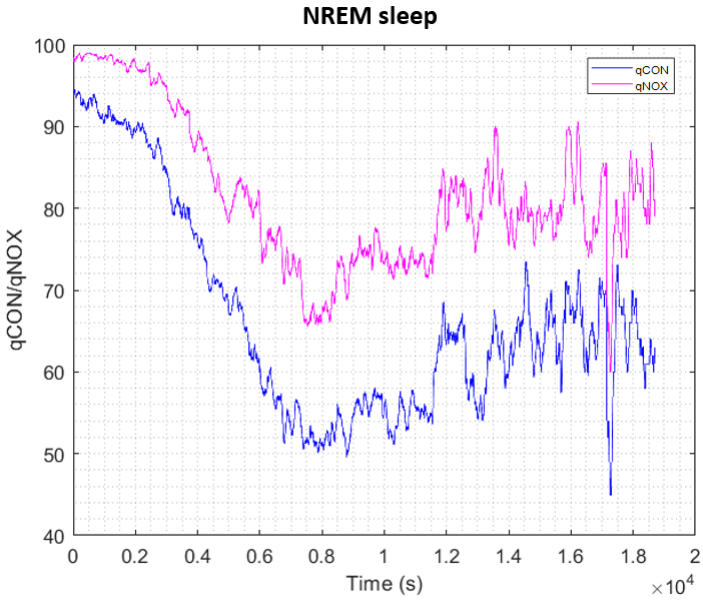


Figure 5: Mean qCON and qNOX for all volunteers during NREM interval. The indices have been aligned at the beginning of NREM interval of each subject.

The trends of the mean qCON and qNOX values for NREM and REM are shown in Figure 5 and Figure 6 respectively. The data of the volunteers has been aligned at the beginning of the REM interval that was around 5000 seconds for all cases. It can be noted from Figure 5 that in NREM sleep qCON and qNOX decrease until the minimum in the first part of the recording and then slightly increase

in the last part, showing more variability. During REM sleep (Figure 6) qCON and qNOX time evolution decreases in the first part of the recording to a stable value that is maintained until the last part, where the index values are lower.

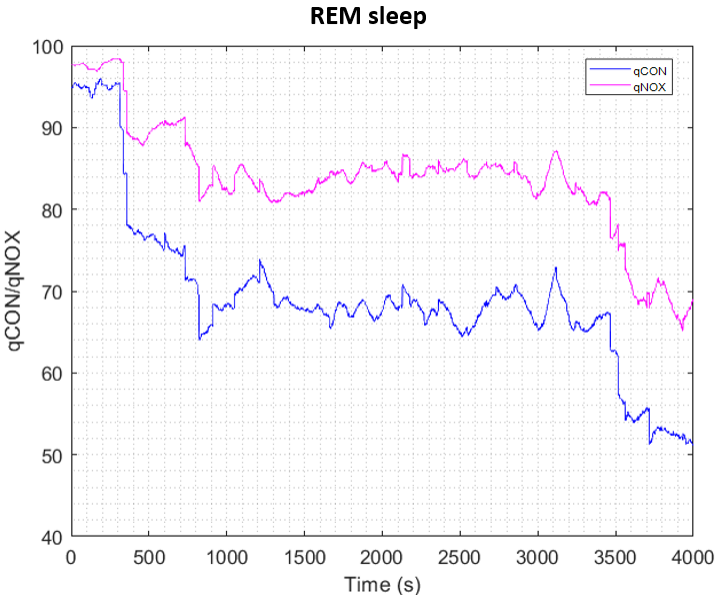


Figure 6: Mean qCON and qNOX for all volunteers during REM interval. The indices have been aligned at the beginning of REM interval of each subject.

4 Discussion

In this study it was found that qCON and qNOX indices behave differently during general anesthesia than compared to natural sleep. During anesthesia, anesthetic and analgesic agents are administrated to reduce both consciousness and nociception levels [6], whereas during night sleep, loss of consciousness is not accompanied by a similar nociception reduction.

Literature suggests that an increase in the energy of Beta is correlated with arousal during sleep [7], as higher values in this band are associated with an awake state. The increases in Beta during sleep found in this study could be related with sympathetic overactivity, which is associated with fragmented sleep [7].

Different NREM and REM sleep stages during natural sleep are characterized by changes in energy distribution in the frequency domain of EEG such as slow (Delta) and fast (Beta 2) oscillation patterns and also other physiological parameters. Indeed, in medical practice, polysomnography is considered the gold standard in sleep analysis because it takes into account several physiological signals such as electrooculogram (EOG) to identify rapid eye movements and electromyogram (EMG) to detect lack of muscular tone, which are characteristic physiological parameters of REM sleep [8]. Conox monitor performance is based on the processing of the EEG, hence it does not take into consideration as many parameters as a polysomnography. Despite that, from the results obtained in this study, qCON index has been able to correctly identify REM intervals

Other studies have been performed about depth of anesthesia monitors being used to monitor sleep since they provide real time EEG data that may be of some value for assessing sleep. A study performed with Sedline monitor (Masimo, Irvine CA) showed higher values during awake and the first parts of the sleep recording, as entering into deeper sleep stages the index of the monitor decreased considerably [5]. From the results obtained with the Conox it has been observed this same tendency, higher qCON values when awake and a decrease as the subject falls asleep. Nevertheless, qCON showed an increase during REM intervals, whereas the Sedline index continued at low values. Another sleep study performed with Bis Vista (Covidien, CO, USA) showed similar results than qCON during REM stage, with a mean (SD) BIS value of 75 (10) [9]. The mean (SD) value of qCON during REM interval was 64.81 (12.09), it is possible to differentiate this interval from an awake state where qCON values are higher.

In regard with nociception, qNOX index showed higher values during sleep than those observed in general anesthesia. Since qNOX index decreases with the administration of analgesic drugs, it could be expected to obtain higher values during sleep than in surgical procedures, where no analgesics are involved.

5 Conclusions

Analysis presented here allows to conclude that natural sleep is characterized by a reduction in consciousness which can fit into ranges similar to general anesthesia, especially during NREM sleep (qCON below 60), but that nociception (qNOX index) tends to remain higher, revealing that an asleep person can still respond to external painful stimuli. It has also been observed that when the subject is losing consciousness, the nociception index decreases slower than the consciousness index.

When studying the energy of the different EEG frequency bands, no significant differences were encountered in the low frequencies between the awake and asleep periods. Nevertheless, higher frequencies did show an increase in their energy when the subjects are awake.

By analyzing time series of energy contained in Delta and Beta 2 bands, one can identify cycles during night sleep. After studying qCON and qNOX indices evolution during these cycles, it can be concluded that qCON has greater values in REM intervals, whereas qNOX does not show significant differences between sleep stages. The significant differences of qCON values between NREM and REM cycles suggest that this index can also be used to differentiate these two stages of night sleep. The Conox monitor does not consider as many sleep parameters as the gold standard, polysomnography. Nevertheless, the present investigation is a preliminary study that might open the gate to use depth of anesthesia devices as Conox to monitor sleep, since it has proven to be able to detect the different sleep cycles by processing the energy of the EEG bands.

For future studies it would be interesting to consider a wider sample of patients, including some subjects with sleep disorders to observe differences in the results against healthy patients.

Conflict of Interest The authors of this work Joana Cañellas, Juan Felipe Ortega, Umberto Melia, Carmen González and Erik Weber Jensen are employees of Quantium Medical. Quantium Medical is the commercial developer of the indices qCON and qNOX.

References

[1] A. Espinoso, J. Cañellas, U. Melia, P. Gambús, C. González, E. Jensen, "Study of the evolution of the anaesthesia indices qCON and qNOX during natural sleep," in 2020 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO), 1–2, IEEE, 2020, doi:10.1109/ESGCO49734.2020.9158018.

[2] S. Chokroverty, et al., "Overview of sleep & sleep disorders," Indian J Med Res, **131**(2), 126–140, 2010, doi:10.17241/smr.2011.2.1.1.

[3] R. Ferri, F. I. Cosentino, M. Elia, S. A. Musumeci, R. Marinig, P. Bergonzi, "Relationship between Delta, Sigma, Beta, and Gamma EEG bands at REM sleep onset and REM sleep end," Clinical Neurophysiology, **112**(11), 2046–2052, 2001, doi:10.1016/S1388-2457(01)00656-3.

[4] C. McPherson, K. Behbehani, D. Dao, J. Burk, E. Lucas, "Characterization of sleep using bispectral analysis," in 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, volume 3, 2216–2219, IEEE, 2001, doi: 10.1109/EGCO49734.2020.9158020.

[5] C. Kapp, N. Punjabi, "Using a Sedation Monitor to Assess Sleep," in B65. SRN: DIAGNOSIS AND MONITORING OF SLEEP AND SLEEP DISORDERS, A3884–A3884, American Thoracic Society, 2019.

[6] E. Jensen, J. Valencia, A. López, T. Anglada, M. Agustí, Y. Ramos, R. Serra, M. Jospin, P. Pineda, P. Gambus, "Monitoring hypnotic effect and nociception with two EEG-derived indices, qCON and qNOX, during general anaesthesia," Acta Anaesthesiologica Scandinavica, **58**(8), 933–941, 2014.

[7] T. B. Kuo, C.-Y. Chen, Y.-C. Hsu, C. C. Yang, "EEG beta power and heart rate variability describe the association between cortical and autonomic arousals across sleep," Autonomic Neuroscience, **194**, 32–37, 2016, doi: 10.1016/S1388-2457(02)00656.

[8] "Polysomnography: Overview, Parameters Monitored, Staging of Sleep." Available: https://emedicine.medscape.com/article/1188764-overview#a2 (visited on 18-12-2020).

[9] M.-R. Benissa, S. Khirani, S. Hartley, A. Adala, A. Ramirez, M. Fernandez-Bolanos, M.-A. Quera-Salva, B. Fauroux, "Utility of the bispectral index for assessing natural physiological sleep stages in children and young adults," Journal of clinical monitoring and computing, **30**(6), 957–963, 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1335-1339 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Quality of Nursing Care in Hospitalized Patients of the Carlos Lanfranco La Hoz Hospital, 2019

Amancio Izquierdo-Príncipe[1], Jaqueline Garcia-Núñez[1], Brian Meneses-Claudio[2,*], Hernan Solis-Matta[1], Lourdes Matta-Zamudio[1]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

ARTICLE INFO

Article history:
Received: 03 August, 2020
Accepted: 08 December, 2020
Online: 16 December, 2020

Keywords:
Quality of care
Patient Satisfaction
Hospitalized Patients
Nursing Care
Hospital

ABSTRACT

***Background**: The quality of care of the nursing professional in relation to the hospitalized patient refers to the set of knowledge and attitudes that the nursing professional must provide quality care, preventing risks and improving patient satisfaction, for which its objective is to determine the quality of nursing care in hospitalized patients at the LanFranco la Hoz Hospital.* ***Method:** it is a quantitative, non-experimental, descriptive, and cross-sectional study with a study population of 143 patients in total hospitalized at the Carlos LanFranco la Hoz Hospital.* ***Results:** The results obtained in the quality of nursing care in hospitalized patients, it can be seeing that 86 patients representing 60.1%, are moderately satisfied, followed by 36 patients that represent 25.2% are satisfied, 12 patients that represent 8.4% are dissatisfied, 7 patients that represent 4.9% are very satisfied and finally 2 patients that represent 1.4% are very dissatisfied. In the dimension, experience of the nurse in the care of hospitalized patients, it can be observed 69 patients that represents 48.3% are moderately satisfied with the experience of the nurse in the care and in the dimension satisfaction with nursing care in hospitalized patients, it can be see 93 patients that represents 65% are satisfied with the quality of care.* ***Conclusions:** It is concluded that the high demand for patients sustains a decrease in quality nursing care due to the scarcity of human resources in the hospital and it is recommended that the hospital have a permanent evaluation system of quality of care to meet the expectations of hospitalized patients.*

1. Introduction

The World Health Organization (WHO) notes that 421 million people are hospitalized annually, these patients suffer approximately 42.7 million adverse events according to the most accurate evaluations and damage to patients during their hospital care are the cause of global mortality [1].

One of the factors of the quality of care is the perception that the hospitalized patient has during its hospital stay under care of hospital health personnel [2], since they fix their gaze on health personnel, easily judge and criticize the care provided, therefore, the position of the hospitalized patient will be evident when it comes to refusing care and advising that they should not be treated in that health facility [3].

Likewise, the care provided by the nurse towards patients has been perceived as deficient, because technology has influenced the nurse, where patient care is managed in a less humanized way, therefore the nurse-patient relationship loses every time trust [4].

*Corresponding Author: Brian Meneses-Claudio, Mr., +51 1 950159924 & bmeneses@uch.edu.pe

Likewise, the relationship of the family members of the hospitalized patient and the nurse who cares is essential for comprehensive health care, the adverse effects generated by negative environments for not providing good information timely and understandable for the family are negative factors for hospitalized patients [5].

The length of stay of hospitalized patients in the hospital is an impact on the accessibility of new patients entering the hospitalization service for the first time, generating demand both in the emergency area and hospitalization in the health sector [6].

In [7], the authors presented a study carried out in Jordan, it indicates that in the health area a topic of quality and safety is needed, this contributes to all personnel being trained responsibly, providing quality care and safety. The presence and permanence in health services are considered as a fundamental factor to promote the safety of hospitalized patients.

In [8], the authors mentioned a study carried out in Colombia, it was possible to find a total quality of care reached in 62%. The

dimension with the highest quality was " monitoring and follow-up" (73.2%) and the lowest was "anticipates" (56.1%). 42% of the patients said they had not been attended and did not recognize the nurse of the service.

In [9], the author described a study carried out in Russia, it refers that the perception, both of the health professional and the users, have a link in the human resource, which is ethical, the evaluations of hospitalized patients showed that are linked to the changeable, the concern of the hospital, and the obligation of public employees to provide quality care.

In [10], the authors presented a study carried out in Mexico, indicates that the quality of nursing professionals care towards patients highlighted 31patients that represents 44.0% reported that the nursing staff managed their work tools (baumanometer, stethoscope, secretion aspirator, among others) as they expected; 22 patients that represents 31.0%, better than they expected, and 32 patients that represents 30.0%, as they did not expect.

The objective of the study is to determine the quality of nursing care in hospitalized patients of the Carlos Lanfranco la Hoz hospital, in which it will allow to observe the quality of care that nursing professionals have towards hospitalized patients. This study is important since it will give relevant and real data about how the nursing care is provided to patients hospitalized for a disease.

In the study, the Spanish-version quality of care data collection instrument (CUCACE) was applied by Alonso R. and the original version The Newcastle Satisfaction with Nursing Scales (NSNS) by Thomas in Scotland has been demonstrated its usefulness to assess the quality of nursing care.

The data collection was processed through the survey of hospitalized patients of the Carlos LanFranco la Hoz Hospital, the data to be entered was carried out in a data matrix that will be designed in the statistical program IBM SPSS Statistics Base 25.0, it was proceeded to its corresponding analysis, in which it will allow a better data processing to make statistical tables and graphs so that can be described and interpreted in results and discussions, respectively.

The following research work is structured as follows: In section II, the development of the hospitalized patient data collection process will be presented, as well as the guidelines to consider so that they are within the research work. In section III, the results will show the quality of care presented by nursing professionals according to the specified dimensions of the instrument in the measurement of the variable. In section IV, it will present the discussions of the research work, in section V, the conclusions and in section VI the recommendations as well as the future work that is intended to be reached with the research work.

2. Methodology

In this part, the type and design of the research will be developed, as well as the population and the sample that will be carried out in the research work, the inclusion and exclusion criteria in detail and finally the technique and the instrument for collecting data.

2.1. Research type and design

The present study, due to its characteristics, the way of collecting data and measurement of the variables involved is a quantitative approach. Regarding the methodological design, it is a non-experimental, descriptive, cross-sectional study [11].

2.2. Population

The study worked with a total population of 143 hospitalized patients from the Carlos LanFranco la Hoz Hospital, which are distributed in 4 areas (Medicine, Surgery, Gynecology and Emergencies).

2.3. Inclusion criteria

- Adult patients with more than 2 days of hospitalization.
- Patients not receiving sedatives as treatment.
- Patients who have signed the informed consent.

2.4. Exclusion criteria

- Adult patients with less than 2 days of hospitalization.
- Patients with an altered state of consciousness, a language disorder, or a risk complication.
- Patient who has not signed the informed consent.

2.5. Technique and instrument

The data collection technique used in the research work was the survey, which important data will be obtained optimally and effectively. The survey will determine the quality of care in hospitalized patients using the (CUCACE) questionnaire, the Spanish version was validated by Alonso R. and the original version is The Newcastle Satisfaction with Nursing Scales (NSNS), validated by Thomas in Scotland. The CUCACE questionnaire has three sections: The first is the experience with nursing care that it has 26 items, the second is the satisfaction with the nursing care that it has 19 items and the third is the sociodemographic variable such as: age, sex, educational level, days of stay in the service and some general questions [2].

Regarding statistical validity, this was obtained using the Kaiser-Meyer-Olkin sample adequacy tests and Bartlett's sphericity test. The results of the sphericity test are significant while the sample adequacy test obtained a value of 0,897 (KMO> 0.6) and the results of the sphericity test obtained a favorable value (X^2 3702,347; g.l. = 990; p = 0,000). Likewise, Cronbach's alpha index obtained a value of 0.939 ($\alpha > 0,6$). All the above tests confirm an acceptable index of validity and reliability of the instrument.

This research work will consider the safety aspects of human research participants regarding to quality in hospitalized patients. The ethical behavior of deontological and moral health personnel. The bioethical principles: autonomy, justice, beneficence and non-maleficence, which constitute the basis for reasoning and orientation to their actions [12].

Medical ethics deals with the foundations that govern professional conduct in the practice of medicine, both in the doctor-patient and interprofessional relationships [13]. For this reason, the ethical authorization was carried out correctly when

coordinating with the hospital authority and with the patient itself and its family to carry out the research work.

Bioethical principles are interpreted into fundamental values and rights of hospitalized patients [14]. In the present research work, the bioethical principles of beneficence, non-maleficence, justice, and autonomy were mainly taken.

2.6. Place and Application of the Instrument

The questionnaire to measure the quality of nursing care in hospitalized patients of the Carlos Lanfranco la Hoz hospital, Lima - Peru was carried out on a programming date with the hospital, the time used in each service was approximately 25 minutes, in each service a document was requested for the hospitalized patient to participate in the research work (according to the criteria of inclusion and exclusion), concluding with good satisfaction when collecting the questionnaires from hospitalized patients since they provided a lot of support to carry out the research work.

It is important to emphasize the presence of relatives of some hospitalized patients at the time of completing the questionnaire, since while the family member has the decision in turn that the hospitalized patient can participate in the research, in which they provided support to complete the questionnaires. At the same time, the presence of the nursing staff, since the questionnaire is based on patient care, which allows to observe if the nursing staff adequately performs the appropriate and/or necessary care for the patient.

3. Results

Below is a summary table of the surveys carried out following the guidelines corresponding to the research work:

Table 1: Sociodemographic Data of Quality of Nursing Care in Hospitalized Patients of the Carlos Lanfranco La Hoz Hospital, 2019 (N = 143)

Participant Information	**Total**	
	N	**%**
Total	143	100
Gender		
Female	99	69,2
Male	44	30,8
Marital Status		
Single	37	25,9
Married	21	14,7
Separated	5	3,5
Cohabiting	79	55,2
Widow	1	0,7
Type of Family		
Nuclear	100	50,2
Extended	26	6,6
Expanded	7	11,1
Single parent	6	17,0
Reconstituted	3	6,2
Single Person	1	0,7
Study Degree		
No study	4	2,8
Incomplete primary	12	8,4
Complete primary	22	15,4
Incomplete secondary	33	23,1
Complete secondary	46	32,2
Incomplete University Studies	7	4,9
Complete University Studies	19	13,3
Occupation Condition		
Stable	14	9,8
Eventual	41	28,7
Without occupation	75	52,4
Retired	2	1,4
Student	5	3,5
Does not apply	6	4,2
Days of hospitalization		
Three days	51	35,5
Four days	54	37,8
Five or more days	38	26,6

In Table 1, it has the sociodemographic data of the study participants in total, there were 143 hospitalized patients from the Carlos Lanfranco La Hoz Hospital. Regarding gender, the female sex predominated with 99 patients representing 69.2%, followed by the male with 44 patients representing 30.8%.

Regarding the study degree, complete secondary school stands out with 46 patients representing 32.2%, followed by incomplete secondary school with 33 patients representing 23.1%, followed by complete primary school with 22 patients representing 15.4%, followed by complete university studies with 19 patients representing 13.3%, followed by incomplete primary school with 12 patients representing 8.4%, followed by incomplete university studies with 7 patients representing 4.9% and finally no study with 4 patients representing 2.8%.

Regarding the days of hospitalization, four days predominate with 54 patients representing 37.8%, followed by 51 days with 51 patients representing 35.5% and finally five or more days with 38 patients representing 26.6%.

Below are figures that show the results about the total quality of nursing care in hospitalized patients and its dimensions according to the data based on the research instrument:

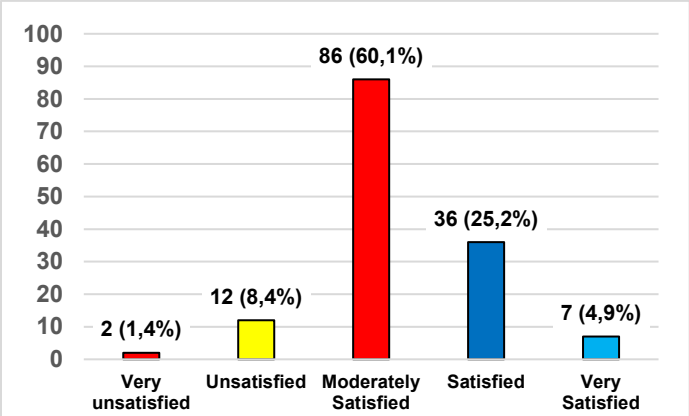


Figure 1: Quality of nursing care in hospitalized patients of the Carlos Lanfranco la Hoz Hospital, 2019 (N = 143)

In Figure 1, the quality of nursing care in hospitalized patients of the Carlos Lanfranco la Hoz Hospital, it can see that 86 patients representing 60.1% are moderately satisfied, followed by 36 patients representing 25.2% are satisfied, 12 patients representing 8.4% are unsatisfied, 7 patients representing 4.9% are very satisfied and finally 2 patients representing 1.4% are very unsatisfied.

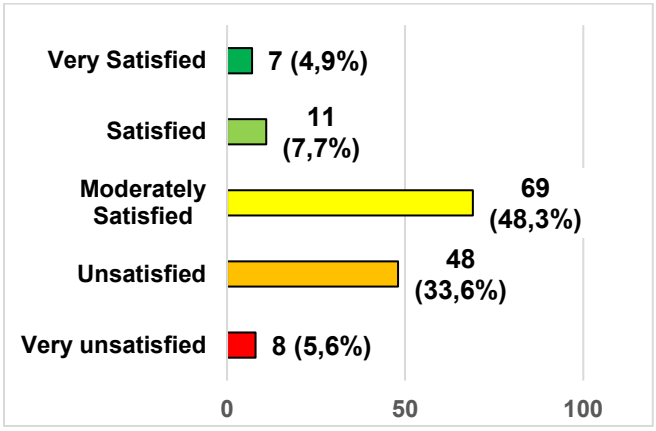


Figure 2: Quality of nursing care according to its dimension experience with nursing care, in inpatients in a public hospital in Lima Norte, 2019 (N = 143)

In Figure 2, the quality of nursing care according to its dimension experience with nursing care, in hospitalized patients of the Carlos Lanfranco La Hoz hospital, it can see that 69 patients representing 48.3% are moderately satisfied, followed by 48 patients representing 33.6% are unsatisfied, 11 patients representing 7.7% are satisfied, 8 patients representing 5.6% are very unsatisfied and finally 7 patients representing 4.9% are very satisfied.

In Figure 3, the quality of nursing care according to its dimension of satisfaction with nursing care in hospitalized patients of the Carlos Lanfranco La Hoz hospital, it can see that 93 patients representing 65% are satisfied, followed by 32 patients representing the 22.4% are moderately satisfied, 13 patients that represent 9.1% are very satisfied, 5 patients that represent 3.5% are unsatisfied and finally 0 patients that represent 0% are very unsatisfied.

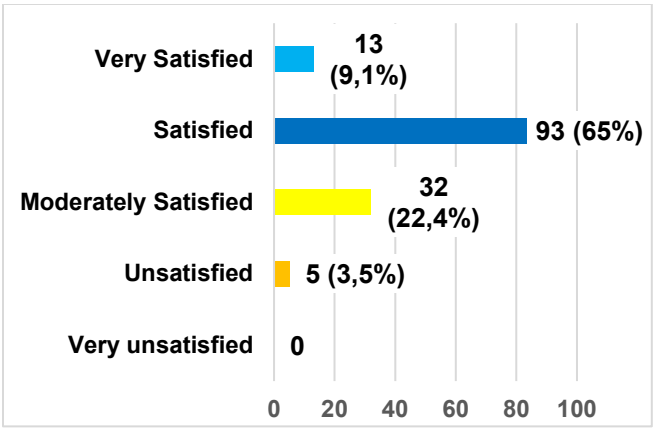


Figure 3. Quality of nursing care according to its dimension of satisfaction with nursing care, in hospitalized patients of the Carlos Lanfranco La Hoz hospital, 2019 (N = 143)

Knowledge of these results is important since it will allow to show if the nursing professional is adequately caring for hospitalized patients, in the same way, it will allow to show if the hospitalized patient in turn collaborates with the staff of nursing for its care during its hospital stay.

4. Discussion

The quality of nursing care constitutes a fundamental tool that values nursing care in hospitalized patients in all aspects and dimensions that are involved in the health of the person and the population.

We agree with what was indicated by Kol E. et al., who affirm that the fundamental objective of measuring the quality of health-related nursing care is to assess the care provided in health services [15].

Regarding the quality of nursing care in hospitalized patients of a public hospital, 60.1% predominantly are moderately satisfied, followed by 25.2% are satisfied, 8.4% are unsatisfied, 4.9% are very satisfied and finally 1.4% are very unsatisfied. These findings can be explained due to the perception that hospitalized patients have, during their stay in the hospital, to the care of health personnel since they fix their gaze and easily judge, criticize the care provided.

Our findings coincide with those of Weldetsadik A. et al., they determine that the satisfaction of hospitalized patients symbolizes the care provided, if it meets their expectations. This magnitude is not necessarily directly related to the degree of scientific-technical quality of the care provided [16].

The quality of nursing care is related to the dimension of experience with patient care mentioned by Amanda G. et al., they highlight the experience that hospitalized patients will have with the health professional, they are linked to the time that is provided during their hospital stay [17].

According to Becker R., he mentions that the wisdom of human resources is safe in medical practice, not only the care of patients is required and the attempt to provide quality care, the gift of expression is needed so that hospitalized patients can recover and reintegration into their family [18].

According to Allen A. et al., they point out the importance of technology as management support, but that it is not essential if there are trained or capable human resources. In many cases, a technological defect or the improper use of the equipment can lead to a failure in clinical analyzes that can lead to severe consequences for the life of hospitalized patients [19].

5. Conclusions

It is concluded that the use of a questionnaire CUCACE is important to evaluate the care received from hospitalized patients by the nursing professional, to evaluate the dimensions of experience and satisfaction with nursing care.

It is concluded that the high demand for patients sustains a decrease in quality nursing care due to the scarcity of human resources in the hospital.

It is concluded that the experience and work capacity of the nursing professional improves the quality and care of hospitalized patients.

It is concluded that it is essential to continue executing studies of this category, to analyze the outcome found and compare with current research, as well as facilitate its validity and reliability of CUCACE to measure the quality of nursing care in hospitalized patients.

6. Recommendations

It is recommended to carry out activities that promote communication with patients, since this interaction improves the nurse-patient relationship.

It is recommended that the hospital have a permanent quality of care evaluation system to meet the expectations of hospitalized patients.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Organización Mundial de la Salud, 10 Datos Sobre Seguridad Del Paciente., OMS, 2018.

[2] C. Lenis, F. Manrique, "Calidad del cuidado de enfermería percibida por pacientes hospitalizados.," Aquichan, **15**(3), 413–425, 2015, doi:10.5294/aqui.2015.15.3.9.

[3] P. Albujar, "Cuando el medico es el Paciente.," Fondo Editorial Comunicational, **32**(3), 1–164, 2015, doi:10.35663/amp.2015.323.103.

[4] P. Ramírez, C. Müggenburg, "Relaciones personales entre la enfermera y el paciente.," Enfermería Universitaria, **12**(3), 134–143, 2015, doi:10.1016/j.reu.2015.07.004.

[5] S. Geravandi, F. Soltani, M. Mohammadi, R. Alizadeh, A. Valipour, A. Hoseini, B. Rastegarimehr, A. Yari, A. Karimyan, A. Ghomeishi, "The effect of education on the nursing care quality of patients who are under mechanical ventilation in ICU ward.," Data in Brief, **16**, 822–827, 2018, doi:10.1016/j.dib.2017.11.090.

[6] C. Wang, M. Zhao, T. Zeng, "Progress in applying patient experience in nursing quality improvement.," Chinese Nursing Research, **4**(1), 1–4, 2017, doi:10.1016/j.cnre.2017.03.001.

[7] K. Al-Awamreh, M. Suliman, "Patients' satisfaction with the quality of nursing care in thalassemia units.," Applied Nursing Research, **47**, 46–51, 2019, doi:10.1016/j.apnr.2019.05.007.

[8] E. Amed, G. Villareal, C. Alvis, "Calidad del cuidado de Enfermería brindado por profesionales en una ciudad colombiana. Revista Ciencia y Cuidado[revista en Internet] 2019 [acceso 01 de julio de 2019]; 16(2): 108-119.," Revista Ciencia y Cuidado, **16**(2), 108–119, 2019, doi:10.22463/17949831.1615.

[9] V. Sarkisova, "Quality of nursing care, availability and accessibility of treatment and prevention of communicable and chronic diseases.," International Journal of Nursing Sciences, **4**(4), 350–351, 2017, doi:10.1016/j.ijnss.2017.10.016.

[10] M. Marín, C. Álvarez, J. Valenzuela, "Percepción de la atención de enfermería y satisfacción laboral en un Hospital de Seguridad Social.," RICSH Revista Iberoamericana de Las Ciencias Sociales y Humanísticas, **6**(11), 287, 2017, doi:10.23913/ricsh.v6i11.119.

[11] C. Fernández, P. Baptista, Metodología de la Investigación. 6ta ed. México: Mc Graw-Hill/Interamericana., 2015.

[12] O. Vera, "Aspectos bioéticos en la atención de los pacientes de las unidades de cuidados intensivos.," Revista Medica La Paz, **21**(1), 61–71, 2015.

[13] J. Martínez, C. Laucirica, E. Llanes, "La ética, la bioética y la investigación científica en salud, complementos de un único proceso.," Revista Médica Electrónica, **37**(4), 310–312, 2018.

[14] L. Mora, "Los principios éticos y bioéticos aplicados a la calidad de la atención en enfermería.," Revista Cubana de Oftalmología, **28**(2), 228–233, 2015.

[15] E. Kol, F. Arıkan, E. İlaslan, M. Akınci, M. Koçak, "A quality indicator for the evaluation of nursing care: determination of patient satisfaction and related factors at a university hospital in the Mediterranean Region in Turkey.," Collegian, **25**(1), 51–56, 2018, doi:10.1016/j.colegn.2017.03.006.

[16] A. Weldetsadik, T. Gishu, T. Mekonnen, A. Mekonnen, L. Girma, T. Demas, "Quality of nursing care and nurses' working environment in Ethiopia: Nurses' and physicians' perception.," International Journal of Africa Nursing Sciences, **10**, 131–135, 2019, doi:10.1016/j.ijans.2019.03.002.

[17] M. Trincado, E. Fernández, "Calidad en enfermería.," Rev Cubana Enfermer, **11**(1), 1–5, 1995.

[18] R. Becker, "Recordando a Sir William Osler a 100 años de su fallecimiento: ¿qué podemos aprender de su legado?.," An Fac Med, **76**(1), 71–6, 2015, doi:10.15381/anales.v76i1.11078.

[19] A. Allen, J. Robinson, M. Stewart, "Healthcare Quality: A Concept Analysis.," Nursing Forum, **52**(4), 377–386, 2017, doi:10.1111/nuf.12207.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1343-1359 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Event Modeller Data Analytic for Harmonic Failures

Futra Zamsyah Md Fadzil*, Alireza Mousavi, Morad Danishvar

System Engineering Research Group, CEDPS, Brunel University London, UB8 3PH, UK

ARTICLE INFO

Article history:
Received: 20 November, 2020
Accepted: 09 February, 2021
Online: 28 February, 2021

Keywords:
Event Modeller
Event Tracker
Event Clustering
Key Performance Indicator
Industrial Internet of Things
Data Analytics
Power Quality Disturbance

ABSTRACT

The optimum performance of power plants has major technical and economic benefits. A case study in one of the Malaysian power plants reveals an escalating harmonic failure trend in their Continuous Ship Unloader (CSU) machines. This has led to a harmonic filter failure causing performance loss leading to costly interventions and safety concerns. Analysis of the harmonic parameter using Power Quality Assessment indicates that the power quality is stable as per IEEE standards; however, repetitive harmonic failures are still occurring in practice. This motivates the authors to explore whether other unforeseen events could cause harmonic failure. Usually, post-failure plant engineers try to backtrack and diagnose the cause of power disturbance, which in turn causes delay and disruption to power generation. This is a costly and time-consuming practice. A novel event-based predictive modelling technique, namely, Event Modeller Data Analytic (EMDA), designed to inclusive the harmonic data in line with other technical data such as environment and machine operation in the cheap computational effort is proposed. The real-time Event Tracker and Event Clustering extended by the proposed EMDA widens the sensitivity analysis spectrum by adding new information from harmonic machines' performance. The added information includes machine availability, utilization, technical data, machine state, and ambient data. The combined signals provide a wider spectrum for revealing the status of the machine in real-time. To address this, a software-In-the-Loop application was developed using the National Instrument LabVIEW. The application was tested using two different data; simulation data and industrial data. The simulation study results reveal that the proposed EMDA technique is robust and could withstand the rapid changing of real-time data when events are detected and linked to the harmonic inducing faults. A hardware-in-the-Loop test was implemented at the plant to test and validate the sensitivity analysis results. The results reveal that in a single second, a total of 2,304 input-output relationships were captured. Through the sensitivity analysis, the fault causing parameters were reduced to 10 input-output relationships (dimensionality reduction). Two new failure causing event/parameter were detected, humidity and feeder current. As two predictable and controllable parameters, humidity and feeder velocity can be regulated to reduce the probability of harmonic fluctuation.

1 Introduction

This paper is an extension of work originally presented in the conference Proceedings of the 2019 IEEE/SICE International Symposium on System Integration (SII) [1]. Power Quality (PQ) monitoring has been the focus of research and development for many years. Recently, [2], [3] addressing the PQ issues in various industry and proposing different methods in tackling this issues.The main focus is to protect the equipment and minimize the losses while increasing the levels of operational safety. However, systems are becoming more complex as they are being further developed for improvement. Huge numbers of control drive and other non-linear loads have been installed to satisfy the demands of modern lifestyles. It has led to instability of the power system, creating high noise levels in the system grids and decaying the electrical distribution system. Even worse, climate change has an impact on the environment that the system operates in. In some cases, particularly for hot countries, the electrical distribution system requires air conditioning to protect the electrical equipment in the substation. In normal circumstances, if the air conditioning fails, it may surpass the permitted set point and may not protect the equipment. When the global temperature rises, the potential rate of failure also increases. Research has shown that the long-term average global temperature is increasing, which has a significant negative impact on the performance of the machines [4].

*Corresponding Author: Futra Fadzil, Brunel University London, Email: Futra.Fadzil@brunel.ac.uk

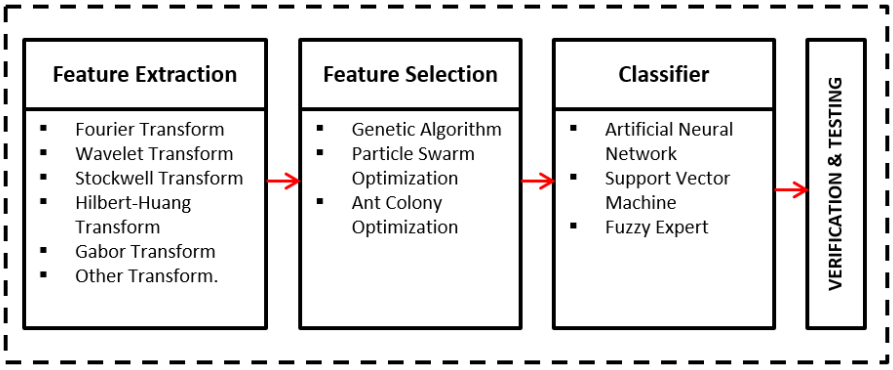


Figure 1: Various techniques of Power Quality Classification

The arrangement of the three principal stages in PQ Classification: (1) Feature extraction - a key transformation of the fundamental signals into time-frequency based information. A broad selection of Fourier Transform, Wavelet Transform, Stockwell Transform, Hilbert-Huang Transform, Gabor Transform and others. (2) Feature Selection - a key method in selecting the most suitable features from the feature extraction stage while discarding the redundant features. Few main optimizers include Genetic Algorithm, Particle Swarm Optimization and Ant Colony Optimization. (3) Classifier – a selection of artificial intelligence tools in classifying PQ disturbances such as artificial neural network, support vector machines, and fuzzy expert.

Earlier, conventional methods use manual configurations and visual inspection to monitor the quality of the power supply in the system. This method was too difficult to interpret and time-consuming. Later, an automated classification system, which uses the signal processing technique was developed and currently advancing with various Machine Learning and Deep Learning approach. Various combinations of features extraction and classifiers have proven to locate PQ disturbances accurately while using synthetic data. However, this is not the case with real industry data. Real industry data are complex and much more complicated when there are external or environmental factors that could potentially influence the state of the system. This creates a gap in existing techniques, which motivates authors to investigate the relevance of the environmental parameter to the issue of PQ disturbances. In addition, with the current emergence of big industrial data, system engineer must ensure that the system is robust and capable of performing ML/DL techniques in real-time applications. Disturbance data exists in the order of microseconds, which greatly enhances the record data [5].In consequence, it will burden learning classifiers to run unimportant parameters that result in high computation. The authors, therefore, suggest a method that could minimize the dimensional by selecting the most relevant data to be trained automatically, while incorporating a new unknown parameter that was previously considered irrelevant in the system state.

The aim of this paper is to introduce event-based analysis as a technique for detecting harmonic failure in real time. The technique is capable of grouping together high correlation system parameters, forming an input-output relationship that is not limited to internal parameters only. As such, the homogenized correlation system will indicate the possible root cause of harmonic failures, while eliminating non-important parameters in real-time [6]. This paper presents a real-time simulation of CSU machines, which shows the changes in the output parameters of the input activity. The data will be used to measure the suitability and applicability of event modelling techniques in the power system environment. It is then extended to the actual plant data, exploring the integration of this dynamic platform with the Key Performance Indicator (KPI) in order to recognize the system operating pattern. This data analytics technique expects to predict the homogenized system parameter that could present the current state of the system in real time.

1.1 Power Quality Disturbance

A number of PQ disturbance types are implicated: voltage sag, voltage swell, transients, harmonics, fluctuations, flickering, voltage imbalance, interruption, DC offset and notches. A high level of engineering expertise [7] is needed to effectively diagnose these PQ disturbances. Preventing PQ disturbances is critical to minimize power interruptions between the power utility and the end user. With modern technology, a substantial amount of research has been devoted to alleviating these problems using signal processing technique. This technique has three principal stages which incorporate feature extraction, feature selection and classification.

Feature extraction is a key transformation of the fundamental signals into time-frequency based information. It can be extracted directly from the original measurement either from a transformed domain or from the signal model parameters [8]. Several approaches include short-time Fourier transform [9], wavelet transform [10], wavelet packet transform [11], Hilbert Huang transform [12], Stockwell transform [13], Gabor-Wigner transform [14], and other hybrid transform-based [15].

Feature selection is a key method in selecting the most suitable features from the feature extraction stage while discarding the redundant features. Based on the extraction data, optimization techniques such as Ant Colony Optimisation [16], Genetic Algorithm [17], and Particle Swarm Optimisation [18] are the common algorithm used in locating PQ disturbances.

Classification is the process of predicting the class of data points. It uses an algorithm to implement classification by approximating the mapping function from the input variable to the output variable. One of the most popular and influential types of machine learning algorithms is the Artificial Neural Network. Other machine learning includes Support Vector Machine [19] and fuzzy expert system [20]. Figure 1 illustrates the three main phases of the PQ classification.

[21]–[22] provide a detailed review of these techniques.

1.2 Power Quality Standards

Harmonic assessments of utility systems require procedures in place to ensure that the efficiency of the voltage supplied to all consumers is adequate [23]. However, most of the harmonic issues arise at the end of the consumer. Their devices contain non-linear loads resulting in resonance conditions [23]. These non-linear loads are the current harmonic sources. The system voltage appears stiff to individual loads, and the loads draw distorted current waveforms. It is therefore important to maintain the PQ International standard, to provide the guidelines and limits for the acceptable levels of compatibility, between consumer equipment and the system utilities.

The International Standard IEEE 519-1992 sets the limit for both harmonic voltages and currents at the Point of Common Coupling (PCC) between the end-user and the utility supplier. It limits Voltage Total Harmonic Distortion (THD), defined as the ratio of the RMS value of the harmonic voltage to the RMS value of the fundamental (50Hz) voltage, to a maximum of 5%. Individual voltage harmonic magnitudes are limited to 3% of the fundamental voltage value [24]. The International Standard IEC 61000-3-4 sets the limits for the emission of harmonic currents in low-voltage power supply systems for the requirement with rated current greater than 16A [25].

1.3 Environmental Parameters

Compliance with the International PQ Standards is the policy for both utilities and customers. In order to maintain a good PQ system, an enormous amount of research and PQ analysis has been undertaken to ensure it complies with the standards. However, in some cases, the machine still fails, although it complies with the standard. It is important to note that most of these research focuses on the fundamental of the electrical system and leave behind the external parameters that may have a substantial effect on the problem. In [26], the authors suggest that there could be further environmental and operational factors that could also affect the performance of power systems.

1.4 Problem Statement

In this article, we intend to investigate what possible factors that could link a well-maintained machines with a proper operation to a frequent harmonic failure incidents. While PQ parameters are eliminated by its healthy measurement, an alternative technique to solve this harmonic failure mystery is required to evaluate if there is an existence of a different parameter that could lead to this failure. Therefore, the authors are proposing the Event Modeller techniques which could link the internal and external parameters together, to create a cause-effect relationship, while updating the system status in near real-time before the machine fails. The implementation of the event modeller technique recently has shown positive practical findings in various area and application such as high-speed causal prediction modelling [27, 28], real-time Remaining Useful Life (RUL) estimation [29], and as a middleware to highly coupled input variable in zero-defect manufacturing [30], and predicting N_2O Emission [31].

1.5 Contribution & Advantages

The first contribution of this paper is to proposed a robust system engineering tools that be able to formulate a piece of new knowledge information in the diagnosis of PQ disturbances in real-time. This tool is found robust, workable with the dynamic and autonomous environment while being able to withstand real industry data which is non-linear and highly influenced with various factors. The second contribution of this paper is that the proposed technique could visualize the group of highly correlated variables in the form of occurrence matrices in real-time, proposing a mathematics equation that represents the current state of the system. This is optimum to system operators who can make a quick decision before the system fails.

1.6 Challenges

However, implementing this technique in solving PQ problems has its challenges. The harmonic readings (THD) is taken from the Multi-Function meter, which is installed at the electrical switchgear. The reading will not be as accurate as of the conventional method, but the measurement is sufficient to prove that the PQ parameters can be eliminated. Another challenge is to determine the threshold setting. In general, a scenario of events in the actual system varies according to the system characteristic and its system deterministic. An event could happen instantaneously, or it may have some delay before it reached to the target output. For example, the scenario of a room temperature does not change immediately when the heater or air conditioner was switched on. In comparison, a dark room will be immediately bright when the switch was turned on. Both of the scenarios have made changes to the system state at a different pace. In order to detect this, a data benchmarking is required which were decided based on historical data and system expert point of view.

2 Event Modeller Technique

The event modeller technique is designed to investigate the association between the observable events (Output Data) with the causal events (Input Data) through a data mapping concept [30]. It clusters the system parameters with the highest correlation in the form of matrices and places them into mutually exclusive blocks. This creates an input-output relationship, which considers both internal and external factors. One significant difference between the proposed event modeller and other traditional data modelling techniques is how the input data assumes its output data. The traditional method assumes the input-output relationship as a true representation of a known data series, while the event modeller technique makes no assumption about it (unbiased) [32]. This fundamentally makes the approach non-unbiased vis-à-vis input-output relationship. The proposed Event Modeller Data Analytic (EMDA) is designed to find any interceptions between the known influential parameters (reported in previous literature) and new operational and environmental data by expanding the spectrum and reassessing the relationship between influential parameters on harmonic behaviour. It is a proven, computationally effective method of analyzing harmonic behaviour without bias. Borrowing from the Event Tracker [33] and Event Clustering [34] method, the proposed EMDA extends

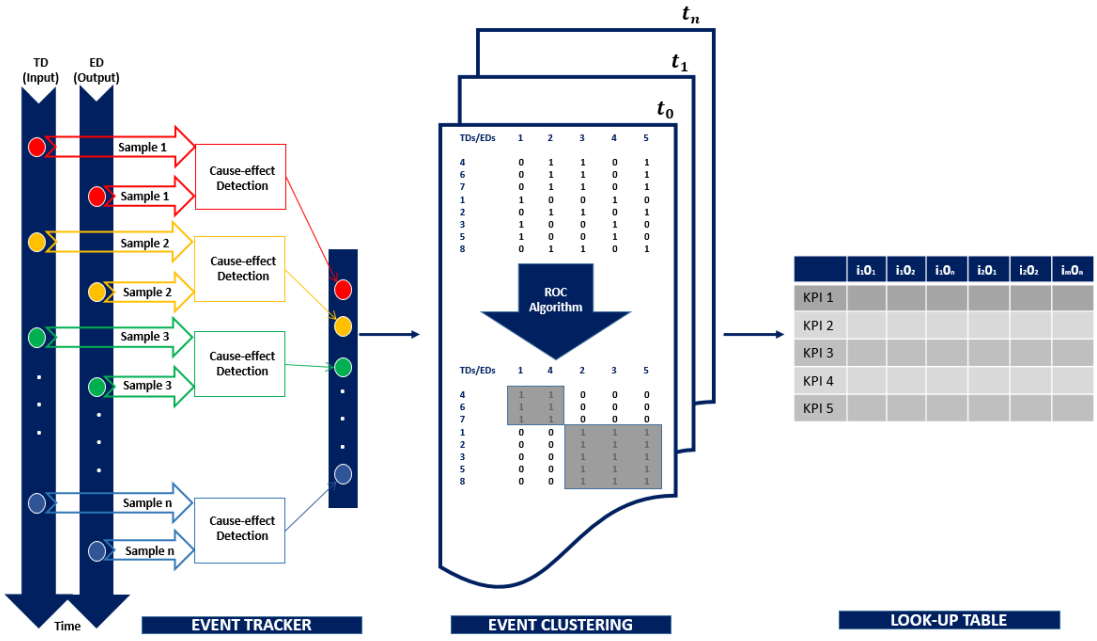


Figure 2: Event Modeller Techniques

The arrangement of the three principal stages in Event Modeller: (1) Event Tracker – A collection of input-output cause-effect detection of a system state with highly sensitivity index in real-time. (2) Event Clustering – A rank order grouping technique based on the interpretation of the changes of input-output data in the previous stage. (3) Look-up Table - A population table that links the Key Performance Indicator with the Sensitivity Indexes.

the previous authors' correlation map by adding the information from the machine performance metrics. Information such as machine availability, utilization and performance are combined with other technical data, to group the highly correlated individual data together, revealing the machine's status in real-time. Figure 2 illustrates the event modeller techniques, including Event Tracker, Event Clustering and a Look-up table. Details of each stage will be briefly discussed in the following subsection.

2.1 *EventTracker Algorithm*

EventTracker algorithm was introduced [35] to construct a discrete event framework for online sensitivity analysis (SA). This algorithm was designed to be applied to highly volatile environments with real-time scalar data acquisition capabilities (e.g. Internet of Things (IoT). It uses a cause-effect detection algorithm to dynamically track events' triggers and their interrelationship with one another in a given system. It generates a sensitivity index that measures the relationship between the triggered data and the event data pairs. High sensitivity index score will be indicated while the lesser impact relationship will be eliminated. This will reduce the computational effort while achieving dimensional reduction. The Event Tracker is a computationally efficient technique that focuses on the state changes of the involved system components. It merely takes a snapshot of the system states, which helps engineers observe system performance [6]. Further details on the description and four EventTracker functional parameters (Search Slot, Analysis Span, Event Threshold and Trigger Threshold) can found in [35].

2.2 *Event Clustering Algorithm*

Event Clustering Algorithm (EventiC) [34] is designed to improve the real-time sensitivity analysis, by automatically re-arranging the input-output relationship in rank order of its importance and relevance. The algorithm applied the Rank Order Clustering (ROC) technique, which was initially introduced in [36]. This technique interprets the changes in input-output data's value at the given level, detecting the coincidence and finally groups it as a related event. The process of calculating the number of coincidence occurs at a specified scan rate time interval, to ensure a relationship weight is established for modelling and control purposes [32]. Despite filtering the unimportant relationship between the input-output relationship, event clustering can identify new influencing parameters that were previously thought irrelevant, making it unique and interesting to improve the data quality. One key advantage of EventiC compared to EventTracker is that EventiC can assess multiple input/output relationship in every single scan, whilst EventTracker considers multiple input single output relations. A detailed explanation of EventiC algorithm and the basic concept is discussed in [34]. Algorithm 1 shows the main procedure of Event Modeller.

2.3 *Look-up Table*

A look-up table is an array of data to map input values to output values. It is used to transform the input data into a more desirable output format. The look-up table allows replacing run-time computation or logic circuitry with a simpler array indexing operation [37]. Retrieving a value from a look-up table is faster rather than examining it from the whole database. With the advantage of the

logical separation of data, it makes it relevant to prepare the data for machine learning purposes. This EMDA technique's novelty can be found in this look-up table, thus a rapid response to changes in the system's stability (i.e. the KPIs).

Algorithm 1: Event Modeller

```
Set Event Modeller Limit (EML);
Set Threshold Setting (Th);
Populate ULTh and LLTh;
Populate All Input Data (TD1, TD2...TDn);
if Triggered Data = Dynamic TD then
    Compute TDx = TDn - TDn-1;
    Compare TDx with ULTh and LLTh;
    if (LLTh<TDx<ULTh) then
        TDx = 0;
    else
        TDx = 1;
    end
else
    Triggered Data = Static TD;
    TDx = TD1,TD2...TDn.;
end
repeat
    repeat
        Populate All Output Data (ED1,ED2...EDn);
        Compute EDx = EDn - EDn-1;
        Compare EDx with ULTh and LLTh;
        if LLTh<EDx<ULTh then
            Set EDx = 0;
        else
            Set EDx = 1;
        end
        Populate the models input-output event coincidence
          matrix with binary weighting values of exclusive
          NOR function;
        Average each input-output event coincidence;
        Sort rows of the resultant binary matrix into
          decreasing order of their decimal weights;
    until for every column;
until position of each element in each row and column does
  not change;
Calculate the weight for each row i and column j (in a m by
  n matrix) using Equation 1 and Equation 2.
```

$$Row : W_i = \sum_{j=1}^{m} a_{ij} 2^{m-j} \tag{1}$$

$$Column : W_j = \sum_{i=1}^{m} a_{ij} 2^{n-i}. \tag{2}$$

2.4 Key Performance Indicator

The use of existing shop-floor in modern manufacturing to measure and monitor industrial KPIs has been a trend. In [38], the author used Object Linking and Embedding (OLE) which integrates with Discrete Event Simulation (DES) modelling capabilities to measure the KPIs for brewery industry. In [39] the author used data-driven scheme of KPIs prediction and diagnosis for hot strip mill industry. In [40], the author proposes an analytics solution for calculating statistical KPIs in the Human Machine Interface (HMI) layer. There is a need to translate a suitable KPI that suit the type of operation. Basic KPIs are calculated directly from the output operation data, and they serve as the foundation for Overall performance KPIs [41].

Time-based KPIs are data related to time duration, defining activities associated with production and maintenance. The calculation of KPI has to consider all these factors to reflect accurate metrics. Hence it is important to understand the KPI interrelationship. A hierarchical structure for KPI categorization is proposed by [41]. The time-based KPIs studied in this paper are Availability, Instantaneous Utilization, Schedule Utilization and Performance.

To complement the Time-Based KPIs, it is useful to look at the energy consumption and emission contribution that could potentially harm the environment. In general, energy consumption can be calculated by multiplying the motor rating (kW) with the duration it operates. There are three states of motor known as run, idle and stop. During running state, the motor is capable of operating at its 100% loading while in idle state, the motor operates at 25% of its loading. Obviously, there is no loading when it stops. Having the total energy consumption, it is easier to calculate the carbon footprint; Carbon Dioxide Emissions ($kgCO_2$, $kgCO_2e$), Methane ($kgCH_4$) and Nitrous Oxide ($kgNO_2$) using the formula in [42].

3 A Case Study

A Continuous Ship Unloader (CSU) is one of the leading bulk material handling machine. In a coal-fired power plant, this machine is used to transport coal from the vessel to the pulverized boiler through a series of belt conveyors. It has been reported that the CSU machine in one of the power plants in Malaysia has a frequent harmonic failure. The repetitive incidents lead to catastrophic failure, which harms the electrical devices. Besides having a vast replacement cost for the faulty parts, it is also affecting the plant availability, which concerned the management team. Even worse, it could pose a potential hazard, to the personnel working in the area if it is happening again in the future. A thorough Power Quality (PQ) assessment within the electrical distribution system has been assessed, but the results have not given any indication of internal disturbance or fault. Table 1 shows the assessment result for the CSU machine. To mitigate the problem, an effort to analyze both internal and environment parameter that has a significant impact on the system is highly desirable. The authors are keen to embrace the event modeller techniques, to evaluate the significant correlation between the output and input, which could cause harmonic failures in the system. A real-time simulation which incorporates the CSU machine parameters and the event modeller algorithm was developed using National Instruments LabVIEW.

3.1 Experiment 1: CSU Real-time Data Simulation

The purpose of a real-time simulation is to measure the suitability and applicability of the event modeller technique in the power system environment. To ensure the system incorporates the industrial

Table 1: Assessment Results for CSU Machine

Parameters	Units	Min	Max	Avg
Voltage	V_{ab}	385.9 V	424.0 V	411.2 V
	V_{bc}	216.8 V	423.8 V	382.4 V
	V_{ca}	241.8 V	424.1 V	381.5 V
Current	I_a	32 A	1718 A	525.8 A
	I_b	46 A	1870 A	639.9 A
	I_c	30 A	1568 A	217.6 A
Voltage Unbalance	% Unbalance	0.10 %	41.60 %	2.60 %
Current Unbalance	% Unbalance	8.80 %	16.70 %	52.80 %
Total Harmonic Distortion Voltage (THDV)	$THDv_{ab}$	0.40 %	1.60 %	1.00 %
	$THDv_{bc}$	0.30 %	1.50 %	1.00 %
	$THDv_{ac}$	0.50 %	1.90 %	1.10 %
Total Harmonic Distortion Current (THDi)	$THDi_a$	1.90 %	20.80 %	4.50 %
	$THDi_b$	1.80 %	25.80 %	3.70 %
	$THDi_c$	1.90 %	82.20 %	12.00 %
Total Demand Distortion Current (TDDi)	$TDDi_a$	0.10 %	2.60 %	1.30 %
	$TDDi_b$	0.10 %	1.80 %	1.20 %
	$TDDi_c$	0.10 %	2.20 %	0.80 %
Frequency	f	49.77 Hz	50.13 Hz	49.98 Hz

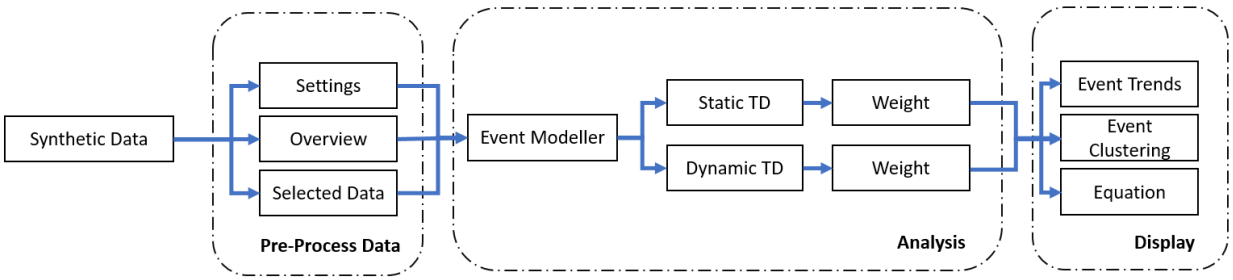


Figure 3: Experiment Strategy for Event Modeller Simulation

The arrangement of the three principal stages in the Experiment Strategy for Event Modeller Simulation: Pre-process Data, Analysis and Display.

case study, the requirement specification has been set based on the following:

1. 8 Event Data (ED's). This synthetic event data is simulated using Normal Distribution which represents the CSU Machine Output Data which includes Voltage, Humidity, Harmonics, Slewing Movement, Luffing Movement, Travel Movement, Temperature and Wind Speed;
2. 8 Triggered Data (TD's). This synthetic TD's is simulated using Normal Distribution which represents the CSU Machine Input data which includes Busy Slew, Busy Luff, Busy Travel, Busy Bucket Elevator, Slewing Motor Run Bit, Hydraulic Motor Run Bit, Travel Motor Run Bit and Bucket Motor Run Bit;
3. Threshold Setting. For the purpose of examining the Trigger Threshold, an arbitrary 5% threshold setting was set, that could later be adjusted;
4. Event Modeller Limit. This setting reduces the complexity of an input-output relationship by its correlation confidence level. An arbitrary 80% was set, that could later be adjusted;

3.1.1 *Experiment Strategy*

Sixteen input/output system parameters of a CSU Machine were sampled from the Allen Bradley PLC-SCADA system using the FactoryTalk view. Once it is proven to be suitable, it will be extended to 120 parameters, which will be discussed further in Section 3.2. Data collection was conducted over a single day shift of a machine operation period, collecting approximately 43,200 lines of 12-hour data samples.

The sampling frequency follows [43] by applying the right bandwidth, time constant and settling time. Having this data in hand, synthetic data was constructed to have the same nature and properties as the real CSU machine. The overview of the experimental methodology is shown in Figure 3. Details of the dataset arrangement will be discussed in the following section.

Table 2: CSU Real-time Data Simulation Results (Weight) Based on Disturbance

Description	Voltage		Humidity		Harmonic		Temperature		Wind Speed	
	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic
Overall Maximum	0.8387	0.9996	0.8364	0.9996	0.8370	0.9997	0.8360	1.0000	0.8340	1.0000
Overall Minimum	0.5000	0.6288	0.5000	0.6371	0.5000	0.6360	0.5000	0.6312	0.5000	0.6325
k-Disturbance (Pre)	0.8345	0.9890	0.8287	0.9790	0.8181	0.9791	0.8306	1.0000	0.8424	1.0000
k-Disturbance (Dur)	0.7615	0.8065	0.7781	0.8304	0.7637	0.8137	0.7977	0.8976	0.7848	0.8103
k-Disturbance (Post)	0.7813	0.8667	0.8011	0.8789	0.7897	0.8650	0.8087	0.9310	0.8035	0.8654

3.1.2 Dataset Arrangement

ED is defined as a series of data that represent the state of the system at a given time [35]. In this simulation, the ED consists of simulated voltage, humidity, harmonics, machine positioning (slewing, luffing and travelling), temperature and wind speed. The simulation data is based on the actual operation and environment data of the CSU machine events by considering the assessment result in Table 1 and the location of the power plant located in Malaysia. The plant is located near to the seaside, which is hot and humid throughout the year and may tend to have strong winds. On the other hand, the machine movements are simulated based on the 3-axis movement, which includes slew (x-axis), luffing (y-axis) and travel (z-axis).

In Discrete Event System, any input variable whose value transition register as an event is defined as a TD [35]. In this simulation, the TD consists of machine status (slewing, luffing, travel and bucket) and motor run feedback (slewing, hydraulic, travel and bucket). For comparison purposes, two types of TD are presented here known as Static TD and Dynamic TD. Static TD registered the original TD signal from the source while Dynamic TD multiply the changes of TD with a threshold setting defined by the system engineer. Both types of TDs were used in this experiment to investigate which source of data to be selected. This will ensure the event modeller algorithm provides an accurate sensitivity index or weight output.

Meanwhile, the threshold level could be adjusted based on the expert point of view. In this simulation, the threshold level is set at 5% (0.05); thus, the Upper Limit and Lower Limit will automatically set to 1.05 and 0.95 consecutively. The limits will be multiplied to the individual data computationally, to calculate the changes of the current data to the previous data and reflects the algorithm for weight score using X-NOR logic. The event modeller limit is the desired weight limits which also could be adjusted based on the expert point of view. In this simulation, the event modeller limit is set at (0.8). The weights who score above the Event Modeller Limit is shaded in the ROC Output table, indicates a significant correlation between the ED and the TD.

The sequence of TD's and ED's is updated every second and are re-arrangeable according to the weighted score. The weighted score can take a value of 0.5 and 1. The value is 1 when both or none of the input/output are triggered. Otherwise, it is 0.5. The weighted score is then averaged based on the number of iteration. Having the weight score in real-time, system engineers could easily notify the management team if there is any disturbance occurs in the system state by looking at the sequence and the weights. For trending purposes, a waveform chart is presented to improve the visibility of the data.

3.1.3 Disturbance Signal

The purpose of this simulation is to test the applicability of the Event Modeller algorithm in the handling of real-time data, thus observing the reaction of the system to the abnormal events. To ensure the system is sensitive to this abnormality, a k-Disturbance signal is introduced to the system. In this simulation, the data is simulated in 3 stages known as pre-disturbance, k-disturbance and post-disturbance. The pre-disturbance refers to the warm-up stage, which represents the machine normal steady state. The k-disturbance refers to the fluctuation of the k-event data, in such generating disturbance to the system. The post-disturbance refers to the reaction of the abnormal system back to normal steady state. A 5 minutes time interval is selected for each stage, which accumulates to 15 minutes of sampling time.

In this simulation, 5 ED's signal, which represents the internal and environment parameter, has been chosen to be disturbed. This includes voltage, humidity, harmonic, temperature and wind speed. The signal data are simulated based on random fluctuation with 5% disturbance limit from normal operation data capability.

3.1.4 Simulation Results

Table 2 compares the results obtained from the simulation of the CSU machine against the k-disturbance signal. To present the findings, the performance is measured from the score of the sensitivity index from two main perspectives, which includes the overall maximum and minimum weight, and the average score for each stage. What stands out in the table is that the value of the weight has changed drastically between the pre-disturbance and during disturbance for all case. For e.g., when the k-disturbance is applied to the machine voltage using Static TD, the weight value drops from 0.83450 to 0.76154. However, when the k-disturbance is removed, the weight value rose to 0.78132. These indicate that the system parameters are sensitive to the disturbance signal introduced in the system.

It is interesting to note that all ten cases of this study have a common finding. The transition of average sensitivity index (weight) from Pre-Disturbance to k-Disturbance have the same reduction trend with an average of 8.2% reduction for Static TD and 15.94% reduction for Dynamic TD. Alternatively, when the k-Disturbance recovers, there is also a trend of increased weight with an average of 3.218% increment for Static TD and 6.023% increment for Dynamic TD. This finding confirms the interrelation between the triggered data and the k-disturbance event data, where different TD types have different performance. The Dynamic TD has a higher percentage of changes compared to Static TD.

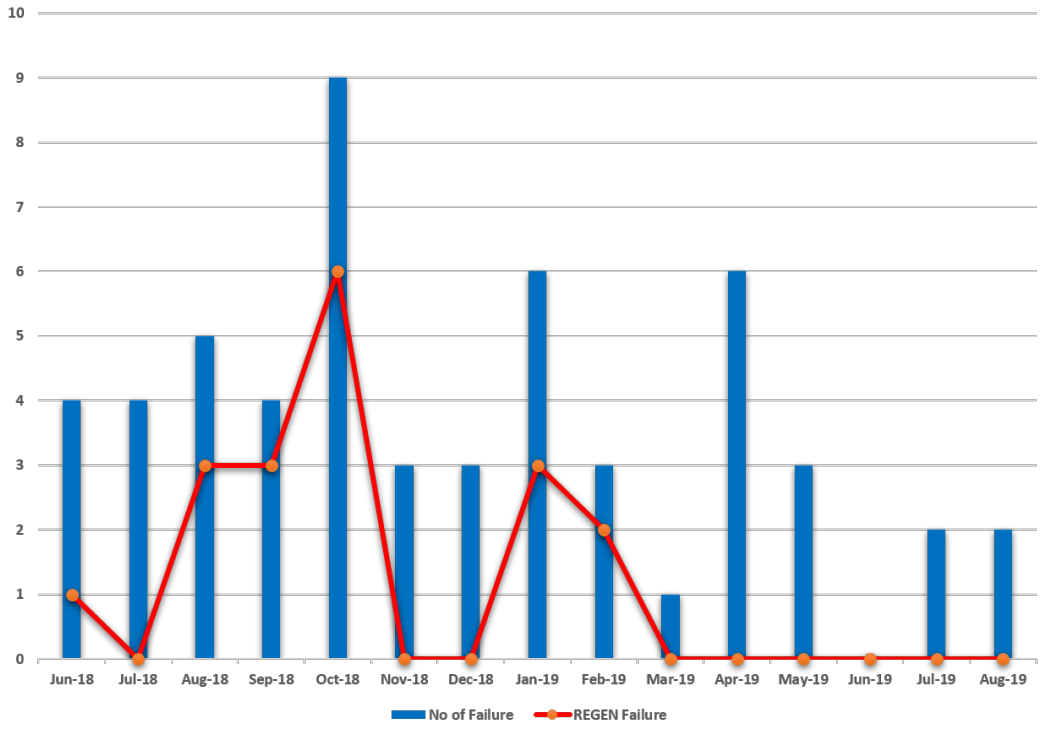


Figure 4: Tabulation of CSU Machine Failure vs REGEN Failure

The bar chart reveals that the highest number of REGEN faults was reported in October 2018, followedby August 2018, September 2018 and January 2019 consecutively.

On the other hand, it has been observed that the minimum weight for Static TD is fixed at 0.5000 while Dynamic TD has different minimum weight scores for each experiment. A possible explanation for this might be that the Dynamic TD captured the changes of input sensitivity rather than a direct Boolean signal from the input itself.

3.2 *Experiment 2: CSU Real-time Machine Failure Analysis with KPIs*

Harmonic filter failure can be influenced by various factors, such as faulty device, operator handling, and environmental factors in which the machine operates. It has also occasionally been affected by the combination of these factors. Based on the CSU machine expert point of view, the harmonic problem is associated with the occurrence of Regenerative Drive (REGEN) fault. In order to access the REGEN fault events, the abnormal dataset was selected. The frequency of the REGEN failure was compared with other types of fault, as presented in Figure 4. The bar chart reveals that the highest number of REGEN faults was reported in October 2018, followed by August 2018, September 2018 and January 2019. Although the highest number of REGEN failures was recorded in October 2018, no report indicated that the REGEN drives had been replaced. However, the January 2019 reports indicated that the maintenance team had replaced the REGEN during this month due to a fault. The January dataset was therefore chosen to further explain the relationship of the variables using the proposed technique.

3.2.1 *Experiment Strategy*

Building a solution capable of translating engineering data automatically into high-level management information is the principal motivation for this experiment. The KPIs shall be calculated based on the operation of the machine during the selected period. The method used in [1] is borrowed to run a fault dataset with a time interval of 5 minutes. The first 5 minutes are the events before the fault occurs, followed by the fault, and end with the last 5 minutes that accumulate up to 15 minutes of sampling time. The reason why this method was adopted is to prove that the event modeller could instantly solve a complex system without having to run the system for a longer period. It is easy to sample data in a simulation environment, as the user can decide when the fault may occur, but this is not the case with industrial data. The dataset was therefore analyzed to determine which data was almost equivalent to the setting. The experiment strategy is summarised in Figure 5. The results are presented in two phases. First, the event modeller was used as a tool to reduce the dimensionality of the data being observed. Second, the relationship between the observed data and the KPIs has been compared. The results will be explained as follows.

3.2.2 *Dataset Arrangement*

In this example, a homogenous dataset representing REGEN failure of the predictive model population was sampled for 16 minutes. The dataset of 30^{th} January 2019 was sampled between 12:47 pm and 13:03 pm. The dataset arrangement indicated that there were 96 TDs to be grouped against 24 EDs in the form of matrices to show

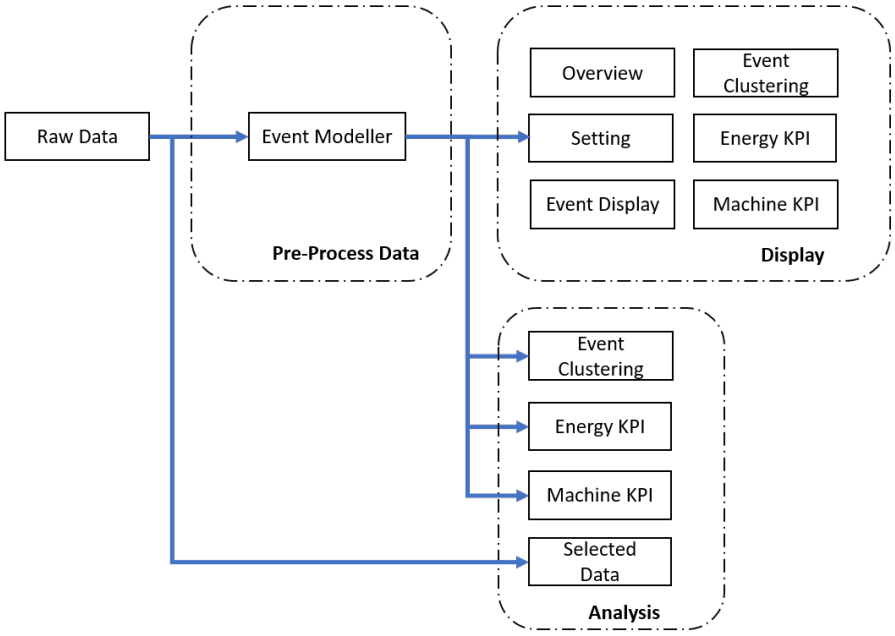


Figure 5: Experiment Strategy for KPIs Analysis

The arrangement of the three principal stages in the Experiment Strategy for KPI Analysis: Pre-process Data, Analysis and Display.

which system parameters were highly correlated. It is noteworthy that the 96 X 24 matrices are too complex for a correlation analysis; therefore, the 96 TD was divided into 4 clusters. For each TD cluster, 24 X 24 matrices were analyzed using the event modelling technique.

The sensitivity index of each input-output relationship has been measured and updated every second. The ROC pattern was observed in real-time and captured every minute. As a result, four-quadrant windows representing four key states were presented: Initial, Pre-Fault, Fault and Post Fault for each TD cluster. This analysis intends to examine how TDs influence the pattern of the ROC. It is necessary to clarify what happens during this 16-minute time stamp before presenting the chart. Based on the report, the machine is running at its optimum level before it is fully phased out after 10 minutes. The reason why machine trips are unknown, which left the proposed method to be concluded further.

3.2.3 *Dimensionality Reduction using Event Modeller*

One advantage of using Event Modeller is to reduce the dimensionality of a complex system. As the main objective is to identify the root cause of the failure, this technique is capable of highlighting the ROC pattern based on its correlation analysis. One important finding was that the first four EDs were the same for all clusters. They comprised of E-House Humidity (ED24), Portal Conveyor Current (ED18), Wind Speed (ED 12) and Portal Speed Drive (ED14). This indicates that one of these EDs would have been the cause of failure. However, it has been noted that Cluster 2 and Cluster 3 consistently remain the same pattern with or without fault happened. Therefore, we have eliminated Cluster 2 and Cluster 3 from the analysis. Moving to the ROC patterns for Cluster 1 and Cluster 4, major changes were observed after the fault in this quadrant. From this observation, it is suggested that the TDs could have originated from these clusters. Equation 3 indicates the possible root cause of Cluster 1. If we overlap both relationships, the new equation for Cluster 1 is expressed in Equation 4.

$$\rho_{0.95} = (ED18) \times (TD6, TD7, TD11, TD8, TD9, TD10) + (ED24) \times (TD8, TD9, TD10, TD24, TD21, TD2, TD22, TD23, TD20, TD13, TD12) \quad (3)$$

$$\rho_{0.95} = (ED18 + ED24) \times (TD8, TD9, TD10) \quad (4)$$

On the other hand, the following equation 5 indicates the possible root cause of Cluster 4. If we overlap both relationships, the new equation for Cluster 4 is expressed in Equation 6.

$$\rho_{0.95} = (ED18) \times (TD15, TD18, TD16, TD17, TD20) + (ED24) \times (TD16, TD17, TD14, TD13, TD19, TD12) \quad (5)$$

$$\rho_{0.95} = (ED18 + ED24) \times (TD16, TD17) \quad (6)$$

Combining Equation 4 with Equation 6, the sensitivity index is expressed in Table 3. With that, Event Modeller has reduced the dimensional to a smaller scale that will help the system engineer

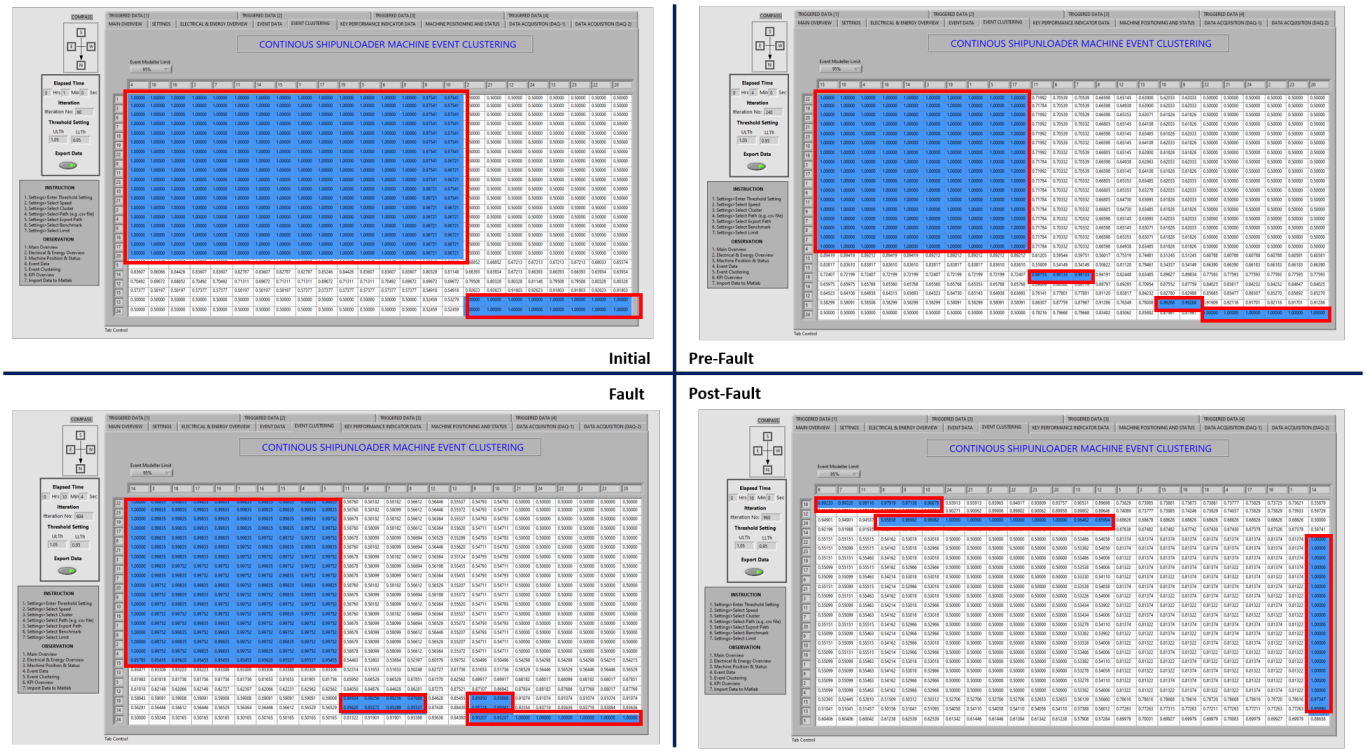


Figure 6: Rank Order Clustering - Cluster 1

ROC Pattern on four quadrant: Initial, Pre-Fault, Fault and Post-Fault for Cluster 1. The blue shaded region with red boxes represent the cluster of a high correlation value between 0.95 to 1.00.

Table 3: Summary of Sensitivity Index

Description	TD8	TD9	TD10	TD16	TD17
ED18	0.97919	0.97138	0.96878	0.98337	0.96881
ED24	0.95838	0.96982	0.996982	0.95842	0.97089

decide which relationship is liable to this fault. The following section will be discussed on the relationship between the observed data and the KPIs to validate the findings of the Event Modeller. Figure 6 and Figure 7 represent the pattern of the ROC system state for Cluster 1 and Cluster 4, respectively.

3.2.4 Relationship between the observed data against Key Performance Indicators

The KPIs are the translated index of the machine operation. It provides operational information for the decision-making of both the machine and the system operator. Four main KPIs have been adapted in this experiment known as Availability (A), Instantaneous Utilization (IU), Schedule Utilization (SU) and Performance (P). Having these four KPIs in line with CSU machine data, the possible root cause of frequent harmonic failure could be determined.

The following steps have been taken to analyze the data: (1) Select fault dataset from the Predictive Model Dataset. (2) Identify the fault observed data location. (3) Distinguish observed data according to Environment Variables, Electrical Variables and Motor Variables. (4) Plot Environment Variables data against KPIs. (5) Plot Electrical Variables data against KPIs. (6) Plot Motor Variables data against KPIs.

The easiest way to identify fault events is by exploring the main incomer trip data. Theoretically, the entire electrical system will be tripped when there is a REGEN fault. Although it may conflict with other events in which the operator may trip the main incomer on purpose, the system could automatically distinguish between the actual fault and the simulated fault using the translated KPIs data.

For example, when a fault occurred during the operation, the data of the KPIs will hold the value in percentage. This verifies that the fault that occurred is genuine. If the machine is stopped or not operated, the value is either zero or hold to a specific value. In this example, the same dataset used in the previous experiment was chosen. This is to compare the relationship between the data observed using both techniques. The dataset dated 30th January 2019 was therefore sampled between 12:47 pm and 13:03 pm. The target data was then divided into three and plotted against the KPIs as follows:

Figure 8 shows the observed electrical data against the KPIs when electrical faults occurred. This graph is quite revealing in several ways. First, the main fault has been observed. The chart reveals that the main fault occurred at 09:50. This is the turning point of all events. Next, the four KPIs are observed. The perfor-

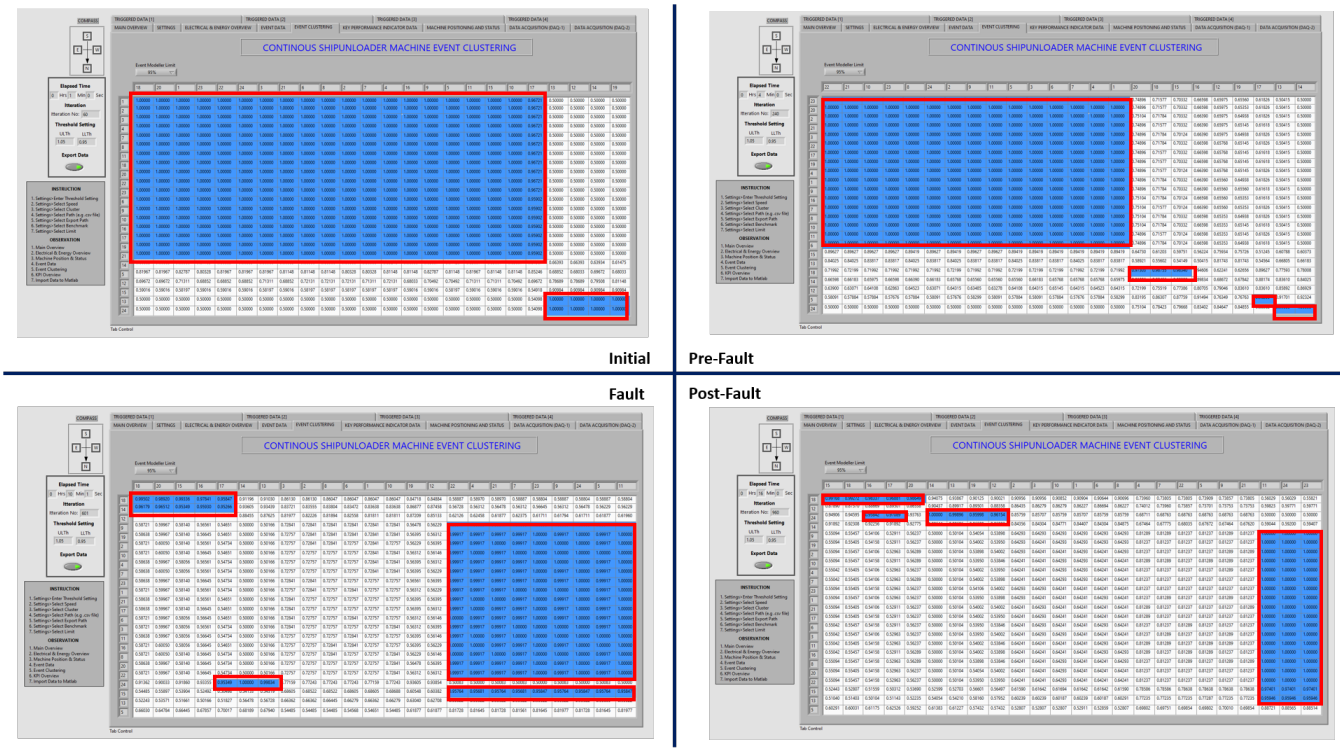


Figure 7: Rank Order Clustering - Cluster 4

ROC Pattern on four quadrant: Initial, Pre-Fault, Fault and Post-Fault for Cluster 4. The blue shaded region with red boxes represent the cluster of a high correlation value between 0.95 to 1.00.

Table 4: Summary of Electrical Variables

Description	ED1	ED2	ED4
Min	427.33 V	9.66 A	1.84 %
Max	427.47 V	242.93 A	2.19 %
Average	427.44 V	73.13 A	2.00 %

mance was initiated at 100% and remained for 1 minute. It was then exponentially reduced to 16% when the main fault occurred. This indicates that the unloading of coal has stopped immediately after 1 minute.

On the other hand, both Instantaneous Utilization and Schedule Utilization started at 58% and gradually decreased to 35% at 04:24. At this point, Schedule Utilization bounces back to 39%, but Instantaneous Utilization remains reduced to 30% until the main fault event occurs. The availability started at 95% and remained until 04:24 before it gradually decreased when the main fault occurred. These trends have shown that there is a relationship between the main fault and the KPIs. Table 5 summarised the KPI results.

Moving to electrical variables, the main input voltage (ED1) remains at 427V for the entire time. Although the fault occurred, the main input voltage does not respond to this. This validates that the fault is not caused by power disturbances from the supply, such as transient, interruption, under-voltage or over-voltage. The Main Incomer Current (ED2) pattern was then observed. Before the fault, ED2 shows the current drawing of the electrical distribution. The pattern is based on loads of the motor during machine operation.

However, when the fault occurred, trends show that ED2 is responding to the event. An instantaneous drop was discovered during this period. ED2 does not fall to zero because the energy was still drawn from the maintenance feeder for essential loads. The main incoming THD (ED4) has also been observed. ED4 has been generated assuming that there will be no harmonic distortion during these events. The ED4 pattern remains at an average of 2.0% throughout the sampling time, indicating that the THD remains at IEEE 519 Standards. Table 4 summarised the electrical variables.

Meanwhile, to further explain the relationship between the motor variables and the KPIs, Figure 9 shows the current drawing of five main processes in the CSU machine. These include Bucket Elevator Current (ED17), Portal Conveyor Current (ED18), Travel Current (ED19), Boom Conveyor Current (ED20) and Hydraulic Power Pack (ED21). In order to validate the ED2 from the previous graph, this chart reveals that ED2 has a linear relationship with all motor variables within this graph. For example, when ED17 was drawn at the beginning of the sampling and all other motor loads were added during that period the ED2 measurement reached its peak value. Some attention is given to the small circle in the chart. It is clear that there was some instability happening at 01:48, which could be linked to the fault.

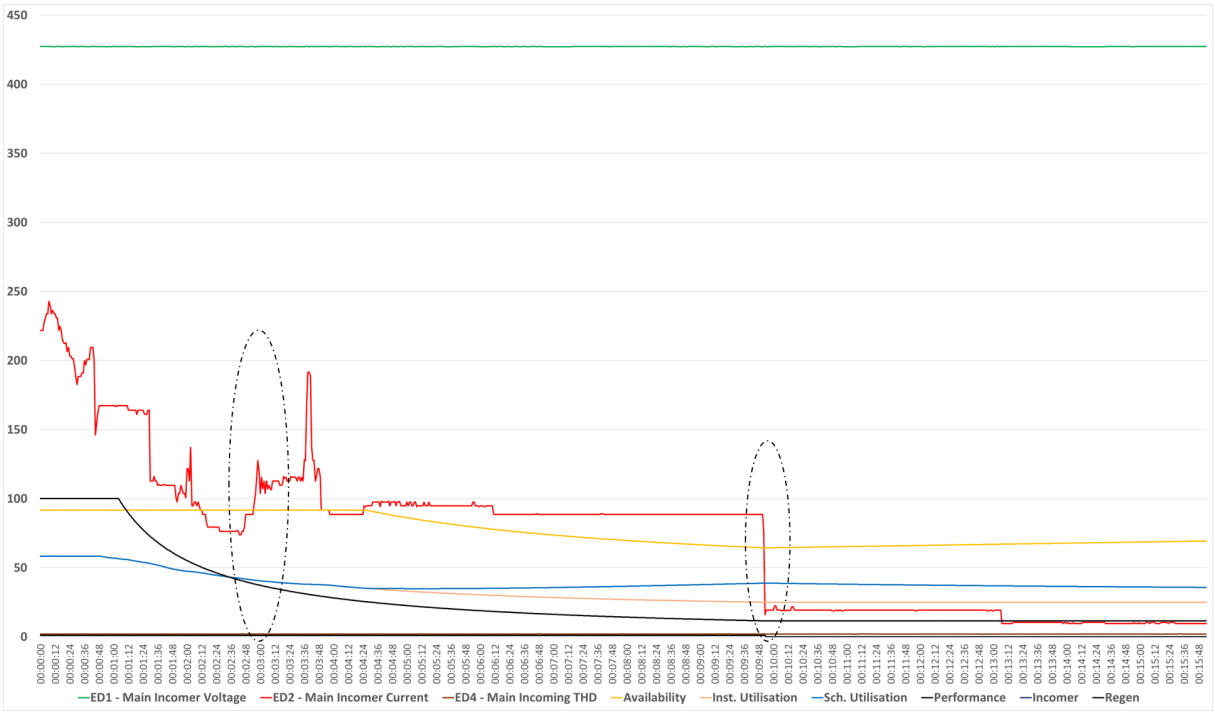


Figure 8: Electrical Variables vs KPIs

This figure illustrates the relationship between the Machine KPIs (Availability, Instantaneous Utilization, Schedule Utilization, and Performance) against the observed electrical data (Voltage, Current, Harmonic, Incomer and Regen) when the same event of electrical faults occurred at Time: 09:50.

Table 5: Summary of KPIs

Description	A	IU	SU	P
Min	64.37 %	25.01 %	34.82 %	11.52 %
Max	91.67 %	58.33 %	58.33 %	100.00%
Average	76.43 %	32.48 %	39.59 %	26.72 %

ED18 had an impulsive signal and remained zero after that event. On the other hand, ED19 has also revealed some dramatic changes. Just after ED18 had an impulsive signal, ED19 experienced some negative distortion for 12 seconds, followed by an interruption for 48 seconds and a further positive distortion for 60 seconds completely before it stopped. Another attention has to be made at ED17. In the beginning, ED17 had drawn some fluctuation current and had spiked for a few seconds before stopping at 00:40. The trend shows that ED17 is struggling to recover, but ends up stopping when the main fault has occurred. Otherwise, all other motor variables are operating in their normal state. Table 6 summarised the motor variables.

Figure 10 displays the environment data observed against the KPIs when the same electrical faults have occurred. Five environmental data were reported, including Panel Humidity (ED3), Panel Temperature (ED5), E-House Temperature (ED6), Wind Speed (ED12) and E-House Humidity (ED24). Attention must be given to the humidity and temperature of the electrical room where the electrical distribution panel is located. The pattern has shown an instant increase of ED24 at 02:36. The humidity builds up and reaches its highest peak at 58% and slowly decreased with two spikes before the fault occurred. On the other hand, ED3 has also been increased but not as much as ED24. ED5 has an average temperature of 33.13 degrees Celsius, while ED5 has an average temperature of 17.82 degrees Celsius. Eventually, ED12 fluctuates within the normal range. Table 7 summarised the environment variables.

3.2.5 *Discussion*

The main outcome of the experiment is to show how the proposed technique can be used as a tool to reduce the dimensions of a highly complex system. In this case study, 96 TDs were analyzed against 24 EDs in real-time to help system engineers visualize any changes to the system state. The EMDA implementation will group high correlation system parameters in the form of matrices and place them in mutually exclusive blocks. Having a huge matrix, however, will make things difficult. The matrix needs to be presented as a square for clear visibility. Thus, the 96 TDs were divided into 4 clusters, where each cluster has a relationship with 24 EDs. The example accessed a 16-minute dataset that belongs to a fault population. The result for each cluster was shown in a four-quadrant to show the pattern change. In this example, there is a significant change in the pattern observed in Cluster 1 and Cluster 4. Further analysis of the two clusters suggested a new formula representing

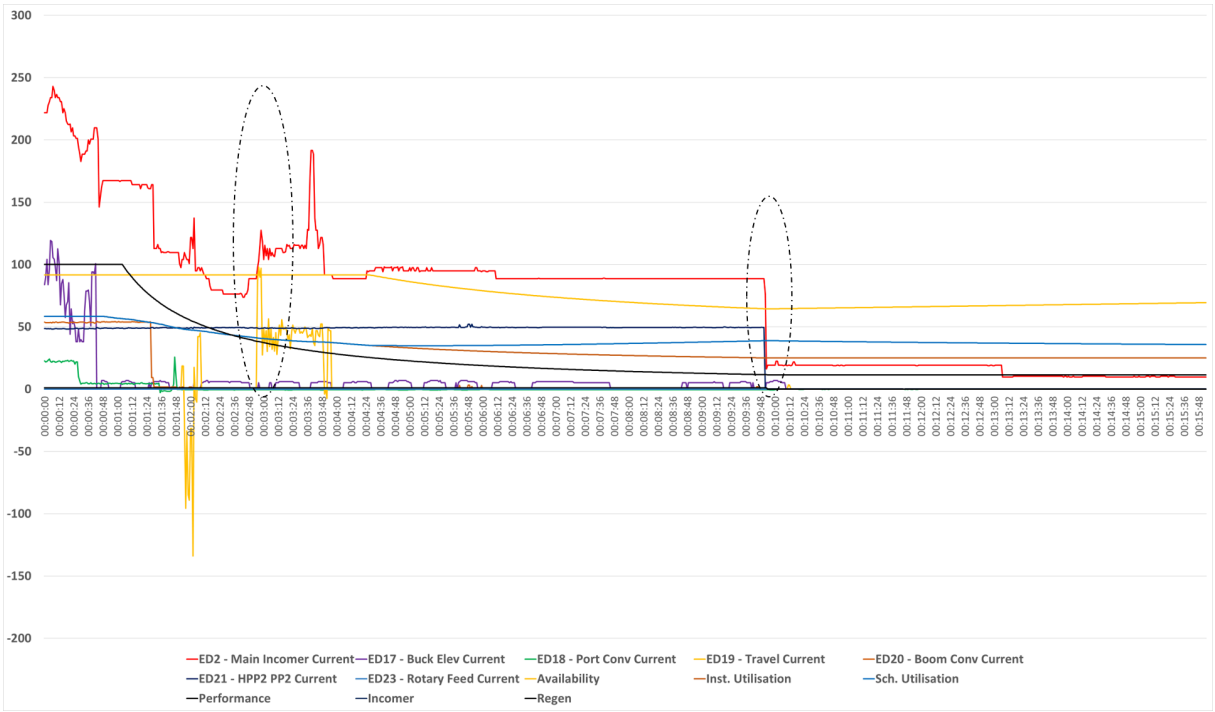


Figure 9: Motor Variables vs KPIs

This figure illustrates the relationship between the Machine KPIs (Availability, Instantaneous Utilization, Schedule Utilization, and Performance) against the Five Main Electrical Motors (Bucket Elevator Current, Portal Conveyor Current, Travel Current, Boom Conveyor Current and Hydraulic Power Pack) when the same event of electrical faults occurred at Time: 09:50.

Table 6: Summary of Motor Variables

Description	ED17	ED18	ED19	ED20	ED21
Min	0.00 A	-2.96 A	-133.97 A	0.00 A	0.05 A
Max	119.31 A	25.67 A	97.87 A	54.41 A	52.05 A
Average	5.41 A	0.69 A	2.49 A	5.03 A	30.60 A

the state of the system. After the fault event, the new formula for Cluster 1 and Cluster 4 is expressed in Equation 4 and Equation 6, respectively. As a result of these equations, the main culprit of the fault was identified. The ED variables were reduced from 24 EDs to 2 EDs (a decrease of 91.76%) while the TDs variable was reduced from 96TDs to 5TDs (a decrease of 94.79%). As listed in Table 3, the possible root causes of the fault are the combination of:

1. Portal Conveyor Current (ED18) with Rotary Feeding Table Motors (TD8), Bucket Elevator Motors (TD9), Bucket Elevator Brake Motors (TD10), Busy Rotary Table (TD16) and Busy Bucket Elevator (TD17) or

2. E-House Humidity (ED24) with Rotary Feeding Table Motors (TD8), Bucket Elevator Motors (TD9), Bucket Elevator Brake Motors (TD10), Busy Rotary Table (TD16) and Busy Bucket Elevator (TD17)

Further analysis with the KPIs will validate the existence of unknown events that could be linked to a harmonic problem. In normal circumstances, electrical variables such as voltage, current and THD should have a linear relationship with the KPIs. As the KPIs responded to the fault, the electrical variables should also respond. However, the voltage and the harmonic measurement remain constant after the fault has occurred. Although current measurements have responded to this fault, trends show that the current value is not zero. It indicates that there was still a current drawing for the essential loads. Therefore, this eliminates the possibility of machine failure due to electrical variables.

Moving to the motor variables, as explained in the result, ED17, ED18 and ED19 showed some undiscovered events that could lead to fault. When comparing the performance trend to ED17, the performance measurement falls in line with the activity of the bucket elevator. Although the trends for ED17 fluctuate before it stops, this could be explained from the perspective of uneven coal inside the bucket. On the other hand, ED18 has experienced an impulsive signal that leads to an immediate stop. However, looking at the timing of this event, it took about 9 minutes for the fault to happen. The same goes for ED19, which experienced negative and positive distortions right after ED18. This can be claimed as a result of an event. However, when it comes to the type of operation, ED19 is a travel motor that operates in two directions. Although some distorted measurements were made, this could be explained by the

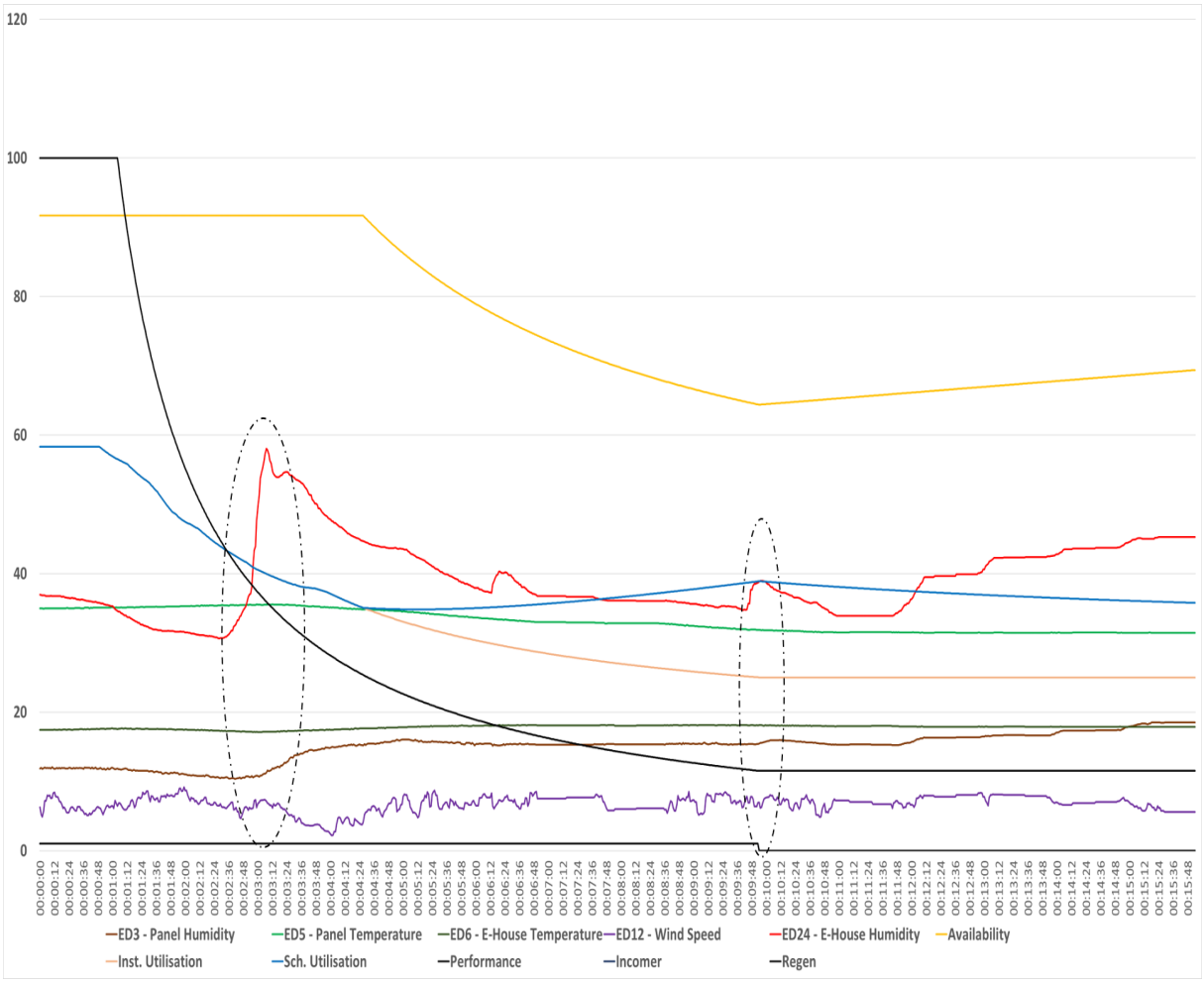


Figure 10: Environment Variables vs KPIs

This figure illustrates the relationship between the Machine KPIs (Availability, Instantaneous Utilization, Schedule Utilization, and Performance) against the Environment Variables (Panel Humidity, Panel Temperature, E-House Temperature, Wind Speed and E-House Humidity) when the same event of electrical faults occurred at Time: 09:50.

number of motors that were operating. There were ten motors in service working at different efficiency levels. Having said that, the motor variables are also excluded from being the root cause of the fault.

Turning to environment variables, the trends in Figure 10 have shown some indications that the humidity could be related to the fault. The rise in humidity up to 58.06% shows that the state of the system has been compromised. Although high humidity patterns started 7 minutes before the fault, the measurement continuously retained above 40 % after the fault. This type of phenomena takes a fixed amount of time to reach the fault level. To explain this phenomenon, we take an example of preheating an oven. When a person decides to bake a cake, the oven needs to be heated to the right temperature before putting the cake inside the oven. Although it takes a few seconds to turn the oven on, it may take a few minutes to get the right temperature. This phenomenon is referred to as a fixed time delay deterministic event. However, if the person warms up the oven, without shutting the oven door, it may take longer to get to the right temperature. This phenomenon is referred to as a deterministic sequence of different input events. The heat produced in the oven was cooled by the amount of cold air present in the surroundings, which would delay the oven pre-heating process.

When we relate this to our case study, the power electronic component such as choke, rectifier, REGEN and control drives produce an excessive amount of heat during operation. To ensure that the heat is kept at certain limits, two units of air conditioners were designed to combat this heat. However, after considering the amount of heat generated in the room along with the hot and humid climate of Malaysia throughout the year, experts have reviewed and decided to install two additional air conditioners to cater for this amount of heat. As a result, when CSU machine is fully operational, the four-unit air conditioners are sufficient to cool down the electrical room. However, when the machine stops at intervals, the air conditioners are found oversized to the power electronic component. As a consequence, the humidity increases; the component becomes wet, resulting in machine failure. The humidity in the electrical switchgear room must be maintained at an elevated temperature relative to the ambient inside the room. Condensation is usually considered a problem only if the humidity in the switchgear room is 65% or greater [44]. Although the humidity in this example does

Table 7: Summary of Environment Variables

Description	ED3	ED5	ED6	ED12	ED24
Min	10.35 %	31.41 degC	17.12 degC	2.22 m/s	30.64 %
Max	18.52 %	35.54 degC	18.16 degC	9.19 m/s	58.06 %
Average	14.95 %	33.13 degC	17.82 degC	6.74 m/s	39.21 %

not reach 65%, considering the machine has been running for a decade, the efficiency of the machine may decrease as the humidity may be the contributing factor. This finding suggests that there is a relationship between the humidity and the fault.

As mentioned earlier, this paper's main outcome is to show how the EMDA can be used as a tool to reduce the dimensions of a highly complex system. 96 Triggered Data were scanned against 24 Event Data, which then formulated an equation representing the system state in a single second. This sort of like a "scanner" that will keep converting a complex relationship into a simple equation (a huge reduction of dimensionality) in a low computational effort. The results in Table 3 shows that the complex relationship has been reduced to 2 EDs against 5 TDs. Further analysis with the KPIs has validated the existence of unknown events that linked to the frequent harmonic failure. The rise in humidity level shown in Figure 10 has proven that the system state has been compromised. In comparison, other variables such as electrical variables, motors and other environment do not show any evidence that compromises the system state. Besides, the relationship between the observed data and the KPIs appears to be consistent. The suggested finding was tested at site, and the operators confirmed the importance of humidity factor (environmental factor) in the machine's performance. A humidity regulator has been suggested to be installed at the site. The control of the regulator was linked to the plant automatic control system.

4 Conclusions

Managing repetitive harmonic failures in an industrial application is a recursive, costly and time-consuming exercise. Industrial data are multidimensional and random in nature. The problem with analyzing this data is that many events do not necessarily repeat often enough to have statistical significance. Still, they cause minor or major faults and need to be captured and understood. Identical machines perform differently in different environments and conditions. In most cases, PQ engineers focus on the fundamental problem, transforming the power system signal into various signal processing methods to locate the fault. An Event Modeller technique, which is low in the computational effort, has been considered to solve this problem. This technique, which is extremely computational effective, managed to reduce system state definitions' complexity into a simpler formulation. Also, the performance metrics translated from the machine operation are now automated and require minimum operator intervention. It reduces wrong interpretations and possible mistakes in decision making. However, one of the method's weaknesses is its reliance on technical expert judgement on adjusting the threshold settings. One of the ongoing research endeavours of the authors is to automate the trigger detection phenomena.

In this present work, authors focused on reducing dimensionality while collectively including the modelling equation's performance metrics parameter to verify the system state. Two approaches were taken to validate and verify the propose EMDA algorithm. First was the simulation process (SiL). It considered the voltage, humidity, harmonic, slewing movement, luffing movement, travelling movement, temperature, and wind speed as system output variables. Two types of triggering data, i.e. Static TD and Dynamic TD were considered. K-disturbance signal was generated for the five (excluding movements) output variables. The results show that, as the disturbance is induced, the sensitivity indices were reduced by an average of 8.2% for Static TD and 15.94% for Dynamic TD. However, when the disturbance was removed, the sensitivity indices regained by an average of 3.218% and 6.023% for Static TD and Dynamic TD consecutively. As a result, the Dynamic TD was chosen for the field experiments. In the field experiment (HiL), 914 data series in real-time were acquired from the SCADA. Each dataset is containing 96 Input vs 24 Output parameters. The results show that two new parameters were detected to significantly impact Harmonic Filters; humidity and portal conveyor current (velocity of the conveyor). The tests proved that humidity as an environmental factor is linked to the frequent harmonic problems. This revelation was also verified by looking up the historical evidence of humidity data against the frequency of failures and the variation in the feeder's speed (conveyor). In the future, the proposed EMDA will be integrated as a vehicle for training supervised machine learning models to predict the remaining useful life of the filter as a predictive

Acknowledgement This work has received financial support from Malaysia's government sponsorship, MARA. The authors would like to acknowledge the editor and anonymous referees.

References

[1] F. Z. M. Fadzil, A. Mousavi, M. Danishvar, "Simulation of Event-Based Technique for Harmonic Failures," in 2019 IEEE/SICE International Symposium on System Integration (SII), 66–72, IEEE, 2019, doi:10.1109/SII.2019.8700381.

[2] N. N. Gomaa, K. Y. Youssef, M. Abouelatta, "On Design of IoT-based Power Quality Oriented Grids for Industrial Sector," Advances in Science, Technology and Engineering Systems Journal, **5**(6), 1634–1642, 2020, doi: 10.25046/aj0506194.

[3] U. Singh, S. N. Singh, "Application of fractional Fourier transform for classification of power quality disturbances," IET Science, Measurement & Technology, **11**(1), 67–76, 2017, doi:10.1049/iet-smt.2016.0194.

[4] M. Panteli, P. Mancarella, "Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies," Electric Power Systems Research, **127**, 259–270, 2015, doi:10.1016/j.epsr.2015.06.012.

[5] M. K. Saini, R. Kapoor, "Classification of power quality events–a review," International Journal of Electrical Power & Energy Systems, **43**(1), 11–19, 2012, doi:10.1016/j.ijepes.2012.04.045.

[6] Z. Huang, M. Li, A. Mousavi, M. Danishva, Z. Wang, "EGEP: An Event Tracker Enhanced Gene Expression Programming for Data Driven System Engineering Problems," IEEE Transactions on Emerging Topics in Computational Intelligence, **3**(2), 117–126, 2019, doi:10.1109/TETCI.2018.2864724.

[7] M. Morcos, W. A. Ibrahim, "Electric Power Quality and Artificial Intellignece: Overview and Applicability," IEEE Power engineering review, **19**(6), 5–10, 1999, doi:10.1109/MPER.1999.768508.

[8] D. Saxena, K. Verma, S. Singh, "Power quality event classification: an overview and key issues," International Journal of Engineering, Science and Technology, **2**(3), 186–199, 2010.

[9] D. C. Robertson, O. I. Camps, J. S. Mayer, W. B. Gish, "Wavelets and electromagnetic power system transients," IEEE Transactions on Power Delivery, **11**(2), 1050–1058, 1996, doi:10.1109/61.489367.

[10] S. Santoso, E. J. Powers, W. M. Grady, P. Hofmann, "Power quality assessment via wavelet transform analysis," IEEE transactions on Power Delivery, **11**(2), 924–930, 1996, doi:10.1109/61.489353.

[11] J. Barros, R. I. Diego, "Analysis of harmonics in power systems using the wavelet-packet transform," IEEE Transactions on Instrumentation and Measurement, **57**(1), 63–69, 2008, doi:10.1109/TIM.2007.910101.

[12] S. Shukla, S. Mishra, B. Singh, "Empirical-mode decomposition with Hilbert transform for power-quality assessment," IEEE transactions on power delivery, **24**(4), 2159–2165, 2009, doi:10.1109/TPWRD.2009.2028792.

[13] R. G. Stockwell, "A basis for efficient representation of the S-transform," Digital Signal Processing, **17**(1), 371–393, 2007, doi:10.1016/j.dsp.2006.04.006.

[14] S.-H. Cho, G. Jang, S.-H. Kwon, "Time-frequency analysis of power-quality disturbances via the Gabor–Wigner transform," IEEE transactions on power delivery, **25**(1), 494–499, 2010, doi:10.1109/TPWRD.2009.2034832.

[15] C. Norman, J. Y. Chan, W. H. Lau, L. L. Lai, "Hybrid Wavelet and Hilbert Transform With Frequency-Shifting Decomposition for Power Quality Analysis." IEEE Trans. Instrumentation and Measurement, **61**(12), 3225–3233, 2012, doi:10.1109/TIM.2012.2211474.

[16] B. Biswal, P. K. Dash, S. Mishra, "A hybrid ant colony optimization technique for power signal pattern classification," Expert Systems with Applications, **38**(5), 6368–6375, 2011, doi:10.1016/j.eswa.2010.11.102.

[17] K. Manimala, K. Selvi, R. Ahila, "Optimization techniques for improving power quality data mining using wavelet packet based support vector machine," Neurocomputing, **77**(1), 36–47, 2012, doi:10.1016/j.neucom.2011.08.010.

[18] R. Ahila, V. Sadasivam, K. Manimala, "An integrated PSO for parameter determination and feature selection of ELM and its application in classification of power system disturbances," Applied Soft Computing, **32**, 23–37, 2015, doi:10.1016/j.asoc.2015.03.036.

[19] C. Cortes, V. Vapnik, "Support-vector networks," Machine learning, **20**(3), 273–297, 1995, doi:10.1007/BF00994018.

[20] L. A. Zadeh, "Fuzzy sets," in Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh, 394–432, World Scientific, 1996.

[21] S. Khokhar, A. A. M. Zin, A. S. Mokhtar, N. A. M. Ismail, N. Zareen, "Automatic classification of power quality disturbances: A review," in 2013 IEEE Student Conference on Research and Developement, 427–432, 2013, doi:10.1109/SCOReD.2013.7002625.

[22] D. Granados-Lieberman, R. Romero-Troncoso, R. Osornio-Rios, A. Garcia-Perez, E. Cabal-Yepez, "Techniques and methodologies for power quality analysis and disturbances classification in power systems: a review," IET Generation, Transmission & Distribution, **5**(4), 519–529, 2011, doi:10.1049/iet-gtd.2010.0466.

[23] M. McGranaghan, "Overview of Power Quality Standards," PQ Network Internet Site, http://www. pqnet. electrotek. com/pqnet, 1998.

[24] I. . W. Group, et al., "IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems," IEEE STD, 519–1992, 1992.

[25] B. IEC, "61000-3-4," Electromagnetic compatibility (EMC)-part, 3–4, 1998.

[26] P. Muchiri, L. Pintelon, "Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion," International journal of production research, **46**(13), 3517–3535, 2008, doi:10.1080/00207540601142645.

[27] V. C. Angadi, A. Mousavi, D. Bartolomé, M. Tellarini, M. Fazziani, "Regressive Event-Tracker: A Causal Prediction Modelling of Degradation in High Speed Manufacturing," Procedia Manufacturing, **51**, 1567–1572, 2020, doi:10.1016/j.promfg.2020.10.218.

[28] V. C. Angadi, A. Mousavi, D. Bartolomé, M. Tellarini, M. Fazziani, "Causal Modelling for Predicting Machine Tools Degradation in High Speed Production Process," IFAC-PapersOnLine, **53**(3), 271–275, 2020, doi:10.1016/j.ifacol.2020.11.044.

[29] M. Danishvar, V. C. Angadi, A. Mousavi, "A PdM Framework Through the Event-based Genomics of Machine Breakdown," in 2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM), 1–6, IEEE, 2020, doi:10.1109/APARM49247.2020.9209530.

[30] Z. Huang, V. C. Angadi, M. Danishvar, A. Mousavi, M. Li, "Zero Defect Manufacturing of Microsemiconductors–An Application of Machine Learning and Artificial Intelligence," in 2018 5th International Conference on Systems and Informatics (ICSAI), 449–454, IEEE, 2018, doi:10.1109/ICSAI.2018.8599292.

[31] M. Danishvar, V. Vasilaki, Z. Huang, E. Katsou, A. Mousavi, "Application of Data Driven techniques to Predict N 2 O Emission in Full-scale WWTPs," in 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), 993–997, IEEE, 2018, doi:10.1109/INDIN.2018.8472075.

[32] M. Danishvar, A. Mousavi, P. Broomhead, "EventiC: A Real-Time Unbiased Event-Based Learning Technique for Complex Systems," IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, doi:10.1109/TSMC.2017.2775666.

[33] S. Tavakoli, A. Mousavi, S. Poslad, "Input variable selection in time-critical knowledge integration applications: A review, analysis, and recommendation paper," Advanced Engineering Informatics, **27**(4), 519–536, 2013, doi:10.1016/j.aei.2013.06.002.

[34] M. Danishvar, A. Mousavi, P. Sousa, "EventClustering for improved real time input variable selection and data modelling," in 2014 IEEE Conference on Control Applications (CCA), 1801–1806, 2014, doi:10.1109/CCA.2014.6981574.

[35] S. Tavakoli, A. Mousavi, P. Broomhead, "Event tracking for real-time unaware sensitivity analysis (EventTracker)," IEEE Transactions on Knowledge and Data Engineering, **25**(2), 348–359, 2013, doi:10.1109/TKDE.2011.240.

[36] J. R. King, "Machine-component grouping in production flow analysis: an approach using a rank order clustering algorithm," International Journal of Production Research, **18**(2), 213–232, 1980, doi:10.1080/00207548008919662.

[37] M. R. Kumar, G. P. Rao, "Design and implementation of 32 bit high level Wallace tree multiplier," Int. Journ. of Technical Research and Application, **1**, 86–90, 2013.

[38] A. Mousavi, H. Siervo, "Automatic translation of plant data into management performance metrics: a case for real-time and predictive production control," International Journal of Production Research, **55**(17), 4862–4877, 2017, doi:10.1080/00207543.2016.1265682.

[39] S. X. Ding, S. Yin, K. Peng, H. Hao, B. Shen, "A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill," IEEE Transactions on Industrial Informatics, **9**(4), 2239–2247, 2012, doi:10.1109/TII.2012.2214394.

[40] K. Ragunathan, S. Ravindranathan, S. Naveen, "Statistical KPIs in HMI panels," in 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), 838–843, IEEE, 2015, doi:10.1109/ICGCIoT.2015.7380579.

[41] N. Kang, C. Zhao, J. Li, J. A. Horst, "Analysis of key operation performance data in manufacturing systems," in 2015 IEEE International Conference on Big Data (Big Data), 2767–2770, IEEE, 2015, doi:10.1109/BigData.2015.7364078.

[42] G. UK, "Greenhouse Gas Reporting: Conversion Factors 2017," GOV. UK [Online]. Available: https://www. gov. uk/government/publications/greenhouse-gas-reporting-conversion-factors-2017., 2017.

[43] Y. Zhu, "Multivariable system identification for process control", Elsevier, 2001.

[44] J. E. Bowen, M. C. Moore, M. W. Wactor, "Application of switchgear in unusual environments," in 2010 Record of Conference Papers Industry Applications Society 57th Annual Petroleum and Chemical Industry Conference (PCIC), 1–10, IEEE, 2010, doi:10.1109/PCIC.2010.5666870.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1356-1365 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Improved Design and Recommendations for Street Lighting in Gitega City

Ntawuhorakomeye Noel[1,*], Ndayiragije Leonidas[2], Belov Mikhail Petrovich[1]

[1]*Department of Robotics and Industrial Automation, Saint- Petersburg Electrotechnical University ETU"LETI", 197227, Russia*

[2]*Department of Electromechanical Engineering, Burundi University, 1550, Burundi*

ARTICLE INFO

Article history:
Received: 07 September, 2020
Accepted: 01 December, 2020
Online: 16 December, 2020

Keywords:
Solar mini power plant
Gitega climate
Public lighting
Smart lamp

ABSTRACT

The article discusses the possibility of using solar energy for street lighting systems in the city of Gitega, in the Republic of Burundi. Analysis of weather and climate conditions of the city was carried out and an effective street lighting system based on solar mini-power plants using an intelligent control system developed. Calculations of technical and economic efficiencies are made and the principle of operation of the developed system is described. A comparison of the methods of transmitting control signals and the performance of different types of lamps was made as well as an assessment of the battery discharge process. The different attachment types of solar panels to their supports were highlighted, the illumination received on a public space according to the installation height of the luminaire and the used LED power as well as the requirements in terms of illumination and brightness depending on the place to be lit were discussed. The system developed consists of two solar panels, connected in parallel, one battery and six LED lamps with adjustable power. At present, the problem of modernizing the lighting of public places and roads of regional and local significance in Gitega city remains relevant. With the aim of promoting clean energy and reducing energy consumption due to global concern about climate change, the use of modern, cost-effective lighting systems based on the use of LED technology and is one of the most dynamic trends in the world at present and in the near future for the well-being of the population.

1. Introduction

Light is a fundamental element for human activity, the day is lit by the sun but when night comes artificial light becomes mandatory. Since solar energy is available all over the globe, the city of Gitega, now the political capital of the Republic of Burundi, located in the center of the country, is privileged to have a strong solar deposit. This work is an extended version of the report [1] presented in 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), where a public Lighting by Mini Solar Power Plants in the Republic of BURUNDI is proposed. With a duration of insolation exceeding 3000 hours per year and a considerable annual average daily irradiation, this region can be a pole of excellence for the use and promotion of photovoltaic solar energy [1, 2]. In this work, we propose a design and dimensioning of autonomous systems for street lighting using solar photovoltaic energy as the primary source. Obtaining information about the intensity of solar radiation in a specific geographic and climatic zone gives an idea of the possibility of using solar energy to generate a certain amount of electricity. The climatic conditions of the city of Gitega make this value almost constant throughout the year [2]. This is an important parameter when calculating any type of solar power plant (SPP). However, in order to improve the energy efficiency of photovoltaic lighting systems, it is necessary to use both high-efficiency photovoltaic modules as well as efficient batteries and charge regulators for energy storage and management; but also it is essential to use luminaires with high energetic efficiency (number of lumens/Watt) [3–7].

Nowadays, artificial lighting has become more than a means of obtaining light; it is an indispensable element of human life in general and in urban areas in particular [8]. The lighting of public places, traffic lanes, monuments and sites in cities and the countryside has now become a major concern for electricity

*Corresponding Author: Ntawuhorakomeye Noel, noentaw1@mail.ru

distributors and administrative and political authorities. Public lighting plays a very important role in various places of attendance both economically and socially [8, 9]. It is above all a measure aimed at increasing the safety of public spaces, whether in traffic areas or in relaxation and recreation areas. The objective is not to provide light only for safety purposes, but also to create a pleasant nighttime atmosphere giving a new vision and approach to the spaces [9]. To produce light, there are three main techniques [1, 8]: incandescent (classical and halogen), fluorescence and LED (Light-emitting diode). Due to the high efficiency and the absence of infrared radiation, LED lamps practically do not heat up during operation. Currently existing LED lamps consist of dozens, and sometimes hundreds of LEDs. The quality of colored light is still an unrivaled achievement of LED technology. The main difference between LED luminaires (lamps) and all others is their unprecedented long service life (up to 100,000 hours) [5, 8]. For electric lighting systems that operate at night, the period of electricity accumulation during daylight hours is a very important. A 120 amp-hours (120 Ah) accumulator battery with two 200-watt solar panels (200 W) were used in the design of this mini SPP to power the six public lighting LEDs. The developed intelligent system is a solution for remote control of street lighting, which has the ability to control the lamps and the lighting level of each street lamp. Increasing the reliability of street lighting networks will ensure energy saving, uninterrupted power supply, and as a result, ensure road safety as the number of accidents and illegal actions is significantly reduced when the city is well lit. The use of an intelligent street lighting control system will ensure safer traffic conditions, pedestrian safety, and significantly improve the city's architectural, tourist, and commercial aspects [10, 11]. All this with the aim of promoting clean energy, and reducing energy consumption due to global concern on climate change.

2. Characteristics and Features of Gitega City Climate

Gitega city is located in the center of the country about 100 km east of the economic capital Bujumbura. Its geographical coordinates are 3° 25' South, 29° 55' East according to Meteonorm 7.3, Sat = 100% of PVSYST. It is part of Africa in the southern hemisphere. This region is characterized by hills composed of metamorphic rocks such as schists, quartzites and phyllites. The Gitega province is located in the Central Plateau area with an altitude ranging from 1600 to 2000 m. It has a tropical climate tempered by altitude. Under normal conditions, the climate of the Gitega city is characterized by two (2) well-marked seasons namely the dry season, which begins in May and continues until October with minimas in June and July as well as the rainy season. The average precipitation is 1200 mm and the average temperature over the year 22.6°C (As shown in Table 1). The main nearby river is the Ruvubu with two main tributaries namely the Ruvyironza and the Mubarazi. Several districts make up the town of Gitega for about 10 km, along and around the RN2 (national road № 2) that crosses it: Magarama, Nyamugari (Swahili district), the commercial district, the administrative district, Musinzira, Shantanya, Rutonde, Nyabiharage, Nyabututsi and Mushasha (religious district). Located on an important relief, there is significant solar irradiation whose average value received annually is estimated at 1800 kWh/m^2/year. The following Table shows the average temperatures, solar radiation and wind velocity, by months according to Meteonorm 7.3, Sat = 100% of PVSYST.

Table 1: Temperature, Average irradiation, Wind velocity and Relative humidity at Gitega

	Average temperature (^{0}C)	Average irradiation (kWh/m^2/mth)	Wind velocity (m/s)	Relative humidity (%)
January	23.3	149.5	3.30	64.0
February	24.3	134.3	3.49	57.5
March	24.4	156.9	3.50	61.1
April	23.2	147.6	3.39	68.6
May	22.7	151	3.70	66.5
June	21.3	151.1	3.80	65.7
July	20.7	158.5	3.80	64.9
August	21.0	156.2	3.80	63.6
September	21.6	149.7	3.99	62.5
October	22.8	154	3.80	62.0
November	22.7	136.0	3.10	68.3
December	23.4	146.5	3.19	65.8
Year	**22.6**	**1791.3**	**3.6**	**64.2**

Based on the abovementioned and presented data in the table, we can conclude that the climatic conditions of the city of Gitega are ideal for the effective use of electricity from solar panels (SP). The Figure1 shows the temperature and irradiation variation according to the months of the year.

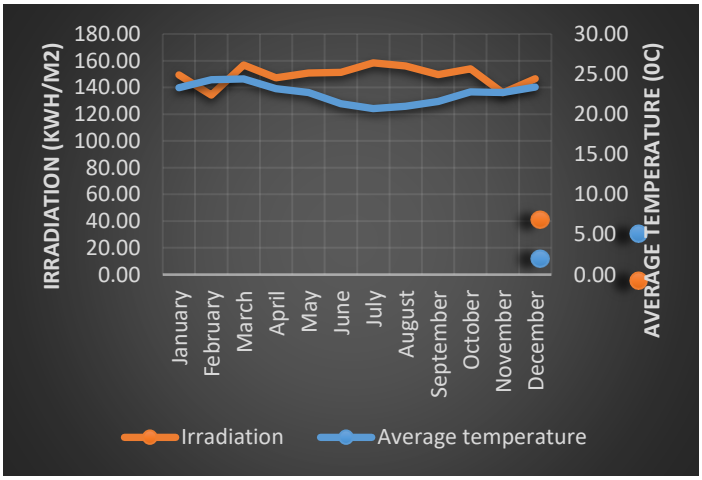


Figure 1: Temperature and Irradiation variation

Figure 2 Shows the Sun Path and Horizon Line respectively at Gitega. It is noted that the city of Gitega sees 9 hours of sunshine per day. Large solar irradiation begins to be observed at 10 o'clock, and begins to decrease at 15 o'clock with a maximum at 12 o'clock [1].

Knowing that there is no public lighting in general in this city and that the only light that the streets and public squares receive comes from the exterior lights of the surrounding houses and buildings, the development of a public lighting system is very important for the well-being of residents and passers-by, especially at night.

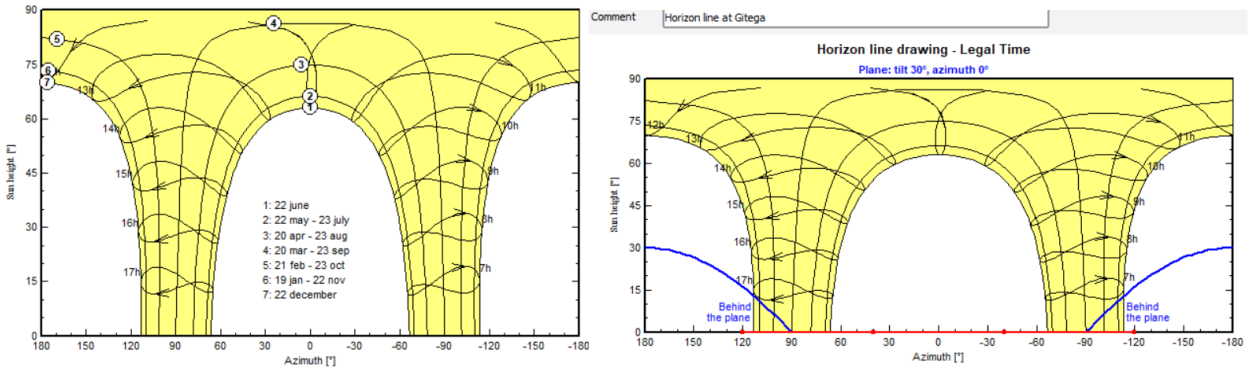


Figure 2: Sun Path Shape Horizon Line

Table 2: Requirements for Roads and Streets Lighting

Lighting Category	Streets and Roads	The Highest Traffic Intensity in Both Directions, units / h	Coating Average Brightness, Cd /m^2	Coating Average Horizontal Illumination, lx
A	Main roads, main streets of citywide significance	More than 3000	1,6	20
		1000 to 3000	1,2	20
		500 to 1000	0,8	15
B	Main streets of district significance	More than 2000	1,0	15
		1000 to 2000	0,8	15
		500 to 1000	0,6	10
		Less than 500	0,4	10
C	Local streets and roads	500 and more	0,4	6
		Less than 500	0,3	4
		Single cars	0,2	4

Table 3: Horizontal Illumination of Some Places

Places to light	Average horizontal illumination, Lx
Main pedestrian streets, impassable parts of squares of categories A and B, and pre-factory squares	10
Pedestrian streets: -within community centers	6
-in other territories	4
Sidewalks separated from the roadway on streets	2-4
Public transport boarding areas on all categories of streets	10
Pedestrian bridges, Children's playgrounds where outdoor games equipment is located	10
Pedestrian tunnel: -in the daytime	100
- evening and night	50
Pedestrian tunnel stairs in the evening and at night	20
Household sites and waste collection sites	2

3. Main Requirements for Public Lighting Systems

When developing solar street lighting, it is possible to achieve maximum efficiency of the system only using, so-called "smart" technical means, which would not only control operation modes, but also control the operation of the entire system in automatic mode. Additional maintenance of street lighting should be carried out in case of violation of the recommended lighting standards (lamps fail, stop shining or Shine dimly). The working hours of

lamps are regulated by municipal services, which approve the schedule for switching on and off lighting.

In cities and rural settlements, the requirements for lighting streets and roads vary significantly. Table 2 shows the standards for artificial lighting of roads and streets in cities and rural settlements. At the same time, the average brightness of a highway's coverage is 1.6 cd/m^2 in cities and at least 1.0 Cd/m^2 outside of cities at the main entrances to airports, river and sea ports, regardless of traffic intensity [8, 12].

The average horizontal illumination at the level of coverage of impassable parts of streets, roads and squares, boulevards and squares, pedestrian streets and territories of micro districts in urban settlements should be taken according to table 3.

At night, lighting level of city streets, roads and squares can be reduced to a standardized average brightness of more than 0.4 cd/m^2 or an average illumination of more than 4 lux by switching on no more than half of the lamps. It is allowed to reduce the lighting level with the regulator in order to obtain additional energy savings in the evening and morning hours of the day [3, 6]: by 30% with a decrease in traffic intensity to 1/3 of the maximum value; and by 50% with a decrease in traffic intensity to 1/5 of the maximum value.

On streets and roads at standardized values of an average brightness of 0.3 cd/m^2 or an average illumination of 4 lux or less, on pedestrian bridges, parking lots, pedestrian walkways and roads, internal, service and fire-fighting passages, as well as on rural streets and roads settlements, partial or complete switching off of lighting at night is not allowed. The following Table (Table 4) shows the variation of the illumination received (lux) on a public space according to the luminaire installation height and the power of the used Led (W) [13].

4. Automatic Control Systems of Public Lighting

Switching street lights on and off as needed is an excellent measure to save energy consumption. However, the various systems are not equally suited to this rapidly evolving technology. It is particularly interesting for public spaces used infrequently, such as public toilets, public waste collection points, and rural roads with little traffic. For areas with high safety sensitivity, it is not recommended (e.g. for busy roads, areas of heavy road or pedestrian traffic, and areas requiring overnight surveillance). On the other hand, for autonomous solar lighting systems, a regular and homogeneous consumption cycle is ideal. Uneven battery discharge due to erratic interruptions leads to rapid degradation and the need to replace them before the end of their normal lifespan. Therefore, measures to interrupt the nighttime lighting duration of such systems are only suitable to a limited extent. For solar systems, regular long-term interruptions are feasible to be managed through programmable astronomical clocks. Motion detectors are only recommended for small isolated spotlights. Automatic street lighting control systems usually operate under the control of a zone controller or server. Depending on the control algorithm, the controller generates a signal, for example, turning on a group of streetlights [10, 13, 14]. To transmit this signal to actuators, the following means are used: -low-current signal lines (twisted pair, RS-485, Ethernet, etc.); -radio channel; -GSM-channel; -signal transmission via power cable.

Table 4: Examples of illumination data for a lighting unit

Installation height of the luminaire	**Illuminated area**	**Maximum illumination (lux) received on a road**					
		LED 18 W	LED 30 W	LED 50 W	LED 65 W	LED 80 W	LED 120 W
6 m	8×20 m	10 Lux	19 Lux	30 Lux	39 Lux	49 Lux	59 Lux
8m	10×25 m	6 Lux	12 Lux	18 Lux	23 Lux	29 Lux	35 Lux
10m	13×30 m	4 Lux	8 Lux	12 Lux	15 Lux	19 Lux	24 Lux
12 m	16×35 m	2 Lux	5 Lux	8 Lux	11 Lux	14 Lux	17 Lux

Table 5: Comparison of control signal transmission methods

	1. Low Current Control	**2. GSM-Channel**	**3. Power Transmission Lines**	**4.Radio Channel**
Addressing (Economically Feasible)	possible Individual lamp controls	Group control only	Group control only	Group control only
Control Method	Digital control Protocol for example based on a calendar schedule	Phone call or SMS to the controller in the control Cabinet	Control via a power cable connected to the controller in the control Cabinet	Radio signal transmission from the control room to the receiver in the control cabinet
Factors Affecting the Reliability	Accumulation of timing error	Dependence on the workload of the public network of the GSM operator.	Upon failure of the required manual switching cable.	Radio interference may cause the control signal to be unable to be received
Labor Costs	High labor costs	Low labor costs through public use network	With individual lamp control, cable laying is labor-intensive	High labor costs when installing transceivers

Territory Coverage	Snap to city / region	Cellular coverage area	The length of the control power cable cannot exceed 1 km	Control is possible only in the area of confident reception of the radio signal.
Territory Size	City district, small settlement	City and nearest suburb	Limited territory	City and suburbs, area along the motorways
Cost Factors	Individual control unit in each lamp	Subscription and connection fees, messaging	The cost of individual power cables installation	Cost of dispatcher's equipment, relay stations and receivers
Factors Affecting Maintenance Cost	Timer adjustment is constantly needed		High repair costs for electrical equipment	A qualified dispatcher is required

Setting up the public lighting network is easier with a computerized centralized management system. Among other functions, this system normally allows: a) Easy and quick re-setting of the whole or parts of the network, for example during extraordinary events, which require a temporary adaptation of the lighting duration in a commune area. b) Accurate detection of faults and malfunctions in light points and electrical cabinets. c) Monitoring of energy consumption by the lighting network. d) Linking of network operation with climatic conditions and traffic volume. e) Management of equipment and human resources related to network maintenance.

Such a system consists of having controllers integrated into the electrical cabinets with a link to the light points or individual controllers integrated into the light points themselves. The controllers communicate with the remote management station (a computer, a tablet or a smartphone) via a wireless network, for example WIFI or radio. Note that for the control of the lighting time of streetlights, programmable time switches preferably of the so-called astronomical, digital or mechanical clock kind which are normally integrated in the control boxes as well as twilight detectors with a photoelectric cell (sensor) which are either integrated in the equipment (for example a regulator), or they are stand-alone units. Lighting controlled by a twilight switch automatically activates the switching off and switching on of public lighting according to daylight. To ensure the reliability and accuracy of the switch, it is essential to correctly position and orient the sensor so that it is not affected by random light sources or disturbed by shadows. It should also be checked and cleaned regularly to prevent dust accumulation. Twilight switches can be associated with clocks in order to cut off the control during part of the nighttime. The programmable time switches control the switching off and switching on of public lighting at specific times. They are installed in public lighting cabinets. If it is not an astronomical clock, it must be adjusted and updated regularly in order to approximate sunrise and sunset times [8, 10].

5. Design Features of the Developed System and its Operation Principle

The functional scheme of the solar-powered street lighting system developed is shown in Figure 3. It is an economical lighting system with intelligent control operating with an intelligent lamp having a step change of the luminous flux.

1. Solar panel; 2. Block of Sensor; 3. Lamp with Progressive Adjustment of the Luminous Flux; 4. Control Block; 5. Storage Battery; 6. Diagnostic Connector.

The photovoltaic solar panel (1) generates an electric current passing through the control block 4, which includes the charge controller of the storage battery. A "smart lamp" (3) is a lamp with stepwise adjustment of the light output power. Depending on its capacity, a battery to be charged by the solar panel ensures the operation of the entire system without recharging for a period corresponding to the duration of the night. The Block (2) consisting of temperature and light sensors sends accurate temperature and illumination data to the Control Unit 4, thereby controlling the operation mode of the lamp. The light sensors in block 2 will turn on the lamp at nightfall, and turn it off as soon as the first rays of the sun appear. Lamp operates in three different modes namely: maximum power mode, average power and economy mode depending on the time of year, the solar radiation intensity and the battery charge level. In addition, if the system notices a long lamp operation in Economy mode, it evaluates the battery charge level and can switch to timer mode in which the lamp turns on in Economy mode for a time T1, after which it turns off and turns on during time T2 before sunrise. Depending on the battery charge level, the T1 and T2 times may vary in the direction of decrease or increase. A computer is connected to the public lighting system via a serial interface according to the USB 2.0 standard from the diagnostic connector 6 to allow the use of software specially designed to control all the parameters of the system, program its operation, determine the dirt level from the solar panel, change the operating modes etc.

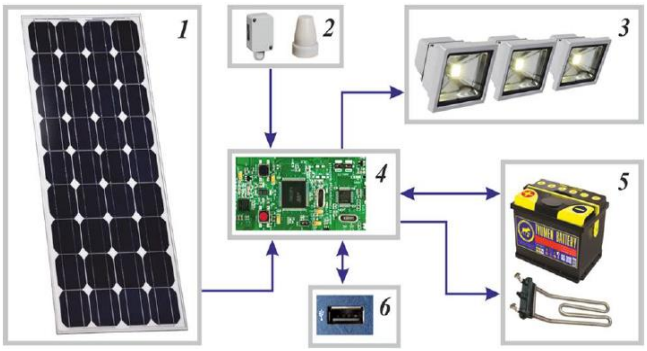


Figure 3: Block Diagram of the Developed Public Lighting

6. Proposed System Description and Operation

The developed street lighting system consists of smart lamps-lamp with a three-stage adjustment of the luminous flux: 10 W-20 W-30 W (the maximum power of each lamp is 30 watts). Each mini-SPP uses two solar panels of 120 W each to power six (6) smart lamps mounted 2 by 2 on three metal poles (As shown in Figure 4). The minimum total power of these lamps will be P_{min} =

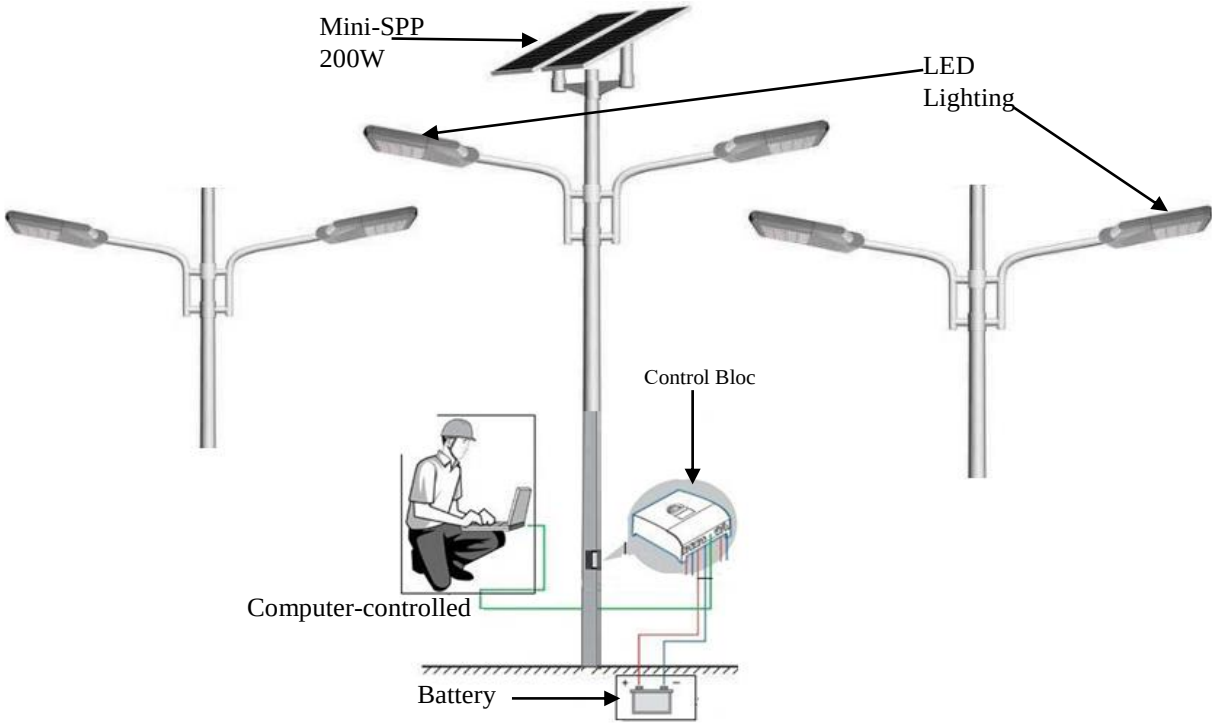


Figure 4: Developed Solar Mini-Power Plant 200 W

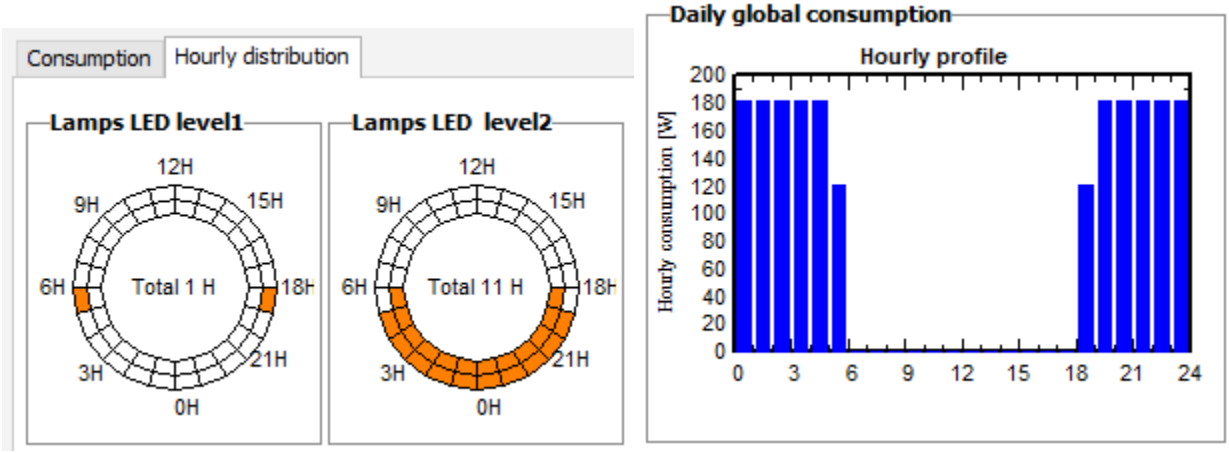


Figure 5: Lighting System hourly Distribution and Daily Consumption

60 W, with a maximum of P_{max} = 180 W. As Shown in the Figure 4, two of the lamps are installed on the central pole supporting the solar panels, the other 4 on the other 2 poles surrounding the first. A voltage of 24 V is transmitted from the battery of the mini solar power plant to the lamps via an underground power line. It should be noted that LED lamps are more advantageous than luminescence or incandescent lamps used in lighting systems in general. The following Table shows the relationship between the power of some lamps (W) and their luminous flux (lumen) [12].

The street lighting system shown in Figure 4 is designed to illuminate boulevards, public areas, etc. The distance between the supports on which the lamps are installed will be about 50 m. In order to reduce the cost of design and installation of the developed system, modernized conventional LED spotlights with systematic adjustment of the luminous flux are used. The following Diagram (As Shown in the Figure 5) shows the use state of this lighting over 24 hours (one day).

The diagram of operation and energy consumption of these lighting lamps shows that from 18: 00 (assumed time of the beginning of the night time),the lamps illuminate under a minimum power of 10 W for 30 minutes after which they illuminate at their maximum power of 30 W each.This minimum power which lasts 30 minutes will also be observed in the morning at the end of the night time so at 05h30 minutes.

7. Solar Panels Orientation and Positioning

In this street lighting system, the electrical energy source is usually located on the same support on which the luminaire is fixed. In this regard, special attention should be paid to the design of attachment of the solar panel to the support, since the amount

of electricity generated will also depend on its orientation. Most often, solar panels are installed at a certain angle (declination) to the sun, in the direction of azimuth to the South or South-East (As Shown in Figure 6) [12, 15].

Table 6: Power Ratio of Various Lamps Types

Incandescent lamp, W	Fluorescent lamp, W	LED lamp, W	Luminous flux, Lm
60	15-16	8-10	700
75	18-20	10-12	1200
100	25-30	12-15	1200
150	40-50	18-20	1800
200	60-80	25-30	2500

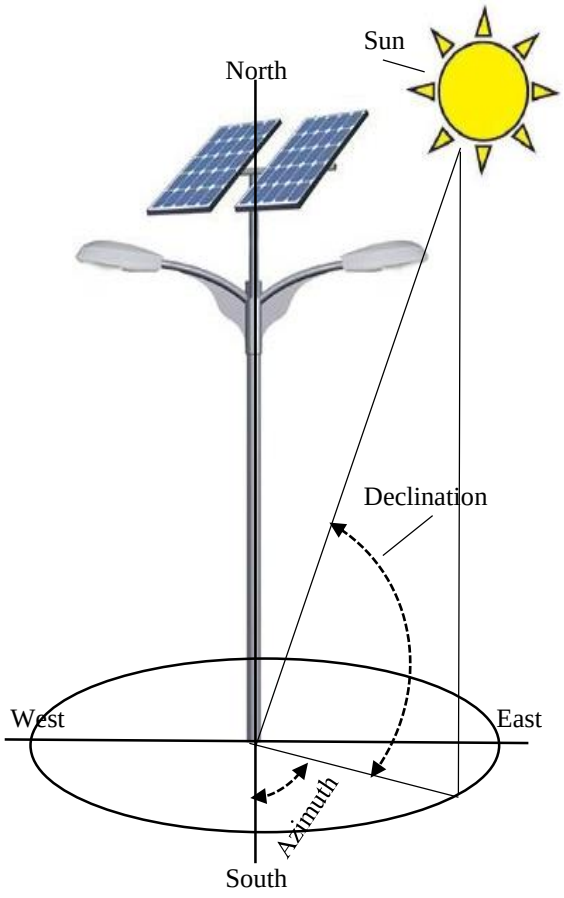


Figure 6: Solar Panel Orientation in Space

It is known that solar panels generate electricity in the absence of direct solar radiation, in cloudy weather, when the sunlight is weak and diffused, the photovoltaic system will produce electricity. The influence of various conditions on the production of solar panels as a percentage of the rated power is shown in Table 7 [12].

The tilt angle of solar panels relative to the earth's surface depends on geographical coordinates. Moreover, to increase the performance of the system, this angle must be changed during the year. As shown in Figure 7, for a fixed orientation throughout the year, the optimal tilt angle for the city of Gitega is 30 degrees.

Table 7: Dependence of Solar Panel Power on Different Conditions

Conditions	**Percentage of the Maximum Radiation (%)**
Bright sun	100
Light cloudy	60-80
Window glass, one layer	91
Overcast weather	20-30
Artificial light indoors	0,2-1,5

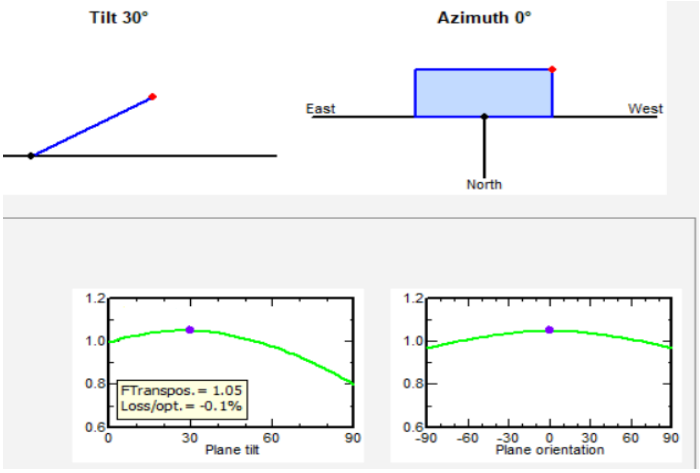


Figure 7: Tilt angle and Azimuth at Gitega

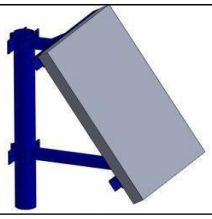
The tilt angle of a solar panel plays an important role, however small changes in the range of 0 to 5 degrees cannot have a serious effect on the system performance. Therefore, the tilt angle for localities (communes) of this region such as Bugendana, Bukirasazi, Buraza, Giheta, Gishubi, Itaba, Makebuko, Mutaho, Nyanrusange and Ryansoro will also be the same.

8. Solar Panels Fixings to Their Supports

Ideally, the luminaire support is fixed to a concrete footing through a metal plate to facilitate its replacement, if necessary. Depending on the function, the aesthetic appearance, the space and the environment, several types of supports can be identified, for example, poles made of steel, aluminum, cast iron, concrete, wooden etc. [12].

At present, metal supports are the most often used. Their manufacturing is more technologically advanced; they have a relatively small mass and retain their operating characteristics for a sufficiently long time. As the metal support is hollow inside, it is possible to run cables through it, connecting the solar panel to the control block and to the storage battery. In general, four (4) fixing points are used to fix a solar panel on a metal support. It is also taken into account that the panel should be oriented vertically, not horizontally. Table 8 shows some standard methods of solar panels fixing.

Table 8: Solar Panels Fixing

Denomination	Solar Module Width	Landing Dimensions	Image
MSU51585 Universal fastening of the solar module 50-150 W to the pipe up to 85 mm	up to 700 mm	Pipe Ø45-85 mm or on the wall	
MSU515220 Universal fastening of solar modules 50150 W to the pipe up to 220mm	up to 700 mm	Pipe Ø80-220 mm	
MSU153085 Universal fastening of solar modules 150-300W on a pipe up to 85mm	up to 1000 mm	Pipe Ø45-85 mm or on the wall	
MSU1530240 Universal fastening of the solar module 150-300 W on the pipe up to 220mm	up to 1000 mm	Pipe Ø80-240 mm	
MSU2x1530280 Universal fastening of two solar modules 150-300W on a pipe (wall) 50-280mm	up to 1000 mm	Pipe Ø 50-280 mm or on the wall	

Note that it is necessary to use a design that facilitates the changeability of the tilt angle of the solar panels according to the time of day or year in order to increase the efficiency of the mini-solar power plant for public lighting.

9. Solar Panels Power Calculation

The energy generated by a solar panel with a power Pw for a selected period is calculated as follows [16, 17]:

$$W = \frac{k \cdot E \cdot P_w}{1000}$$

E – insolation value for the selected period; Pw – solar panel power, k - coefficient equal to 0.5 and 0.7 in summer and winter periods, respectively. This factor takes into account the solar panels heating by the sun, as well as the incidence angle during the day. The insolation value E = 4.91 kWh/day.

The solar panel power to be used is calculated taking into account the daily consumption of all the LED lamps operating during the night hours. Remembering that they are of the adjustable power type and therefore their minimum power $P_{\Sigma lamp}$ will be 60watts, their maximum 180 W. Their maximum daily energy consumption $W_{cons.}$ is given by (1) as follow:

$$W_{cons.} = P_{\Sigma lamp} \cdot t_{\max} \quad (1)$$

где $t_{\max}$ – the operating time of the lamps

The latitude of the city of Gitega and the sun declination on which the average longitude of the day depends, give an interval of the day equal to 12 hours. Accordingly, the maximum daily energy consumption of the system will be 2040 Wh/day according to (1). Since the average number of hours of sunshine is estimated at 9 hours (9h) per day, to determine the daily energy production (generation) of solar panels, it is necessary to multiply their rated power P_n by their operating time per day: $W_{gen} = P_n \cdot t_{op.}$

Where $t_{op.}$ – Operating time (9h), W_{gen} – Daily energy generation. This daily energy produced by the solar panels will be 2016 Wh/day. The 120 W 24 V solar panels that will be chosen from the ranges produced by different manufacturers must have an optimal operating current of at least 6.7 A.

10. Battery Capacity Assessment and Calculation

The developed public lighting systems are intended to work during the night hours, while during this time the solar panels do not generate the electric energy, the electricity stored in the batteries is used. The capacity of the battery to be used, i.e. the amount of energy it can store, plays a big role. Their total capacity depends on the number of storage batteries, and therefore the power of the solar power plant in general. Storage batteries in a solar power plant must meet a number of requirements. They must withstand a large number of charge-discharge cycles. When calculating the parameters of the storage battery, it is necessary to take into account the energy losses during its storage and transformation. Conventional car batteries do not withstand a large number of cycles and have significant self-discharge. Most often, special storage batteries are used for solar power plants. Table 9 shows the states of charge (Battery charge levels) of a 12-volt battery depending on the temperature [12, 18].

Table 9: 12V- Battery Charge Status Assessment

Battery charge level, %	Electromotive force (EMF) , depending on temperature		
	+20…+25°C	+5…–5°C	–10…–15°C
100	12,70 – 2,120	12,80 – 13,00	12,9 – 13,10
75	12,55 – 12,65	12,65 – 12,75	12,75– 12,85
50	12,20 – 12,30	12,30 – 12,40	12,40 – 12,50
25	12,00 – 12,10	12,10 – 12,20	12,20 – 12,30
0	11,70 – 12,00	11,80 – 12,00	11,120 – 12,10

The energy generated and stored in the battery during the day is used at night to power the load (streetlights) of public lighting. This energy shows the capacity of the battery; i.e. how long it can power the load if it is fully charged. The change in battery capacity ΔC during the load supply time $\Delta t_{nig.}$ is determined by the expression (2):

$$\Delta C = \frac{P_n}{U_n} \cdot \Delta t_{nig.} = \frac{P_n}{U_n} \cdot (24 - \Delta t_{dayl.}) \qquad (2)$$

где P_n – rated load power; U_n – rated load voltage; $\Delta t_{nig.}$ – the night time interval (12hours); $\Delta t_{dayl.}$ – day time interval

Battery manufacturers define the final discharge voltage, once reached, the battery must be disconnected from the load and charged, this because deep discharge can damage it. A battery should not be more than 70-80% discharged for it to work for a long time. The battery discharge level is determined by the expression [16]:

$$S_r = \frac{C_n - C_{\min}}{C_n} \cdot 100\% = \frac{\Delta C}{C_n} \cdot 100\% \qquad (3)$$

The required capacity C_n of the accumulator battery is obtained from (3) while taking into account (2):

$$C_n = \frac{100}{S_r} \cdot \frac{P_n}{U_n} \cdot \Delta t_{nig.} \qquad (4)$$

In the city of Gitega, taking into account that the nighttime interval is 12hours, with the 200-Watt solar panels as it was said above, a 120 Ah will be ideal for this mini solar power plant consisting of six LEDs of street lighting. Its energy capacity will be 2880 Wh. The higher the voltage supplied to the load, the lower the capacity can be, since the discharge current $I_r = \frac{P_n}{U_n}$ of the storage battery will be lower. For conventional lead acid batteries, the maximum discharge current is limited to a value that, in amperes, is between 5 and 25 times the battery capacities. The lower the discharge current, the lower the power loss, and therefore the higher the efficiency of the system. It is the increase in voltage and decrease in current that explains the use of voltage 24 V instead of 12 V, when placing "smart lights" at a significant distance from the solar mini-power plant.

11. Conclusions

The task of our research was to develop and study the public lighting system of the city of Gitega using mini solar power plants as a source of renewable energy. This project was designed taking into account the current state of the lighting in Gitega, the current global trends, as well as the great importance of public lighting in general and through photovoltaic solar power plants in particular such as: increasing the citizens' safety, reducing the number of crimes and offenses and increasing their detection potential, reducing cabling costs for transmission systems, reducing the burden on emergency services and electrical technicians. This public lighting system with an automated solar power plant will regulate energy consumption, monitor the integrity of the equipment and quickly send a signal to operating personnel on emergencies on the network. The current state of public lighting in the city of Gitega and its surroundings is poor due to the lack of financial resources needed using traditional energy sources. The development of such solar street lighting systems that would not require large investment costs, promote clean energy, reduce energy consumption and which maintenance would be carried out once or twice a year, is very relevant.

Conflict of Interest

The authors declare no conflict of interest regarding the publication of this paper.

References

[1] N. Ntawuhorakomeye, M.P. Belov, "Development and Analysis of Public Lighting by Mini Solar Power Plants in The Republic of BURUNDI," in 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), IEEE, St. Petersburg, Russia: 785–790, 2020, doi:10.1109/EIConRus49466.2020.9039030.

[2] Ministry of Energy and Mines of the REPUBLIC OF BURUNDI, Opportunités dans le secteur des énergies renouvelables au Burundi Ministère de l'Energie et des Mines REPUBLIQUE DU BURUNDI, Bujumbura,Burundi, 2012.

[3] D. V. Shatilov, Проект «Программы внедрения автономных систем энергоснабжения уличного освещения на основе использования экономичных светодиодных светильников и солнечных батарей для стимулирования применения методов энергосбережения в жилищно-коммунальном хозяйстве, Svetorezerv LLC, Moscow, 2008.

[4] C. Muchunku, K. Ulsrud, D. Palit, W. Jonker-Klunne, "Diffusion of solar PV in East Africa: What can be learned from private sector delivery models?," WIREs Energy and Environment, **7**(3), 2018, doi:10.1002/wene.282.

[5] I.A. Suvorov, Электротехнологические промышленные установки и освещение (2009) — книга автора Суворова И.А. | НЭБ, Kirov, 2009.

[6] R. Ciriminna, F. Meneguzzo, L. Albanese, M. Pagliaro, "Solar street lighting: a key technology en route to sustainability," Wiley Interdisciplinary Reviews: Energy and Environment, **6**(2), e218, 2017, doi:10.1002/wene.218.

[7] V.K. Satyakar Veeramallu, S. Porpandiselvi, B.L. Narasimharaju, "Analysis and implementation of soft-switched bidirectional buck-boost DC-DC converter for solar PV-fed LED street lighting systems," International Journal of Circuit Theory and Applications, **47**(12), 1990–2018, 2019, doi:10.1002/cta.2706.

[8] G.P. Kolesnik, Электрическое освещение Основы проектирования Учебное пособие.pdf, Vladimir, 2006.

[9] Metode penelitian Nursalam, 2016, A.. Fallis, "GUIDE TECHNIQUE MODULE 3 : ECLAIRAGE PUBLIC," Journal of Chemical Information and Modeling, **53**(9), 1689–1699, 2013.

[10] N. Khatavkar, A.A. Naik, B. Kadam, "Energy efficient street light controller for smart cities," in 2017 International Conference on Microelectronic Devices, Circuits and Systems, ICMDCS 2017, Institute of Electrical and Electronics Engineers Inc.: 1–6, 2017, doi:10.1109/ICMDCS.2017.8211714.

[11] R.C. Marques, R.R. Geddes, "THE USE OF PPP ARRANGEMENTS IN STREET LIGHTING: A WIN-WIN OPTION?," in Annals of Public and Cooperative Economics, Wiley-Blackwell Publishing Ltd: 311–327, 2019, doi:10.1111/apce.12229.

[12] A.S. Stolkov, Разработка эффективной системы уличного освещения с питанием от солнечной мини - электростанции, Altai State Technical University, 2018.

[13] Р.Ф.Д.А. Медведева, Автоматизированные системы управления наружным освещением. Опыт Москвы - ЭнергоСовет.ru, 2011.

[14] M. Wadi, A. Shobole, M.R. Tur, M. Baysal, "Smart hybrid wind-solar street lighting system fuzzy based approach: Case study Istanbul-Turkey," in 2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), IEEE: 71–75, 2018, doi:10.1109/SGCF.2018.8408945.

[15] E. Alsakkaf, "An Automatic Off-Grid Solar Street Lighting System," IJIREEICE, **4**(3), 1–3, 2016, doi:10.17148/IJIREEICE.2016.4301.

[16] M.P.Belov. N. Ntawuhorakomeye, "TECHNICAL METHOD OF CALCULATING POWER OF SOLAR POWER PLANTS IN TROPICAL CLIMATES AS IN THE REPUBLIC OF BURUNDI," Izvestiya SPbGETU "LETI," **8**, 66–74, 2019.

[17] N. Ntawuhorakomeye, M.P. Belov, "Replacement of Liquid Fuel by Solar Electrical Installation for the Production of Electric Energy by Diesel-Generators of the «REGIDESO» Industrial Group in the Republic of Burundi," in 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), IEEE: 1021–1024, 2019, doi:10.1109/EIConRus.2019.8656909.

[18] M.D. Vijay, K. Shah, G. Bhuvaneswari, B. Singh, "LED based street lighting with automatic intensity control using solar PV," in 2015 IEEE IAS Joint Industrial and Commercial Power Systems / Petroleum and Chemical Industry Conference (ICPSPCIC), IEEE: 197–202, 2015, doi:10.1109/CICPS.2015.7974074.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1360-1368 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Towards a Hybrid Probabilistic Timing Analysis

Haoxuan Li[*,1], Ken Vanherpen[2], Peter Hellinckx[3], Siegfried Mercelis[3], Paul De Meulenaere[2]

[1]*Faculty of Applied Engineering, CoSys-Lab, University of Antwerp, Antwerp, 2000, Belgium*

[2]*Faculty of Applied Engineering, CoSys-Lab, University of Antwerp - Flanders Make, Antwerp, 2000, Belgium*

[3]*Faculty of Applied Engineering, IDLab, University of Antwerp - imec, Antwerp, 2000, Belgium*

ARTICLE INFO

Article history:
Received: 21 December, 2020
Accepted: 09 February, 2021
Online: 28 February, 2021

Keywords:
Hybrid Timing Analysis
Worst-Case Execution Time
Probabilistic Timing Analysis
Real-Time Systems
Hybrid Probabilistic Timing Analysis

ABSTRACT

Real-time embedded systems are widely adopted in applications such as automotive, avionics, and medical care. As some of these systems have to provide a guaranteed worst-case execution time to satisfy the time constraints, understanding the timing behaviour of such systems is of the utmost importance regarding the reliability and the safety of these systems. In the past years, various timing analysis techniques have been developed. Probabilistic timing analysis has recently emerged as a viable alternative to state-of-the-art deterministic timing analysis techniques. Since a certain degree of deadline miss is still tolerable for some systems, instead of deriving an estimated worst-case execution time that is presented as a deterministic value, probabilistic timing analysis considers execution times as random variables and associates each possible execution time with a probability of occurrence. However, in order to apply probabilistic timing analysis, the measured execution times must be independent and identically distributed. In the particular case of hybrid timing analysis, since the input and the initial processor state of one software component are influenced by the preceding components, it is difficult to meet such prerequisite. In this article, we propose a hybrid probabilistic timing analysis method that is able to (i) reduce the dependence in the measured execution times to facilitate the application of extreme value theory and (ii) reduce the dependence between software components to make it possible to use convolution to calculate the probabilistic WCET of the overall system.

1 Introduction

Due to the demand for better performance and the increased number of functionalities in embedded systems, acceleration features, such as multi-cores and caches, are added to the modern processors [1]. However, because of these acceleration features, timing analysis has become more complicated and time-consuming.

Over the past years, various researches have been carried out to develop timing analysis techniques that can derive a sound and safe result within a reasonable cost. With "sound", it is meant that a timing analysis result is accurate and not overly optimistic concerning its timing behaviour. The word "cost" describes the time and effort spent on deriving a timing analysis result. In the plethora of techniques, probabilistic timing analysis (PTA) has recently emerged as a viable alternative to state-of-the-art deterministic timing analysis techniques.

In real-time systems, failing to meet the deadline does not necessarily imply a failure of the system. For instance, in video conferences, an occasional deadline miss of the arrival of one frame can degrade the service but is not a failure of the system. Therefore, the timing requirement does not always have to be strictly guaranteed, and a certain degree of deadline misses is still tolerable. Therefore, Instead of deriving an estimated worst-case execution time (WCET) that is presented as a deterministic value, PTA considers execution times as random variables and associates each possible execution time with a probability of occurrence.

Figure 1 is the examples of the probabilistic WCET (pWCET) and the deterministic WCET, as well as the terminologies we use through this paper. The authors suggested that since the upper timing bound derived using the probabilistic worst-case execution time (pWCET) can be adjusted according to the safety requirement of the system, probabilistic timing analysis (PTA) is supposed to be more flexible and potentially less pessimistic compared with the classical deterministic approaches [2]. Therefore, we developed our approach based on the work in [3, 4] with measurement-based probabilistic timing analysis (MBPTA).

[*]Corresponding Author: Haoxuan Li, Groenenborgerlaan 171, 2000 Antwerp, Belgium, +32465934733 & Haoxuan.Li@uantwerpen.be

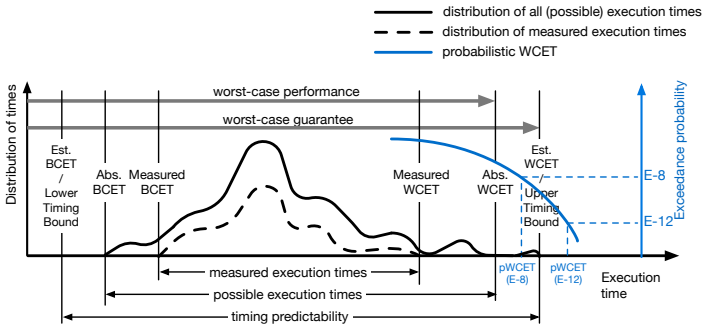


Figure 1: Examples of pWCET analysis and deterministic WCET timing analysis (based on [5]).

One prerequisite of applying MBPTA is that the measured execution times must be independent and identically distributed (i.i.d.). However, in hybrid timing analysis, the input and the initial processor state of one code block/component are influenced by the preceding code/component. Therefore, it is very challenging to collect execution times following the independent and identical distribution.

In this article, we present a hybrid probabilistic timing analysis (HPTA) approach that can facilitate in conducting MBPTA on software blocks/components and safely construct the overall execution of the complete program by reducing the dependence in the measured execution times and the dependence between the components. The goal of this approach is to provide more flexibility in the timing analysis results, which can be used by experienced engineers as references when designing non safety-critical systems that can tolerate a certain degree of deadline misses.

The rest of this paper is organised as follows. Section 2 gives the background of PTA. Section 3 explains the details of the extreme value theory (EVT) and how to apply it in MBPTA and HPTA. The proposed HPTA method is presented in Section 4. Section 5 demonstrates the experiment setup and the case studies, followed by the results and discussion in Section 6.

2 Background

The acceleration features of the hardware, the different input values, and the paths taken through the software introduce a certain degree of variability to the execution time of a system [6]. Because of such variability, PTA considers each observed execution time as an event with a probability of occurrence. Thus the worst-case execution time (WCET) can be viewed as an event that is more extreme than any previously observed event [7].

PTA derives a cumulative distribution function (CDF) using the observed execution times in the measurement. Since PTA aims to provide the probability that one execution time is going to exceed a given time bound, the result is usually presented as a complementary cumulative distribution function (CCDF) [8, 9, 10].

Figure 2 is an example of the results in measurement-based probabilistic timing analysis (MBPTA). The EVT takes the measured execution times to generate the CCDF. This CCDF-function then represents the probability that an event will have an execution time larger than a given pWCET value. This way, every pWCET is associated with such an exceedance probability. For example, in the given figure, the pWCET shows that for every execution of the program, the probability of the execution time exceeding 9.6ms is 1E-16.

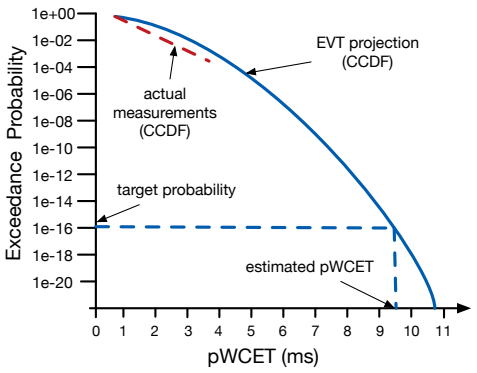


Figure 2: Example of the analysis result of one MBPTA shown as CCDF [11].

For instance, to calculate the pWCET of a task with the probability that the execution time will be 1 time unit is 0.15, the probability that the execution time will be 2 is 0.4, the probability that the execution time will be 4 is 0.4 and the probability that the execution time will be 8 is 0.05. The discrete probability distribution function of execution times of this task can be expressed in the form of a probability mass function (PMF), which is also known as execution time profile (ETP) in the context of PTA:

$$X = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0.15 & 0.4 & 0.4 & 0.05 \end{pmatrix} \tag{1}$$

The CDF, which indicates the probability that one execution time is less than x, is:

$$CDF_X(x) = F_X(x) = \mathrm{P}(X \leq x) = \begin{cases} 0 & if \quad x \in [0,1) \\ 0.15 & if \quad x \in [1,2) \\ 0.55 & if \quad x \in [2,4) \\ 0.95 & if \quad x \in [4,8) \\ 1 & otherwise \end{cases} \tag{2}$$

Eventually, the CCDF is generated to represent the analysis results because pWCET analysis rather focuses on the probability that an execution time exceeds a particular level:

$$\begin{aligned} CCDF_X(x) &= \bar{F}_X(x) = \mathrm{P}(X > x) \\ &= 1 - F_X(x) = \begin{cases} 1 & if \quad x \in [0,1) \\ 0.85 & if \quad x \in [1,2) \\ 0.45 & if \quad x \in [2,4) \\ 0.05 & if \quad x \in [4,8) \\ 0 & otherwise \end{cases} \end{aligned} \tag{3}$$

3 Extreme Value Theory

EVT is a branch of statistics that is used to assess the probability of the occurrence of rare events based on previous observations

[8]. It has been widely applied in many disciplines such as finance, earth sciences, traffic prediction, and geological engineering. In MBPTA, EVT is used to analyse the tail behaviour of the execution time distribution based on the execution times observed during the measurement.

3.1 Independent and identical distribution

The prerequisite of EVT requires the observations to be independent and identically. *Independent* means that the occurrence of one event does not impact the occurrence of the other event. *Identical* indicates that each measured execution time is random and has the same probability distribution as the others.

In the context of timing analysis, i.i.d. implies that the measured execution time is not affected by the past events of the same task. For example, task A is executed twice consecutively on a processor with a first-in-first-out (FIFO) cache. The measured execution time is x for the first time and y for the second time. Because the cache is loaded with data and instructions needed after the first time the task is executed, y is expected to be much smaller compared with x. In this case, the execution time x and y are said to be dependent, which does not meet the i.i.d. requirement.

Previous research has proposed the adoption of randomised architectures for measurement-based probabilistic timing analysis[12, 13]. If the processor of the target system has a deterministic architecture, randomisation needs to be applied to randomise the measurement samples to satisfy the i.i.d. precondition [9, 14, 15].

If two events are independent, the joint probability is equal to the product of their individual probabilities:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B) \tag{4}$$

For example, given a a task A with probabilistic execution time:

$$\begin{pmatrix} 5 & 10 \\ 0.1 & 0.9 \end{pmatrix} \tag{5}$$

and task B with probabilistic execution time:

$$\begin{pmatrix} 1 & 2 \\ 0.4 & 0.6 \end{pmatrix} \tag{6}$$

the probabilistic execution time of the both tasks is:

$$\begin{pmatrix} 5 & 10 \\ 0.1 & 0.9 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 0.4 & 0.6 \end{pmatrix} = \begin{pmatrix} 6 & 7 & 11 & 12 \\ 0.04 & 0.06 & 0.36 & 0.54 \end{pmatrix} \tag{7}$$

The CCDF is therefore:

$$\bar{F}_{A\cap B}(x) = \begin{cases} 1 & if \quad x \in [0,6) \\ 0.96 & if \quad x \in [6,7) \\ 0.90 & if \quad x \in [7,11) \\ 0.54 & if \quad x \in [11,12) \\ 0 & otherwise \end{cases} \tag{8}$$

3.2 Applying extreme value theory in measurement-based probabilistic timing analysis

Algorithm 1 demonstrates how to apply EVT in MBPTA, which includes three main steps: (1) extreme value selection; (2) fitting the distribution; (3) compare and convergence.

Algorithm 1: How to apply EVT in MBPTA

```
Result: pWCET
select extreme values;
fit the first distribution;
estimate the first CDF;
while first round or not converged do
    collect more extreme values;
    fit a new distribution;
    estimate the new CDF;
    compare;
end
```

3.2.1 Extreme value selection

EVT makes assumptions on the tail of the data distribution based on the extreme observations obtained from the measured data. Therefore, the measured execution times of the program must be converted into a set of extremes before applying the EVT. Two approaches can be used to collect the observed extreme execution times: block maxima (BM) and peak over threshold (POT).

The BM approach divides the collected execution times into subgroups/blocks. The longest execution time in every subgroup is one sample of extremes. The POT approach defines a threshold. All the execution times longer than the threshold are used as the extremes to estimate the pWCET.

The selection of a suitable size of the subgroups in the BM or a suitable threshold value in the POT plays a crucial role. Studies have been carried out to determine the impact of these decisions on the estimated results [2, 16, 17]. Neither an optimal size of the subgroups nor the best threshold value exists. Generally speaking, a smaller subgroup size or a lower threshold value can decrease the pessimism in the results. However the risk of obtaining an overly optimistic pWCET that is not safe to use increases.

3.2.2 Fitting the distribution

In this step, a model distribution that best fits the extremes is estimated. If BM was used in the previous step, the distribution must be Gumbel, Fréchet, or Weibull distribution. These three distributions are combined into a family of continuous probability distributions known as the generalised extreme value (GEV) distribution.

The GEV is characterised by three parameters: the location parameter $\mu \in \mathbb{R}$, the scale parameter $\sigma > 0$, and the shape parameter $\xi \in \mathbb{R}$. The value of the shape parameter ξ decides the tail behaviour of the distribution and to which sub-family the distribution belongs: Gumbel ($\xi = 0$), Fréchet ($\xi > 0$) or Weibull ($\xi < 0$). The CDFs of these three distributions are given below.

- Gumbel or type I extreme value distribution ($\xi = 0$):

  $F_{(\mu,\sigma,0)}(x) = e^{-e^{-(x-\mu)/\sigma}}$

- Fréchet or type II extreme value distribution ($\xi > 0$):

$$F_{(\mu,\sigma,\xi)}(x) = \begin{cases} e^{-y^{-\alpha}} & y > 0 \\ 0 & y \leq 0 \end{cases}, (\xi = \alpha^{-1} \text{ and } y = 1 + \xi(x-\mu)/\sigma)$$

- Reversed Weibull or type III extreme value distribution ($\xi < 0$):

$$F_{(\mu,\sigma,\xi)}(x) = \begin{cases} e^{-(-y)^{\alpha}} & y < 0 \\ 1 & y \geq 0 \end{cases}, (\xi = -\alpha^{-1} \text{ and } y = -(1 + \xi(x - \mu)/\sigma)).$$

If POT was used in the previous step, the distribution needs to fit the generalised Pareto distribution (GPD), which is a family of continuous probability distributions. The GPD is specified by the same three parameters as GEV: location parameter μ , scale parameter σ , and shape parameter ξ. The CDF of the GPD is:

$$F_{(\mu,\sigma,\xi)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0 \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{for } \xi = 0 \end{cases}, (x \geqslant \mu \text{ when } \xi \geqslant 0, \text{ and } \mu \leqslant x \leqslant \mu - \sigma/\xi \text{ when } \xi < 0).$$

3.2.3 *Compare and convergence*

Once the first data set of extremities is collected, the EVT uses the data to estimate the first CDF of the pWCET. In every subsequent round of the estimation, more extremities are collected and added to the data set to derive a new CDF. The CDF in the current round is compared with the CDF of the previous round using metrics such as the continuous ranked probability score (CRPS) [18]. In our case, we use the following convergence criteria to compare the respective accuracy of the two CDFs:

$$Score = \sum_{x=0}^{+\infty} [F_{current}(x) - F_{previous}(x)]^2 \quad (9)$$

Where F is the cumulative distribution. The score indicates the level of difference between the two cumulative distributions. The lower the score is, the closer the CDF is going to reach the point of convergence. Once the score reaches the desired value (such as 0.1), the pWCET can be derived using the converged CDF.

4 The proposed hybrid probabilistic timing analysis

Figure 3 is the illustration of the proposed HPTA approach. In our research, a system can be divided at different decomposition levels according to the requirement of the user. Because of the block isolation technique described in [3, 4], the software components can be executed independently from the rest of the program, and the processor state can be adjusted before every execution. This allows us to overcome two major challenges in applying the HPTA approach: (1) the i.i.d. data requirement of the EVT; (2) the calculation of the upper bound of the overall system.

In the section, we will describe the concept of decomposition level, and explain the challenges and how we solved these challenges using the proposed HPTA.

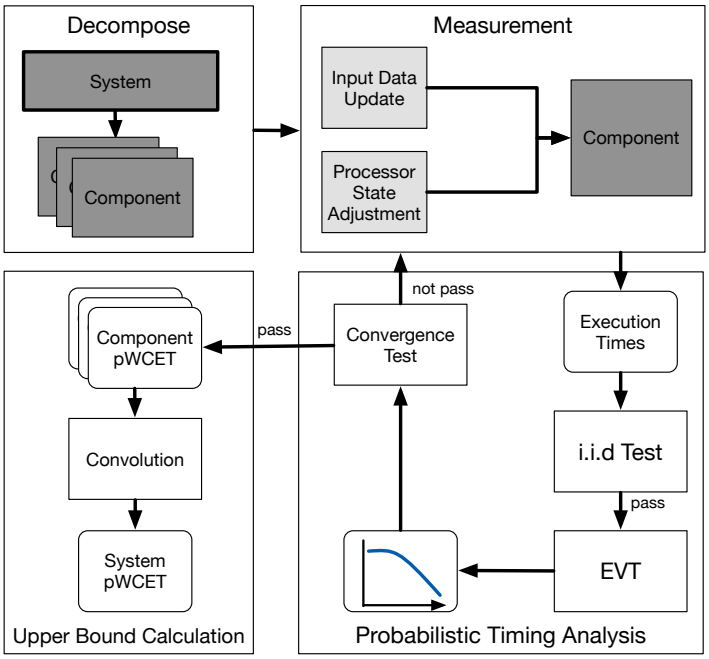
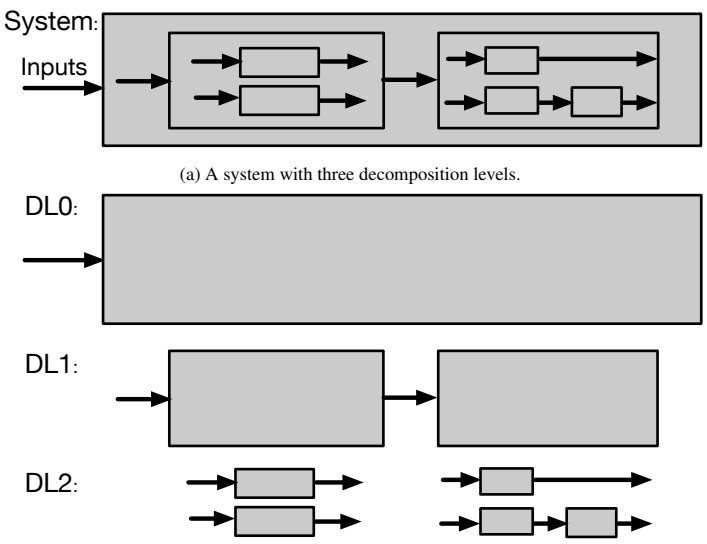


Figure 3: The proposed HPTA approach.



(a) A system with three decomposition levels.

(b) The components of the system at *DL*0, *DL*1 and *DL*2. DL: decomposition level.

Figure 4: A system and the decomposition of the system at three decomposition levels.

4.1 *Decomposition level*

Figure 4a shows a system with components at three decomposition levels. The decomposition of the system is demonstrated in Figure 4b. Every grey rectangle represents one component, and the solid black arrows represent the data flow of the components. By our definition, *DL*0 is the lowest degree of decomposition because the complete system is measured at this level. The complete system is considered as one abstract component from which no details of the subsystems can be retrieved. At *DL*1, the mask of the system at *DL*0 is opened. Measurement can be conducted on the subsystems that are revealed at this level. As the decomposition level increases,

the system is decomposed in further detail, more components are revealed. The decomposition level can increase until the last subsystem is opened. The granularity of the timing analysis increases with the decomposition level.

4.2 Challenges

4.2.1 Independent and identical data requirement

Although EVT requires the samples to be identically distributed, the adoption of GEV opened up the applicability of MBPTA to more cases because it can tolerate up to a certain degree of dependence in the measured execution times[10]. The disadvantage is that such dependence can corrupt the reliability of the derived pWCET [19]. In some domains, techniques such as re-sampling and de-clustering can be applied to create independence between measurements [20]. Nevertheless, forcing independence into dependent execution time measurements can result in overly optimistic pWCET that is not safe anymore [2].

For instance, the calculation in the example given in Section 3.1 is only valid when task A and B are independent, otherwise the result might be overly optimistic. Suppose dependence exists between the two events such that the execution time of task B is always 2 when the execution time of task A is 10. The probabilistic execution time when A and B are both executed would be:

$$\begin{pmatrix} 6 & 7 & 12 \\ 0.04 & 0.06 & 0.9 \end{pmatrix} \tag{10}$$

The probabilistic execution time of the both tasks in this case is:

$$\bar{F}_{A\cap B}(x) = \begin{cases} 1 & if \quad x \in [0,6) \\ 0.96 & if \quad x \in [6,7) \\ 0.90 & if \quad x \in [7,12) \\ 0 & otherwise \end{cases} \tag{11}$$

Given a deadline of 11, the probability of missing the deadline is 0.54 when A and B are independent (8), and 0.90 when A and B are dependent (11). In this case, overlooking the dependence between the tasks resulted in overly optimistic probabilities of a deadline miss.

4.2.2 Upper bound calculation

In deterministic hybrid timing analysis, the results are presented in absolute values. Therefore, the WCET of the complete system can be calculated simply by adding up the WCETs of every individual component. In statistics, convolution may be usually used to calculate the joint probability of two random variables. Unfortunately, this is not applicable in this context. The reason is that the probability distribution of the sum of two or more random variables is equal to the convolution of their distributions only when the variables are independent.

Since hybrid timing analysis is conducted on the components that construct the complete system, the execution time of one component often impacts the execution time of another component. This dependence makes the calculation of the overall pWCET much more complicated.

4.3 Solutions

Based on the information provided in the previous text, one can conclude that in order to overcome the challenges, we need to find a solution to the dependence in the measured execution times and the dependence between individual components.

4.3.1 Dependence in the observed execution times

In MBPTA and HPTA, dependence exists in the data samples if two or more of the observed execution times interfere with each other. Generally speaking, such dependence can come from:

1. *Software dependence*: The dependence between inputs and the paths taken by the program. For example, if input/path A is always followed by input/path B, A and B are dependent.
2. *Hardware dependence*: The hardware state left by the previous execution, which becomes the initial hardware state of the current execution.

To reduce the software dependence, we choose the *per-path* approach to collect the execution times for EVT. The other approach is know as the *per-program* approach [7]:

1. *Per-program*: After the measurement, the EVT uses all the measured execution times as one sample group. EVT is applied to the sample group to estimate the pWCET for the complete program.
2. *Per-path*: After the measurement, all the measured execution times are divided into sample groups. One group includes all the execution times measured from one program path. EVT is applied to each sample group to estimate the pWCET of the program path the sample group represents. The pWCET of the complete program is then calculated by taking the upper bound of the pWCETs of all the paths.

For instance, a system has two paths with the execution times at 10/15 (10 or 15 time units) and 60/65. The probabilities of the paths taken by one randomly chosen single execution are 0.3 and 0.7 respectively. Suppose the different execution times observed in the same path are due to a hardware component, which may generate 0 extra time units with a probability of 0.4 or 5 extra time units with a probability of 0.6. The probability execution time profile (ETP) of the system is:

$$\begin{pmatrix} 10 & 15 & 60 & 65 \\ 0.12 & 0.18 & 0.28 & 0.42 \end{pmatrix} \tag{12}$$

The pWCET derived using the per-program approach is:

$$pWCET_{per-program} = \begin{cases} 1 & if \quad x \in [0,10) \\ 0.88 & if \quad x \in [10,15) \\ 0.7 & if \quad x \in [15,60) \\ 0.42 & if \quad x \in [60,65) \\ 0 & otherwise \end{cases} = \begin{bmatrix} 0 & 10 & 15 & 60 & 65 \\ 1 & 0.88 & 0.7 & 0.42 & 0 \end{bmatrix} \tag{13}$$

The pWCET derived using the per-path program is:

$$pWCET_{per-path} = \begin{bmatrix} 0 & 60 & 65 \\ 1 & 0.6 & 0 \end{bmatrix} \quad (14)$$

As can be seen, compared with the per-path approach, the per-program approach was able to produce a tighter result. The pWCET derived using the per-program approach is valid to upper bound randomly chosen single executions of the target program. However, it is not applicable to be used in convolution when calculating the overall pWCET of multiple executions of the target program. The reason is that the probability distribution of the sum of two or more random variables is equal to the convolution of their distributions only when the variables are independent. For example, if path 60/65 is forced to be taken when path 10/15 is observed in the previous execution, the pWCET of two consecutive executions of the program is:

$$\begin{bmatrix} 0 & 70 & 75 & 80 & 120 & 125 & 130 \\ 1 & 0.952 & 0.808 & 0.7 & 0.588 & 0.252 & 0 \end{bmatrix} \quad (15)$$

The pWCET calculated using the probability derived from $pWCET_{per-program}$ is:

$$\begin{bmatrix} 0 & 70 & 75 & 80 & 120 & 125 & 130 \\ 1 & 0.9424 & 0.8428 & 0.6412 & 0.4116 & 0.1764 & 0 \end{bmatrix} \quad (16)$$

Under the dependence introduced by the execution condition, the probability that the two consecutive executions exceed 125 time units is 0.252. However, the pWCET using the convolution on the probability derived from $pWCET_{per-program}$ gives a probability of 0.1764, which is too optimistic.

In the per-path approach, the measured execution times are grouped according to the program paths. Therefore, the influence of the input data is eliminated within every group. The variations in the execution time within the same group are purely due to the hardware. By taking the upper bound of the pWCETs of all the paths, the pWCET generated by the per-path approach can be used safely in convolution for estimating any overall pWCET in which the execution of this task is executed consecutively. In the above example, the pWCET calculated by applying convolution to the probability derived from $pWCET_{per-path}$ is:

$$\begin{bmatrix} 0 & 120 & 125 & 130 \\ 1 & 0.84 & 0.36 & 0 \end{bmatrix} \quad (17)$$

Therefore, using the per-path approach, the proposed HPTA method can upper bound the $pWCET$ under the dependence introduced by the execution condition.

To reduce the hardware dependence, we try to push the hardware to the worst-case state at the beginning of every execution so that every execution can start with nearly the same initial state. For example, the pipeline is flushed, and "trash code" is executed to fill the caches with irrelevant data and instructions before every execution.

4.3.2 Dependence between components

In order to calculate the pWCET of the overall system using convolution, the independence between the pWCETs of the software components must be guaranteed. Such kind of dependence can be found in:

1. *Software dependence*: The output of the precedent components, which is the input of the current component.
2. *Hardware dependence*: The hardware state left by the previous component, which becomes the initial hardware state of the current component.

The software dependence can be removed together with the dependence in the observed execution times by the per-path approach. This is because the path-approach automatically selects the input(s) leading to the longest execution paths that upper bound all the other inputs. The hardware dependence no longer exists because the components execute in isolation from the rest of the program in the proposed method. Hence there are no precedent components.

5 Case study and experiment setup

5.1 Case Study

For the case study, we adopted a Simulink power window model. Figure 5 is the power window model [21, 22] at *DL*0, *DL*1 and *DL*2. A power window is an automatic window in a car that can be raised and lowered by pressing a button or a switch instead of a hand-turned crank handle. The complete application of a power window normally consists of four windows that can be operated individually. The three passenger-side windows can be operated by the driver with higher priority. In this paper, we only use the front passenger-side window and the driver-side window.

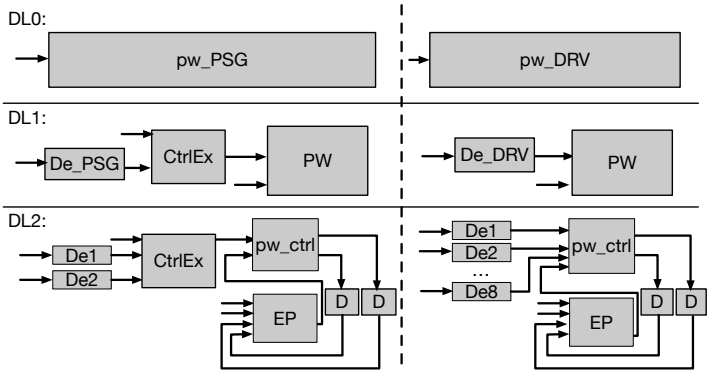


Figure 5: Power window passenger (left) and power window driver (right) at decomposition level 0, 1 and 2. DL: decomposition level; pw_PSG: powerwindow passenger, De_PSG: debounce passenger, CtrlEx: control exclusion, PW: powerwindow main component, De_DRV: debounce driver, De: debounce, pw_ctrl: powerwindow control, EP: end of range and pinch detection, D: delay.

The C code of the power window was generated using Embedded Coder® [23] targeting an ARM Cortex microprocessor with a discrete fixed-step solver at a fixed-step size of 5 ms. The BM approach with a block size of 20 was used in the extreme value selection process to collect the extremities for the EVT. The exceeding probability of 1E-10 was used for the pWCET.

The Kolmogorov-Smirnov (KS) test was used to verify the identical distribution requirement, and the runs-test was used to verify

Table 1: The p-values for the independent and identical distribution tests. NoN: number of nop instructions; pw_DRV: powerwindow driver; pw_PSG: powerwindow passenger; R: runs test; KS: Kolmogorov-Smirnov test.

NoN	pw_DRV		pw_PSG		NoN	pw_DRV		pw_PSG	
	R	KS	R	KS		R	KS	R	KS
0	0.7495	0.9999	1	0.4973	8	1	0.7710	0.8486	0.1349
1	0.4019	0.2753	1	0.4973	9	0.3894	0.9999	0.6822	0.2753
2	0.9764	0.9999	0.4736	0.1349	10	0.2703	0.9655	0.3894	0.7710
3	0.3441	0.9655	0.8769	0.4973	11	0.9080	0.9655	0.3337	0.2753
4	0.7939	0.9655	1	0.1349	12	0.1065	0.7710	0.8052	0.7710
5	0.8046	0.7710	0.2003	0.2753	13	1	0.9999	0.0882	0.7710
6	0.5634	0.9655	0.1789	0.4973	14	0.5643	0.9999	0.3634	0.9999
7	0.6744	0.9655	0.2176	0.2753	15	1	0.2753	0.9687	0.4973

Table 2: The measured deterministic WCETs and the estimated probabilistic WCETs The exceeding probability of 1E-10 was used for the pWCET. of the complete powerwindow driver benchmark and the complete powerwindow passenger benchmark with 16 different alignments. NoN: number of nop instructions; pw_DRV: powerwindow driver; pw_PSG: powerwindow passenger; cc: clock cycles; dWCET: measured deterministic WCET; pWCET: estimated probabilistic WCET.

NoN	pw_DRV (cc)		pw_PSG (cc)		NoN	pw_DRV (cc)		pw_PSG (cc)	
	dWCET	pWCET	dWCET	pWCET		dWCET	pWCET	dWCET	pWCET
0	2516	2538	1638	1666	8	2660	2797	1678	1953
1	2656	2712	1660	1688	9	2650	2698	1674	1681
2	2514	2515	1664	1717	10	2586	2608	1650	1681
3	2764	2849	1672	1725	11	2700	2762	1648	1683
4	2712	2768	1640	1699	12	2748	2813	1646	1679
5	2708	2762	1642	1700	13	2646	2746	1646	1684
6	2726	2794	1666	1764	14	2620	2735	1670	1701
7	2714	2797	1670	1689	15	2568	2569	1632	1658

the requirement for independence. $p > 0.05$ indicates the data follows the desired distribution.

The case study is designed to test the following assumptions:

1. Can the proposed HPTA method reduce the dependence in the measured execution times and facilitate the collection of i.i.d. execution times?

2. Can the proposed HPTA method reduce pessimism at a higher decomposition level?

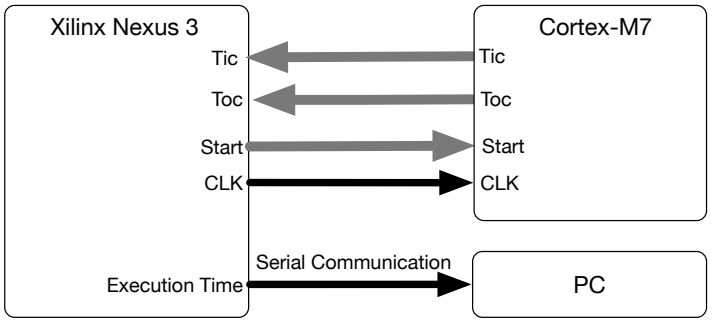


Figure 6: Experiment setup.

5.2 *Experiment Setup*

The experiment setup is shown in Figure 6. The processor is an ARM Cortex-M7 processor that contains a four-way set-associative data cache and a two-way set-associative instruction cache. Both caches use the random replacement policy. The Xilinx Nexus 3 FPGA is used to measure the execution time of the code on the ARM-processor. It is driven by the *CLK* signal sent from the processor to synchronise the clocks. The measurement process starts by sending the *Start* signal from the FPGA to the processor. Two GPIO pins of the processor are used to set the *Tic* and *Toc* flags. Before the code starts to execute, the *Tic* pin is toggled to start counting the number of clock cycles that have passed. After the code is executed, the *Toc* pin is set to stop the timer. Eventually, the value in the timer (*Execution Time*) is collected through a serial port.

6 Results and discussion

6.1 *Reducing dependence in the measured execution times*

Table 1 shows the results of KS and runs test. As can be seen, all the p-values are significantly higher than 0.05, which means that the i.i.d. requirement is fulfilled and the collected data can be applied in EVT to estimate the pWCET. The NoN indicates the number of the *nop* instructions added to emulate different memory alignments [3]. These p-values confirm our assumption that the proposed HPTA can reduce the dependence in the executions to facilitate the collection of i.i.d. data required by the EVT.

Table 3: The estimated pWCET calculated using sum and convolution at different decomposition levels. DL: decomposition level; pw_DRV: powerwindow driver; pw_PSG: powerwindow passenger; cc: clock cycles; dWCET: measured deterministic WCET; $pWCET_{Sum}$: estimated probabilistic WCET calculated by adding the pWCETs of the components up; $pWCET_{Conv}$: estimated probabilistic WCET derived from the convolution of the pWCETs of the components.

	pw_DRV (cc)			pw_PSG (cc)		
	dWCET	$pWCET_{Sum}$	$pWCET_{Conv}$	dWCET	$pWCET_{Sum}$	$pWCET_{Conv}$
DL0	2764	2849		1678	1953	
DL1	2652	3852	3465	1838	2949	2690
DL2	4820	6025	5167	2596	3778	3484

Table 2 shows the WCETs of the complete powerwindow-driver benchmark and the complete powerwindow-passenger benchmark with 16 different alignments. The longest measured deterministic WCET was 2764 reached by NoN = 3 in the driver-side window and 1678 reached by NoN = 8 in the passenger-side window. The estimated pWCET in the driver-side window was 2849 reached by NoN = 3 and 1953 reached by NoN = 8.

Across all the alignments, the estimated pWCET is more pessimistic at the selected exceeding probability compared with the deterministic WCET. However, in some cases, such as NoN = 15 in pw_DRV, the difference between the two WCETs is relatively small. The reason could be that the execution time contributed by independent random elements in the hardware is reduced by imposing a worst-case execution scenario on the longest path. Since random events in the hardware rarely happen, the measured execution times can be very close to the estimated pWCET.

6.2 *Reducing pessimism in high decomposition levels*

Table 3 shows the estimated pWCETs of the benchmarks calculated using sum and convolution at three decomposition levels. The highest pWCET of the 16 different alignments were used as the pWCET of every block. As can be seen, the WCETs increase with the decomposition levels. The highest WCET was reached by $pWCET_{Sum}$ at $DL2$ with 6025 clock cycles for the driver and 3778 clock cycles for the passenger. The difference between the deterministic WCET and pWCETs also became more prominent with the decomposition level, which means a higher decomposition level is associated with larger pessimism.

The pWCET calculated using convolution ($pWCET_{Conv}$) is in general lower compared with the $pWCET_{Sum}$. Nevertheless, $pWCET_{Conv}$ still leaves a reasonable margin with the deterministic WCET. This observation confirms our assumption that simply adding up pWCETs in HPTA will lead to overly pessimistic results. Using convolution can help to reduce such pessimism and produce a reasonably sound result.

7 Conclusion

In this article, we proposed a HPTA method that is able to (i) collect the i.i.d. data required by EVT by reducing the dependence between the measured execution times and (ii) reduce the dependence between blocks to make it possible to use convolution to calculate the pWCET of the overall system. We are aware that forcing independence into dependent execution time measurements can make the data no longer representative of the dependent system behaviour. Therefore, several approaches, such as using the per-path approach to upper bound the pWCETs of all the paths, were taken to ensure the derived pWCET is not overly optimistic. However, this might have introduced too much pessimism into the estimated pWCET.

Another novelty of the proposed HPTA method is that it can be applied at different granularities by adjusting the decomposition level. Performing the analysis at a higher decomposition level can help to reduce pessimism. On the other hand, with the increased decomposition level, the cost to derive the pWCET also increases.

Compared with the deterministic HTA approach proposed in [3], this HPTA approach is more pessimistic with the exceedance probability of 1E-10. On the other hand, instead of giving a deterministic value, this approach gives an exceedance probability to each possible execution time. We hope the results of this HPTA approach can be used as a reference to avoid unnecessary waste on the hardware cost in the hand of experienced engineers when designing a non safety-critical system that can tolerate a certain degree of deadline misses.

References

[1] L. Cucu-Grosjean, L. Santinelli, M. Houston, C. Lo, T. Vardanega, L. Kosmidis, J. Abella, E. Mezzetti, E. Quiñones, F. J. Cazorla, "Measurement-based probabilistic timing analysis for multi-path programs," Proceedings - Euromicro Conference on Real-Time Systems, 91–101, 2012, doi:10.1109/ECRTS.2012.31.

[2] F. Guet, L. Santinelli, J. Morio, "On the representativity of execution time measurements: Studying dependence and multi-mode tasks," in OpenAccess Series in Informatics, **57**, 31–313, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017, doi:10.4230/OASIcs.WCET.2017.3.

[3] H. Li, K. Vanherpen, P. Hellinckx, S. Mercelis, P. De Meulenaere, "Component-based Timing Analysis for Embedded Software Components in Cyber-Physical Systems," in 2020 9th Mediterranean Conference on Embedded Computing, MECO 2020, 1–8, IEEE, 2020, doi:10.1109/MECO49872.2020.9134177.

[4] H. Li, P. De Meulenaere, K. Vanherpen, S. Mercelis, P. Hellinckx, "A hybrid timing analysis method based on the isolation of software code block," Internet of Things, **11**, 100230, 2020, doi:10.1016/j.iot.2020.100230.

[5] R. Wilhelm, J. Engblom, A. Ermedahl, N. Holsti, S. Thesing, D. Whalley, G. Bernat, C. Ferdinand, R. Heckmann, T. Mitra, F. Mueller, I. Puaut, P. Puschner, J. Staschulat, P. Stenström, "The worst-case execution-time problem-overview of methods and survey of tools," Transactions on Embedded Computing Systems, **7**(3), 36:1–36:53, 2008, doi:10.1145/1347375.1347389.

[6] L. Santinelli, F. Guet, J. Morio, "Revising measurement-based probabilistic timing analysis," in Proceedings of the IEEE Real-Time and Embedded

Technology and Applications Symposium, RTAS, 199–208, IEEE, 2017, doi: 10.1109/RTAS.2017.16.

[7] R. Davis, L. Cucu-Grosjean, "A Survey of Probabilistic Schedulability Analysis Techniques for Real-Time Systems," LITES: Leibniz Transactions on Embedded Systems, **6**(1), 1–53, 2019.

[8] J. Hansen, S. Hissam, G. A. Moreno, "Statistical-based WCET estimation and validation," OpenAccess Series in Informatics, **10**(January), 123–133, 2009, doi:10.4230/OASIcs.WCET.2009.2291.

[9] F. J. Cazorla, T. Vardanega, E. Quiñones, J. Abella, "Upper-bounding program execution time with extreme value theory," in OpenAccess Series in Informatics, volume 30, 64–76, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2013, doi:10.4230/OASIcs.WCET.2013.64.

[10] L. Santinelli, J. Morio, G. Dufour, D. Jacquemart, "On the sustainability of the extreme value theory for WCET estimation," OpenAccess Series in Informatics, **39**(July), 21–30, 2014, doi:10.4230/OASIcs.WCET.2014.21.

[11] F. J. Cazorla, J. Abella, J. Andersson, T. Vardanega, F. Vatrinet, I. Bate, I. Broster, M. Azkarate-Askasua, F. Wartel, L. Cucu, F. Cros, G. Farrall, A. Gogonel, A. Gianarro, B. Triquet, C. Hernandez, C. Lo, C. Maxim, D. Morales, E. Quinones, E. Mezzetti, L. Kosmidis, I. Aguirre, M. Fernandez, M. Slijepcevic, P. Conmy, W. Talaboulma, "PROXIMA: Improving Measurement-Based Timing Analysis through Randomisation and Probabilistic Analysis," Proceedings - 19th Euromicro Conference on Digital System Design, DSD 2016, 276–285, 2016, doi:10.1109/DSD.2016.22.

[12] S. Trujillo, R. Obermaisser, K. Grüttner, F. J. Cazorla, J. Perez, "European Project Cluster on Mixed-Criticality Systems," in Design, Automation and Test in Europe (DATE) Workshop 3PMCES, 2014.

[13] M. Schoeberl, S. Abbaspour, B. Akesson, N. Audsley, R. Capasso, J. Garside, K. Goossens, S. Goossens, S. Hansen, R. Heckmann, S. Hepp, B. Huber, A. Jordan, E. Kasapaki, J. Knoop, Y. Li, D. Prokesch, W. Puffitsch, P. Puschner, A. Rocha, C. Silva, J. Sparsø, A. Tocchi, "T-CREST: Time-predictable multi-core architecture for embedded systems," Journal of Systems Architecture, **61**(9), 449–471, 2015, doi:10.1016/j.sysarc.2015.04.002.

[14] L. Kosmidis, J. Abella, E. Quinones, F. J. Cazorla, "A cache design for probabilistically analysable real-time systems," Proceedings -Design, Automation and Test in Europe, DATE, 513–518, 2013, doi:10.7873/date.2013.116.

[15] J. Jalle, L. Kosmidis, J. Abella, E. Quinoñes, F. J. Cazorla, "Bus designs for time-probabilistic multicore processors," Proceedings -Design, Automation and Test in Europe, DATE, 1–6, 2014, doi:10.7873/DATE2014.063.

[16] I. Fedotova, B. Krause, E. Siemens, "Applicability of Extreme Value Theory to the Execution Time Prediction of Programs on SoCs," in Proceedings of International Conference on Applied Innovation in IT, volume 5, 71–80, Anhalt University of Applied Sciences, 2017.

[17] Y. Lu, T. Nolte, J. Kraft, C. Norström, "Statistical-based response-time analysis of systems with execution dependencies between tasks," in Proceedings of the IEEE International Conference on Engineering of Complex Computer Systems, ICECCS, 169–179, IEEE, 2010, doi:10.1109/ICECCS.2010.26.

[18] H. Hersbach, "Decomposition of the continuous ranked probability score for ensemble prediction systems," Weather and Forecasting, **15**(5), 559–570, 2000, doi:10.1175/1520-0434(2000)015⟨0559:DOTCRP⟩2.0.CO;2.

[19] F. Guet, L. Santinelli, J. Morio, "On the Reliability of the Probabilistic Worst-Case Execution Time Estimates," European Congress on Embedded Real Time Software and Systems 2016 (ERTS'16), 10, 2016.

[20] J. Segers, "Automatic declustering of extreme values via an estimator for the extremal index Automatic declustering of extreme values via an estimator for the extremal index," (April 2002), 2014.

[21] H. Li, P. De Meulenaere, P. Hellinckx, "Powerwindow: A multi-component TACLeBench benchmark for timing analysis," in Lecture Notes on Data Engineering and Communications Technologies, volume 1, 779–788, Springer, 2017, doi:10.1007/978-3-319-49109-7_75.

[22] H. Li, P. De Meulenaere, S. Mercelis, P. Hellinckx, "Impact of software architecture on execution time: A power window TACLeBench case study," International Journal of Grid and Utility Computing, **10**(2), 132–140, 2019, doi:10.1504/IJGUC.2019.098216.

[23] "Embedded Coder," 2020.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1366-1371 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Investigating Students' Computational Thinking through STEM Robot-based Learning Activities

Sasithorn Chookaew[*,1,2], Suppachai Howimanporn[1,2], Santi Hutamarn[1,2]

[1]*Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, 10800, Thailand*

[2]*Industrial Robotics Research and Development Center, King Mongkut's University of Technology North Bangkok, 10800, Thailand*

ARTICLE INFO

Article history:
Received: 26 September, 2020
Accepted: 09 December, 2020
Online: 16 December, 2020

Keywords:
Educational robotics
STEM education
Computational thinking

ABSTRACT

Nowadays, robots are an important learning tool in education and are increasingly used inside and outside the classroom to foster the development of students' knowledge, skills, and attitudes connecting to a real-world situation. In past years, using robots in STEM (science, technology, engineering, and mathematics) education has been proposed as one of the strategies to integrate the different fields in order to construct more effective projects and innovations. Especially, engineering is the combination of math and science in solving a problem. Nevertheless, students could not understand and appreciate how to apply the knowledge of interdisciplinary integration to operate certain tasks by the engineering method. This research article presents STEM robot-based learning activities (STEM-RoLA) with sixty high-school students who were studying science and engineering. The obtained results show that the proposed STEM-RoLA is beneficial for students, especially when compared with high and low robotics performance students. The results found that high robotics performance students have higher computational thinking than low robot performance students, and they have a positive engagement response in the learning activity.

1. Introduction

At present, there is a proposal to increase robot usage in educational establishments undoubtedly because the robots have been integrated into the learning process to encourage students to think and connect to a real-world situation. Hence, using robots for education has gained a lot of attention from researchers and educators. This paper is an extension of work originally presented in the 7th International Congress on Advanced Applied Informatics (IIAI-AAI 2018) [1].

Nowadays, many studies focus on educational robots that help students develop essential skills inside and outside the classrooms that are significant in connecting the concept to real-life. Moreover, the learning activities were designed in consideration of promoting STEM discipline (Science, Technology, Engineering, and Mathematics) with a challenging learning mission in different situations [2]. Computational thinking (CT) is one of the essential core skills for students in the 21st century that is often regarded as the basic skill of computer science. CT is a kind of analytical thinking that uses the general way of thinking mathematically to solve real-life situations with scientific thinking [3]. The robot kit is a learning tool for use in the educational system, it is easy to use for students and enables them to have hands-on experience of activities and allows for a number of challenges for students to give them opportunities of robotics experience [4].

In the education context of Thailand, the students in high school have learned various subjects independently, such as Science (Physics, Chemistry, Biology, Mathematics), and Computer Programming. However, they may not understand how to apply the knowledge of interdisciplinary integration to perform or operate certain tasks in the engineering process [5].

The Thai government has attempted to push STEM education to enhance learning seamlessly for a new generation of students. Also, with the growing interest in STEM, educational robots have several benefits. Many educators have indicated that STEM integration would help enhance engineering students' learning performance with a flexible learning process. Besides, using

[*]Corresponding Author Sasithorn Chookaew, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand, Email: sasithorn.c@fte.kmutnb.ac.th

appropriate robotics-based STEM activities can be an effective way to introduce students' thinking to operate tasks systematically [6].

With the swift development of the field of robotics, it was proposed that robotics not only engage in active learning environments but also develop computational thinking knowledge and skills for students [7]. Of course, educational robotics was popular to bring an adaptive instrument for learning. Therefore, the robotics has been to promote STEM content, innovation, and creativity among students through an educational robotics competition [8].

This learning activity has integrated the benefits of robotics with the STEM education framework to promote senior high school students' computational thinking with the learning engagement process of students that was held outside the classroom for three days. This study aims at addressing the following research study:

- Do students who participate in STEM robot-based learning activities have computational thinking?
- What are the students' engagements in the STEM robot-based learning activities?

2. Literature Review Literature Review

2.1. Using Robots in Education

Robotics is a branch of engineering that includes a variety of subject matters that combine with STEM disciplines (science, technology, engineering, and mathematics). A robot is a machine applied in a variety of tasks such as industrial systems, medicinal purposes, and educational tools. Too many researchers state that educational robotics is a core element when STEM is a significant tool to integrate into the curriculum for students in the 21st century. A recent study has found that using a robot in education is increasingly being defined as an excellent instrument of teaching and learning that integrates STEM education, which is crucial for the future success of students [9].

Many studies proposed the benefit of an educational robot able to motivate students' learning with authentic learning activities based on real-world problems [10-11]. In [12], the author proposed robotic cooperative learning to promote students' critical thinking, and STEM interest indicates that students participating in robotic service-learning can benefit from this experience. Besides, robot hands-on activities can be easily adopted in classrooms to connect interdisciplinary core ideas for engineering design and STEM learning [13-14]. A robot in education can be used not only inside and outside of the classroom but also in meaningful learning activities for students as well.

Robotics makes the students able to connect and find relevance in the experience, together with the complexity of the global situation. So, it can be suggested that this is how to apply the concept of a robot for enhancing computational thinking skills.

2.2. Computational Thinking

The computational thinking basis concept of mathematics education research by Seymour Papert [15] proposed the aspects of computation that are: how to use computation to create new knowledge; how to use computers to enhance thinking and how to change patterns of access to knowledge. There is the basic concept that using technology can provide new ways to learn and teach with the logical thinking process. After that, computational thinking was described often by many researchers worldwide. In the last twenty years, computational thinking has become a fundamental skill for everyone in every field, not just for computer scientists, especially for students in the 21st-century who need skills such as analytical ability, reading, writing, and arithmetic to be processed systematically.

Based on the core concepts of computational thinking as a fundamental skill for everyone's analytical ability along with reading, writing and arithmetic, in [16], the author proposed that these should include decomposition, pattern recognition, abstractions, and algorithm design. In [17], the author classified computational thinking core components into three dimensions, including computational concepts which focus on the concepts that students employ when they are learning. Computational practices focus on problem-solving practices that occur in the learning process. Computational perspectives focus on the students' understanding of themselves and can connect learning with the world around them.

Measuring students' computational thinking is complicated, but it is a necessary task for understanding the effectiveness of the STEM-RoLA project. Our investigation focused on the implications of claims about three computational thinking dimensions consisting of computational concepts: the scientific concepts that students employ to learn and understand during activities, computational practices: problem-solving practices that occur in the learning process, and computational perspectives: students' understanding of themselves and the relationships between group members.

3. Methodology

3.1. Participants

The study was implemented in a public school in central Thailand. We selected a curriculum that prepares all students to be a scientist or an engineer. The participants in this study were sixty high school students, grades 10-12 (all males). The students were separated into twenty groups, so that each group had three members, and they learned and worked together in those groups.

3.2. Instruments

- The items in this observation checklist were developed to the three dimensions of computational thinking that cover all concepts of the robot activities including fourteen items: computational concept (4 items), computational practices (5 items), and computational perspectives (4 items). For this observation checklist, the mentors have rated the students on a 5-point Likert scale (1 = Failed, 2 = Passed, 3 = Acceptable, 4 = Good, 5 = Excellent). The scale validity was evaluated by five professors and teachers of educational robotics with more than seven years of experience. The internal consistency for the overall scale was 0.82.

- The items in this engagement questionnaire were adopted from a revised version in [5]. There are three dimensions of students' engagements after attending the robot learning activity that consist of eleven items asking about students' experiences toward STEM-RoLA. Three dimensions included behavioral engagement (3 items), cognitive engagement (4 items), and emotional engagement (4 items). The Likert scale is a five-point scale that allows the individual to express how much they agree or disagree with each item (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = agree, 5 = Strongly Agree). The internal consistency for the overall scale was 0.79.

3.3. Robot-based learning activity

The commercial robotic kits chosen in this study were mBot platform that operates and engages the students' learning experience of the mechanics, electronics, control systems, and computer programming [18]. The mBot is the easiest educational robot kit for student learning in robotics. The robot shape is easy to assemble and contains approximately 45 pieces in total, so students have a sense of achievement when they are able to assemble it quickly.

Normally, the components of mBot included the body, main control board, sensors, communication (Bluetooth), and power supply (battery). Also, it has three pre-set control modes of supporting student learning, consisting of obstacle avoidance mode, line-follow mode, and manual control mode. The students could feel that they had achieved and feel proud when they could program the mBot easily without writing difficult codes and language.

In addition, mBlock is a block-based coding platform for Scratch programming. It is easy-to-use software that makes teaching and learning coding fun and interactive. Scratch 2.0 is a very popular graphical logical programming software for teachers and students' learning. It has been proved to be one of the easiestto use graphical programming tool. Therefore, we have appled mBlock software to use Scratch coding to program Arduino and robots. Also, graphical programming is a visual programming technique where visual block connections are used for a student to code instead of texts, making it easy for non-coders to implement algorithms. It plays a meaningful role in enhancing computational thinking in the learning process.

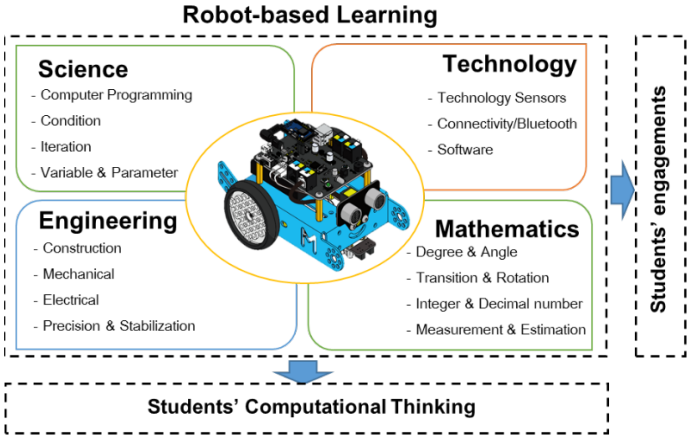


Figure 1: The STEM robot-based learning framework

Figure 1 shows the overall research of STEM-RoLA. We attempted to design the learning activity with STEM integrated. The students have learned about the mBot kits in a stepwise manner that describes its components with scientific understanding.

To achieve the tasks in this learning activity, students are encouraged to employ knowledge, skills, experience, and a teamwork process based on three dimensions of computational thinking:

- Computational thinking concepts, the students employ the knowledge when they code or control the robot for operating the mission and competition.
- Computational thinking practices, the students solve the problems that occur in activities: experimenting, iterating, and testing.
- Computational thinking perspectives, the students' understanding of themselves and their relationships with a group member.



Figure 2: The steps of STEM robot-based learning activity

To be an effective activity, we provide mentors who are pre-service engineering teachers. The mentor will help students while they work in a group during the activity. The mentors can provide

technical consultation and can be a source of support when students need it most. For the time being, each group has to come up with better solutions by integrating the knowledge of STEM and robotics, while their computational thinking is promoted accordingly.

The eight-steps to facilitate their learning process with hands-on experience lasts approximately twenty-one hours as follows:

- *Step 1: Robot introduction*: the first step addresses motivation to student-related robots that are important in daily life. After that, they introduce the functions of the robot and the introduction of the robot's components based on the engineering design process. At this moment, the students have to get acquainted with the robot components through several mini-labs such as the main control board, sensors, communication options (USB cable, Bluetooth, Wireless), and power supply (battery). In addition, the robot has three different modes of functionalities.
  - Mode A – Remote control driving or mobile device: the robot can be controlled by arrow buttons on the device to drive it forwards and backwards as well as to turn left or right.
  - Mode B – Obstacle avoidance: the robot is driven forward on its own until it detects an obstacle. If the robot detects an obstacle, it turns then continues driving forward.
  - Mode C – Line following: the robot drives forward while following a line
- *Step 2: Robot assembly:* The students were separated into twenty groups (3 students per group) then students of each group worked together. They can be learning at their own pace with support from the group mentor. This step requires students to get acquainted with each part and start to join each part together by robot function. Then, they learn three pre-set modes consisting of IR control mode, obstacle avoidance mode, and line-following mode. After that, they get the robot started by pressing on-board and push the button to change the mode.
- *Step 3: Primary programming*: the students can learn an easy programming language to control the robot. They use mBlock based on Scratch 2.0 software that is graphical programming for writing on a laptop or personal computer. The student receives the coding questions for programming practice. For example, how to control the robot to move forward or to turn left or right. They can drag the module block to move the robot area. During this step, the students in each group must analyze the facing proposition and situation in which they require to test-run-revise the code in a stepwise manner.
- *Step 4: Robot in the field:* the students from each group work together on the given tasks, a range of tasks designed to test the robot in the field. Each group has to adapt the physical robot as well as the programming instructions to overcome different challenges. This step is the basis for improving reliable robot working. The robot was assembled by students who programmed it to perform certain tasks and then there is the testing to operate, prove system functionality, and gain confidence that a system will perform as expected. Testing in the field also allows the student to learn the weaknesses, problems, and constraints of the robot and improve upon them.
- *Step 5: Robot mission:* the students learned to analyze and solve the problems in a logical way by taking the knowledge integration of STEM. The adjustment of one point will affect the related points. In this step, each group is faced with different problems in their robot's settings and programming. For example, the robot starts to move at a certain speed. When there is a barrier in front, the robot will turn right; when there is not, the robot keeps moving forward.
- *Step 6: Sharing:* The students have opportunities to share knowledge. What they have learned about each other helps build positive thinking. The representative student presents ideas to organize and control the robot to carry out each task in the missions. The students in each group help to justify the approach to solving the problems and present their ideas.
- *Step 7: Practice before the battle:* the students in each group have a rehearsal before the competition. They have analyzed and discussed opportunities and weaknesses to compete in the competition. This will increase their team's confidence in the final robot, while the final configuration may be needed. They receive the robot tasks for learning line-follow and obstacle avoidance and find a solution and strategy for solving the problem. For example, the students have learned how to program robots to take one step at a time. Meanwhile, students get to know the mechanical structure and understand how to control a robot.
- *Step 8: The robot competition:* the robot competition requires students to control the robot that makes learning robot programming simple and fun. The student members in each group are encouraged to apply knowledge of what they have learned to accomplish the goal effectively on the robot competition. There are two separate modes, including an autonomous robot following a task and a control robot via a mobile application on the Balloon Battle Game. They have to use the Makeblock open-source to control the robot with a mobile device, as shown in Figure 3.

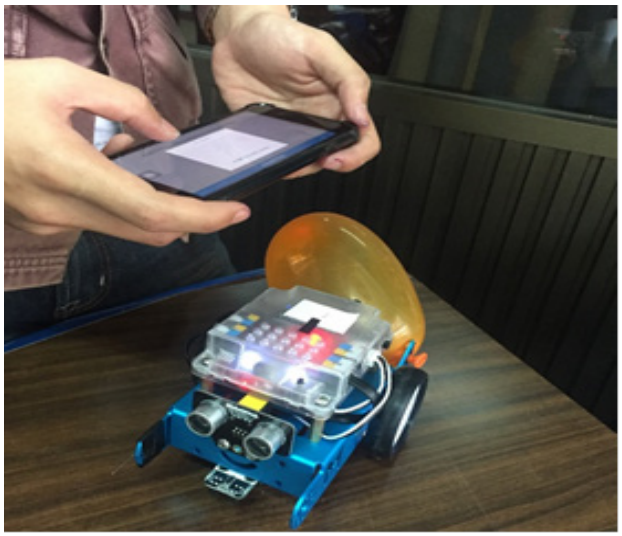


Figure 3: Illustration of a robot controlled by a mobile device

This moment is vital to each group since the robot's performance during rehearsal may be ineffective when competing with others. For this step, each robot has a needle and a balloon attached to the front and back of its body. Then each group's robot tries to stab the balloon's competitor while maintaining their

attached balloon. The loser is eliminated from the competition. For the final of the competition, each qualified group has ten minutes to come up with a better strategy for controlling the robot in a battle while maintaining their balloons. The last group to have a balloon at the end of the competition is the winner.

4. Findings

To understand better the effects of this part of the study, we compared the results with two separate groups of students' data consisting of a tasks score (30 points) and the robot competition score (70 points) total 100 points. Therefore, twenty groups demonstrated the difference between a high robotics performance group (HIRP) for the top ten groups with M=79.00 (SD =12.10) and a low robotics performance group (LORP) for the bottom ten groups. These were elaborated with M= 29.50 (SD =2.69). The findings related to the research questions were demonstrated.

4.1. The students' computational thinking

To analyze the differences between the students' robotics performances in two groups of three computational thinking components included computational thinking concepts (CTC), computational thinking practices (CTP), and computational thinking perspectives (CTPP), and these are shown in Table 1.

Table 1: Means and SDs of the students' computational thinking

CT Component	LORP		HIRP	
	M (SD)	***Remark***	***M (SD)***	***Remark***
CTC	4.01 (0.72)	Good	4.13 (0.22)	Good
CTP	2.88 (0.74)	Acceptable	4.75 (0.43)	Excellent
CTPP	3.25 (0.83)	Acceptable	4.13 (0.54)	Good

As the descriptive statistics in Table 1 show for the proposed STEM-RoLA in computational thinking concepts (CTC), both the HIRP and LORP students have concepts of the robot at a good level with M = 4.01 (SD = 0.72), and M = 4.13 (SD = 0.22), respectively.

Based on the results in computational thinking practices (CTP), the HIRP students performed tasks in each mission at an excellent level (M = 4.75, SD = 0.43), better than the LORP students at an acceptable level (M = 2.88, SD = 0.74).

Furthermore, in computational thinking perspectives (CTPP), the HIRP students tended to have positive perspectives at a good level (M = 4.13, SD = 0.22) while those who were LORP students, it was at an acceptable level (M = 3.25, SD = 0.83).

4.2. The term of students' engagements

The students' engagement results were analyzed based on quantitative and qualitative data for proposing STEM integrated robotics learning. We have presented three different aspects consisting of behavioral, cognitive, and emotional engagement as presented in Table 2.

In the first dimension, the HIRP students revealed behavioral engagement in that they could manage the function of members and have confidence in learning (M=4.63, SD=0.53). The entry indicated that the HIRP students were good at time management and working well together, as they expressed during the interviews, for example:

Table 2: Means and SDs of the students' engagements

Component	LORP		HIRP	
	M (SD)	***Remark***	***M (SD)***	***Remark***
Behavioral	4.32 (0.77)	Agree	4.63 (0.53)	Strongly Agree
Cognitive	4.47 (0.64)	Agree	4.45 (0.67)	Agree
Emotional	4.40 (0.71)	Agree	4.62 (0.51)	Strongly Agree

"Through this activity, we helped in my group about the coding program and explained to each other. I think I have learned a lot about cooperation and teamwork. I attend a robot activity with my friends, it made me have so much fun." (Student A)

"I am so excited about the robot because it is admirable to have this experience that I can construct a robot by myself. Besides, we have an opportunity to cooperate and share the ideas for working." (Student B)

While the LORP students had trouble with cooperating and working together, they did not enjoy learning in a group with members very much, and they had less time because they could not manage time (M=4.32, SD=0.77), for example:

"In my group, we have many members, so somebody does not involve in some activities. We can't organize the tasks and events on time." (Student C)

"I have less time to work on many tasks. It must spend a lot of time to solve the problems. Sometimes, I think I like to work without my friends." (Student D)

For cognitive engagement, the entry indicated that the HIRP students can reflect higher thinking skills on applications in their daily lives (M=4.45, SD=0.67), for example:

"I apply multiple tasking skills, and I can try when not sure. It is important to know how to employ the basic engineering principles for this activity." (Student A)

"For this activity, I think I have a thinking process and much knowledge I play it again when I come back home. Unfortunately, while I have learned a lot of programming, I never took any classes on robotics" (Student B)

While the LORP students just reflected what they had experienced from the activity by putting more effort before the success (M=4.47, SD=0.64), for example:

"I am upset about how to complete the work. However, I think I attempt to find the solution myself." (Student C)

"In the activity, I do not plan, it's not a success. I have attempted, but still failed, again and again, so I consulted the mentor." (Student D)

In the end, the HIRP students have emotional engagement. They enjoyed the learning activities (M=4.62, SD=0.51), especially the students in HIRP who have positive responses such as:

"I think it could be an interesting experience. I like to control a robot to follow my instructions. Each task, I am quite confident that a robot can do some pretty fun stuff." (Student A)

"I very enjoyed the activities in the activity because I think a robot can be a good assistant in the future. It helps me to be more productive in performing daily activities." (Student B)

On the other hand, the LORP students did not concentrate on some activities. However, they exposed their emotions towards the assistance of peer members in the group that could encourage them to proceed with the tasks. So, they have an attitude towards this (M=4.44, SD=0.71), as they expressed during the interviews, for example:

"I think a robot is difficult to use because it was always programmed. I'm running into a lot of infrared sensor issues. I'm attempt to do the typical line follower activity and also go ahead and stop when it finds a black line." (Student C)

"This activity was held during the semester. I have a lot of homework, so I don't concentrate on some activities. I think it should be held in my free time." (Student D)

There was unanimous agreement from the students present that, as a result of this robot activity, the students had significantly developed their computational thinking to solve the problems and challenges, their ability to work as a group, and their resilience. Also, students said they greatly enjoyed participating in the challenges and wanted more opportunities to participate in such STEM-RoLA.

5. Conclusions

Within Thailand's educational context, the integration of interdisciplinary knowledge to enhance crucial skills thinking, and a positive attitude are also important to realize goal. So the robotics STEM was integrated through active learning and the new learning tool has an effective strategy. The robotics challenges allow students to grasp the fundamentals of basic programming and develop their computational thinking with STEM education in an engaging way. The STEM-RoLA was held to motivate the senior high school students in three days (21 hours), consisting of eight steps in total, from robot introduction, assembly, programming, testing in the field, robot mission, sharing, practice before the battle, and final robot competition. The students established their robotics knowledge with hands-on experience based on the STEM strategy. Therefore, the obtained results show that the proposed STEM-RoLA is beneficial for the students, especially when compared with high and low robotics performance students. The results find that high robotics performance students have higher computational thinking than low robotics performance students. Also, they have positive engagement responses to the learning activity.

Nevertheless, the learning activities have shown that there are some limitations due to the availability of the school context, the information, and the educational context students already have before they learn robot information that might affect effectiveness. Also, the number of robot kits are quite expensive and inflexible to learning in some contexts. Future studies may explore whether there are specific learning strategies and learning tools that are particularly useful based on the attitudes of the students. Based on the result, the present research involved all males. So, future studies should clarify the results that can be described by gender differences in learning outcomes and assess the long-term effects on attitudes.

Acknowledgment

This study is supported in part by King Mongkut's University of Technology North Bangkok.

References

[1] S. Chookaew, S. Howimanporn, P. Pratumsuwan, S. Hutamarn, W. Sootkaneung, & C. Wongwatkit, "Enhancing High-School Students' Computational Thinking with Educational Robotics Learning" in 2018 International Congress on Advanced Applied Informatics (IIAI-AAI), 204-208). 2018, doi:10.1109/IIAI-AAI.2018.00047.

[2] T. J. Kopcha, et al. "Developing an integrative STEM curriculum for robotics education through educational design research" Journal of Formative Design in Learning, **1**(1), 31-44, 2017, doi.org/10.1007/s41686-017-0005-1

[3] J. M. Wing, "Computational thinking" Communications of the ACM, 49(3), 33-35, 2006.

[4] B. Zhong, Y. Wang, "Effects of roles assignment and learning styles on pair learning in robotics education" International Journal of Technology and Design Education, 1-19, 2019, doi.org/10.1007/s10798-019-09536-2

[5] S. Hutamarn, S. Chookaew, C. Wongwatkit, S. Howimanporn, T. Tonggeod, S. Panjan, "A stem robotics workshop to promote computational thinking process of pre-engineering students in Thailand: STEMrobot" in 2017 International Conference on Computers in Education, 514-522, 2017.

[6] C. Chalmers "Robotics and computational thinking in primary school" International Journal of Child-Computer Interaction, **17**, 93-100, 2018, doi.org/10.1016/j.ijcci.2018.06.005

[7] E. B. Witherspoon, R. M. Higashi, C. D. Schunn, E. C. Baehr, R. Shoop, "Developing Computational Thinking through a Virtual Robotics Programming Curriculum," ACM Transection on Computing Education, **18**(1), 1–20, 2017, doi.org/10.1145/3104982.

[8] A. Eguchi, "RoboCupJunior for promoting STEM education, 21st century skills, and technological advancement through robotics competition," Robotics and Autonomous Systems, **75**, 692–699, 2016, doi.org/10.1016/j.robot.2015.05.013.

[9] T. Larkin, "Topic Order in Introductory Physics and its Impact on the STEM Curricular Ladder" International Journal of Engineering Pedagogy, **7**(1), 136-150, 2017, https://www.learntechlib.org/p/207332/.

[10] C. Julià, J. Ò. Antolí, "Impact of implementing a long-term STEM-based active learning course on students' motivation" International Journal of Technology and Design Education, **29**(2), 303-327, 2019, doi.org/10.1007/s10798-018-9441-8.

[11] M. Barak, M. Assal, "Robotics and STEM learning: Students' achievements in assignments according to the P3 Task Taxonomy—practice, problem solving, and projects" International Journal of Technology and Design Education, **28**(1), 121-144, 2018, doi.org/10.1007/s10798-016-9385-9.

[12] P. Mosley, G. Ardito, L. Scollins, "Robotic cooperative learning promotes student STEM interest" American Journal of Engineering Education, **7**(2), 117-128. 2016.

[13] S. Ziaeefard, M. H. Miller, M. Rastgaar, N. Mahmoudian, "Co-robotics hands-on activities: A gateway to engineering design and STEM learning" Robotics and Autonomous Systems, **97**, 40-50, 2017, doi.org/10.1016/j.robot.2017.07.013.

[14] S. Chookaew, C. Silawatchananai, S. Hutamarn, S. Howimanporn, W. Sootkaneung, C. Wongwatkit, "An investigation of vocational students' attitude towards STEM robotic activities" in 2019 International Conference on Computers in Education (ICCE) 134–138, 2019.

[15] S. Papert, "An exploration in the space of mathematics educations" International Journal of Computers for Mathematical Learning, **1**(1), 95-123, 1996, doi: https://doi.org/10.1007/BF00191473.

[16] J M. Wing, "Computational thinking and thinking about computing" Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, **366**(1881), 3717-3725, 2008, doi.org/10.1098/rsta.2008.0118.

[17] K. Brennan, M. Resnick, "New frameworks for studying and assessing the development of computational thinking" in 2012 Annual Meeting of the American educational research association, Vancouver, Canada 1, 1-25, 2012.

[18] P. P. Merino, E. S. Ruiz, G. C. Fernandez, and M. C. Gil, "A Wireless robotic educational platform approach," in 2016 International Conference on Remote Engineering and Virtual Instrumentation, 2016, 145–152.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 6, No. 1, 1369-1377 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Texture Based Image Retrieval Using Semivariogram and Various Distance Measures

Rajani Narayan*, Anjanappa Sreenivasa Murthy

Department of ECE, University Visvesvaraya College of Engineering, Bengaluru, 560001, India

ARTICLE INFO

Article history:
Received: 03 January, 2021
Accepted: 09 February, 2021
Online: 28 February, 2021

Keywords:
Image Retrieval
Lag distance
Semivariogram
Similarity measurement
Texture Feature

ABSTRACT

In a content-based image retrieval system(CBIR) feature classification,identification, and extraction play an important role. The retrieval of images using a single feature is a challenging task in CBIR systems. The high retrieval rates are reported based on combining multiple features, multiple algorithms and preprocessing steps, feature classification, and segmentation because the image retrieval are mainly based on the content in an image. This paper presents a texture feature extraction for the image retrieval system from semivariogram and robust semivariogram technique. A semivariogram is a statistical approach that provides the textural information based on the lag distance 'h'.The proposed method is tested on various standard image databases such as Corel-1k, Corel-10k, and Coil-100 database. The semivariogram and robust semivariogram methods are tested for the Corel 1k database using four distance metrics i.e. Euclidean, Manhatten, Canberra, and Chord distance to check which distance measure is appropriate for the CBIR system.The proposed method is also tested on three types of databases to investigate the performance of the CBIR system. The Matlab simulation results show that the effective performance of the system with Euclidean distance.

1 Introduction

The capturing of an image is started back in the past one and half a century, and the collection of images is rapidly growing from the past thirty years of growing digital technology and internet usage. The collection of images is also increasing leading to a extremely large database. In many applications, the extraction of equivalent images from the large database is essential and useful. There are two trends currently using the image search from the image collection data set namely text-based and content-based. The text-based search uses meta-data like keywords, description of the image to get back the similar images. It is the main disadvantage of this system. In contrast content based uses existing property (color, texture, and shape) to find the equivalent images from the large database. This paper is an extension of work originally presented in conference [1].

Many researchers have come up with the idea of combining basic image features to get the most comparable images as required by the query image. The color and texture feature is taken together to get the image information by applying wavelet [2]. The wavelet transform decomposes the images into orthogonal components for the better localization of the spatial information in an image. The method applied to extract color and texture information is a color histogram and wavelets respectively. .

It is noticeable that based on these work the retrieval rates are improved by combining multiple features, multiple algorithms and preprocessing steps, classification, segmentation, etc. In this work, a good CBIR system is designed by using the texture feature. In an image retrieval system all the three features color, texture, and shape play an important role [3]. Texture identification is an important part of the research in image retrieval and pattern recognition. It is categorized into two classes i.e. regular and irregular, directional, and non-directional based on the observations of real life. In general, the texture is characterized as two main approaches: Statistical and Structural.

1)Structural approach: the texture analysis is done based on the arrangement of preserving texture pixel in some pattern or repetitive pixel. 2) Statistical approach: The texture feature can be analyzed by studying numerical data of the image. It provides the relative estimation of the arrangements of the intensities in an image region [4]. The most commonly used and widely accepted statistical approach for texture, classification is the GLCM method [5]. The recently developed method for the texture classification and analysis includes block truncation coding [6], Markov random fields [7] to construct the image patches using neighboring pixel, local energy histogram [8], hidden Markov model [9], shearlets and linear regression [10]. Many of the texture algorithms and methods have been applied to the Brodatz database.

*Corresponding Author: Rajani Narayan, Bengaluru, +91-9740798868 & rajni.n25@gmail.com

In this paper, geostatistical parameter called semivariogram and robust semivariogram are applied to the image to get the texture information to the image databases. The semivariogram method is basically applied to remote sensing images to characterize the texture property. It is a distance metric that calculates the absolute differences between the semivariogram of the two images to determine the class of test images. The various distance measures are used to verify the effective retrieval of images.

1.1 Motivation

The main idea of the designing image retrieval system is to provide the desired images as of user interest. This leads to finding the solution to set up a good image retrieval system that will be able to find similar images from the huge image data set. The image search is basically with the use of content in the image.

In this study, a new CBIR system is proposed based on semivariogram. In general the semivariogram is applied for the remote sensing images in the area of geo-statistics.These images are captured by satellites that will be having high spatial resolution.To analyse the characteristics of this data semivariogram is one of the best method.Like remote sensing image CBIR system works with large image database. These databases contain many categories of images that will have a non-uniform distribution of the pixel intensities and dissimilar backgrounds.Since semivariogram is a statistical approach, it really works well for the image to characterise the texture intensities.This is the case the idea of adapting semivariogram method for the CBIR system.The main focus of the work is to use the minimal feature from the image and reduce the computational time required for the entire image database.

1.2 Organisation of the paper

The remainder of this article is organized as follows. Section 2 discusses the related work and section 3 gives the mathematical model of the proposed system. In section 4 feature extraction is discussed with an example. Section 5 briefly explains the experimental design and results. Section 6 gives the distance metrics for similarity measurements. Section 7 discusses the experimental design and Section 8 gives performance evaluation and conclusion followed by section 8.

2 Related Work

From the discrete wiener-Levy process in [11], it is proved that a semivariogram gives a spatial correlation of the pixels than the covariance method. The random process called the Levy process is a Brownian motion mathematical representation. The covariance fails to extract spatial information with distortion whereas the semivariogram strongly characterizes the spatial features. In geostatistics, the features analysis has to be done by fitting the experimental semivariogram to a theoretical semivariogram. The model examines the structure and continuity of the spatial relationship in a random field.

The experimental semivariogram model can be fit into the theoretical semivariogram model. There are two methods have been used in [12]. One is by manual fitting and the second is by automatic model fitting. The automatic model fitting uses methods like maximum likelihood or least squares [13].In manual fitting, the model was chosen based on picture perception of the experimental semivariogram. For a given waveband information reproduction of the pattern will be clear and strong explanation of the pattern in an image can be provided by the semivariogram.

In general remote sensing images for the land, consolidation is having a high spatial resolution. This is a limitation for the pixel-based analysis and land cover classification in remote sensing images concerning its spectral behaviour. To overcome this problem texture extraction is done using fixed window size and varying window size depending on semivariogram result. Texture classification in microwave images [14] is executed using the semivariogram method. The statistical parameters are determined and considered as image features. They are first and second-order element difference, first and second-order element inverse difference, entropy and, uniformity.

The accurate object-oriented classification is presented in [15] for Quick Bird images by applying a method called the set of semivariogram texture feature(STF) based on the mean square root pair difference(SRPD). The parameter like direction, lag distance, and moving window size are considered for semivariance analysis to classify the texture feature.

A new set of 2-D RCWF(Rotated complex wavelet filter) is designed to improve the retrieval accuracy using complex wavelet filter coefficients. It produces strong texture information of the oriented images in six different directions. This is non-separable and oriented relatively improves the classification of oriented textures. Also, they have used the combination of dual-tree RCWF and DFWCT to extract the texture features in 12 different directions [16].

The new algorithm based on Gabor filter and EHD(Edge Histogram Descriptor) has been proposed for the texture feature extraction. These two are fused to get the good performance of the system. The Gabor and EHD are applied to every 40 images and then six bins are considered from each of the images. This improves the efficiency of the EHD. As a result, the dimension of the character is reduced to 40X6 as mentioned in [17].

Perceptual texture features are considered for the representations and retrieval of images. These features contain contrast, directionality, and coarseness and dryness. Where busyness has given less importance compared to the other three features. They have used these four perceptual features for the experiment. The experiment result shows appreciable scores for the broadtz image database retrieval [18].

The image retrieval system uses a combination of color and texture features. As a color feature color autocorrelation of the HSV color model is used and for texture feature BDIP(Block difference of inverse probabilities) and BVLC(Block variation of the local correlation coefficient)are used. These two texture features effectively measure the texture smoothness. The image is divided into 2X2 blocks and these moments are computed in the wavelet transform domain. This gives better accuracy when compared with conventional texture extraction methods [19].

Palm print retrieval is proposed in [20] using texture features. Texture energy is presented to identify the global and local texture features of the palm print. The four masks which are defined in the work are used to extract the line segment in horizontal and verti-

cal directions and is referred to as local feature. The local feature provides detailed information and global feature provides overall properties of the palm print. The steps involved in global feature calculations are image alignment, energy computation, boundary tracing, a center of gravity calculation, and mapping the key points.

The content-based image retrieval system is an active research area for the past two decades. Many researchers have come up with different ideas to enhance the performance of the system by extracting texture features. The most used texture extraction methods are GLCM, covariance method and co-occurrence matrix, and so on. In this work, the semivariogram method is applied for texture extraction. The semivariogram is widely used to analyse the pattern of the remote sensing images. The content-based image retrieval databases will have distinct images with different types of patterns. The function of the semivariogram is to find the spatial distance between each pattern in the image. The semivariogram approach and spectral distortion measures are applied to image texture retrieval in [21]. The experiment is conducted using the Broadtz texture database. The spectral distortion methods are combined with a semivariogram to retrieve similar images. The effectiveness of the work is tested on Illinois at Urbana-Champaign texture database. The semivariogram method does not need the knowledge of the mean value of the function for the estimation of the spatial relationship. The probability of the random function is not required for the semivariogram methodology as in the GLCM method and co-occurrence method.

A hierarchical retrieval system is proposed in [22] using the visual contents(color,shape and texture).The fusion of histogram gradients, adaptive tetrolet transform and auto color correlogram methods are applied to extract the image features and to make the image retrieval more effective.

The proposed method estimates the pattern similarity by finding the distance between the pixels as it plays an important role in recognition of the similar images from the large database.

3 Mathematical Model for Semivariogram and Robust semivariogram

Variance is a measure of data spread between numbers in a data set. It measures the data on how far each number in the set is from the mean and also its variability from the average mean. A small variance shows how close the data to the mean and high variance indicates the data points are how far from the mean.The variogram is a statistical measure can be represented graphically in a manner which characterizes the spatial continuity i.e. roughness of a dataset.

Semivariogram is one of the standard geo-statistical parameters is used to find spatial relationships between two values in that location as a function of the distance. The lag distance h is taken for paired location in an image.

Let X, z, and h be a random function, a spatial location, and a lag distance in the sampling space, respectively. The variogram of the random function is defined as:

$$2\gamma(h) = Var[Z(x) - Z(x+h)] \tag{1}$$

where $\gamma(h)$ is called the semivariogram of the random function. Let $Z(x_i)$ be the pixel value at the location , i = 1, 2, . . ., n, where n is the sampling size. The variable z is the intensity value at location of the image. The experimental semivariogram, of the random function is expressed as in [26]

$$\hat{\gamma(h)} = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \tag{2}$$

The reason for introducing the robust semivariogram is because the semivariogram is not robust with respect to skewness and distorted information.The robust semivariogram denoted as $\gamma_R(h)$ is defined as in [23] ,[24]

$$\gamma_R(h) = \delta[\frac{1}{N(h)} \sum_{i=1}^{N(h)} \mid Z(x_i) - Z(x_i + h) \mid^{1/2}]^4 \tag{3}$$

where

$$\delta = \frac{1}{2(0.457 + \frac{0.494}{N(h)})} \tag{4}$$

4 Extraction of Texture Feature of an Image

The texture feature can be extracted using the semivariogram.The block diagram of the proposed method is as shown in fig 1.Let us consider an image as shown in fig 2. To extract texture information the image is resized into 640X640 pixels. Then the image is divided into 9 sub-images of size 215X215. The lag distance 'h' value is taken as 20. The variogram is calculated for each sub-image and its results in a feature vector size 20X9. The pictorial representation of the semivariogram features and robust semivariogram features is as shown in fig 3 and fig 4.The variogram shape near the origin is linked to the smoothness of the phenomenon. If the shape is continuous and differentiable it shows smoothness of the image, if it is not differentiable it shows roughness of the image. This process is applied to all the images in the query database and search database and these features are stored to find the similarity distance.

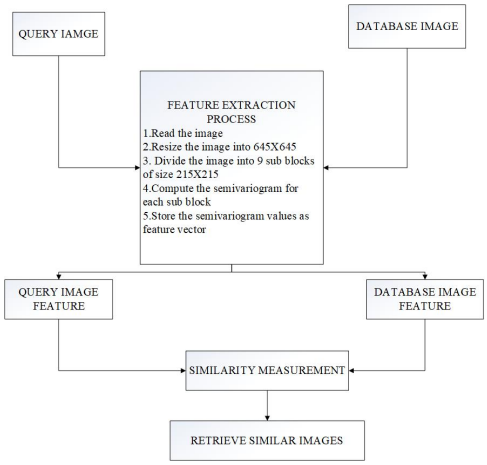


Figure 1: Block diagram of the proposed method



Figure 2: Bus Image

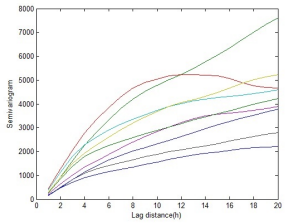


Figure 3: Experimental Semivariogram for the Bus image

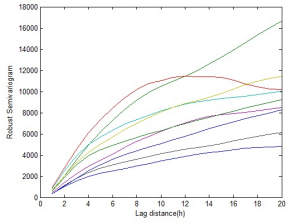


Figure 4: Robust Semivariogram for the Bus image

5 Distance Metrics for similarity measurements

Euclidean distance: If S (x1, y1) and T (x2, y2) are the two feature points of the image then the Euclidean distance between the two points is given by,

$$EUD(S,T) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{5}$$

label1.4

Canberra distance: The Canberra distance is a numerical measure of the distance between pairs of points in a vector space. The Canberra distance (CBD) between two vectors is given by ,

$$CBD(S_i, T_i) = \sum_{i=1}^{N} \frac{|\, x_i - y_i \,|}{|\, x_i \,| + |\, y_i \,|} \tag{6}$$

Manhattan distance: If S (x1, y1) and T (x2, y2) are two points then the Manhattan distance between two points S and T is given by,

$$MHD(S,T) = |\, x_1 - x_2 \,| + |\, y_1 - y_2 \,| \tag{7}$$

Chord distance: It is used to find the correlation between the two images.If a and b are referred as a array of the two images then the chord distance is given by.

$$C = 1 - abs[\frac{a(i) * b(j)}{norm(a(i)) * norm(b(j))}] \tag{8}$$

6 Experimental Design and Discussion

6.1 Experiment with three databases using Euclidean distance

The image retrieval experiment is conducted using the natural and object-based image database namely Corel 1K, Corel 10K, and Coil 100. The images in these databases are taken to investigate the performance of the proposed method. The Corel 1K database contains 10 categories and each category has 100 images such as people, beach, bus, dinosaur, roses, and so on. Each image of size 256X384 and 384X256.This is divided into query database of 300 images i.e 30 images from each category and search a database of 700 images i.e. 70 images from each category.

The proposed method is examined using the Corel 10K database [6]. This database contains 10000 natural image collection of 100 categories. Each image of size 126X187 and 187X126. The images in this database are having a dissimilar and unstructured pattern. Some categories of images are dolls, doors, sunset, mountains, etc. Each category has 100 images which are divided into 3000 images of query database i.e 30 images from each class and 7000 images of search database i.e. 70 images from each category.

The COIL 100 dataset is used to investigate the performance of the proposed image retrieval system. This database contains 100 categories of 7100 images. Each image is of size 128X128. This database is having object-based images like a hammer, car, toy, fruits, etc with uniform background and these images are most similar in pattern. This database is divided into a query database of 1000 images 10 images from each class and a search database of 6100 images 61 images from each class.

The system is designed to show the top ten retrievals of similar images for each query image. The similarity matching is based on the shortest distance score measured using Euclidean distance metrics between the query image and the database image. If the system retrieves similar images according to query, then we say the system is retrieved the target images else the system fails to retrieves the target images. The performance of the system is examined using average precision and average recall.Table 1 shows the three databases' precision and recall values using semivariogram method and lag distance is set to 20.

Table 1: Prescision and recall values using Experimental semivariogram on Image Databases.

sl No	Name of the Database	Precision	Recall
1	Corel 1K	0.77	0.109
2	Corel 10K	0.323	0.045
3	Coil 100	0.858	0.126

6.2 *Experiment with Corel-1k database using four distance measures*

The retrieval rates are computed using the complete Corel 1k dataset using four distance metrics. The retrieval test is done using the semivariogram and robust semivariogram method by setting the lag distance value h is equal to 20.

The distance measures compared using semivariogram and robust semivariogram methods. Table 2 and Table 3 show the average precision and average recall values for four distance measures.The examples for retrieval of these measure is as shown in fig 12,fig 13, fig 14 and fig 15. The fig 11 shows the query images.

Table 2: The average Precision values for Four distance metric

sl No	Distance metric	Semivariogram	Robust Semivariogram
1	Euclidean	0.77	0.78
2	Manhatten	0.52	0.51
3	Canberra	0.58	0.72
4	Chord	0.38	0.54

Table 3: The average recall values for four distance metric

sl No	Distance metric	Semivariogram	Robust Semivariogram
1	Euclidean	0.109	0.109
2	Manhatten	0.061	0.072
3	Canberra	0.071	0.102
4	Chord	0.176	0.057

6.3 *Experiment with lag distance 'h' for the Corel-1k database using Euclidean distance*

The semivariogram is a function of variable lag distance h. The semivariogram increases the dissimilarity with the lag distance h. If the distance is more it fails to identify the patterns of the image. The semivariogram is calculated based on the lag distance. In this experiment, the distance chosen is 20 and 30. As the distance increases the textural information is lost.if the distance is less than strong textural information can be extracted. The table 4 shows the time taken for the calculation of semivariogram and robust semivariogram for a single image.Table 5 shows the precision and recall values for the retrievals of corel 1k database with lag distance 20 and 30.From the table 5 we can notice that for the lower lag distance the performance of the system is effective.

Table 4: The computation time for different lag distance

sl No	Lag distance(h)	Semivariogram	Robust Semivariogram
1	30	1.070 sec	2.88 sec
2	50	1.483 sec	4.25 sec
3	100	2.136 sec	6.59 sec
4	150	2.448 sec	7.566 sec
5	200	2.479 sec	7.91 sec

Table 5: The computation of Image retrievals for the lag distance h=20 and h=30 using semivariogram

sl No	Lag distance(h)	Precision	Recall
1	20	0.77	0.109
2	30	0.74	0.06

From the Table 4, we can see that as distance increases computation time increases, and the amount of information decreases. As the distance taken is small the rate of information increases and results in more retrieval of similar images i.e. lower the distance higher the continuity. For the lag distance h the proposed method results in an effective retrieval of similar images.



Figure 5: Query Images from COIL 100 dataset

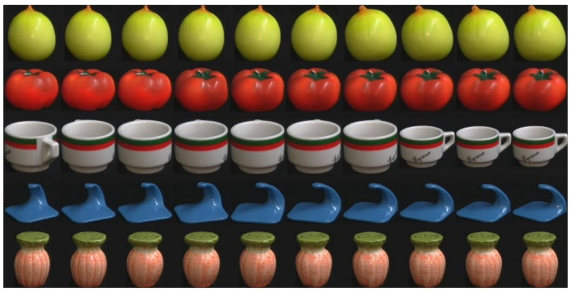


Figure 6: Retrieved images for the queries show in figure 5.

6.4 *Retrieval results*

The feature vectors are stored and are tested for four distance metrics. The result shows that Euclidean distance is best suited for the retrieval of similar images from the image database. The robust semivariogram shows still good performance comparing with the semivariogram method. This is because the semivariogram is not robust against the distribution and distortion of the pixel values. The Euclidean distance measure results in good retrieval of similar images when compared to the other three distance measures. Results

are tabulated for the first top ten retrieved images. For the Dinosaurs, Rose, and Horse class Ten images of the same class are retrieved at the top ten images. The results by using the semivariogram method for the image databases are tabulated in Table 1.



Figure 7: Query image from COREL-1K database



Figure 8: Retrival using Euclidean distance for the COREL-1K database



Figure 9: Query image from COREL-10K database



Figure 10: Retrival uisng Euclidean distance for the COREL-10K database

The top ten retrieved images for the Coil-100, Corel 1K, and Corel 10K database are shown in fig 6, fig 8, and fig 10 respectively. The Query samples for these retrievals are shown in fig 5, fig 7, and fig 9. For the COIL 100 database, the proposed technique is effective in retrieving similar images from the database for all the categories present in the database. The semivariogram method works well for the Corel 1k and Corel 10k database in which these databases are having the most dissimilar pattern of images.

7 Performance Evaluation and comparison

In content-based image retrieval, most of the researchers use precision and recall as the evaluation metrics. Precision is the proportion of relevant images retrieved over the total number of images retrieved and recall is the relevant images retrieved over a total number of relevant images are there in the database. The precision and recall are defined,

Precision = $\frac{R}{R1}$

Recall = $\frac{R}{R2}$.

where, R: Number of relevant images retrieved.
R1: Total number of images retrieved.
R2: Total number of relevant images in the database.

The proposed method gives 77% of efficiency for the corel 1k database,32.3% for corel 10k and 85.8% efficiency for the coil-100 database as shown in the Table 1.

The experiment is conducted to investigate which distance measure is best suited to find the query image's closeness with the database image.We have compared the performance of the CBIR systems in terms of precision and recall values using Euclidean distance, Canberra distance, Manhattan distance, and Chord distance. The performance results are as tabulated in the table shows that Euclidean distance is best suited for retrieving images from the large database for the proposed approach. Whereas Chord distance shows very low performance compared to Euclidean distance,Canberra distance and Manhattan distance.

7.1 Comparison with existing methods

The experiment is compared with some of the existing methods under the Corel 1k database. It can be seen that the proposed method performs better in comparison with the existing CBIR methods as in [12]–[25]. The proposed methods yield the average precision value 0.77 and 0.78 concerning semivariogram and robust semivariogram.Table 6 shows the comparison of Corel-1K database with existing methods.

For Corel 10000 database, the experiments were conducted as similar to Corel 1000 database. In this work, 100 categories are used to check the system performance. The work carried in [26] uses the 20 image categories to test the system performance. As similar to the experimental set-up in [27], [28]the number of retrieved images is set to 10.Table 7 exhibits the proposed method's comparisons over the existing methods for Corel 1k and Corel 10k database.

The COIL-100 database is compared with existing methods as shown in Table 8. The former schemes show good results by considering only a few categories of the COIL-100 database. In our work, the results are tabulated for the entire categories present in the COIL-100 database.

The Table 7 and 8 shows the average precision values for the best retrievals of image categories. The authors from [29], [30]

presents the system performance which gives best results for some categories of the image database.The number categories does not play any role for the system performance ,it mainly depends on right choice of the feature and the similarity measurement. The proposed work is tested for all the categories of the image database and it is capable of retrieving most similar images for all the categories except for the Corel 10k database.



Figure 11: Query Image from Corel-1K database



Figure 12: Retrieval using Euclidean distance



Figure 13: Retrieval using Manhatten distance



Figure 14: Retrieval using Canberra distance



Figure 15: Retrieval using Chord distance

Table 6: Comparison with Existing CBIR methods under Corel-1k Database.

Sr. No	Method	Average precision
1	[31]	0.396
2	[32]	0.526
3	[33]	0.532
4	[34]	0.533
5	[35]	0.566
6	[36]	0.565
7	[25]	0.0665
8	[26]	0.717
9	[29]	0.725
10	**Proposed method**	**0.77**

Table 7: Comparison with Existing CBIR methods under corel-10k Database.

Sr. No	Method	Average precision
1	EOAC[27](65 Query images)	0.210
2	CSD3[26](10 categories)	0.290
3	MTH[37](100 categories)	0.310
4	TCM[28](100 categories)	0.204
5	EDBTC[38](20 categories)	0.798
6	**Proposed method** (100 categories)	**0.323**

Table 8: Comparison with Existing CBIR methods under COIL-100 Database.

Sr. No	Method	Average precision
1	[39](10 categories)	0.86
2	[40](15 categories)	0.88
3	[41](15 categories)	0.78
4	[42](10 categories)	0.86
5	[43](10 categories)	0.93
6	**Proposed method**(100 categories)	**0.858**

8 Conclusion

The semivariogram and robust semivariogram techniques are applied to the color images for the feature extraction and retrieval processes from the large database.These techniques gives promising texture feature for the images.Both methods are analysed and compared with different distance metrics. The experiment tested using three databases having images of natural scenes, color, texture, and objects. The results are computed in terms of precision and recall. The precision rate achieved for the Coil-100 database is 85% and the Corel-1K database is 76%. The improved performance is observed with Euclidean distance measure over three distance measures. The experiment results show that the semivariogram is more effective for the content-based image retrieval system but not limited to remote sensing images.

References

[1] N. Rajani, A. Sreenivasa Murthy, “Semivariogram based feature extraction for content based image retrieval,” in ACM International Conference Proceeding

Series, 2019, doi:10.1145/3332340.3332348.

[2] M. Singha, K. Hemachandran, "Content based image retrieval using color and texture," Signal & Image Processing, **3**(1), 39, 2012.

[3] D. Srivastava, S. Goel, S. Agarwal, "Pipelined technique for image retrieval using texture and color," in 2017 4th International Conference on Power, Control & Embedded Systems (ICPCES), 1–6, IEEE, 2017, doi:10.1109/ICEE.2017.31.

[4] H. Kebapci, B. Yanikoglu, G. Unal, "Plant image retrieval using color, shape and texture features," The Computer Journal, **54**(9), 1475–1490, 2011.

[5] R. M. Haralick, K. Shanmugam, I. H. Dinstein, "Textural features for image classification," IEEE Transactions on systems, man, and cybernetics, (6), 610–621, 1973.

[6] J.-M. Guo, H. Prasetyo, "Content-based image retrieval using features extracted from halftoning-based block truncation coding," IEEE Transactions on image processing, **24**(3), 1010–1024, 2014, doi:10.1109/IEES.2014.31.

[7] T. Ružić, A. Pižurica, "Context-aware patch-based image inpainting using Markov random field modeling," IEEE transactions on image processing, **24**(1), 444–456, 2014, doi:10.1109/TRANS.2014.31.

[8] Y. Dong, J. Ma, "Wavelet-based image texture classification using local energy histograms," IEEE Signal Processing Letters, **18**(4), 247–250, 2011, doi:10.1109/LSP.2011.2111369.

[9] Y. Qiao, L. Weng, "Hidden markov model based dynamic texture classification," IEEE Signal Processing Letters, **22**(4), 509–512, 2014, doi: 10.1109/SDF.2014.2111369.

[10] Y. Dong, D. Tao, X. Li, J. Ma, J. Pu, "Texture classification and retrieval using shearlets and linear regression," IEEE transactions on cybernetics, **45**(3), 358–369, 2014, doi:10.1109/TCFG.2014.21369.

[11] R. A. Olea, "The Semivariogram," in Geostatistics for Engineers and Earth Scientists, 67–90, Springer, 1999.

[12] N. Cressie, "Statistics for spatial data New York," 1993.

[13] A. Balaguer, L. A. Ruiz, T. Hermosilla, J. A. Recio, "Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification," Computers & Geosciences, **36**(2), 231–240, 2010.

[14] K. Kitada, K. Fukuyama, "Land-use and land-cover mapping using a gradable classification method," Remote Sensing, **4**(6), 1544–1558, 2012.

[15] J. R. Carr, F. P. De Miranda, "The semivariogram in comparison to the co-occurrence matrix for classification of image texture," IEEE Transactions on geoscience and remote sensing, **36**(6), 1945–1952, 1998.

[16] A. Yue, C. Zhang, J. Yang, W. Su, W. Yun, D. Zhu, "Texture extraction for object-oriented classification of high spatial resolution remotely sensed images using a semivariogram," International Journal of Remote Sensing, **34**(11), 3736–3759, 2013.

[17] X. Wu, J. Peng, J. Shan, W. Cui, "Evaluation of semivariogram features for object-based image classification," Geo-spatial Information Science, **18**(4), 159–170, 2015.

[18] M. Kokare, P. K. Biswas, B. N. Chatterji, "Texture image retrieval using new rotated complex wavelet filters," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), **35**(6), 1168–1178, 2005.

[19] H. Zhang, X. Jiang, "An improved algorithm based on texture feature extraction for image retrieval," in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 1270–1274, IEEE, 2015.

[20] M. Jian, L. Liu, F. Guo, "Texture image classification using perceptual texture features and Gabor wavelet features," in 2009 Asia-Pacific Conference on Information Processing, volume 2, 55–58, IEEE, 2009.

[21] H. J. So, M. H. Kim, N. C. Kim, "Texture classification using wavelet-domain BDIP and BVLC features," in 2009 17th European Signal Processing Conference, 1117–1120, IEEE, 2009.

[22] N. Varish, A. K. Pal, R. Hassan, M. K. Hasan, A. Khan, N. Parveen, D. Banerjee, V. Pellakuri, A. U. Haqis, I. Memon, "Image retrieval scheme using quantized bins of color image components and adaptive tetrolet transform," IEEE Access, **8**, 117639–117665, 2020.

[23] W. Li, J. You, D. Zhang, "Texture-based palmprint retrieval using a layered search scheme for personal identification," IEEE Transactions on Multimedia, **7**(5), 891–898, 2005.

[24] T. D. Pham, "The semi-variogram and spectral distortion measures for image texture retrieval," IEEE Transactions on Image Processing, **25**(4), 1556–1565, 2016.

[25] T.-C. Lu, C.-C. Chang, "Color image retrieval technique based on color features and image bitmap," Information processing & management, **43**(2), 461–472, 2007.

[26] F.-X. Yu, H. Luo, Z.-M. Lu, "Colour image retrieval using pattern co-occurrence matrices based on BTC and VQ," Electronics letters, **47**(2), 100–101, 2011.

[27] F. Mahmoudi, J. Shanbehzadeh, A.-M. Eftekhari-Moghadam, H. Soltanian-Zadeh, "Image retrieval based on shape similarity by edge orientation autocorrelogram," Pattern recognition, **36**(8), 1725–1736, 2003.

[28] C.-H. Lin, D.-C. Huang, Y.-K. Chan, K.-H. Chen, Y.-J. Chang, "Fast color-spatial feature based image retrieval methods," Expert Systems with Applications, **38**(9), 11412–11420, 2011.

[29] M. Subrahmanyam, Q. M. J. Wu, R. P. Maheshwari, R. Balasubramanian, "Modified color motif co-occurrence matrix for image indexing and retrieval," Computers & Electrical Engineering, **39**(3), 762–774, 2013.

[30] G.-H. Liu, J.-Y. Yang, "Image retrieval based on the texton co-occurrence matrix," Pattern Recognition, **41**(12), 3521–3527, 2008.

[31] M. R. Gahroudi, M. R. Sarshar, "Image retrieval based on texture and color method in BTC-VQ compressed domain," in 2007 9th International Symposium on Signal Processing and Its Applications, 1–4, IEEE, 2007.

[32] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," Image and Vision Computing, **22**(14), 1211–1220, 2004.

[33] P.-W. Huang, S. K. Dai, "Image retrieval by texture similarity," Pattern recognition, **36**(3), 665–679, 2003.

[34] T.-W. Chiang, T.-W. Tsai, "Content-based image retrieval via the multiresolution wavelet features of interest," Journal of Information Technology and Applications, **1**(3), 205–214, 2006.

[35] S. Silakari, M. Motwani, M. Maheshwari, "Color image clustering using block truncation algorithm," arXiv preprint arXiv:0910.1849, 2009.

[36] M. Saadatmand-Tarzjan, H. A. Moghaddam, "A novel evolutionary approach for optimizing content-based image indexing algorithms," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), **37**(1), 139–153, 2007.

[37] G. H. Liu, L. Zhang, Y. K. Hou, Z. Y. Li, J. Y. Yang, "Image retrieval based on multi-texton histogram," Pattern Recognition, **43**(7), 2380–2389, 2010, doi:10.1016/j.patcog.2010.02.012.

[38] J.-M. Guo, H. Prasetyo, J.-H. Chen, "Content-based image retrieval using error diffusion block truncation coding features," IEEE Transactions on Circuits and Systems for Video Technology, **25**(3), 466–481, 2014.

[39] J. Kavya, H. Shashirekha, "A Novel approach for image retrieval using combination of features," International Journal of Computer Technology & Applications, **6**(2), 323–327, 2015.

[40] K. Velmurugan, L. D. S. S. Baboo, "Content-based image retrieval using SURF and colour moments," Global Journal of Computer Science and Technology, 2011.

[41] A. Bahri, H. Zouaki, "A SURF-color moments for images retrieval based on bag-of features," European Journal of Computer Science and Information Technology, **1**(1), 11–22, 2013.

[42] A. Mukherjee, S. Chakraborty, J. Sil, A. S. Chowdhury, "for Content-Based Image Retrieval," 79–87, doi:10.1007/978-981-10-2104-6.

[43] K. Kavitha, M. V. Sudhamani, "Object based image retrieval from database using combined features," Proceedings - 2014 5th International Conference on Signal and Image Processing, ICSIP 2014, **1**(12), 161–165, 2014, doi: 10.1109/ICSIP.2014.31.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1372-1380 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Effect of Starch Oxidation Degree on the Properties of Hydrogels from Dialdehyde Starch and Polyvinyl Alcohol

Jahel Desire Carrera[1], Daniela Alejandra Viteri Narváez[1], Marco Leon[2], José Francisco Alvarez-Barreto[*,1]

[1]*Universidad San Francisco de Quito, Department of Chemical Engineering, Institute for the Development of Alternative Energies and Materials, IDEMA, Biomaterials Laboratory, 170901, Ecuador*

[2]*Universidad San Francisco de Quito, Department of Mechanical Engineering, 170901, Ecuador*

ARTICLE INFO

Article history:
Received: 30 September, 2020
Accepted: 28 November, 2020
Online: 16 December, 2020

Keywords:
Dialdehyde starch
Hydrogels
Polyvinyl alcohol
Swelling
Controlled drug release

ABSTRACT

Starches have been applied as biomaterials due to their wide availability and biocompatibility. These have also been modified by oxidation, resulting in dialdehyde starch (DAS), to improve their stability in water and mechanical properties. Cassava starch with a low oxidation degree has been introduced into hydrogels based on polyvinyl alcohol (PVA) to improve their properties. However, the behavior of these materials with starch at different oxidation levels has not been previously explored. In the present work, the effect of the oxidation degree of cassava starch on the physical and chemical properties of DAS-PVA hydrogels was evaluated. To modify the degree of oxidation, different concentrations of H_2O_2 were used, and a high degree of oxidation was achieved by incorporating copper sulfate II as a catalyst. Oxidation was confirmed by quantification of carbonyl groups and Fourier Transformed Infrared Spectroscopy. Hydrogels with low and medium oxidation DAS displayed greater swelling, but also lower stability over time. Similarly, scanning electron microscopy confirmed greater porosity in them. On the other hand, hydrogels with high oxidation DAS had lower water absorption capacity, but greater stability over time. Regarding the controlled release of ibuprofen, as a model drug, hydrogels formulated with low and medium oxidation DAS presented a greater and faster release, compared to the formulations with high oxidation DAS. These results showed that the degree of starch oxidation, for the PVA-DAS hydrogel synthesis has a significant effect on the behavior of the polymeric network.

1. Introduction

Hydrogels are three-dimensional polymeric networks that are capable of containing significant amounts of water. They have been used in biomedical applications, such as tissue engineering and controlled drug release [1]. Furthermore, hydrogels have tunable biocompatibility, and respond to external stimuli such as pH, temperature, and concentrations of species, among others [2]. Thus, their porous structures are suitable for the encapsulation of drugs for a subsequent controlled release [3].

Hydrogels can be synthetized from virtually any water-soluble polymer, either natural or synthetic, such as alginic acid, pectin, carrageenan, dextran sulfate, chitosan, poly-lysine, and agarose [4]. Poly(vinyl) alcohol (PVA), is a hydrophilic, synthetic biodegradable polymer that has been widely used, due to its proved biocompatibility [5]. An additional benefit is that this polymer can be crosslinked physically, without the need of potentially toxic chemical agents . This can be achieved through freezing / thawing cycles, where the polymer in aqueous solution depending on its concentration and freeze cycles, form polymer microcrystals [1]. Nevertheless, since it is a petroleum derivative, PVA has limited availability. At the same time, in order to obtain stable structures, PVA concentrations must be high, up to 35% w/v, with molecular weights of 98000 Da or greater [2]. Therefore, its use poses an economical limitation for the creation of new technologies, especially in developing countries.

Starch, on the other hand, is one of the most abundant and cheapest polysaccharides [6]. It is comprised of two main molecules: amylose, a linear macromolecule composed of 250-300

[*]Corresponding Author: José Francisco Alvarez-Barreto, Email: jalvarezb@usfq.edu.ec

α(1→4)-linked-D-glucose, amylopectin, a less hydrophilic and branched macromolecule, and constitutes about 80% of the granular structure. To form starch hydrogels, it is possible to chemically modify its structure in order to have a greater number of functional groups [4,7]. Oxidation, for example, breaks the long glucose chains of the polymer molecules, replacing the hydroxyl groups for carbonyl and carboxyl groups . This results in reduced viscosity [8]. In drug-release applications, integrity and stability in a humid environment are required over a relatively long period of time. It has been shown that the use of dialdehyde starch (DAS), resulting from oxidation, exhibits a substantially lower solubility in water, making it more suitable for this application [9].

In [9], the authors developed protein-based films with dialdehyde starch (DAS), PVA and feather keratin. They found that if DAS was not added, an increase of the PVA concentration in the blended films resulted in a slight decrease of the film's solubility in water. By adding 5% of DAS, the total solubility mass of the film was reduced, which indicates an increase in the degree of cross-linking.

The most commonly used modifier in starch oxidation is hypochlorite, but despite being chemically efficient, it results in the formation of chlorinated products, that could potentially be toxic [10]. Several comparative studies have been carried out on the properties of starches at different degrees of oxidation with hypochlorite[10]; nevertheless, no comparative study is found in the literature on the effect that hydrogen peroxide has on different degrees of oxidation in cassava starch, and thereby in the formulations of hydrogels [11].

In [12], the authors synthesized hydrogels based on PVA and modified cassava starch as carriers for controlled drug delivery. In this research, the oxidation of starch was done with sodium hypochlorite solution containing 5% active chlorine. The hydrogels synthetized with DAS/PVA, showed an enhanced stability in water; nevertheless, the oxidation degree was maintained constant through the entire research. However, it has been demonstrated that, for other biomaterial applications, starch oxidation degree has important implications, such as bioplastics, in material behavior [13,14]. For this reason, the present study proposes to determine the effects of the oxidation degree of cassava starch on the physical and chemical properties of DAS-PVA hydrogels.

2. Materials and Methods

2.1. Chemical materials

Cassava starch was obtained from Ecuador's local stores. Hydroxylamine hydrochloride (ACS Reagent Grade, 5470-11, ACROS Organics), Poly vinyl alcohol (PVA, Mowiol 28-99, MW 145000, 99.0-99.8 mol% hydrolysis, Millipore Sigma, Cat. 10849), Sodium Phosphate, Dibasic, Anhydrous (3828-01, J.T.Baker), Potassium chloride (6858, Mallinckrodt), di-Potassium hydrogen phosphate, anhydrous GR for analysis (1.05104.1000, MERCK), Ibuprofen (≥98% GC, Millipore Sigma, Cat. I4883) were used as received.

2.2. Starch Oxidation

Starch oxidation was performed with hydrogen peroxide (H_2O_2), through adaptation of the methodology proposed by [15]. Briefly, H_2O_2 solutions at 3, 5 and 10% v/v were prepared in distilled water. For the chemical modification, a 2% w/v solution of starch in distilled water was gelatinized at 80ºC for 30 minutes, under moderate stirring. Later, the solution was cooled until room temperature, and its pH was adjusted to 7. For every 2g of gelatinized starch, 6.25mL of H_2O_2 solution were added,, drop wise during 1h at 25ºC, under continuous stirring, maintaining pH at 7.0. When the oxidation process was done, an equal volume of 70% v/v ethanol was added, and starch precipitation was allowed for 1 hour. The supernatant was discarded, and the remaining mixture was centrifuged at 1000g for 5 minutes. Five rinsing cycles were carried out with diH_2O, in a similar fashion. The resulting dialdehyde starch (DAS) was dried at 40ºC until constant weight.

To obtain a high degree of oxidation DAS, copper (II) sulfate was used as a catalyst, as proposed by [16]. Starch was gelatinized with the same methodology described above. Before adding the H_2O_2 solution at 10% v/v, the catalyst was incorporated at 0.1% w/v, and the reaction was allowed to take place for 1 hour at 25 ºC. Subsequently, precipitation and washing were done as previously mentioned.

2.3. Starch characterization

Carbonyl groups were quantified based on the method proposed by [17], with some modifications. Briefly, 1g of DAS was mixed with 25mL of distilled water, gelatinized at 80ºC under constant stirring, and then, cooled at room temperature. The pH was adjusted to 3.2, and, immediately, 3,75mL of hydroxylamine hydrochloride (25g of hydroxylamine hydrochloride in 100 mL of NaOH 0.5M in a 500mL solution) were added to the gelatinized solution, with subsequent incubation at 40ºC for 4h. A potentiometric titration followed, with 0.1M HCl until pH 3.2. The amount of carbonyl groups per 100 glucose units (CO/100GU), was calculated through equation 1:

$$\frac{CO}{100GU} = \frac{(Vb-Vs)\times M\times 0{,}028\times 100}{W} \quad (1)$$

where, Vb and Vs are the volumes (mL) of HCL used to titrate the blank (native starch) and the oxidized starch, respectively; M is the molar concentration of HCl (mol/L), and W is the weight of the sample on a dry basis (g).

On the other hand, carboxyl groups were determined using the protocol proposed by [15]. Briefly, 1g of DAS was suspended in 60mL of distilled water, gelatinized and cooled, as previously mentioned. Later, a potentiometric titration took place until pH 8.2, using NaOH 0.01M. The amount of carboxyl groups per 100 glucose units (CO/100GU), was calculated through equation 2:

$$\frac{COOH}{100GU} = \frac{(Vs-Vb)\times M\times 0{,}045\times 100}{W} \quad (2)$$

where, Vs and Vb are the titration volumes (mL) of the oxidized starch and the blank (native starch), respectively; M is the molarity of NaOH (mol/L), and W is the sample weight (g).

2.4. Hydrogel preparation

Hybrid hydrogels with Polyvinyl Alcohol (PVA) and dialdehyde starch were synthetized through a freezing/thawing method, by adaptation of the protocol reported by [2]. The

common protocol for the preparation was to completely dissolve PVA in distillated water under continuous stirring at 95ºC, at 20% w/v. In parallel, a starch suspension, at a given concentration, was gelatinized as previously mentioned. Both solutions were then mixed at a 1:1 ratio in a vortex homogenizer, with alternation with heating in a water bath at 70 ºC, for a total time of 30 min. Then, 5 mL of the polymeric mixture were poured into 1-inch petri-dishes.

Crosslinking was induced through freezing/thawing cycles; the polymeric mixture in the petri dish was frozen at -20ºC for 12 h, followed by thawing at 25ºC for 3h. Table 1 specifies the variables that were modified for the different conditions of hydrogel preparation in order to identify the formulations that could be suitable for the subsequent ibuprofen encapsulation and release. It was decided to vary the number of freeze/thawing cycles because through every cycle, there is a formation of microcrystals through the polymeric network, allowing a high crosslinking degree. Consequently, after thawing, the mixture was frozen again, and this process was repeated according to the chosen number of cycles. Since one of the objectives of this research is to partially substitute the amount of PVA in hydrogels formulations, the final concentration of the latter was maintained constant at 10% w/v. Hydrogels with native starch were used as controls.

Table 1: Variables for the synthesis of hybrid DAS/PVA hydrogels

Variable	**Values (nomenclature)**		
Oxidation degree ($\%H_2O_2$)	3 (3DAS)	5 (5DAS)	10, 10* (10DAS, 10DASC)
DAS concentration (% w/v)	5	10	15
Crosslinking cycles	1	2	3

*Copper (II) sulfate was added to obtain high oxidation DAS

According to the protocol described by [18], after crosslinking, hydrogels were cut in 5mm diameter discs, and placed in 24-well plates with 2mL of 1X phosphate buffered saline (PBS, pH 7.3), in an incubator, at 37ºC. The hydrogel weight was registered at 0, 5, 15, 60 min, 2, 3, 4, 5, 24, 48 and 72h. Water-uptake was calculated according to equation 3.

$$Water-uptake\ (\%) = \frac{(W_t - W_1)}{W_1} \times 100 \quad (3)$$

where W_1 is the initial weight of the hidrogel, and W_t its weight at a given time.

2.5. FTIR Analysis

Analyses were conducted in a Fourier transformed infrared spectrometer; model Cary 630 FTIR-ATR (Agilent Technologies) for the native and oxidized starches, as well as for the different hydrogel formulations [15]. Spectra were acquired in the region of 4000-500 cm-1.

2.6. Scanning Electron Microscopy (SEM)

Hydrogel morphological analyses were carried out in a JEOL JSM-IT300 Scanning Electron Microscope (Tokyo, Japan) with lyophilized hydrogels. Samples were cut and placed on metallic stubs with carbon tape. Images were obtained at 5kV and 50 Pa, at different magnifications..

2.7. Ibuprofen encapsulation and release

Ibuprofen was incorporated into the hydrogels at a concentration of 10 mg/mL, as proposed by [18], by adding the drug into the PVA solution at 50ºC. The same process described above for hydrogel preparation was followed. The release was performed as previously described [18]. DAS/PVA hydrogels with ibuprofen were cut into 5 mm diameter and placed in 2 mL of PBS at 37 ºC. Later, 400 µL of supernatant were collected and replaced with fresh PBS at 15, 30, 60, 120, 180, 240, 300, 360 min, 24, 48 and 72 h. Four samples of each hydrogel formulation were tested for Ibuprofen release.

Each supernatant was placed in 2.6 mL of PBS, in a quartz cuvette, and the optical density of this final solution was determined at 222 nm in a CECIL CE 2041 UV-VIS spectrophotometer. The sample concentration was calculated with a calibration curve done with known solutions of ibuprofen in PBS. The release kinetics was expressed as the percent cumulative release of ibuprofen over time.

2.8. Statistical analyses

All assays were performed with three samples of each formulation, and four in the case of ibuprofen release. The results are expressed as the average ± standard deviation. A two-way analysis of variance (ANOVA) was also applied, and multiple pairwise comparisons were carried out using Turkey-HSD, with a 95% confidence level ($p<0.05$).

3. Results and discussion

3.1. Oxidized starch characterization

Carbonyl and carboxyl contents of native and oxidized starches (dialdehyde starch, DAS) with low and high degrees of oxidation are shown in Table 2. It can be observed that carbonyl was the primary functional group produced during the oxidation process, whereas, a significantly minor amount of carboxyl groups was formed. It has been reported that the reaction mechanism of hydrogen peroxide with starch proceeds via a radical chain reaction [19] . It is more evident in dialdehyde catalyzed starch with copper sulfate, since, in the presence of the metal catalyst, the H_2O_2 is decomposed into a hydroxyl radical. Then, this radical rapidly reacts with the carbohydrate by eliminating hydrogen from a C-H group on the sugar ring, forming a R*CHOH radical, which further undergoes acid or base catalyzed rearrangement resulting in the cleavage of the glycosidic bond and formation of a carbonyl group [19].

One assumption that the authors in [19] proposed, is that the reaction between starch and hydrogen peroxide is so rapid that the major quantity of reagent was consumed for hydrogen abstraction, affording high amounts of carbonyl groups during the early period of time; therefore, less oxidant was available for the oxidation of

carbonyl to carboxyl groups. At the same time, the reaction time on the current research was stablished at one hour; which is in accordance with the literature since it has been shown that the oxidation of hydroxyl groups in starch molecules by H_2O_2 was almost completed within the first 30 minutes.

Table 2: Carbonyl and carboxyl contents of dialdehyde cassava starch at different oxidation levels

H_2O_2 Concentration (%v/v)	Carbonyl content [CO/100 GU]	Carboxyl content [CO/100 GU]
3	0.255 ± 0.091	0.046±0.010
5	0.419 ± 0.124	0.034±0.009
10	0.209 ± 0.108	0.052±0.010
10 (Catalyzed)	2.875 ± 0.089	0.025±0.003

GU: Glucose units

Oxidation degree is determined by carbonyl and carboxyl content. The relatively bulky carbonyl and carboxyl groups are introduced with a partial depolymerization of the starch chains; for this reason, the determination of these values shows the degree of oxidation degree. A major carbonyl content is an indication of a higher oxidation degree [20].In this study, a low degree of oxidation will be considered as having a carbonyl content of less than 1CO/100GU; thus, all starches oxidized without catalyst will be within this category.

3.2. Hydrogel morphological analyses

It was important to evaluate hydrogel morphology as it can have important implications in swelling behavior and controlled released, for which a porous structure is desired. Lyophilized hydrogels with the same compositions, but different DAS are presented in Figure 1. Natural porous materials have no uniform porosity; rather, it is distributed allowing to maximize the overall performance of the structure [21]. For drug delivery applications, a balance between adequate porosity and mechanical stability is required.. At comparing the different images, it is possible to recognize that hydrogels with low oxidation degree have flatter and larger pores; while, in high oxidation DAS, the pores seem to be smaller and rounded. There is clear indication that pore arrangement is altered by the incorporation of starch, and that the oxidation degree further influences pore morphology and structure. Additionally, PVA only hydrogels (without starch) appeared to less porous.

Hybrid hydrogels composed with hydrophilic-hydrophobic polymers have been reported previously in literature. It has been shown that one of the major issues in obtaining porosity is due to the intrinsic immiscibility of hydrophilic (i.e. PVA) and hydrophobic (i.e. DAS) polymers, which, in turn, affect pore structure, and potentially, distribution..

3.3. FTIR analysis

The Fourier transformed infrared absorption spectra of lyophilized hydrogels composed with 10% w/v PVA and 10% w/v starch is presented in Figure 2. In the case of FTIR spectra for different hydrogel formulations with oxidized starch, there is a decreased intensity. As previously reported in the literature, this lower intensity corresponds to crosslinked polymer structures, probably associated with a major bond vibration restriction [22]. A stretching band can be observed at 2850 cm^{-1} due to C-H broad alkyl. Another lower intensity stretching is because of the strong-OH group band for free unreacted alcohol and hydrogen bonded bands at 3600-3200 cm-1, which are associated with freeze-thawed PVA [23]. As expected, the peak of dialdehyde starch appears at 1630 cm^{-1} - 1710 cm^{-1}, due to C=O and C=C stretching frequency groups that are present in DAS [24].

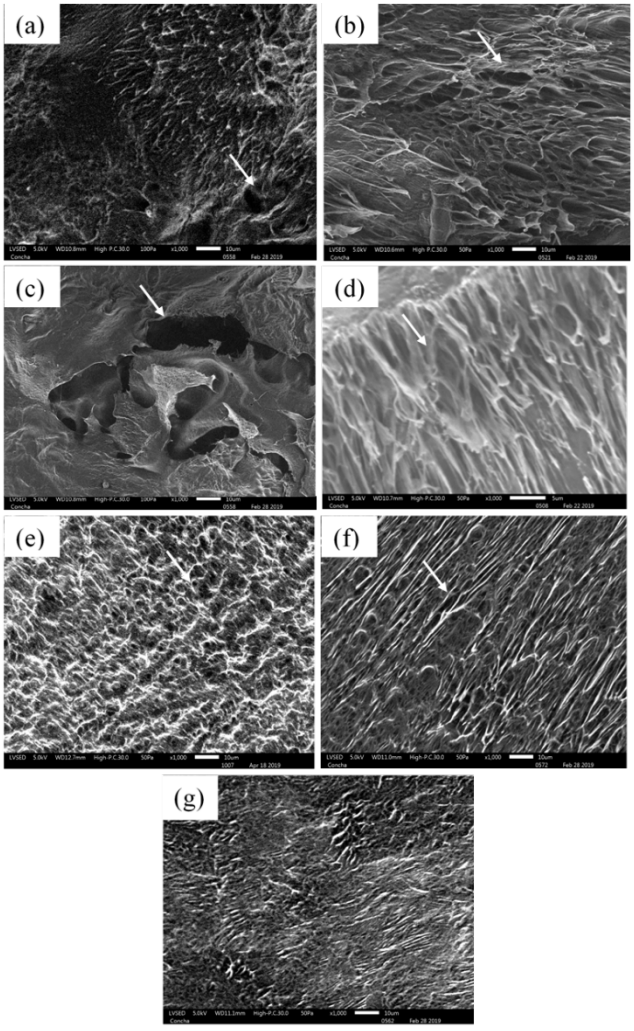


Figure 1: Scanning electron micrographs of lyophilized hydrogels composed of 10% w/v PVA, 10% w/v Starch and 1 crosslinking cycle: (a) NS, (b) 3DAS, (c) 5 DAS, (d) 10 DAS, (e) 10DASC, (f) Solely 20% w/v PVA and (g) Solely 10% w/v PVA. Calibration bar: 10 µm. Figure 1d was reported at 3000X (Calibration bar: 5 µm) in order to better observe the porous structure.

When comparing the spectra of hydrogels composed solely with PVA to hybrid hydrogels, there is lower intensity absorbance in the different bands; nevertheless, those associated with -OH groups, could be related to a strong interaction between both polymers forming a tight network, as suggested by [22]. PVA addition enhances the elasticity of hydrogels, which is corroborated by the increased intensity of the C-H stretching

absorption band at 2031 cm-1, and the C-C stretching vibrational band at 1200-1400 cm^{-1} [25].

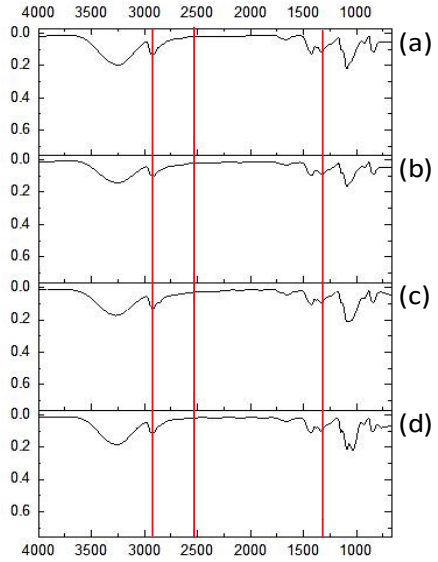


Figure 2: FT-IR spectra of lyophilized hydrogels formulations with 10% w/v PVA, 10% w/v starch and 1 crosslinking cycle: (a) Native starch, (b) 5DAS, (c) 10DAS and (d) 10DASC.

3.4. Hydrogel Water-uptake

Hydrogel water-uptake, or swelling ratio, is one of their most important characteristics, since this parameter is directly related to drug release mechanisms, as well as structure stability in aqueous media [1]. Water uptake capacity (%) of the different hydrogel formulations specified in table 1 are presented in Figure 3 through 5, where PVA composition remained constant at 10%. All assays where performed at 37ºC in PBS to emulate physiologic conditions. Since one of the objectives of this research was to partially substitute the weight/volume percentage of PVA in hydrogels formulations, it was necessary to compare the water-uptake percentage vs. time, according to hydrogels composed solely with PVA. In general, it was observed that 1 crosslinking cycle presented the largest swelling capacity; nevertheless, they showed the lower stability in time, with important swelling reduction toward the end of the assay. Two and three crosslinking cycles exhibited similar behaviors. At the same time, by increasing the quantity of PVA, hydrogels displayed greater stability (data not shown).

According to the analysis of variance, hydrogel formulation (both starch content and oxidation degree), number of crosslinking cycles and time had significant effects on water uptake ($p<0.05$). Figure 3a shows the swelling kinetics of hydrogels with 15% starch. There were no statistical differences between native and low oxidation DAS (3, 5 and 10DAS), which displayed fast water uptake, with important fluctuations over time with one crosslinking cycle. This could be due to degradation effects that, ultimately, become relevant, as the swelling decrease after a maximum at 48h. When increasing the number of cycles (Figure 3a and b), however, these fluctuations become less apparent, and decreases in water uptake are lower. These changes could indicate an improvement in hydrogel stability. In the case of 10DASC, on the other hand, had rapid swelling and seemed to reach equilibrium after 15h. This high oxidation DAS resulted in significantly lower swelling than the other formulations; moreover, this parameter tends to decrease with greater number of crosslinking cycles.

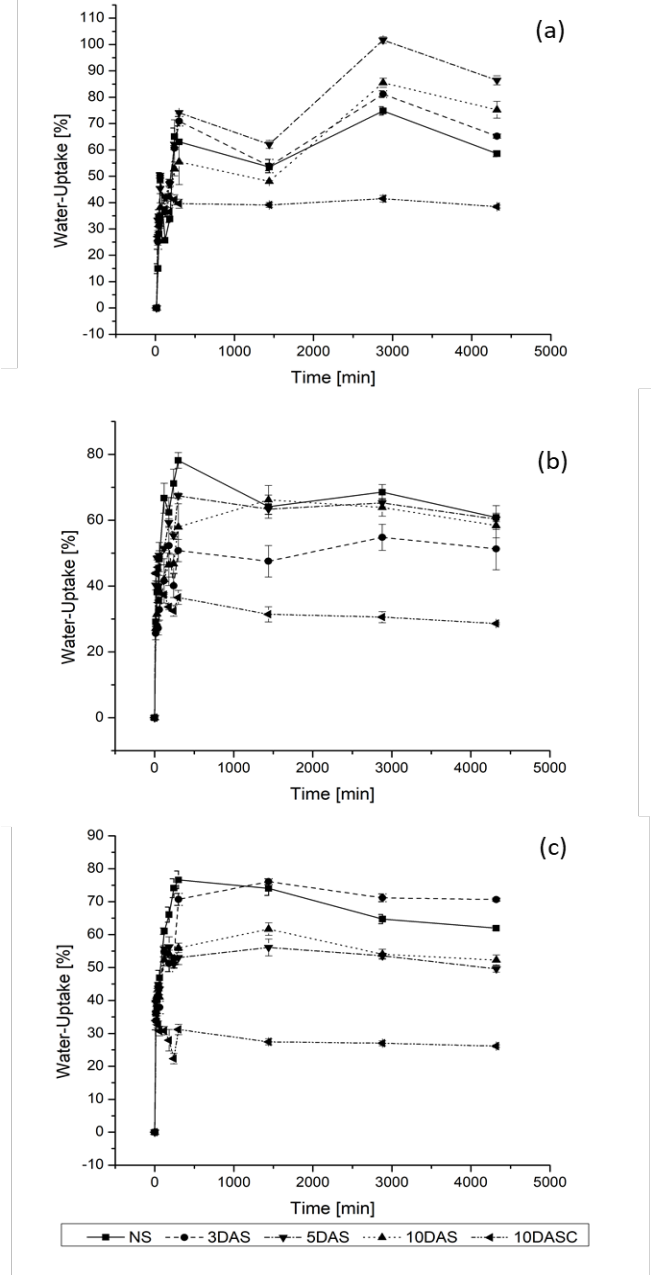


Figure 3: Swelling kinetics, expressed as water uptake percentage in time, for hydrogels composed of 10% w/v PVA and 15% w/v starch, either native (NS) of dialdehyde (3DAS, 5DAS, 10DAS and 10DASC). Formulations with (a) one crosslinking cycle, (b) two crosslinking cycles, and (c) three crosslinking cycles.

Figure 4 shows water-uptake capacity for hydrogels with 10% starch, in which, after an analysis of variance, it is determined that formulation and time affect water-uptake percentage ($p<0.05$). Once again, the major water-uptake is obtained with one freeze-thaw cycle. The stability in this case is significantly reduced in comparison with Figure 3(a), where hydrogel degradation, regardless of formulation, begins at 240 min until 72h. After the

statistical analysis, it was confirmed that oxidation degree had a significant effect on swelling ratio ($p<0.05$); nevertheless, there is not a clear pattern in the way the amount of carbonyl groups affects the swelling degree.

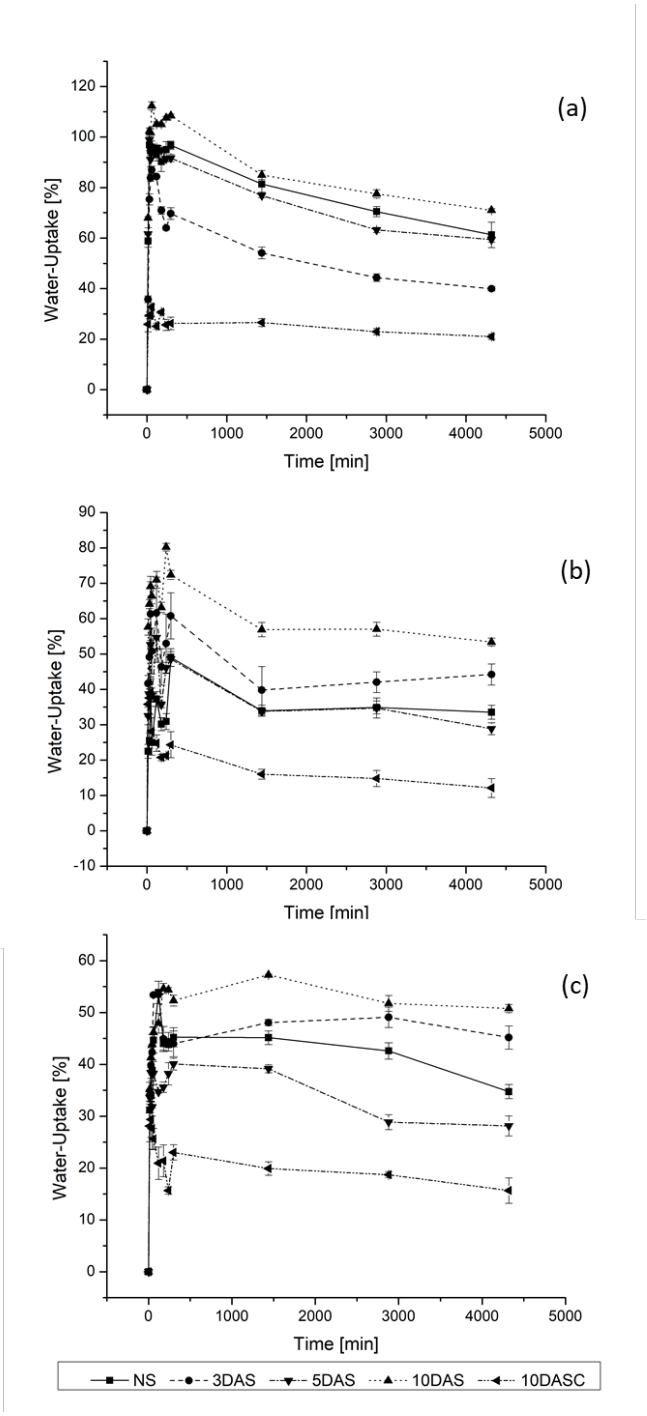


Figure 4: Swelling kinetics, expressed as water uptake percentage in time, for hydrogels composed of 10% w/v PVA and 10% w/v starch, either native (NS) of dialdehyde (3DAS, 5DAS, 10DAS and 10DASC). Formulations with (a) one crosslinking cycle, (b) two crosslinking cycles, and (c) three crosslinking cycles.

Figures 4(b) and (c) show a more accurate pattern within the effect of DAS, where starch with lower carbonyl content presents the greater swelling degree. After a Tuckey comparison on values of Figure 4(b), all the formulations are significantly different between them; nevertheless, the greatest water uptake was obtained in formulations with 3DAS and 10DAS, which agrees with lower carbonyl content. At three freeze-thaw cycles, Figure 4(c), Tukey analysis shows that there is not a difference between formulations with 3% DAS and 10% DAS. In all cases, the way time affects swelling degree, establishes an equilibrium tendency up to 72h. Finally, the number of cycles does affect hydrogel response, being inversely proportional to the swelling degree. However, the swelling degree within a time greater than 24 h, is not statistically different between formulations with 2 and 3 freeze-thaw cycles.

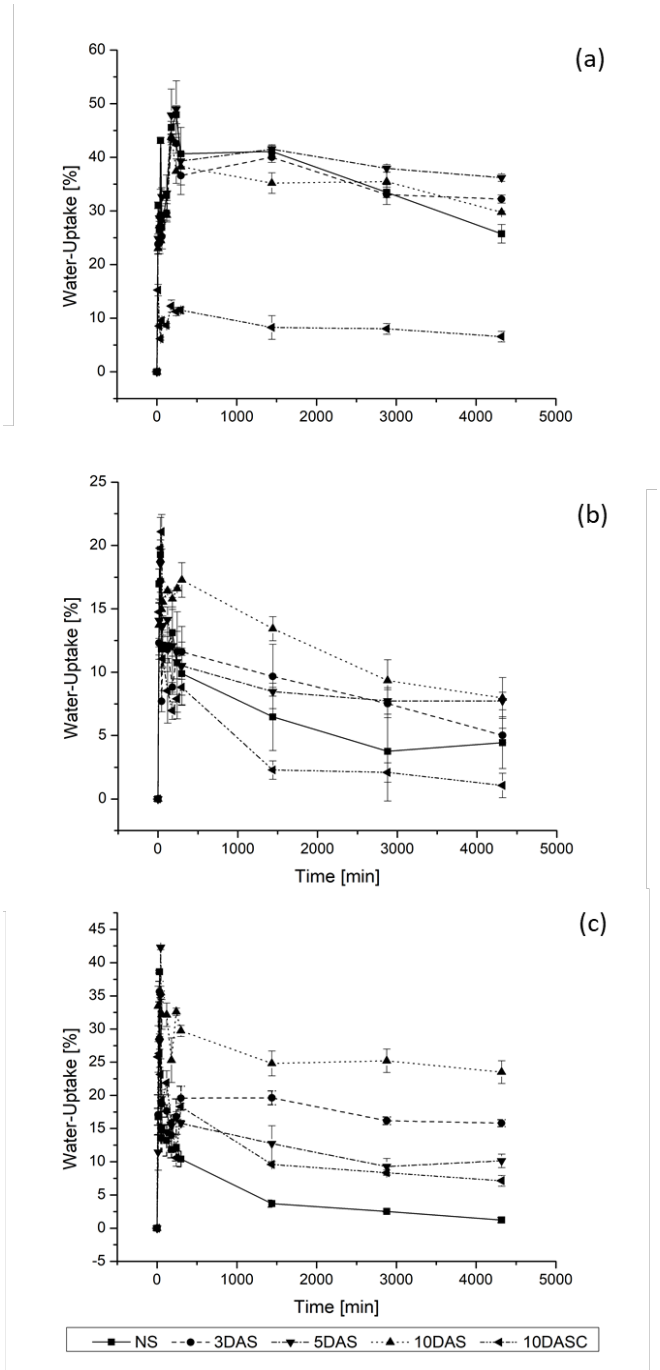


Figure 5: Swelling kinetics, expressed as water uptake percentage in time, for hydrogels composed of 10% w/v PVA and 5% w/v starch, either native (NS) of dialdehyde (3DAS, 5DAS, 10DAS and 10DASC). Formulations with (a) one crosslinking cycle, (b) two crosslinking cycles, and (c) three crosslinking cycles.

By comparing hydrogels with 15% and 10w/v starch (Figures 3 and 4, respectively), there are clear differences. The greater water-uptake in the first case is presented at 300 min; while, for the latter, this value is obtained at 180 min. There is a major hydrogel degradation within the first 24 h at a lower starch content. These results suggest that oxidized starch content affects swelling.

Finally, swelling kinetics for hydrogels formulations with 5% starch are shown in Figure 5. Considering factors such as formulations and number of cycles, swelling degree in all cases was considerably lower, compared with 15 and 10% starch (Figures 3 and 4), due to the decrease of polymer content in the structure. The statistical analysis resolved that there is not a significant variation on swelling, considering the formulation of the hydrogel as a swelling response. At this low concentration of starch, there is not a clear pattern on the effect of the oxidation degree on swelling, which could be due to a predominant effect of PVA, independent of the starch´s characteristics.

Literature shows that starch oxidation affects the surface in the granule by increasing the porosity and presence of holes in the structure [20,26]. This could explain the reason behind a large water-uptake in hydrogels formulations with dialdehyde starch. Even if there is a significant presence of hydrophobic groups in modified starch, this porosity helps hydrogels to increase not only their swelling degree, but also their stability during a period of time.

After the statistical analysis, it was possible to show that there is a major increasing swelling degree during the first 5 hours in hydrogels with a lower carbonyl content; which agrees with previous literature reported by [27]. According to previous works, higher PVA composition would result in hydrogels with lower swelling ratio due to its higher degree of crystallinity [25]. Nonetheless, the opposite effect occurs after the addition of dialdehyde starch. In [28], the authors reported that oxidation occurs mainly in the amorphous lamella of the semi-crystalline growth rings in starch granules. They evaluated the effect of different degrees of oxidation with sodium hypochlorite on bean starch. A higher carbonyl content in dialdehyde bean starch resulted in a decrease of the crystallinity, suggesting that the amylopectin chains were damaged and therefore, there is a degradation of the crystalline region. The conclusion proposed in the current research behind the large water-uptake of hydrogels composed with PVA and dialdehyde starch, is due to the significant presence of less crystalline polymer.

An important fact about dialdehyde starch at a high oxidation degree was evidenced through a Tukey analysis. The way in which time affects swelling degree, in all cases, follows a tendency in which water uptake within the first hour is rather high. Suddenly, there is a significant de-swelling until 240 min. Beyond this time, hydrogels exhibit stability. In [10], the authors reported that at higher carbonyl content, the crystallinity decreases, and at a certain point, starch structure becomes amorphous. It can be assumed that this de-swelling is due to the less crystallinity arrangement, even though, there is stability during times greater than 240 min.

Finally, swelling degree is related with the porosity of hydrogels, evidenced through the scanning electron micrographs of lyophilized hydrogels. As presented in Figure 1, oxidation degree does affect the morphology of hydrogels, by allowing flatter and seemingly larger pores in the case of modified starch at a low oxidation degree. The behavior according to porosity and swelling degree is as expected, since 10DAS presents not only large porosity, but also greater water-uptake; unlike, 10DASC, which exhibits a lower porosity and so was the swelling degree with these formulations.

3.5. Ibuprofen release

Hydrogels have sparked interest in drug delivery applications due to their porous structures, that assists drug loading in the polymeric matrix, for a subsequent sustained release. In the present work, ibuprofen was used as a model drug that can give insights into potential applications for other non-steroidal anti-inflammatory drugs (NSAIDs). In this case, control hydrogels were done at 20% PVA, since those at 10% were rather unstable and disintegrated quickly. According to ANOVA, ibuprofen release was significantly affected by composition, time and number of freeze-thaw, or crosslinking, cycles. Release assays are presented in Figure 6.

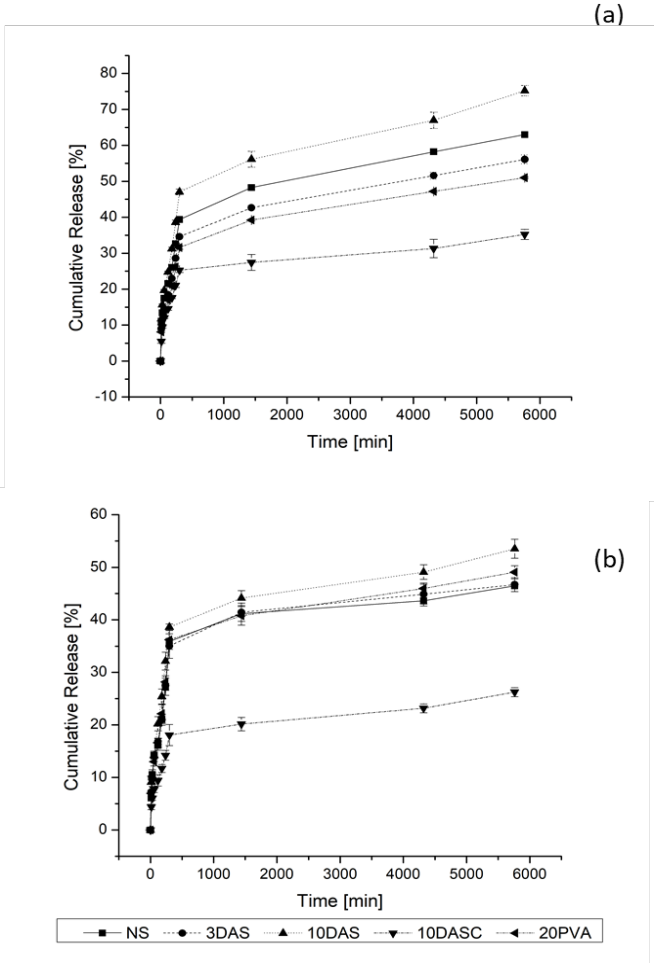


Figure 6: Ibuprofen cumulative release profiles from hydrogels with 10% /v PVA and 10% starch at different oxidation levels. Only PVA at 20% was used as control. (a) One crosslinking cycle and (b) Two crosslinking cycles.

For all formulations, there was a fast release in the first 6 hours, followed by small, not statistically significant, increases, indicating a tendency toward equilibrium. In hydrogels with one crosslinking cycle (Figure 6a), there is a very similar ibuprofen

profile in every hydrogel formulation. After Tukey comparison, it was shown that formulations with native starch, 3DAS, 10DAS and solely 20% PVA have no significant differences between them. Interestingly, hydrogels with native starch and low oxidation DAS displayed greater swelling, and some degradation that played an important role in the release. Hydrogels with 10DASC, on the other hand, did present significantly lower cumulative release at every time point. All these previous formulations exhibit the same behavior; the largest drug release was performed within the first 6 hours of contact with PBS; and, there is an equilibrium tendency at 72h.

Values of ibuprofen release in hydrogels with two crosslinking cycles, Figure 6b, were significantly affected by formulations and time. Again, Tukey comparisons indicated that formulations with 3DAS, 10DAS and solely 20% PVA are similar among them, representing the formulations with largest drug release percentage. Similar to what occurred with one cycle, 10DASC resulted in hydrogels with the lowest levels of release. This time, nonetheless, values of cumulative release were lower than those found with one crosslinking cycle, which relates to the lower swelling capacities found at greater number of freeze-thaw cycles. Moreover, there was a tendency to a more sustained release, possible due to the greater stability provided by more crosslinking

When comparing Figures 6a and b, it is possible to affirm that as presented in literature, the high-water content of hydrogels results in an increased release of drugs from the polymeric matrix [3]. This is confirmed by the statistical analysis. The tendency in both cases is that hydrogels with 10DAS exhibit the largest drug release, and the lowest is attributed to those with 10DASC.

In hydrogels, the rate of drug diffusion through the polymeric structure depends on the extent of cross-linking. Converting the ensemble of individual polymer chains into a macromolecular network allows the water diffusion into the interstices of the network. Comparing Figures 6a and b, there is indication that the extent of cross-linking minimizes the swelling and, thereby, the diffusion of the drug. The reason behind this behavior is due to the fact that ibuprofen is an entrapped drug, and the rate of diffusion depends on the interchain separation and the size of the diffusing drug [29].

4. Conclusions

The oxidation degree of cassava starch has important effects on the properties of hydrogels with PVA, particularly at larger carbonyl contents. At a low oxidation degree obtained with 10% v/v H_2O_2 and 3% v/v H_2O_2, hydrogels showed the major suitable characteristics of hydrogels for drug delivery applications, since the water-uptake percentage was higher, as it was drug release. Considering hydrogels formulations with DAS catalyzed with copper sulfate (II), and the largest carbonyl content, there was a significant reduction of water-uptake through the polymeric structure, due to strong intermolecular bonds, which resulted in a reduced swelling degree, but, as expected, those showed the major stability during time, proving an improved stability in physiological environment. It is possible that, for the latter formulation, the hydrogels could be more appropriate for tissue engineering because, given that they do not swell significantly, their mechanical properties could be superior to those that presented higher swelling ratios. Additionally, being more stable in time, they could potentially meet mechanical demands for longer periods, and we recommend doing this assessment in future studies.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was funded by the PoliGrant program at Universidad San Francisco de Quito. Publication of this article was funded by the Academic Articles Publication Fund of Universidad San Francisco de Quito USFQ

References

[1] A.S. Hoffman, "Hydrogels for biomedical applications," Advanced Drug Delivery Reviews, **64**(SUPPL.), 18–23, 2012, doi:10.1016/j.addr.2012.09.010.

[2] J.L. Holloway, A.M. Lowman, G.R. Palmese, "The role of crystallization and phase separation in the formation of physically cross-linked PVA hydrogels," Soft Matter, **9**(3), 826–833, 2013, doi:10.1039/c2sm26763b.

[3] T.R. Hoare, D.S. Kohane, "Hydrogels in drug delivery: Progress and challenges," Polymer, **49**(8), 1993–2007, 2008, doi:10.1016/j.polymer.2008.01.027.

[4] E.A. Kamoun, "N-succinyl chitosan-dialdehyde starch hybrid hydrogels for biomedical applications," Journal of Advanced Research, **7**(1), 69–77, 2016, doi:10.1016/j.jare.2015.02.002.

[5] S. Jiang, S. Liu, W. Feng, "PVA hydrogel properties for biomedical application," Journal of the Mechanical Behavior of Biomedical Materials, **4**(7), 1228–1233, 2011, doi:10.1016/j.jmbbm.2011.04.005.

[6] Y. Zhu, C. Romain, C.K. Williams, "Sustainable polymers from renewable resources," Nature, **540**(7633), 354–362, 2016, doi:10.1038/nature21001.

[7] H. Hu, F.J. Xu, "Rational design and latest advances of polysaccharide-based hydrogels for wound healing," Biomater Sci., **8**(8), 2084–2101, 2020, doi:10.1039/d0bm00055h.

[8] A.R.G. Dias, E.D.R. Zavareze, E. Helbig, F.A. De Moura, C.G. Vargas, C.F. Ciacco, "Oxidation of fermented cassava starch using hydrogen peroxide," Carbohydrate Polymers, **86**(1), 185–191, 2011, doi:10.1016/j.carbpol.2011.04.026.

[9] Y. Dou, B. Zhang, M. He, G. Yin, Y. Cui, I.N. Savina, "Keratin/polyvinyl alcohol blend films cross-linked by dialdehyde starch and their potential application for drug release," Polymers, **7**(3), 580–591, 2015, doi:10.3390/polym7030580.

[10] Y.R. Zhang, X.L. Wang, G.M. Zhao, Y.Z. Wang, "Preparation and properties of oxidized starch with high degree of oxidation," Carbohydrate Polymers, **87**(4), 2554–2562, 2012, doi:10.1016/j.carbpol.2011.11.036.

[11] V.M. de Oliveira Cardoso, B. Stringhetti Ferreira Cury, R.C. Evangelista, M.P. Daflon Gremião, "Development and characterization of cross-linked gellan gum and retrograded starch blend hydrogels for drug delivery applications," Journal of the Mechanical Behavior of Biomedical Materials, **65**, 317–333, 2017, doi:10.1016/j.jmbbm.2016.08.005.

[12] A.P. Erazo Bastidas, J. Alvarez Barreto, Synthesis and Characterization of Hydrogels based on Poly (vinyl alcohol) and Modified Starch as Carriers for Controlled Drug Delivery, Quito, 2018.

[13] Y.-R. Zhang, X.-L. Wang, G.-M. Zhao, Y.-Z. Wang, "Influence of oxidized starch on the properties of thermoplastic starch," Carbohydrate Polymers, **36**(1), 358–364, 2013, doi:10.1016/j.carbpol.2013.03.093.

[14] Y. Yu, Y. Wang, W. Ding, J. Zhou, B. Shi, "Preparation of highly-oxidized starch using hydrogen peroxide and its application as a novel ligand for zirconium tanning of leather," Carbohydrate Polymers, **174**, 823–829, 2017, doi:10.1016/j.carbpol.2017.06.114.

[15] P. Parovuori, A. Hamunen, P. Forssell, K. Autio, K. Poutanen, "Oxidation of Potato Starch by Hydrogen Peroxide," Starch - Stärke, **47**(1), 19–23, 1995, doi:10.1002/star.19950470106.

[16] Y. Lu, Q.M. Kong, R. Jing, X. Hu, P.X. Zhu, "Solid state oxidation of polyvinyl alcohol by hydrogen peroxide-Cu (II)," Polymer Degradation and Stability, **98**(6), 1103–1109, 2013, doi:10.1016/j.polymdegradstab.2013.03.022.

[17] R.J. Smith, Characterization and analysis of starches, Academic Press, New

York and London: 620–625, 1967.
[18] P.M. Castro, P. Fonte, A. Oliveira, A.R. Madureira, B. Sarmento, M.E. Pintado, "Optimization of two biopolymer-based oral films for the delivery of bioactive molecules," Materials Science and Engineering C, **76**, 171–180, 2017, doi:10.1016/j.msec.2017.02.173.
[19] K. Sangseethong, N. Termvejsayanon, K. Sriroth, "Characterization of physicochemical properties of hypochlorite- and peroxide-oxidized cassava starches," Carbohydrate Polymers, **82**(2), 446–453, 2010, doi:10.1016/j.carbpol.2010.05.003.
[20] L.M. Fonseca, J.R. Gonçalves, S.L.M. El Halal, V.Z. Pinto, A.R.G. Dias, A.C. Jacques, E. da R. Zavareze, "Oxidation of potato starch with different sodium hypochlorite concentrations and its effect on biodegradable films," LWT - Food Science and Technology, **60**(2), 714–720, 2015, doi:10.1016/j.lwt.2014.10.052.
[21] Q.L. Loh, C. Choong, "Three-dimensional scaffolds for tissue engineering applications: Role of porosity and pore size," Tissue Engineering - Part B: Reviews, **19**(6), 485–502, 2013, doi:10.1089/ten.teb.2012.0437.
[22] O. Moreno, J. Cárdenas, L. Atarés, A. Chiralt, "Influence of starch oxidation on the functionality of starch-gelatin based active films," Carbohydrate Polymers, **178**, 147–158, 2017, doi:10.1016/j.carbpol.2017.08.128.
[23] E.R. Kenawy, E.A. Kamoun, M.S. Mohy Eldin, M.A. El-Meligy, "Physically crosslinked poly(vinyl alcohol)-hydroxyethyl starch blend hydrogel membranes: Synthesis and characterization for biomedical applications," Arabian Journal of Chemistry, **7**(3), 372–380, 2014, doi:10.1016/j.arabjc.2013.05.026.
[24] A.L. Da Róz, M.D. Zambon, A.A.S. Curvelo, A.J.F. Carvalho, "Thermoplastic starch modified during melt processing with organic acids: The effect of molar mass on thermal and mechanical properties," Industrial Crops and Products, **33**(1), 152–157, 2011, doi:10.1016/j.indcrop.2010.09.015.
[25] S.C. Pang, S.F. Chin, S.H. Tay, F.M. Tchong, "Starch-maleate-polyvinyl alcohol hydrogels with controllable swelling behaviors," Carbohydrate Polymers, **84**(1), 424–429, 2011, doi:10.1016/j.carbpol.2010.12.002.
[26] D. Kuakpetoon, Y.J. Wang, "Locations of hypochlorite oxidation in corn starches varying in amylose content," Carbohydrate Research, **343**(1), 90–100, 2008, doi:10.1016/j.carres.2007.10.002.
[27] A. V. Reis, M.R. Guilherme, T.A. Moia, L.H.C. Mattoso, E.C. Muniz, E.B. Tambourgi, "Synthesis and characterization of a starch-modified hydrogel as potential carrier for drug delivery system," Journal of Polymer Science Part A: Polymer Chemistry, **46**(7), 2567–2574, 2008, doi:10.1002/pola.22588.
[28] N.L. Vanier, E. Da Rosa Zavareze, V.Z. Pinto, B. Klein, F.T. Botelho, A.R.G. Dias, M.C. Elias, "Physicochemical, crystallinity, pasting and morphological properties of bean starch oxidised by different concentrations of sodium hypochlorite," Food Chemistry, **131**(4), 1255–1262, 2012, doi:10.1016/j.foodchem.2011.09.114.
[29] L. Djekic, M. Martinović, V. Dobričić, B. Čalija, Đ. Medarević, M. Primorac, "Effect of Bioadhesive Polymers on Stability and Drug Release Kinetics of Biocompatible Hydrogels for Topical Application of Ibuprofen.," J Pharm Sci., **108**(3), 1326–1333, 2019, doi:10.1016/j.xphs.2018.10.054.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1381-1386 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Image Tag Recommendation based on Ranked Categorical Nearest Neighbors and Weighted Tag Features

Anupama D. Dondekar*, Balwant A. Sonkamble

Department of Computer Engineering, Pune Institute of Computer Technology, Pune, 411042, India

ARTICLE INFO

Article history:
Received: 16 October, 2020
Accepted: 24 November, 2020
Online: 16 December, 2020

Keywords:
Tag Recommendation
Categorical Nearest neighbor
Weighted Tag Frequency
Flickr

ABSTRACT

The huge number of images on the image sharing websites poses challenges for classification and retrieval of the images. On many image sharing websites, tags can be assigned by the users to an image that describes the contextual and visual description of an image. However, ambiguous or incorrect tags have appeared in frequent tags that affect the performance of an image retrieval system. Thus, assigning appropriate tags to the images plays a very important role in image retrieval and classification. In this paper, the ITR-WTF image tag recommendation method is proposed which explores tags from ranked nearest neighbors of each category. For a given input image, the method first determines the neighbors from training images of each category and ranks the neighbors according to the distance from the input image. In the second step, the weight is assigned to each tag based on the vote from each neighbor. Finally, the weighted tag frequency is determined to recommend appropriate tags to a given image. The experimentation is done on two datasets self-generated and NUS-WIDE. The results obtained using the proposed method ITR-WTF gives good results as compared with the existing methods of tag recommendation.

1. Introduction

The rapid development of advanced technology and the high usage of social media have created a large repository of images which poses many challenges for an image retrieval system. On many image sharing social websites, the users are uploading images for faster communication or to find the people with the same interest. The images on the image sharing social websites are associated with the tags which are assigned by the users. The tags describe the visual content of the images along with the context information such as the location and time the images are captured. These tags are used indexing during an image search. Thus, tags play a very important role in tag-based image retrieval systems and classification.

Many times, users assign tags that are imprecise and irrelevant to the image. According to the survey, only 50% of the Flickr tags describe the content of the images [1]. The presence of irrelevant tags affects the performance of the classification and image retrieval system. Hence, it is necessary to design an algorithm that assigns correct tags to the images. It improves the accuracy of the image retrieval system by suggesting proper tags for the images. The advantages of tag suggestions are: it reduces the cost of manual annotation of the images, spelling mistakes. Several studies have been done on tag recommendation using visual contents, tags and metadata. Still, the performance of tag recommendation is not satisfactory due to personalized tags as shown in Figure 1. The tags may not describe the content of the images and may be assigned as per the user perspective. The tags Denmark, 2011 does not describe visual content of an image as shown in Figure 1 (a).

(a) 

Tags:

Denmark, 2011, nature, D60

(b) 

Tags:

Fish, Canon 50mm, Canon EOS 5D Mark III

Figure 1: Tag based Images

*Corresponding Author: Anupama D. Dondekar, Email: agphakatkar@pict.edu

In this paper, the ITR-WTF image tag recommendation method is proposed which contributes in the following way:

- Obtained ranked visual neighbor of each category for each feature
- The tag score is calculated by combining weighted tf-idf value and vote from rank categorical nearest neighbor to recommend top k tags

The paper is organized as follows: existing methods for tag recommendation are described in section 2. Section 3 explains the ITR-WTF proposed method for image tag recommendation. Section 4 and 5 describe the metric used for evaluating the performance of the proposed method and dataset used for experimentation. Section 6 describes the experimental results. Finally, in section 7 the conclusion and future work is given.

2. Work Done

The methods for tag recommendation are categorized into three methods: classification, semantic and nearest neighbor based.

In classification-based methods, the features of the images are extracted and classified into a category using a classification algorithm. The classifier trained for multiple class result in a multi-class problem. Finally, the tags are recommended based on the category of an input image. In [2] the method was proposed labeling of the images. The images were segmented into regions and identified salient region. Based on particle swarm optimization, the SVDD trained to assign labels to the images by assigning more weights to the salient regions. In [3] the method was proposed using probability and weight based SVM classifier for annotation of images. Given an image with tag, the method identifies the related and unrelated images using majority of voting from SVM in [4]. The system named 'SheepDog' was developed to which identifies the suitable group for the inclusion of photos and suggest appropriate tags to the users on Flickr dataset in [5].

In the semantic method, the tags are recommended based on the joint distribution of image and tag features. The method was proposed in [6] for annotation of images using KCCA framework by constructing semantic space in which the correlation was built between image feature and tag features. In [7] an approach for retagging of social images with diverse semantics was presented. Both the relevance of a tag and the semantic compensation to the already determined tags was fused to determine the final tag list for a given image. The method was proposed in [8] to recommend tags for geotagged images using unified subspace which correlates the textual and visual features. The hyper-graph-based method was proposed for tag-based image retrieval using image features and tag features simultaneously [9].

The nearest neighbor-based tag recommendation methods are very popular due to its effectiveness. The model is based on the assumption that feature based similar images tend to have the same tags. Given an input image, the method determines k nearest neighbors by combining various features either by using early or late fusion and calculates the tag relevance score by collecting votes of a tag from its nearest neighbors. The advantages of the method are: it is scalable and model building is not needed as it makes an assumption based on training data. The tag recommendation method was proposed in [10] using random walk on bipartite graph constructed based on weighted user and image nearest neighbor. In [11], the Bayesian based image annotation model was proposed based on semantic nearest neighbors. The method was proposed in [12] for annotation of images using a variation of traditional kNN algorithm by defining matrix which shows the relationship between labels and images. In [13] the method was proposed in which the given an image the similar images were determined using k nearest neighbor and tag graph was created from tags of neighbors and clustered to assign label to an image. The personalized image tag recommendation method was proposed based on neighbor voting scheme by building tripartite graph to show relationship between user, tags and images in [14]. The VS-KNN method was proposed in [15] for image labeling by exploring image features and label features simultaneously as a maximum posterior estimation. Given an image with label l, the method was proposed which identify the images labeled with l using kNN and denote it by set S [16]. Finally, the labels were assigned to the images by calculating the similarity between the input image and set S. The tag relevance method was proposed by assigning weight to each neighbor based on distance using kNN [17]. The method was proposed in [18] to suggest tags for an image based on visual features and tag correlations using neighbor voting scheme. Photo tagging method was proposed in [19] using history of the users. The method finds the geographical, visual and time neighbors for a given image and recommends tags by counting accumulating votes for each tag from three types of neighbors. In [20], the method was proposed to suggest tags to the images with and without labels. The method first identifies a set of k images using feature-based k nearest neighbor. Finally, assign or recommends the tags by counting the difference between tag frequency count from the entire database and k neighbor. The image annotation algorithm in [21] identifies rank based and weight based nearest neighbor and suggests the label for an input image using a probabilistic model.

In this paper, the proposed method recommends tags based on the nearest neighbor method. Compared with the classification and semantic based method, the nearest neighbor method is popular due to its effectiveness and scalability. Also, it does not require any training.

However, the performance of the existing nearest neighbor-based tag recommendation methods depends on the number of k neighbors and may be affected due to the equal weight and voting irrespective of the class. This observation motivated us to develop a method which improves the performance of tag recommendation. The method first determines the weighted categorical nearest neighbors and suggest the tags based on the ranked categorical neighbor and weighted tag frequency.

3. Research Methodology

In this section, ITR-WTF proposed tag recommendation algorithm is described. The main objective of the proposed algorithm is to improve the accuracy of tag recommendation by identifying distance based nearest neighbor and rank the tags by combining image score and tag score.

3.1. ITR-WTF Tag Recommendation Method

The block diagram of the proposed method for tag recommendation is shown in Figure 2. The proposed method consists of three main modules: feature representation, classification and tag recommendation.

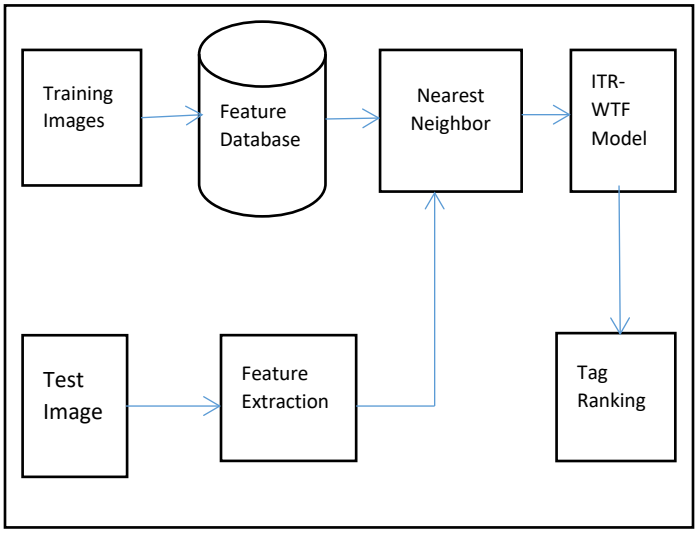


Figure 2: System Diagram of Proposed Method

3.1.1. Feature Representation:

Features play a very important role in image representation. Several researchers have worked on feature extraction methods for image retrieval using color, texture and shape features. Extracting an effective feature and represent them efficiently is very important.

Color is the most used feature in an image retrieval system. Using color features, a human can recognize most images and objects included in the images. Also, the color features are invariant to scaling, translation and rotation of an image. Another important feature is texture. Texture measures look for visual patterns in images and how they are spatially defined.

During the training phase, the features are extracted using color moment and wavelet packet transform [22, 23]. The first, second and third moments are extracted as color features in L*a*b* color space by segmenting an image into two by two sub-block along with a centralized sub-block of the same size resulting into 9 features for each sub-block. For texture features are extraction using wavelet packet transform, an image is decomposed into sub-bands up to level three using daubechies wavelet. The energy and standard deviation of each band are determined as texture features using eq.1 and 2.

$$M_i = \frac{1}{HT*WT} \sum_{h=1}^{HT} \sum_{w=1}^{WT} |Coeff_i(h,w)| \quad (1)$$

$$SD_i = \sqrt{\frac{1}{HT*WT} \sum_{h=1}^{HT} \sum_{w=1}^{WT} (|Coeff_i(h,w)| - M_i)^2} \quad (2)$$

where i = 1 to 4^L and L=3, $Coeff_i(h,w)$ represents the coefficient values of i^{th} decomposed sub-band image at level L, WT and HT are the width and height of the decomposed sub-band image $Coeff_i$.

Each feature has a different range of values. To avoid the influence of one feature due to variation in a range of values, the features are normalized between 0 and 1 range using min-max normalization method.

There are two methods used for combining features: early and late fusion. In early fusion, the individual features are combined before image similarity score calculation. In late fusion, the image similarity score determines for each feature and combine the individual score to calculate the final score. The late fusion method needs more cost of computation [24]. For this reason, the early fusion technique is used to determine the image similarity score.

3.1.2. Nearest neighbor:

After feature extraction, the nearest neighbor classifier determines the neighbors of a test image. In the nearest neighbor classifier, the neighbors for each class/category are determined as shown in Figure 3 and combined to form the final neighbors. The harmonic mean of the neighbors of each category is determined and finally predicts the category with the smallest mean [25].

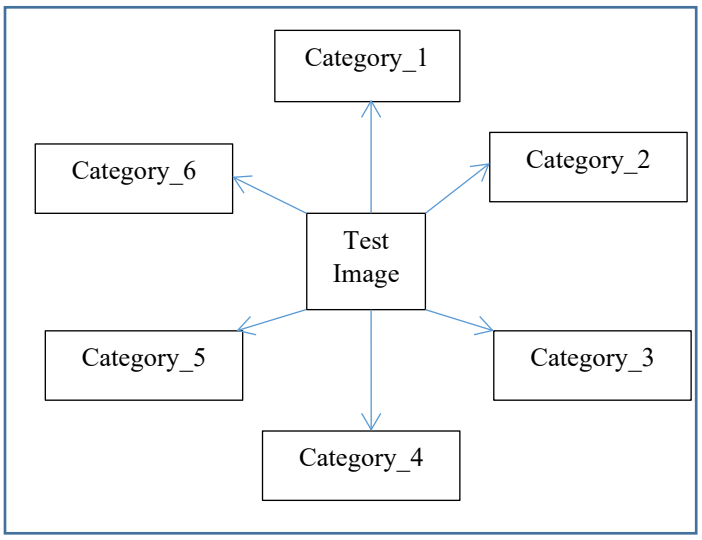


Figure 3: Nearest Neighbor from Each Category

To improve the effectiveness of feature-based neighbors, the weight is assigned to each neighbor so that nearby training image will get more weight and the training images which are farther away will be assigned less weight.

3.1.3. ITR-WTF Model:

Once the feature based neighbors are determined, the method assign an importance to each tag based on image similarity score as follow:

$$P(t) = \sum_{x_i \in NN} sim(I_q, x_i).\delta(x_i, t) \quad (3)$$

where $sim(I_q, x_i)$ denotes the importance of image x_i in predicting tag t according to their visual similarity; $\delta(x_i, t)$ is equal to one if tag is associated with an image x_i, otherwise $\delta(x_i, t)$ is equal to zero. Finally, the score of tag t is calculated as follow:

$$w - tfidf(t) = \left[1 + \log(N_t) * \log\left(\frac{NN}{N_t}\right) * P(t)\right] \quad (4)$$

where N_t represents the number of images associated with t and NN represents the total number of nearest neighbors.

3.2. Algorithm

Algorithm 1: The proposed ITR-WTF Algorithm

Input:

CV: Color feature vector of an input image
TV: Texture feature vector of an input image
$TFV_j = (y_i^j, C_i^j)_1^{NF_j}$: j[th] class texture feature vector of training images
$CFV_j = (x_i^j, C_i^j)_1^{NF_j}$: j[th] class color feature vector of training images
M = $C_1, C_2, \ldots\ldots C_M$: the number of class labels
$T_{mn} = \{T_{1n}, T_{2n}, T_{3n}, \ldots\ldots\ldots T_{mn}\}$: Tag associated with n[th] image

Output: Ranking of Tags

Step 1: Calculate the distance between CV and CFV_j using eq. (3). The set of m number of nearest neighbor for each class C_j is denoted as $CD_{mj}^{NF} = [x_{1j}^{NF}, x_{2j}^{NF}, \ldots\ldots, x_{mj}^{NF}]$

$$dist(CV, x_i^j) = \sqrt{\sum_{l=1}^{m}(CV_l - x_{il}^j)^2} \quad (5)$$

where $x_i^j \epsilon\, CFV_j$

Step 2: Calculate the distance between TV and TFV_j using eq. (4). The set of n number of nearest neighbor for each class C_j is denoted as $TD_{nj}^{NF} = [y_{1j}^{NF}, y_{2j}^{NF}, \ldots\ldots, y_{nj}^{NF}]$

$$dist(TV, y_i^j) = \sqrt{\sum_{l=1}^{n}(TV_l - y_{il}^j)^2} \quad (6)$$

where $y_i^j \epsilon\, TVF_j$

Step 3: Merge the nearest neighbors obtained using color and texture features

$NN = \{CD_{ij}^{NF} \cup TD_{ij}^{NF}\}$ where i = 1……M

Step 4: Assign an importance to each tag based on image similarity score as follows:

$$P(t) = \sum_{x_i \in NN} sim(I_q, x_i).\delta(x_i, t)$$

where

$$sim(I_q, x_i) = dist(I_q, x_i) * wt_j^i$$

$$wt_j^i = \frac{1}{j} \qquad i = 1, \ldots\ldots, m$$

Step 5: Calculate the w-tfidf of each tag as follows

$$w - tfidf(t) = \left[1 + \log(N_t) * \log\left(\frac{NN}{N_t}\right) * P(t)\right]$$

Step 6: Rank the tags according to the w- tfidf values and select top k values

4. Performance Metric

To evaluate the performance of proposed method, we use NDGC, Precision, Recall and F1-score.

Given an image with ranked tag list $T_1, T_2, \ldots, T_n$, the NDCG is computed as where r(i) is the relevance level of the i[th] tag and Z_n is a normalization constant that is chosen so that the optimal ranking's NDCG score is 1. After computing the NDCG measures of each image's tag list, we can average them to obtain an overall performance evaluation of the tag ranking method.

$$NDCG@k = \frac{1}{z}\sum_{i=1}^{k}\frac{2^{rel(i)} - 1}{\log(1+i)} \quad (7)$$

rel(i) is a binary indicator, which is equal to one if the i[th] tag in the ranking list is relevant to an input image, and zero otherwise.

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F1 - Score = \frac{2 * precision * recall}{precision + recall} \quad (10)$$

5. Dataset

Two datasets are used: self-generated and NUS-WIDE. The self-generated dataset consists of images collected from Flickr image sharing website belonging different categories such as fish,

actor, aeroplane, butterfly, autumn etc. and each category consists of 300 images. The size of the images is fixed with maximum width or height to 320 pixels. The eight categories of the images from NUS-WIDE dataset are used for experimentation. The images are divided into two groups: training and testing images after 10-fold cross validation.

For self-generated dataset, the tags associated with the images are collected from Flickr using public API. However, some tags are do not describe the image content. Therefore, tags related to year, camera and brands are excluded from the tag list.

6. Experimental Results

The performance of traditional K nearest neighbor and nearest neighbor is shown in Table 1 and 2 respectively.

Table 1: Performance of traditional K nearest neighbor

Method	Self-generated dataset	NUS-WIDE dataset
Precision	84.33 %	65.87 %
Recall	84.50%	60.75 %
F1-Score	84.00%	61.37 %

Table 2: Performance of Nearest Neighbor

Method	Self-generated dataset	NUS-WIDE dataset
Precision	87.73 %	69.88 %
Recall	88.66 %	67.00 %
F1-Score	88.00 %	67.13 %

From Table 1 and Table 2 it is observed that the performance of nearest neighbor based on each category achieves good performance as compared to traditional k nearest neighbor.

The Figure 4 shows that the NDGC value is high for tag recommendation result when no of the tags recommended is 5 for self-generated dataset and 15 for NUS-WIDE dataset.

Table 3 shows the performance of the existing and proposed tag recommendation algorithm. The performance of proposed tag recommendation algorithm is better as it can recommend/suggest tags with higher NDGC score.

Tagvoting method [18]: In this method, the feature based similar images are determined and the tags are recommended to an input image based on the frequency of tags that appeared in k visually similar images. The method assigns a uniform weight to each neighbor.

TagProp method [20]: In TagProp method, the weights are assigned each visual neighbor of a query image. The weights are assigned using rank based and distance-based method.

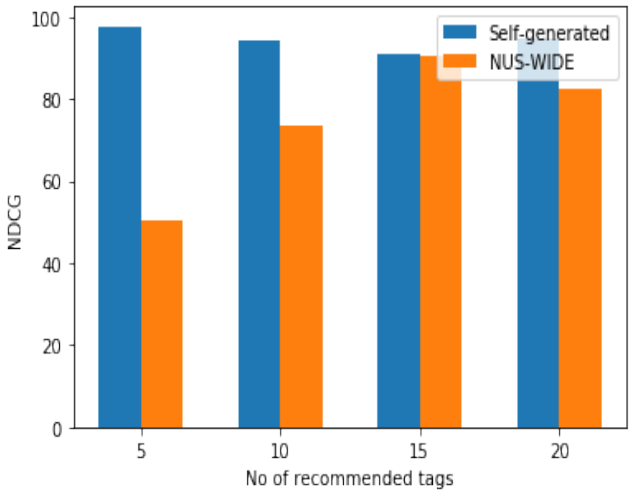


Figure 4: The NDCG values for different number of tags

NVote method [19]: In Nvote method, the tags are recommended based on the difference between global and local tag frequency by assigning equal weight to each neighbor.

Table 3: Performance of different tag recommendation algorithm

Method	Self-generated dataset	NUS-WIDE dataset
Tagvoting	84.45 %	85.20 %
TagProp	89.67 %	68.50 %
NVote	92.07 %	68.96 %
ITR-WTF	95.43 %	88.90%

From Table 3, it is observed that the methods for tag recommendation using the nearest neighbor depend on the value of k which indicates the count of neighbors. The existing methods first identify the uniform/weighted neighbors and then consider only tag information for tag voting which affects the accuracy of tag recommendation. The proposed method improves the accuracy of tag recommendation by combining the image score and tag score.

Table 4: Result of Tag Recommendation

Image	Initial Tags	Recommended Tags
	Brainedge Background Celebration Closeup Clover	Clover Green Nature Leaves Macro

	Bracom (Bram) Bracom Forest Bos Autumn Herfst	Autumn Trees Leaves Path Forest
	Mikepaws London Airport Aircraft Aeroplane	Flying Aircraft Aeroplane Plane Air
	Demerarah Fishalive Tropical-Fish Fish Tropical Fish-Tank Underwater	Fish Aquarium Tropical-Fish Underwater Fishalive

Table 4 shows the result of tag recommendation obtained using the proposed algorithm on self-generated dataset. The initial tag does not describe the entire image content. The proposed algorithm recommends relevant tags to the images. For the second image in table 4 initial tags do not cover the trees, forest and path between the trees which are added by the proposed algorithm.

Therefore, the effectiveness of the proposed method ITR-WTF for tag recommendation is demonstrated using examples in Table 4.

7. Conclusion and future work

In the paper, a method is proposed for tag recommendation of the images by identifying rank neighbors from each category. The method improves the accuracy of tag recommendation by combining the tag frequency score and weighted similarity score of the nearest neighbor images of each category. The experimentation is done two datasets: self-generated and NUS-WIDE dataset. The effectiveness of the proposed method is demonstrated in the experimental results.

The future work will focus on: i) exploring the relationship between tags obtained using from nearest neighbor ii) developing a more optimized approach which works on large dataset iii) exploring metadata associated with the images.

References

[1] E. Spyrou, P. Mylonas, "An overview of flickr challenges and research opportunities," in Proceedings - 9th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP-2014, 2014, doi:10.1109/SMAP.2014.19.

[2] Z. Hao, H. Ge, L. Wang, "Visual attention mechanism and support vector machine based automatic image annotation," PLoS ONE, 2018, doi:10.1371/journal.pone.0206971.

[3] Wu, Wei, Jianyun Nie, and Guanglai Gao, "An improved SVM-based multiple features fusion method for image annotation," Journal of Information & Computational Science **11**(14), 4987-4997, 2014.

[4] X. Li, C.G.M. Snoek, "Classifying tag relevance with relevant positive and negative examples," in MM 2013 - Proceedings of the 2013 ACM Multimedia Conference, 2013, doi:10.1145/2502081.2502129.

[5] H.M. Chen, M.H. Chang, P.C. Chang, M.C. Tien, W.H. Hsu, J.L. Wu, "Sheepdog - group and tag recommendation for flickr photos by automatic search-based learning," in MM'08 - Proceedings of the 2008 ACM International Conference on Multimedia, with co-located Symposium and Workshops, 2008, doi:10.1145/1459359.1459473.

[6] T. Uricchio, L. Ballan, L. Seidenari, A. Del Bimbo, "Automatic image annotation via label transfer in the semantic space," Pattern Recognition, 2017, doi:10.1016/j.patcog.2017.05.019.

[7] X. Qian, X.S. Hua, Y.Y. Tang, T. Mei, "Social image tagging with diverse semantics," IEEE Transactions on Cybernetics, 2014, doi:10.1109/TCYB.2014.2309593.

[8] J. Liu, Z. Li, J. Tang, Y. Jiang, H. Lu, "Personalized geo-specific tag recommendation for photos on social websites," IEEE Transactions on Multimedia, 2014, doi:10.1109/TMM.2014.2302732.

[9] Y. Gao, M. Wang, Z.J. Zha, J. Shen, X. Li, X. Wu, "Visual-textual joint relevance learning for tag-based social image search," IEEE Transactions on Image Processing, 2013, doi:10.1109/TIP.2012.2202676.

[10] L. Zheng, Z. Tianlong, H. Huijian, Z. Caiming, "Personalized tag recommendation based on convolution feature and weighted random walk," International Journal of Computational Intelligence Systems, 2020, doi:10.2991/ijcis.d.200114.001..

[11] Y. Ma, Y. Liu, Q. Xie, L. Li, "CNN-feature based automatic image annotation method," Multimedia Tools and Applications, 2019, doi:10.1007/s11042-018-6038-x.

[12] Q. Ji, L. Zhang, X. Shu, J. Tang, "Image annotation refinement via 2P-KNN based group sparse reconstruction," Multimedia Tools and Applications, 2019, doi:10.1007/s11042-018-5925-5.

[13] V. Maihami, F. Yaghmaee, "Automatic image annotation using community detection in neighbor images," Physica A: Statistical Mechanics and Its Applications, 2018, doi:10.1016/j.physa.2018.05.028.

[14] J. Zhang, Y. Yang, Q. Tian, L. Zhuo, X. Liu, "Personalized Social Image Recommendation Method Based on User-Image-Tag Model," IEEE Transactions on Multimedia, 2017, doi:10.1109/TMM.2017.2701641.

[15] Q. Ji, L. Zhang, Z. Li, "KNN-based image annotation by collectively mining visual and semantic similarities," KSII Transactions on Internet and Information Systems, 2017, doi:10.3837/tiis.2017.09.016.

[16] Y. Verma, C. V. Jawahar, "Image Annotation by Propagating Labels from Semantic Neighbourhoods," International Journal of Computer Vision, 2017, doi:10.1007/s11263-016-0927-0.

[17] S. Lee, W. De Neve, Y.M. Ro, "Visually weighted neighbor voting for image tag relevance learning," Multimedia Tools and Applications, 2014, doi:10.1007/s11042-013-1439-3.

[18] C. Chaoran, Jialie Shen, Jun Ma, and Tao Lian, "Social tag relevance learning via ranking-oriented neighbor voting," Multimedia Tools and Applications **76**(6) 8831-8857, 2017.

[19] X. Qian, X. Liu, C. Zheng, Y. Du, X. Hou, "Tagging photos using users' vocabularies," Neurocomputing, 2013, doi:10.1016/j.neucom.2012.12.021.

[20] L. Xirong, Cees GM Snoek, and Marcel Worring, "Learning social tag relevance by neighbor voting," IEEE Transactions on Multimedia **11**(7), 1310-1322, 2009.

[21] M. Guillaumin, T. Mensink, J. Verbeek, C. Schmid, "TagProp: Discriminative metric learning in nearest neighbor models for image auto-annotation," in Proceedings of the IEEE International Conference on Computer Vision, 2009, doi:10.1109/ICCV.2009.5459266.

[22] D.A. Dondekar and Balwant A. Sonkamble, "Analysis of Flickr Images using Feature Extraction Techniques," 4th IEEE International Conference on Computer and Communication Systems (ICCCS 2019), Singapore, 278-282 2019.

[23] D.A. Dondekar and Balwant A. Sonkamble,"Tag-based Image Retrieval using Hybrid Visual-Tag Feature Extraction Method", International Journal of Advanced Science and Technology, **9**(4), 5931 – 5940, 2020.

[24] X. Li, "Tag relevance fusion for social image retrieval," Multimedia Systems, 2017, doi:10.1007/s00530-014-0430-9.

[25] D.A. Dondekar and Balwant A. Sonkamble, "Harmonic Mean based Classification of Images using Weighted Nearest Neighbor for Tagging" International Journal of Advanced Computer Science and Applications, **11**(11), 2020, doi:/10.14569/IJACSA.2020. 0111131

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1403-1413 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Exposure to Optical Radiation and Electromagnetic Fields at the Workplace: Criteria for Occupational Health Surveillance According to Current European Legislation

Alberto Modenese, Fabriziomaria Gobba*

Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, 41125, Italy

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 16 February, 2021
Online: 28 February, 2021

Keywords:
Non-ionizing radiation
Occupational exposure
Health Surveillance

ABSTRACT

A very large number of workers is occupationally exposed to Optical Radiation (OR) worldwide, while indeed nowadays an exposure to Electromagnetic fields (EMF) can occur in almost all workplaces. OR origin can be natural, including the most relevant source, i.e. the sun, or artificial, that can be further classified in incoherent and coherent, i.e. the LASERs. Solar radiation (SR) exposure, and in particular its most harmful component, the ultraviolet radiation (UVR), is a significant occupational risk in "outdoor workers", including e.g. farmers and construction workers. UVR is mainly absorbed in the eye and the skin, there inducing various short-term and chronic adverse health effects, as burns, cataract and skin cancers. At least in Europe, for SR exposed workers no specific obligations currently exist regarding the Health Surveillance (HS), that is instead required for occupational exposures to artificial OR according to the legislation of the European Union (EU, Directive 2006/25/EC). Considering now EMF, the EU Directive 2013/35/EU provides an obligation for the HS of exposed workers, aimed at the prevention of the possible direct short-term effects, as involuntary contractions or temperature increase of tissues, and indirect effects, as shocks and interference. Conversely, long-term effects are not considered in the Directive as data on causal relationship, including reliable mechanisms, are considered inadequate. Direct short-term and indirect effects can appear solely in case of high exposures, usually occuring only accidentally, but a specific group of workers, defined "at particular risk", exists, and it includes e.g. persons with implanted active medical devices, as cardioverter defibrillators or pacemakers. In these workers, adverse effects can be induced at lower EMF levels. The identification and an adequate protection of the workers at particular risk is one of the main goals of the HS of occupational EMF exposure.
The main HS criteria applicable for workers with exposure to OR and EMF are discussed in this article.

1. Introduction

Exposure to Optical Radiation (OR) and Electromagnetic Fields (EMF), i.e. Non-Ionizing Radiation (NIR), at work represents a relevant occupational risk for many work activities: this paper is an extension of work originally presented in the 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe) [1].

Occupational risks can be classified in chemical, biological, physical risks and others (including ergonomic, psichosocial and safety risks) [2]. NIR exposure is included among physical risks together with e.g. noise and vibrations. As it happens for all the other occupational risks, adequate prevention should be emplaced in the companies in order to warrant appropriate occupational health and safety (OHS) conditions for the exposed workers. This involves a specific evaluation of the risk related to the working activity with NIR exposure, aimed at identifying the most appropriate preventive measures. An important part of the prevention process to ensure OHS is the health surveillance (HS) of the workers exposed to occupational risks that may affect their health, usually performed by trained occupational physicians. HS is defined, according to the International Labour Office (ILO), as a group of "procedures and investigations to assess workers' health,

*Corresponding Author: Fabriziomaria Gobba, v. G. Campi 287 – 41125 Modena (Italy), +390592055463, fabriziomaria.gobba@unimore.it

in order to identify any abnormality possibly associated to a recognized work-related risk in exposed workers", involving "the review of the health records and any other opportune medical examinations needed" [3]. Considering possible implications for workers' health of NIR exposure, it should be noted that NIR are included in the part of the elctromagnetic spectrum with lower energy compared to ionizing radiation, meaning that both OR and EMF have not enough energy to cause ionization, i.e. to remove electrons from atoms and molecules. Accordingly, the interaction mechanisms with the biological tissues, and the penetration ability of NIR are very different from ionizing radiation. Nevertheless, also within the various regions and sub-regions of the NIR spectrum significant differences exist, and therefore the possible health effects and the criteria for the health surveillance of the workers exposed to non-ionizing radiation should be discussed separately. In the following parts of this article these criteria are introduced and discussed.

2. Optical radiation: occupational exposure and health effects

2.1. Occupational exposure to optical radiation

OR includes wavelengths (λ) between 100 nanometers (nm) and 1 millimeter (mm), distributed in three different regions: ultraviolet radiation (UVR), further classified in the subregions UV-A, UV-B and UV-C, visible radiation and infrared radiation (IR), with other three sub-regions: IR-A, IR-B and IR-C [4].

Occupational OR exposure can be possible both in case of artificial OR and solar radiation (SR) exposure.

Artificial sources can emit incoherent radiation or coherent bands of amplificated radiations (i.e. LASERs). Considering UVR, among the main applications relevant for occupational exposure there is the use of UV as sterilizing agents for both surfaces and products, the idustrial processes applying UV rays in the machineries for photo-lithography and printing, the medical applications as e.g. the photo-therapy for the neonatal jaundice, the UV emissions during welding processes, and many others. Also visible light is emitted during welding, and of course it is used for illumination purposes: kind of intense exposures can be possible in the entertainment sector, e.g. in theatres or in television studios, in shopping centers, and in general where it is required to be close to sources emitting intense lights (indicators, signals, LED, projectors, etc). Also visible light is used for industrial, e.g. printing processes, and medical applications, e.g. for photopolymerization of the materials used in dental clinics. Finally, infrared radiation is also emitted during welding activities, and moreover it is emitted as a result of industrial processes generating very high themal energies during the treatment of metals, glass and other materials. Furthermore, IR-based technologies are widely applied e.g. in communication and surveillance systems [4]. LASER technologies can use all the optical radiations. Gas LASERs, as those using argon or krypton, mainly amplify UV or short wavelength visible light, while solid state LASERs, as neodymium YAG LASERs, and semi-conductor LASERs, as those using gallium, mainly amplify infrared or long wavelength visible light [4].

Even if the number of artificial sources emitting OR at the workplaces is huge, the effective number of workers exposed to levels potentially determining relevant health risks is relatively limited, as frequently the sources are appropriately shielded [4–7]. A possible exception is represented by welding activities: welders should be appropriately protected, with adequate clothing and face masks with eye filters for the specific bands of ORs emitted during the welding process [4–6, 8]. For the other sources of artificial OR the major problems may occur in case of malfunctions of the machineries with failures of the protection mechanisms, and/or of workers with inadequate protections, determining work-related accidents possibly causing severe damages [9, 10]. The eye is one of the most frequently involved organ, in particular in these cases [11, 12].

On the other hand, the most important source of occupational exposure to OR, in terms of both the numbers of exposed workers and relevancy of possible adverse effects, is certainly the Sun. Millions of workers have an exposure to solar radiation (SR) worldwide: among the main job categories involved there are farmers [13, 14], construction workers [15–17], fishermen [18], lifeguards [19, 20], gardeners [21, 22], ski instructors/mountain guides [23, 24] and others [25, 26]. These workers, spending a significant part of the working day outside, are collectively identified with the term of “outdoor workers”. Fortunately, SR at the Earth' surface does not include all the components of OR spectrum, and in particular all UVC rays and the large majority of the UVB are blocked in the stratosphere [5, 26]. As UV radiation is a carcinogenic agent, also SR is considered a carcinogenic exposure [27]: it is estimated that SR represents the occupational carcinogenic exposure with the highest amount of workers exposed [28].

2.2. Interaction mechanisms with biological tissues and main effects of optical radiation exposure

OR have a quite low ability of penetrating the biological tissues, and therefore the main targets are the eyes and the skin [5, 26]. Two main mechanisms of interaction with the biological tissues can be identified for OR bands:

- Thermal mechanisms: these effects are typical of IR and of long wavelength visible light, and they are related to the conduction of heat in the tissues, usually requiring intense exposures. In case of insufficiently intense exposure no appreciable effects can be determined, as the heat is dissipated through the surrounding tissues [29]. Accordingly, usually only acute effects, resulting in eye injuries or skin thermal burns can be associated to infrared or visible light exposure with thermal mechanisms [29], even if, mainly in the past, some chronic skin alterations [30] as well as thermal cataract have been reported [31, 32].
- Photo-chemical effects are possible in case of exposure to UVR and to short wavelength visible light [5, 26]. Photons are absorbed by specific targeted molecules (i.e. chromophores) in the biological tissues, releasing energy and inducing chemical reactions, with consequent damages and alterations (e.g. inflammation) [5, 26]. Photochemical effects are possible both in case of short-term intense exposures but also as a consequence of long-term less intense exposures [33].

The most hazardous component of OR is UVR: UV rays have higher energy compared to infrared and visible radiation and they act with photochemical mechanisms, causing both acute and chronic effects in humans [5, 26]. After intense short-term exposures, acute skin effects as skin erythema and photodermatoses, in case of presence of a photo-toxic substance (e.g. a topic drug or a cosmetic) on the skin, can occur (it should be noted that photodermatoses may be also associated with some wavelenghts of visible radiations) [5, 26]. Also photochemical burns of the unprotected eyes are possible, especially in case of intense direct exposures or indirect exposures from surfaces with high reflectance (e.g. snow, polished metals): UV rays can be highly absorbed in the ocular surface, causing an acute inflammation of the cornea and of the conjunctiva called photokeratitis/photoconjunctivities, determining photophobia, redness, pain, burning, tearing and foreign body perception [5, 12, 26].

Considering long-term exposure, UVR is able to induce various skin and eye effects, many of them very severe. UVR is mutagenic for the DNA and cumulative damages over time can induce pre-cancerous and cancerous alterations [5, 26, 27]. In the skin chronic UVR exposure is able to induce photoaging, i.e. a progressive process including functional and aesthetic skin changes with loss of elasticity, cell depletion, pigment accumulation and substitution of normal tissue with fibrous tissue. considered to be a precancerous alteration [5, 26, 34, 35]. Other pre-cancerous lesions, sometimes identified also as *in situ* carcinomas, induced by chronic UVR exposure of the skin are actinic keratoses (AK). AK are disorders of the keratinocytes resulting in thickening of the skin with formation of scaly and/or crusty lesions. Untreated lesions have up to a 20% risk of progression to squamous cell carcinoma (SCC) [5, 26, 27, 35, 36]. In addition to pre-cancerous lesions, UVR is able to induce skin cancers. The most common are the basal cell carcinoma (BCC) and the SCC of the skin, also called as keratinocytes cancers [26, 27, 37]. These tumours are the most frequent malignancies in subjects with fair skin, and accordingly they are the most frequent type of cancer in Europe, Canada, US and Australia. Fortunately these tumors have mainly a local malignancy, resulting only seldom in methastatization processes: accordingly, the prognosis is quite good compared to other tumors, and, if treated in time, the mortality is low, but they can result in quite impressive mutilations, especially considering that they usually appear in solar UV exposed skin areas, as the face [38, 39]. Both BCC and SCC are somehow related to cumulative UVR exposure and there is evidence that outdoor workers have a relevantly increased risk of getting these tumors [40], while only few reports of associations with artificial UVR exposure at work are available [41]. Another skin tumor related to excessive UVR exposure is the malignant melanoma (MM), a cancer originating from the skin melanocytes. MM are less frequent compared to BCC and SCC, but they often methastatize if not diagnosed in time, determining an important burden not only in terms of disability adjusted life years (DALYs) but also in term of mortality [27, 42]. Considering occupational exposure, the relation with SR exposure in outdoor workers is controversial, as some studies reported a possible protective effect of cumulative UVR exposure, while the increased risk of MM is associated with intense exposure in the childhood with repeated sunburns [27, 43]. Some studies reporting a possible increased melanoma risk for welders have also been published [27, 41].

Considering now possible long-term effects in the eyes, UVR can induce chronic alterations of the ocular surface, being absorbed mainly in the cornea and in the conjunctiva, determining a chronic inflammatory stimolous and pathological alterations with abnormal growths of the corneal and conjunctival cells [5, 26, 44]. This may result in various diseases, some of them representing mainly aesthetic alterations with no clinical relevancy, as pinguecula, other more severe with potential affection of the visual function if not treated, as the pterygium, an abnormal growth of the conjunctiva from the nasal angle of the eye to the center, finally covering the cornea [44, 45]. Corneal and conjunctival squamous cell carcinomas (i.e. the same types of cancers occurring in the skin) are also possible: fortunately they are very rare, but with demonstrations of associations with UV exposure [44]. Moreover, also another eye tumor, the melanoma of the eye, more frequent compared to corneal and conjunctival carcinomas, even if still quite rare, has been reported as possibly associated with UV exposure, in particular in welders [44, 46].

Furthermore, a relevant quote of UVR is able to penetrate the ocular surface being absorbed in the lens [5]. There, with photo-oxidation damages contributing to the age-related process of the denaturation of the lens proteins, UVR can accelerate and increase the opacification of the lens [47]: it is estimated that more than the 20% of the cataracts can be attribued to excessive UVR exposure [48], and it should be noted that cataract is still the main cause of blindness worldwide [49]. Not only solar radiation, but also artificial UVR is associated to an increased risk of cataract development [47, 50].

Finally, another possible chronic adverse eye effect that can be associated in particular to long-term SR exposure is macular degeneration, a very frequent multi-factorial degenerative disease of the central part of the retina, currently the third among the major causes of blindness worldwide [49]. Scientific demonstrations of the relation between this eye pathology and SR exposure are mainly epidemiological, rather then experimental, but some hypothesis have been proposed [51]. In this case, the role of UVR is limited to that of a small portion of UV-A radiation, able to reach the retina especially in younger subjects when the lens is more transparent [5]. But also violet/blue visible light can contribute to a chronic photochemical retinal damage through photo-oxidations [52, 53].

3. Prevention of the health risk related to occupational optical radiation exposure

3.1. Occupational optical radiation exposure risk in the workplaces and workers belonging to particularly sensitive risk groups

Currently in Europe there are no specific obligations for the HS of workers exposed to solar radiation, even if, considering the relevancy of the risk and of the adverse health effects to be expected, in many countries an increasing attention to the protection of outdoor workers against UVR and to the occurrence of solar UV-induced occupational diseases has been raised in recent years [5, 26, 38, 54–59]. One of the issues still to be solved is the lack of a recognized occupational limit for solar UVR

exposure at work in Europe: the recognition of such a limit would be helpful to implement risk evaluation processes in the workplaces, defining the specific protections needed. On the other hand, since 2006 a specific European Directive "2006/25/EC" for the prevention of the occupational risk related to artificial optical radiation exposure has been emplaced [4]. In this case, specific occupational exposure limits to be respected are available, and moreover the Directive requires mandatory risk evaluation, definition of technical and organizational preventive measures, specific training of the workers, the forniture of adequate personal protective equipment and, if needed, an appropriate health surveillance.

For all the previously mentioned steps of the process for an adequate prevention of OR risk at the workplaces, the Directive indicates that a special attention has to be deserved to the workers "*belonging to particularly sensitive risk groups*" [4]. No specific indications are included in the EU Directive helping in identifying these workers. We report here below a non-exhaustive list of possible conditions determining a particular susceptibility to OR exposure, increasing the risk of adverse effects, mainly for the eye and the skin, and requiring a special attention [4, 5, 26].

- Conditions related to a possible increased risk of adverse skin effects (mainly related to the UVR component of OR): e.g. albino subjects and all the workers with fair skin photo-types (i.e. at least Fitzpatrick skin photo-type 1 or 2 [60]), who are at an increased risks for developing all the UV-related acute and long-term skin effects; workers affected by skin diseases that can be photo-induced or photo-aggravated, as psoriasis or xeroderma pigmentosum; workers being diagnosed with precancerous or cancerous skin lesions; workers being treated with photosensitizing drugs or in contact with photosensitizing substances as psoralenes.
- Conditions related to a possible increased risk of adverse eye effects: workers with alterations of the eyes structure possibly affecting the quantity of OR absorbed in specific eye regions, as e.g. iris alterations (as coloboma, aniridia), pupil alterations (as tonically dilated pupil, mydriasis), lens alterations (aphakia, artificial intraocular lens); workers being diagnosed with eye diseases possibly being worsened by excessive OR exposure, as e.g. initial lesions of the lens (as presence of small lens opacities or cataract at an initial stage) or of the retina (as the presence of drusen in the macula); subjects with only a monocular vision, as the possible occurrence of OR-related adverse effects in the eye may result in blindness.
- Other conditions: pregnant workers are usually considered among the workers belonging to particularly sensitive risk groups, mainly considering (but not limited to) the possible thermal effects of the infrared component, increasing the body temperature in the abdominal region; underage workers can be also potentiually included, especially considering that the lens is more transparent to UVR, possibly enabling a higher amount of radiation reaching the retina.

3.2. Health surveillance of workers exposed to optical radiations

According to the current European legislation, there are two possible scenarios for the HS of OR exposed workers. The HS is an obligation in case a relevant risk is evaluated in the workplace according to the EU Directive 2006/25/EC for artificial OR exposure, both in case of incoherent sources and of LASER use. On the other hand, considering SR exposure risk of outdoor workers, currently in Europe HS is not mandatory [5, 26], even if, in the light of the numbers of outdoor workers, the high levels of occupational UVR exposure detected in several studies [13–26] and of the documented increased risk for these workers of developing UVR-induced eye and skin effects [25, 26, 40, 44–48, 50, 51, 54, 55, 27, 28, 34–39] an appropriate HS seems absolutely necessary.

The objective of the HS of OR-exposed workers is the prevention and the early diagnosis of adverse effects, mainly of the eyes and of the skin. During HS activities, specific attentions have to be deserved to the workers affected by conditions inducing a particular susceptibility to the risk, as the ones mentioned here in the previous section 3.1. HS programs usually include both pre-employment and periodic medical examinations from trained occupational physicians, who may require, on individual basis, supplementary health controls possibly encompassing other specialists involved in the specific problem detected (e.g. ophthalmologists, dermatologists, etc) [26, 38, 59]. As stated above, the most alarming health risk to be considered during HS is the one related to the UVR component that can induce relevant adverse health effects as skin cancers, currently the most frequent tumors among Caucasian subjects. These cancers, together with other OR-induced eye and skin pathologies, are recognized occupational diseases in many European countries, but not in all [38, 59]. Nevertheless, these cancers are largely underreported: an extension of the HS also to outdoor workers should help in increasing the reporting of OR-induced occupational diseases, allowing workers to receive an adequate medical care and compensation for the diseases, and letting institutional authorities and governments acknowledged of the true magnitude of the phenomenon, and consequently of the need of preventive measures [38, 55, 59].

4. Electromagnetic fields: occupational exposure and health effects

4.1. Occupational exposure to electromagnetic fields

Nowadays, a certain level of EMF exposure is present in almost all the working environments, and there is an extremely large, and increasing, variety of occupational EMF sources. Accordingly, some exposure to EMF is present for almost all workers. But, on the other hand, numerous sources are present also in the general environment. Accordingly, it can seem inappropriate to consider every single worker as an "*exposed worker*". A detailed discussion on this topic is beyond the scopes of this article but, in general, it can be considered appropriate to evaluate as "*exposed*" workers, those with an expected EMF exposure above the levels usually detectable for the general population at home and in public places [61–64]. A peculiar aspect in this context is that in some workers with a particular susceptibility (*"workers at particular risk"* - WPR), also EMF levels permittable for the general public may determine a non-irrelevant residual health risk. This is specifically the case of workers with active implanted medical devices (AIMD), who may be at risk for interference problems when in proximity of an EMF source [65–67]: these aspects will be discussed in detail in the next paragraphs.

When dealing with the problem of occupational exposure to EMF, as it happens also for OR, it is important to identify the specific types of EMFs emitted by the considered sources, as different regions of the EMF spectrum have different biological interaction mechanisms and, consequently, different effects, depending on the frequency bands. For the objectives of this article, we adopt a simplified EMF classification, discussing here the main issues related to the exposure in particular to static magnetic fields (SMF, frequency= 0 Hertz), extremely-low frequency magnetic fields (ELF-MF, frequency= 1 - <300 Hertz), intermediate frequency EMF (IF-EMF, frequency= 300 Hertz - <100 kilohertz) and high frequency EMF (HF-EMF, frequency> 100 kilohertz, including radiofrequency - RF -, microwaves, millimeter and terahertz waves) [61–64, 68]. Considering now specific working situations and occupational sources with potentially relevant emissions of EMF, a useful guide for the prevention of EMF risk at work is the so-called "*non-binding guide*" for the application of the European Directive 2013/35/EU [68]: we present here an extract of the document with a list of work activities and EMF-sources determining occupational exposures (Table 1).

Table 1: Non-exhaustive list of work activities/occupational sources inducing a potentially relevant exposure to electromagnetic fields extracted from the "non-binding guide" for the application of the European Directive 2013/35/EU

Type of equipment or workplace
Construction: various equipment and machineries (e.g. concrete mixers, cranes, vibrators, welding machines, etc)
Electrical supply: electrical installations, generators, inverters
Infrastructure (buildings): base station antennas
Infrastructure (grounds): garden appliances
Heavy industry: industrial electrolysis, furnaces, arc melting, welding machines
Light industry: arc welding processes, dielectric heating and welding, industrial magnetizer/ demagnetizers
Medical: magnetic Resonance Imaging (MRI), other medical equipment using EMF for diagnosis and treatment (for example, short wave diathermy, transcranial magnetic stimulation)
Office: audio-visual equipment containing RF-transmitters, office equipment (e.g. photocopiers, electric staplers, paper shredders)
Transport: motor vehicles and plants, radars (air traffic control, military use, weather monitoring, long range radars), electrically driven trains and trams
Wireless communications Cordless phones (including base stations for cordless phones), mobile phones, wi-fi routers

It should be noted that for the sources and the occupational activities with potentially relevant occupational EMF exposures presented in the Table 1 no specific frequencies ranges of EMF emissions are reported. This is because in many cases the exposure can include various EMF bands, as e.g. in welding activities, where ELF, IF-EMF and RF can be represented in the specific emission spectra depending also on the type of welding procedure, or in MRI activities, where operators are exposed to the action of the SMF as well as to time-varying EMF when they move within the SMF, while the patients are also exposed to RF when the scanner is in function. In other cases, the EMF exposure can be referred to specific EMF bands, as e.g. for electrical installations, generators and inverters, emitting mainly in the ELF range, or for wireless communication systems, with emissions mainly in the RF band [61–64, 68–70]. Another relevant note for an appropriate interpretation of the Table 1 is that not all the potential exposures listed represent, in normal conditions, a relevant risk for the workers; e.g., again in the case of wireless communication systems or in the case of office equipment, the listed sources may determine an occupational risk only for workers with an increased susceptibility to the EMF risk, as subjects with AIMD who have to work close to the sources [61–68], as previously introduced. On the other hand, other sources can be able to usually induce high exposure levels, potentially representing an occupational risk in case of inadequate working conditions, e.g. in the case of close proximity of operators to a MRI scanner, in industrial electrolysis plants, or in working activities involving induction heating systems, soldering devices, or broadcasting systems and devices, etc. [61–64, 68].

4.2. Interaction mechanisms with biological tissues and main effects of electromagnetic fields exposure

According to the above mentioned European Directive 2013/35/EU, currently in Europe for the prevention of the occupational risks related to EMF exposure only effects based on recognized interaction mechanisms with the biological tissues are considered [68]. The EU Directive indicates specifically the short term direct biophysical effects and the indirect effects related to EMF exposure [68]. A necessary punctualization here, considering the quite large number of studies dealing with suggested long-term effects related to EMF (among the main examples there are the suggested associations between leukemia and ELF-MF exposure, and brain cancers with RF exposure), is that, for these effects, the current European legislation does not consider sufficient the available scientific evidence, in particular considering that no adequate pathogenetic mechanisms in humans have been demonstrated [61–64, 68]. The position of the EU Commission is coherent with the opinions of various recognized national and international organizations and institutions [61–64, 71-85].

According to the above mentioned EU approach, we discuss here only: a) direct biophysical effects caused in the human body by its presence within an EMF; and b) indirect effects, induced by the presence of an object in the EMF (object that can also be placed inside the body, as in the case of implanted medical devices) [68].

According to the frequency band of the EMF, different mechanisms are involved for the occurrence of direct effects: non-thermal mechanisms for static and low-frequency fields, and thermal mechanisms for high frequency fields. For the intermediate frequencies both the mechanisms are possible [61–64, 68]. Among non-thermal effects, based on cuurents induction in the biological tissues, there are the following: limb currents,

stimulation of muscles, nerves or sensory organs with induction of temporary annoyance or with a possible detrimental effect on cognition or on other brain or muscle functions, with possible risks for the safety of the workers [68]. Considering thermal mechanisms, the direct effects, related to HF-EMF exposure, and possibly to IF-EMF, are related to an increase of the temperature in the body, that can be usually more relevant at the surface, and accordingly in targets as the eyes, where very intense HF-EMF can theoretically induce a thermal damage of the lens with opacification (cataract), and the skin, where thermal burns can be possible, also in this case after very high exposures levels [64, 68]. A further classification of direct effects according to the 2013/35/EU Directive is the one concerning *"sensory"* and *"health effects"* [68]. Sensory effects are usually rapidly reversible and transient. They can be considered non-pathological effects, usually with no consequences for the individuals, even if they can be associated with an increased risk of work accidents. Sensory effects are mainly related to relevant SMF and ELF-EMF exposures, representing the consequence of an induction of currents, stimulating the nervous systems and/or sensory organs. Examples of sensory effects are vertigo, nausea, perception of a metallic taste in mouth after SMF exposure, magnetophosphenes (i.e. eye lamps perception) and minor reversible changes of some brain functions (e.g. attention, congnitive function) in case of ELF-EMF exposure [61–63, 68]. On the other hand, EMF-related *"health effects"* represent severe alterations, that usually appear only in case of overexposure to EMF, significantly above the occupational limits established e.g. in Europe by the 2013/35/EU Directive. A non-exhaustive list of these health effects, based on non-thermal mechanisms for SMF and ELF-MF (an example can be an alteration of the cardiac rythm) and on thermal mechanisms in case of very intense HF-EMF exposure (e.g. skin burns), is presented in the Table 2. As stated above, for IF-EMF exposure sensory and health effects typical of both low and high frequencies can be possible [61–64, 68].

Table 2: Non-exhaustive list of health effects caused by overexposure to electro-magnetic fields, based on the European Directive 2013/35/EU

EMF type	
Static magnetic field	Alterations of the blood flow in the limbs; alterations of the brain functions; alterations of the cardiac rhythm and of the cardiovascular fuction
Low frequency fields	Pain and/or tingling sensations due to an involuntary contractions of the muscles or to the stimulation of the nervous system; alterations of the cardiac rhythm
Intermediate frequency fields	The effects of both high and low frequencies can be possible
High frequency fields	Excessive increase in temperature and/or thermal burns over the whole body or in specific areas; thermal damage of the eyes (possibility of thermal cataracts), skin, testicles

Considering now indirect effects, these effects can occur, as stated above, when there is an object inside the EMF, together with the human body. All the EMF frequencies can be involved, and, as in the case of medical devices, it should be noted that the object can be also implanted in the human body [68]. Some of these indirect effects mainly represent a risk of work-related injuries: EMF can initiate electro-explosive devices causing fires and explosions. Other indirect effects possibly associated to severe injuries are contact currents, that can be induced in the body by the contact with charged objects within an EMF [68]. The most relevant indirect effects that can happen also in case of usual exposures, and not as a consequence of a malfunction/wrong procedure, are interference problems of EMF with implanted or body worn devices, usually for medical purposes. The major problems can be experienced in case of active devices, as pacemakers or implantable cardioverter defribrillators (ICD): their functions can be altered by the EMF, determining a risk for the workers' health [61–68]. Also, passive devices, as e.g. metal splinters or graft, can be affected by EMF: in particular, SMF may dislocate the devices made of ferromagnetic materials, causing inflammation or other problems in the surrounding tissue. Moreover, IF-EMF and HF-EMF can determine a heating of the metallic parts of the devices, also in this case with a potential damage to the biological tissue [65–68].

5. Prevention of the health risk related to electromagnetic fields exposure

5.1. Occupational electromagnetic fields exposure risk in the workplaces and workers at particular risk

In Europe, the prevention of the occupational risk related to EMF is regulated according to the Directive 2013/35/EU. As it happened also for OR, a specific *"non binding guide"* was then produced by the European Commission, to give practical indications for the application of the Directive [68]. The Directive requires mandatory EMF risk evaluation in the workplaces, defining the technical and organizational preventive measures to be applied, including a specific training of the exposed workers, the adoption of adequate personal protective equipment and, if needed, an appropriate health surveillance. Specific occupational exposure limits are available for the different EMF frequencies, mainly based on the guidelines of the International Commission on Non-Ionizing Radiation Protection (ICNIRP) [61–63]. Different limit values are available not only according to the EMF frequency, but also based on the different effects to be prevented: separate limits for the prevention of the above-mentioned *sensory* and *health effects* are defined in the EU Directive [68]. As previously introduced, the possibility to reach exposure levels able to induce *health effects* is very rare, happening only in case of overexposure, e.g. in case of accidents. Moreover, it is not sufficient to simply exceed the limit value to induce the appaerance of a *health effect*, as usually exposures highly above the limits are needed. Furthermore, also the exceeding of the limits for *sensory effects*, much lower compared to the limits for *health effects*, is not usual, and in principle it has to be avoided in normal conditions [86, 87].

Considering this, one of the main problems for the prevention of the occupational EMF risk is the possible presence at the workplace of WPR, as defined by the Directive 2013/35/EU, for whom the occupational limits set might be not fully protective [68]. Currently, a non-exhaustive list of WPR for EMF exposure includes:

- Workers with active implanted medical devices (AIMD): e.g. pacemakers, implanted cardioverter defibrillators, ear implants, neurostimulators, implanted drug infusion pumps, etc [65–68, 88, 89].
- Workers with passive implanted medical devices containing metals: e.g. artificial joints, plates, surgical clips, stents, heart valve prostheses, intra-uterine devices, etc [68, 88–91].
- Workers with body-worn devices: e.g. drugs infusion pumps, hearing aids, continuous glucose monitoring systems, etc [68, 88, 89].
- Other WPR categories included in the EU Directive: pregnant workers [68, 88, 89].

5.2. Health surveillance of workers exposed to electromagnetic fields

According to the Directive 2013/35/EU, HS of EMF exposed workers is mandatory, when a relevant risk for the workers is evaluated. Moreover, for the HS, as well as for the risk assessment, the problem of the WPR has to be specifically taken into account [68, 89]. As defined in the Introduction, HS is aimed at identifying any abnormality possibly associated to a recognized work related risk in exposed workers [3]. All the investigations chosen for HS have to be proved with an appropriate level of scientific evidence: the contents of the HS can include biological tests and other investigations, when they are chosen for their validity and relevance with respect to the occupational risks, avoiding investigations that don't fulfill these criteria [92]. Considering the specific case of EMF exposure, a reliable application of these criteria may involve some possible issues. In fact, according to the EU Directive the scopes of the HS are the prevention and early detection of the previously defined direct biophysical effects, including the sensory ones (as e.g. the presence of vertigo or of nausea associated to EMF exposure, etc.), based on non-thermal and thermal mechanisms [68]. In general, these effects do not appear for occupational exposure levels below the limit values, as they can be induced only as a consequence of excessive exposures to quite high levels of EMF [86, 87]. It has also to be highlighted here that sensory effects, that are fully reversible and transient, cannot be considered an actual health risk for the workers [93–95], even if some data suggest a possible association with an increased risk of work-related injuries for the exposed subjects [96, 97]. The other main purpose of the health surveillance related to the occupational EMF risk is the prevention of indirect effects, and this specifically involves the identification of the WPR. Among the conditions of particular susceptibility, probably the main issues are related to possible electromagnetic interference (EMI) problems with AIMD. EMI can appear also as a consequence of kind of low exposure levels if a susceptible subject with an AIMD is close to an EMF-source, and the interference can impair the fundamental functions of the devices [61–68].

Considering this, the health surveillance criteria to be considered should mainly address the identification of possible medical signs and symptoms of sensory effects and of the conditions of particular susceptibility to the EMF-risk, typical of the WPR. For these purposes, particular lab tests or other specific clinical investigations are not required on a routine basis, while their opportunity can be judged by the occupational physician on an individual basis after a medical examination. Useful instruments to be periodically administered to the workers are the questionnaires evaluating symptoms and conditions of particular susceptibility [93–95]. When administering questionnaires, important prerequisites are a validation process [98] and a preparation of the tool based on scientific results: unfortunately, to the best of our knowledge, to date there are no examples of questionnaires specifically designed for an application during HS examinations of EMF-exposed workers. A further note to the problem of the EMF-related HS is that, in general, medical examinations for workers in good health conditions, with no conditions of particular susceptibility and working in environnments where the EMF levels are supposed to be low or very low (e.g. offices) have not to be considered as activities to be periodically repeated. In these cases, only a single medical investigation performed before the starting of the job, with an effective information on all the conditions possibly determining a particular susceptibility to the EMF risk (NB: the information has to be periodically repeated in these cases, while the medical examination doesn't), seems advisable for an adequate prevention [68, 89].

Finally, a peculiar situation explicitly mentioned in the EU Directive is the case of an *"extraordinary"* HS of the EMF exposed workers, needed when *"any undesired or unexpected health effect is reported by a worker"* or *"in any event where exposure above the ELVs* (i.e. exposure limit values) *is detected"*. In these situations, an appropriate *"medical examination"* needs to be provided to the workers [68]. Also, in this case, no guidelines or evidence-based indications are available to define valid contents for these medical examinations. Furthermore, as previously noted, situations with exposure levels exceeding the limits do not necessarily result in the occurrence of adverse health effects [86, 87]. Also, in this case, the main problems should be expected for the WPR. Accordingly, both for extraordinary and routine HS the contents of the examinations include an appropriate in-depth medical examination of the workers by an occupational physician with an adequate competence in the field, administrating specific questionnaires to collect information on suspected EMF-related symptoms and possible conditions of particular susceptibility to the risk, eventually integrated by specific medical consultations and/or laboratory tests and/or diagnostic exams to be decided on an individual basis, considering the types of EMF frequency involved, the exposure level, the susceptibility of the workers and the expected effects to be evaluated [68, 89].

6. Conclusions

Occupational NIR exposure, including EMF and OR, is almost ubiquitous: in principle, an evaluation of the occupational health risk should be appropriate in all workplaces. Based on the health risk assessment, HS programs for the exposed workers have to be established, with the objective of preventing or, at least, identifying at an early stage the possible health effects associated with the exposure. Both risk evaluation and HS activities have to specifically consider the possible presence in the workplace of "workers at particular risk", who deserve a specific attention, as they have an increased susceptibility to the risk. The HS needs to be performed by occupational physicians with adequate and updated skills in the prevention of NIR risk, and accordingly with an appropriate specific training in this field. Here below we present

a conclusive summary of the main relevant points to be considered for the health surveillance of NIR exposed workers:

Optical radiation:

i. Occupational physicians have to consider that the exposure can be associated to both artificial and natural sources. In particular, solar radiation is certainly the main occupational exposure source, even if sometimes under-recognized as a specific occupational risk.

ii. The main target organs of optical radiation exposure are the eyes and the skin.

iii. For the prevention and early diagnosis purposes, the possible occurrence of both short-term and long-term eye and skin effects, respectively related to acute and cumulative exposures, has to be adequately considered as a fundamental criterion of HS.

iv. The most harmful component of optical radiation is UVR, able to induce both acute and long-term effects with photochemical mechanisms; among the most severe adverse effects related to chronic UV exposures there are skin cancers, i.e. the most frequent forms of neoplasms in Caucasian subjects. Other diseases identified in UVR exposed workers are: skin erythema, photoaging and actinic keratosis for the skin, photokeratoconjunctivitis, pterygium, lens opacities and possibly macular degeneration for the eyes.

v. During HS activities, specific conditions of particular susceptibility for the exposed workers have to be considered: e.g. a fair skin phototype, eye alterations as aphakia, and others.

vi. Optical radiation related adverse effects are recognized "occupational diseases", even if frequently underreported: physicians have to notify the occurrence of these diseases in exposed workers to the specific workers' compensation Authorities in their Country.

Electromagnetic fields:

i. According to the European legislation, the recognized effects related to EMF exposure to be considered for the health surveillance of the workers are short-term direct biophysical effects and indirect effects.

ii. Direct effects can be related to both non-thermal and thermal mechanisms, and they are usually induced only by high exposures.

iii. Indirect effects, as the electromagnetic interference with AIMD, can occur also as a consequence of lower exposure levels.

iv. To date, there are no agreed and shared criteria for the health surveillance of EMF-exposed workers, and in particular there are no evidence-based indications on the appropriate contents and procedures for the medical examinations to be performed for the monitoring of the health status of the exposed subjects.

v. The trained occupational physicians have to evaluate all the possible symptoms related to the sensory and health effects of EMF exposure, considering specifically the type of frequency, the exposure level and the mechanisms involved.

vi. One of the most important activities during health surveillance is the screening of all the conditions that can induce a particular risk for the exposed workers, as e.g. AIMD: a check for the presence of these conditions needs to be done before the employment by the physician, and an extensive information on these conditions has to be provided to the workers, and periodically repeated.

vii. In case of overexposure situations or in case of workers reporting symptoms associated to the EMF exposure, an appropriate medical examination has to be provided to the workers, with specific contents to be evaluated on an individual basis, again based on the type of EMF-frequency, on the exposure level detected and, on the mechanisms, involved.

Conflict of Interest

The Authors declare no conflict of interest.

Acknowledgment

The work presented in this article was possible with the support of the "Istituto Nazionale Assicurazione Infortuni sul Lavoro" (INAIL) within the research project: "*Protezione dei lavoratori dai campi elettromagnetici: supporto alla valutazione del rischio e indicazioni per la sorveglianza sanitaria, con particolare attenzione alle condizioni di superamento dei limiti di esposizione previste dal D.Lgs. 81/08 e ai lavoratori particolarmente sensibili al rischio*" (BRIC 2016, ID 40/2016), coordinated by the Italian National Health Institute (ISS).

References

[1] A. Modenese, F. Gobba, "Occupational Exposure to Non-Ionizing radiation. Main effects and criteria for health surveillance of workers according to the European Directives, " Proceedings - 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe, EEEIC / I and CPS Europe 2020, 2020, doi:10.1109/EEEIC/ICPSEurope49358.2020.9160831.

[2] E.L. Melnick, Brian S. Everitt, Encyclopedia of Quantitative Risk Analysis and Assessment, Wiley, 2008.

[3] ILO, Occupational Safety and Health Series, No. 72 Technical and ethical guidelines for workers' health surveillance, 50, 1998.

[4] European Commission, Non-binding guide to good practice for implementing Directive 2006/25/EC (Artificial optical radiation), 2011.

[5] G. Ziegelberger, "Icnirp statement-protection of workers against ultraviolet radiation, " Health Physics, **99**(1), 66–87, 2010, doi:10.1097/HP.0b013e3181d85908.

[6] D. Sliney, "Risks of occupational exposure to optical radiation, " Med Lav, **97**(2), 215–20, 2006.

[7] E.A. Talbot, P. Jensen, H.J. Moffat, C.D. Wells, "Occupational risk from ultraviolet germicidal irradiation (UVGI) lamps, " International Journal of Tuberculosis and Lung Disease, **6**(8), 738–741, 2002.

[8] T.D. Tenkate, "Optical radiation hazards of welding arcs, " Reviews on Environmental Health, **13**(3), 131–146, 1998, doi:10.1515/REVEH.1998.13.3.131.

[9] K. Barat, "Laser accidents: Occurrence and response, " Health Physics, **84**(5 SUPPL.), 93–95, 2003, doi:10.1097/00004032-200305001-00013.

[10] J.S. Pierce, S.E. Lacey, J.F. Lippert, R. Lopez, J.E. Franke, M.D. Colvard, "An assessment of the occupational hazards related to medical lasers, " Journal of Occupational and Environmental Medicine, **53**(11), 1302–1309, 2011, doi:10.1097/JOM.0b013e318236399e.

[11] F. Gobba, E. Dall'Olio, A. Modenese, M. De Maria, L. Campi, G.M. Cavallini, "Work-related eye injuries: A relevant health problem. main epidemiological data from a highly-industrialized area of northern Italy, " International Journal of Environmental Research and Public Health, **14**(6),

2017, doi:10.3390/ijerph14060604.
[12] S. Zaffina, V. Camisa, M. Lembo, M.R. Vinci, M.G. Tucci, M. Borra, A. Napolitano, V. Cannatã, "Accidental exposure to UV radiation produced by germicidal lamp: Case report and risk assessment, " Photochemistry and Photobiology, **88**(4), 1001–1004, 2012, doi:10.1111/j.1751-1097.2012.01151.x.
[13] C. Smit-Kroner, S. Brumby, "Farmers sun exposure, skin protection and public health campaigns: An Australian perspective, " Preventive Medicine Reports, **2**, 602–607, 2015, doi:10.1016/j.pmedr.2015.07.004.
[14] A.W. Schmalwieser, A. Cabaj, G. Schauberger, H. Rohn, B. Maier, H. Maier, "Facial solar UV exposure of Austrian farmers during occupation, " Photochemistry and Photobiology, **86**(6), 1404–1413, 2010, doi:10.1111/j.1751-1097.2010.00812.x.
[15] M. Antoine, S. Pierre-Edouard, B. Jean-Luc, V. David, "Effective exposure to solar UV in building workers: Influence of local and individual factors, " Journal of Exposure Science and Environmental Epidemiology, **17**(1), 58–68, 2007, doi:10.1038/sj.jes.7500521.
[16] M.A. Serrano, J. Cañada, J.C. Moreno, "Solar UV exposure in construction workers in Valencia, Spain, " Journal of Exposure Science and Environmental Epidemiology, **23**(5), 525–530, 2013, doi:10.1038/jes.2012.58.
[17] M. Wittlich, S.M. John, G.S. Tiplica, C.M. Sălăvăstru, A.I. Butacu, A. Modenese, V. Paolucci, G. D'Hauw, F. Gobba, P. Sartorelli, J. Macan, J. Kovačić, K. Grandahl, H. Moldovan, "Personal solar ultraviolet radiation dosimetry in an occupational setting across Europe, " Journal of the European Academy of Dermatology and Venereology, **34**(8), 1835–1841, 2020, doi:10.1111/jdv.16303.
[18] A. Modenese, F.P. Ruggieri, F. Bisegna, M. Borra, C. Burattini, E. Della Vecchia, C. Grandi, A. Grasso, L. Gugliermetti, M. Manini, A. Militello, F. Gobba, "Occupational exposure to solar UV radiation of a group of fishermen working in the Italian north adriatic sea, " International Journal of Environmental Research and Public Health, **16**(16), 1–12, 2019, doi:10.3390/ijerph16163001.
[19] P. Gies, K. Glanz, D. O'Riordan, T. Elliott, E. Nehl, "Measured occupational solar UVR exposures of lifeguards in pool settings, " American Journal of Industrial Medicine, **52**(8), 645–653, 2009, doi:10.1002/ajim.20722.
[20] M.A. Serrano, J. Cañada, J.C. Moreno, "Erythemal ultraviolet exposure in two groups of outdoor workers in valencia, Spain, " Photochemistry and Photobiology, **85**(6), 1468–1473, 2009, doi:10.1111/j.1751-1097.2009.00609.x.
[21] M. Boniol, A. Koechlin, M. Boniol, F. Valentini, M.C. Chignol, J.F. Doré, J.L. Bulliard, A. Milon, D. Vernez, "Occupational UV exposure in French outdoor workers, " Journal of Occupational and Environmental Medicine, **57**(3), 315–320, 2015, doi:10.1097/JOM.0000000000000354.
[22] E. Thieden, S.M. Collins, P.A. Philipsen, G.M. Murphy, H.C. Wulf, "Ultraviolet exposure patterns of Irish and Danish gardeners during work and leisure, " British Journal of Dermatology, **153**(4), 795–801, 2005, doi:10.1111/j.1365-2133.2005.06797.x.
[23] G.R. Casale, A.M. Siani, H. Diémoz, G. Agnesod, A. V. Parisi, A. Colosimo, "Extreme UV index and solar exposures at plateau rosà (3500ma.s.l.) in valle d'aosta Region, Italy, " Science of the Total Environment, **512**–**513**, 622–630, 2015, doi:10.1016/j.scitotenv.2015.01.049.
[24] M. Moehrle, B. Dennenmoser, C. Garbe, "Continuous long-term monitoring of UV radiation in professional mountain guides reveals extremely high exposure, " International Journal of Cancer, **103**(6), 775–778, 2003, doi:10.1002/ijc.10884.
[25] M.S. Paulo, B. Adam, C. Akagwu, I. Akparibo, R.H. Al-Rifai, S. Bazrafshan, F. Gobba, A.C. Green, I. Ivanov, S. Kezic, N. Leppink, T. Loney, A. Modenese, F. Pega, C.E. Peters, A.M. Prüss-Üstün, T. Tenkate, Y. Ujita, M. Wittlich, S.M. John, "WHO/ILO work-related burden of disease and injury: Protocol for systematic reviews of occupational exposure to solar ultraviolet radiation and of the effect of occupational exposure to solar ultraviolet radiation on melanoma and non-melanoma skin cancer, " Environment International, 2019, doi:10.1016/j.envint.2018.09.039.
[26] A. Modenese, L. Korpinen, F. Gobba, "Solar radiation exposure and outdoor work: An underestimated occupational risk, " International Journal of Environmental Research and Public Health, **15**(10), 2018, doi:10.3390/ijerph15102063.
[27] International Agency for Reasearch on Cancer, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 100D, Lyon, 2012.
[28] T. Kauppinen, J. Toikkanen, A. Savela, D. Pedersen, R. Young, W. Ahrens, P. Boffetta, M. Kogevinas, J. Hansen, H. Kromhout, V. de La Orden-Rivera, J. Maqueda Blasco, D. Mirabelli, B. Pannett, N. Plato, R. Vincent, "Occupational exposure to carcinogens in the European Union, " Occupational and Environmental Medicine, **57**(1), 10–18, 2000, doi:10.1136/oem.57.1.10.
[29] G. Ziegelberger, "ICNIRP guidelines on limits of exposure to incoherent visible and infrared radiation, " Health Physics, **105**(1), 74–96, 2013, doi:10.1097/HP.0b013e318289a611.
[30] R.J. Kettelhut EA, Traylor J, Erythema Ab Igne, 2020.
[31] A. Dorozhkin, "Cataract of metallurgists, " Vestn Oftalmol, **119**(3), 31–4, 2003.
[32] J.J. Vos, D. Van Norren, "Thermal cataract, from furnaces to lasers, " Clinical and Experimental Optometry, **87**(6), 372–376, 2004, doi:10.1111/j.1444-0938.2004.tb03097.x.
[33] R.W. Bunsen, H.. Roscoe, "Photochemical Researches—Part V. On the Measurement of the Chemical Action of Direct and Diffuse Sunlight, " in Proc. R. Soc, 306–12, 1862, doi:10.1098/rspl.1862.0069.
[34] P. Sartorelli, R. Romeo, V. Paolucci, V. Puzzo, F. Di Simplicio, L. Barabesi, "Skin photoaging in farmers occupationally exposed to ultraviolet radiations, " Medicina Del Lavoro, **104**(1), 24–29, 2013.
[35] K. Grandahl, J. Olsen, K.B.E. Friis, O.S. Mortensen, K.S. Ibler, "Photoaging and actinic keratosis in Danish outdoor and indoor workers, " Photodermatology Photoimmunology and Photomedicine, **35**(4), 201–207, 2019, doi:10.1111/phpp.12451.
[36] R. Schwartz, T. Bridges, A. Butani, A. Ehrlich, "Actinic keratosis: an occupational and environmental disorder, " Journal of the European Academy of Dermatology and Venereology, **0**(0), 080304135428024-???, 2008, doi:10.1111/j.1468-3083.2007.02579.x.
[37] C. Karimkhani, L.N. Boyers, R.P. Dellavalle, M.A. Weinstock, "It's time for 'keratinocyte carcinoma' to replace the term 'nonmelanoma skin cancer, '" Journal of the American Academy of Dermatology, **72**(1), 186–187, 2015, doi:10.1016/j.jaad.2014.09.036.
[38] F. Gobba, A. Modenese, S.M. John, "Skin cancer in outdoor workers exposed to solar radiation: a largely underreported occupational disease in Italy, " Journal of the European Academy of Dermatology and Venereology, **33**(11), 2068–2074, 2019, doi:10.1111/jdv.15768.
[39] A. Modenese, F. Farnetani, A. Andreoli, G. Pellacani, F. Gobba, "Questionnaire-based evaluation of occupational and non-occupational solar radiation exposure in a sample of Italian patients treated for actinic keratosis and other non-melanoma skin cancers, " Journal of the European Academy of Dermatology and Venereology, **30**, 2016, doi:10.1111/jdv.13606.
[40] T. Loney, M.S. Paulo, A. Modenese, F. Gobba, T. Tenkate, D.C. Whiteman, A.C. Green, S.M. John, "Global evidence on occupational sun exposure and keratinocyte cancers: a systematic review, " British Journal of Dermatology, 1–11, 2020, doi:10.1111/bjd.19152.
[41] L.M. Falcone, P.C. Zeidler-Erdely, "Skin cancer and welding, " Physiology & Behavior, **176**(1), 139–148, 2018, doi:10.1117/12.2549369.Hyperspectral.
[42] M. RASTRELLI, S. TROPEA, C.R. ROSSI, M. ALAIBAC, "Melanoma: Epidemiology, Risk Factors, Pathogenesis, Diagnosis and Classification, " In Vivo, **28**, 1005–1011, 2014, doi:10.32388/7xj0gw.
[43] B.K. Armstrong, A.E. Cust, "Sun exposure and skin cancer, and the puzzle of cutaneous melanoma: A perspective on Fears et al. Mathematical models of age and ultraviolet effects on the incidence of skin cancer among whites in the United States. American Journal of Epidemiology 1977; , " Cancer Epidemiology, **48**, 147–156, 2017, doi:10.1016/j.canep.2017.04.004.
[44] J.C.S. Yam, A.K.H. Kwok, "Ultraviolet light and ocular diseases, " International Ophthalmology, **34**(2), 383–400, 2014, doi:10.1007/s10792-013-9791-x.
[45] A. Modenese, F. Gobba, "Occupational exposure to solar radiation at different latitudes and pterygium: A systematic review of the last 10 years of scientific literature, " International Journal of Environmental Research and Public Health, **15**(1), 2018, doi:10.3390/ijerph15010037.
[46] T. Nayman, C. Bostan, P. Logan, M.N. Burnier, "Uveal Melanoma Risk Factors: A Systematic Review of Meta-Analyses, " Current Eye Research, **42**(8), 1085–1093, 2017, doi:10.1080/02713683.2017.1297997.
[47] C.A. McCarty, H.R. Taylor, "A review of the epidemiologic evidence linking ultraviolet radiation and cataracts., " Developments in Ophthalmology, **35**, 21–31, 2002, doi:10.1159/000060807.
[48] T. Tenkate, B. Adam, R.H. Al-Rifai, B.R. Chou, F. Gobba, I.D. Ivanov, N. Leppink, T. Loney, F. Pega, C.E. Peters, A.M. Prüss-Üstün, M. Silva Paulo, Y. Ujita, M. Wittlich, A. Modenese, "WHO/ILO work-related burden of disease and injury: Protocol for systematic reviews of occupational exposure to solar ultraviolet radiation and of the effect of occupational exposure to solar ultraviolet radiation on cataract, " Environment International, **125**, 2019, doi:10.1016/j.envint.2018.10.001.
[49] D. Pascolini, S.P. Mariotti, "Global estimates of visual impairment: 2010, " British Journal of Ophthalmology, **96**(5), 614–618, 2012, doi:10.1136/bjophthalmol-2011-300539.
[50] A. Modenese, F. Gobba, "Cataract frequency and subtypes involved in

workers assessed for their solar radiation exposure: a systematic review, " Acta Ophthalmologica, **96**(8), 2018, doi:10.1111/aos.13734.

[51] A. Modenese, F. Gobba, "Macular degeneration and occupational risk factors: a systematic review, " International Archives of Occupational and Environmental Health, **92**(1), 2019, doi:10.1007/s00420-018-1355-y.

[52] J. Moon, J. Yun, Y.D. Yoon, S. Il Park, Y.J. Seo, W.S. Park, H.Y. Chu, K.H. Park, M.Y. Lee, C.W. Lee, S.J. Oh, Y.S. Kwak, Y.P. Jang, J.S. Kang, "Blue light effect on retinal pigment epithelial cells by display devices, " Integrative Biology (United Kingdom), **9**(5), 436–443, 2017, doi:10.1039/c7ib00032d.

[53] J.Z. Nowak, "Age-related macular degeneration (AMD): Pathogenesis and therapy, " Pharmacological Reports, **58**(3), 353–363, 2006.

[54] T.L. Diepgen, M. Fartasch, H. Drexler, J. Schmitt, "Occupational skin cancer induced by ultraviolet radiation and its prevention, " British Journal of Dermatology, **167**(SUPPL. 2), 76–84, 2012, doi:10.1111/j.1365-2133.2012.11090.x.

[55] S.M. John, M. Trakatelli, C. Ulrich, "Non-melanoma skin cancer by solar UV: The neglected occupational threat, " Journal of the European Academy of Dermatology and Venereology, **30**, 3–4, 2016, doi:10.1111/jdv.13602.

[56] A. Modenese, T. Loney, F.P. Ruggieri, L. Tornese, F. Gobba, "Sun protection habits and behaviors of a group of outdoor workers and students from the agricultural and construction sectors in north-Italy, " Medicina Del Lavoro, **111**(2), 116–125, 2020, doi:10.23749/mdl.v111i2.8929.

[57] C.E. Peters, T. Tenkate, E. Heer, R. O'Reilly, S. Kalia, M.W. Koehoorn, "Strategic Task and Break Timing to Reduce Ultraviolet Radiation Exposure in Outdoor Workers, " Frontiers in Public Health, **8**(August), 1–9, 2020, doi:10.3389/fpubh.2020.00354.

[58] D. Reinau, M. Weiss, C.R. Meier, T.L. Diepgen, C. Surber, "Outdoor workers' sun-related knowledge, attitudes and protective behaviours: A systematic review of cross-sectional and interventional studies, " British Journal of Dermatology, **168**(5), 928–940, 2013, doi:10.1111/bjd.12160.

[59]C. Ulrich, C. Salavastru, T. Agner, A. Bauer, R. Brans, M.N. Crepy, K. Ettler, F. Gobba, M. Goncalo, B. Imko-Walczuk, J. Lear, J. Macan, A. Modenese, J. Paoli, P. Sartorelli, K. Stageland, P. Weinert, N. Wroblewski, H.C. Wulf, S.M. John, "The European Status Quo in legal recognition and patient-care services of occupational skin cancer, " Journal of the European Academy of Dermatology and Venereology, **30**, 2016, doi:10.1111/jdv.13609.

[60] T.B. Fitzpatrick, "The Validity and Practicality of Sun-Reactive Skin Types I Through VI, " Archives of Dermatology, **124**, 869–871, 1988.

[61] International Commission on Non-Ionizing Radiation Protection, "Guidelines on limits of exposure to static magnetic fields., " Health Physics, **66**(1), 100–106, 1994.

[62] I.C. on N.-I.R. Protection, "Guidelines for limiting exposure to time-varying electric and magnetic fields (1 Hz TO 100 kHz), " Health Physics, **99**(6), 818–836, 2010, doi:10.1097/HP.0b013e3181f06c86.

[63] I.C. on N.-I.R. Protection., "Guidelines for limiting exposure to electric fields induced by movement of the human body in a static magnetic field and by time-varying magnetic fields below 1 Hz, " Health Physics, **106**(3), 418–425, 2014, doi:10.1097/HP.0b013e31829e5580.

[64] G. Ziegelberger, R. Croft, M. Feychting, A.C. Green, A. Hirata, G. d'Inzeo, K. Jokela, S. Loughran, C. Marino, S. Miller, G. Oftedal, T. Okuno, E. van Rongen, M. Röösli, Z. Sienkiewicz, J. Tattersall, S. Watanabe, Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz), 2020, doi:10.1097/HP.0000000000001210.

[65] A. Napp, D. Stunder, M. Maytin, T. Kraus, N. Marx, S. Driessen, "Are patients with cardiac implants protected against electromagnetic interference in daily life and occupational environment?, " European Heart Journal, **36**(28), 1798–1804, 2015, doi:10.1093/eurheartj/ehv135.

[66] B. Hocking, K.H. Mild, "Guidance note: Risk management of workers with medical electronic devices and metallic implants in electromagnetic fields, " International Journal of Occupational Safety and Ergonomics, **14**(2), 217–222, 2008, doi:10.1080/10803548.2008.11076763.

[67] M. Tiikkaja, A.L. Aro, T. Alanko, H. Lindholm, H. Sistonen, J.E.K. Hartikainen, L. Toivonen, J. Juutilainen, M. Hietanen, "Electromagnetic interference with cardiac pacemakers and implantable cardioverter-defibrillators from low-frequency electromagnetic fields in vivo, " Europace, **15**(3), 388–394, 2013, doi:10.1093/europace/eus345.

[68] European Commission, Non-binding guide to good practice for implementing Directive 2013/35/EU Electromagnetic Fields, 2015.

[69] F. Gobba, G. Bravo, P. Rossi, G.M. Contessa, M. Scaringi, "Occupational and environmental exposure to extremely low frequency-magnetic fields: A personal monitoring study in a large group of workers in Italy, " Journal of Exposure Science and Environmental Epidemiology, **21**(6), 634–645, 2011, doi:10.1038/jes.2011.9.

[70] F. Gobba, A. Bargellini, G. Bravo, M. Scaringi, L. Cauteruccio, P. Borella, "Natural Killer Cell Activity Decreases in Workers Occupationally, " International Journal Of Immunopathology And Pharmacology, **22**(4), 1059–1066, 2009.

[71] Health Protection Agency, Health Effects from Radiofrequency Electromagnetic Fields, 2012.

[72] ANSES, "Radiofréquences et santé. Mise à jour de l'expertise, " 461, 2013.

[73] ANSES, Extremely low frequency electromagnetic fields Health effects and the work of ANSES, 2020.

[74] ARPANSA, "Review of Radiofrequency Health Effects Research - Scientific Literature 2000 - 2012, " Technical Reports of Australian Radiation Protection and Nuclear Safety Agency, **164**(164), 1–76, 2014.

[75] ARPANSA, Extremely low frequency electric and magnetic fields Extremely low frequency (ELF) electric and magnetic fields exist wherever electricity is generated, transmitted or distributed in powerlines or cables, or used in electrical appliances., 2020.

[76] Comité Científico Asesor en Radiofrecuencias y Salud, Informe sobre radiofrecuencias y salud, 2017.

[77] Food and Drug Administration, Scientific Evidence for Cell Phone Safety, 2020.

[78] Health Council of the Netherlands, Mobile phones and cancer?, The Hague, 2016.

[79] Health Council of the Netherlands, Power lines and health part I: childhood cancer, Dec. 2020.

[80] International Agency for Research on Cancer, "International Agency for Research on Cancer Iarc Monographs on the Evaluation of Carcinogenic Risks To Humans, " Iarc Monographs On The Evaluation Of Carcinogenic Risks To Humansarc Monographs On The Evaluation Of Carcinogenic Risks To Humans, **80**, i-ix+1-390, 2002.

[81] IARC, "Part 2 : Radiofrequency Electromagnetic Fields, " **102**, 2018.

[82] M.R.Scarfì, S. Lagorio, L. Anglesio, G. D'Amore, C. Marino, Radiazioni a riadiofrequenza e tumori: sintesi delle evidenze scientifiche, 2019.

[83] New Zealand Ministry of Health, Interagency Committee on the Health Effects of Non-ionising Fields Report to Ministers 2018, 2018.

[84] Public Health England, Collection Electromagnetic fields Advice on exposure to electromagnetic fields in the everyday environment, including electrical appliances in the home and mobile phones., 2020.

[85] Swedish Radiation Safety Authority's Scientific Council on Electromagnetic Fields, Recent research on EMF and health risk: eleventh report from SSM's scientific council on electromagnetic fields: including thirteen years of electromagnetic field research monitored by SSM's Scientific Council on EMF and health: how has the evidence chang, 2016.

[86] B. Hocking, F. Gobba, "Medical aspects of overexposures to electromagnetic fields, " Journal of Health, Safety and Environment, **27**(3), 185–195, 2011.

[87] IEEE Committee on Man and Radiation, "Medical Aspects Of Radiofrequency Radiation Overexposure, " Health Physics, (September 2001), 0–4, 2002.

[88] HSE, Electromagnetic Fields at Work, 2016.

[89] F. Gobba, "Health surveillance of EMF-exposed workers at particular risk, " G Ital Med Lav Ergon, **41**(4), 285–292, 2019.

[90] F. Gobba, N. Bianchi, P. Verga, G.M. Contessa, P. Rossi, "Menometrorrhagia in magnetic resonance imaging operators with copper intrauterine contraceptive devices (iuds): A case report, " International Journal of Occupational Medicine and Environmental Health, **25**(1), 97–102, 2012, doi:10.2478/s13382-012-0005-y.

[91] A. Huss, K. Schaap, H. Kromhout, "A survey on abnormal uterine bleeding among radiographers with frequent MRI exposure using intrauterine contraceptive devices, " Magnetic Resonance in Medicine, **79**(2), 1083–1089, 2018, doi:10.1002/mrm.26707.

[92] International Commission on Occupational Health, ICOH Code of Ethics, 2012.

[93] F. de Vocht, E. Batistatou, A. Mölter, H. Kromhout, K. Schaap, M. van Tongeren, S. Crozier, P. Gowland, S. Keevil, "Transient health symptoms of MRI staff working with 1.5 and 3.0 Tesla scanners in the UK, " European Radiology, **25**(9), 2718–2726, 2015, doi:10.1007/s00330-015-3629-z.

[94] K. Schaap, Y.C. De Vries, C.K. Mason, F. De Vocht, L. Portengen, H. Kromhout, "Occupational exposure of healthcare and research staff to static magnetic stray fields from 1.5-7 tesla MRI scanners is associated with reporting of transient symptoms, " Occupational and Environmental Medicine, **71**(6), 423–429, 2014, doi:10.1136/oemed-2013-101890.

[95] G. Zanotti, G. Ligabue, L. Korpinen, F. Gobba, "Subjective symptoms in Magnetic Resonance Imaging operators: prevalence, short-term evolution and possible related factors, " Med Lav, **107**(4), 263-270., 2016.

[96] S. Bongers, P. Slottje, L. Portengen, H. Kromhout, "Exposure to static magnetic fields and risk of accidents among a cohort of workers from a

medical imaging device manufacturing facility, " Magnetic Resonance in Medicine, **75**(5), 2165–2174, 2016, doi:10.1002/mrm.25768.

[97] A. Huss, K. Schaap, H. Kromhout, "MRI-related magnetic field exposures and risk of commuting accidents – A cross-sectional survey among Dutch imaging technicians, " Environmental Research, **156**(April), 613–618, 2017, doi:10.1016/j.envres.2017.04.022.

[98] F. Gobba, R. Ghersi, S. Martinelli, A. Richeldi, P. Clerici, P. Grazioli, "Italian translation and validation of the Nordic IRSST standardized questionnaire for the analysis of musculoskeletal symptoms, " Medicina Del Lavoro, **99**(6), 424–43, 2008.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 1, 1414-1421 (2021)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Fusion of Optical and Microfabricated Eddy-Current Sensors for the Non-Destructive Detection of Grinding Burn

Isman Khazi[1,2], Andras Kovacs[1], Ulrich Mescheder[1,3,*], Ali Zahedi[4], Bahman Azarhoushang[4]

[1]*Institute for Microsystems Technology (iMST), Faculty of Mechanical & Medical Engineering, Furtwangen University, 78120 Furtwangen, Germany*

[2]*Department of Microsystems Engineering (IMTEK), University of Freiburg, Georges-Köhler-Allee 103, 79110 Freiburg im Breisgau, Germany*

[3]*Faculty of Engineering, University of Freiburg, Georges-Köhler-Allee 103, 79110 Freiburg im Breisgau, Germany*

[4]*Institute of Precision Machining (KSF), Faculty of Mechanical & Medical Engineering, Furtwangen University, 78120 Furtwangen, Germany*

ARTICLE INFO

Article history:
Received: 24 December, 2020
Accepted: 05 February, 2021
Online: 28 February, 2021

Keywords:
Cylindrical grinding process
Grinding burn
Non-destructive testing (NDT)
Sensor fusion
Optical sensor
Microfabricated microcoils
Eddy-current
Surface roughness
Surface reflection
42CrMo4 steel

ABSTRACT

A sensor fusion concept integrating the optical and microfabricated eddy-current sensor for the non-destructive testing of the grinding burn is reported. For evaluation, reference grinding burn with varying degrees are fabricated on 42CrMo4 tool steel cylinder. The complementary sensing nature of the proposed sensors for the non-destructive testing of the grinding burn is successfully achieved, wherein both the superficial and an in-depth quantitative profile information of the grinding zone is recorded. The electrical output (voltage) of the optical sensor, which is sensitive to the optical surface quality, dropped only by 20 % for moderate degree of grinding burn and by ca. 50 % for stronger degree of grinding burn (i.e., by exclusively considering the superficial surface morphology of the grinding burn). Moreover, a direct correlation among the average surface roughness of the grinding burn, the degree of grinding burn and the optical sensor's output voltage was observed. The superficial and in-depth information of the grinding burn was recorded using a microfabricated eddy-current sensor (planar microcoil with circular spiral geometry with 20 turns) by measuring the impedance change as function of the driving frequency. The depth of penetration of induced eddy-current in the used 42CrMo4 workpiece (with a sensor to workpiece distance of 700 µm) varied from 223 µm to 7 µm on increasing the frequency of the driving current from 1 kHz to 10 MHz, respectively. A very interesting nature of the grinding burn was observed with two distinct zones within the grinding zone, namely, the superficial zone (starting from the workpiece surface to 15 µm in grinding zone) and a submerged zone (>15 µm within the grinding zone). The impedance of the microcoils changed by ca. 8 % and 4 % for the superficial and submerged zone for regions with stronger degree of grinding burn at a frequency of 10 MHz and 2.5MHz, respectively. Furthermore, a correlation between the microhardness of the grinding burn and the impedance change is also observed.

1. Introduction

Grinding technology is considered as the strategic process in the manufacturing technology with its application spanning from the manufacturing of parts for aerospace, defense industry to the production of surfaces with optical quality for the electronics and telecommunications devices, with technology moving towards processing of also the hard ceramic materials [1]. The mechanics involved during the grinding process include the interaction between the grinding wheel with abrasive grains and the

*Corresponding Author: Ulrich Mescheder, Furtwangen University, Robert-Gerwig-Platz 1, 78120, Furtwangen, Germany
+49 7723 920 2232, mes@hs-furtwangen.de
www.astesj.com
https://dx.doi.org/10.25046/aj0601160

workpiece's surface, wherein the material from the workpiece (WP) is removed in a complex manner governed by the interplay among the rubbing, cutting and ploughing stages of material abrasion. However, in case of grinding processes with high material removal rates involving hard steel workpieces, the operating temperatures during the grinding process might easily go beyond the tempering temperature of the WP when the grinding operating conditions are sub-optimal, and might result in an irreversible thermal damage on and within the grinding zone (i.e., the region of interaction between the grinding wheel and WP surface), which is usually referred to as the so-called grinding burn (GB) [1–4]. Among the several defects such as microcracks, pores etc., occurring within the grinding zone (GZ) due to sub-optimal grinding conditions, the occurrence of GB is the most complex defect to detect. The reason behind the complexity in the detection of GB is attributed to the nature of complex microstructural modifications that occur not only on the surface of GZ, but also a few hundred micrometers deep within the GZ [4–6]. As described in [1] and [7], the thermal damage caused by sub-optimal grinding conditions results in tensile residual stresses within the GZ. Moreover, it can lead to grain growth, precipitation, softening, phase transformations resulting in re-hardening, thermal expansion/contraction creating microcracks, and chemical reactions leading to discoloration, i.e., severe oxidation GB; all of which are not only detrimental to the quality of the WP (limiting the performance of the WP), but also lay an enormous financial burden on rejection of ground workpieces having GB [8].

In a recent review paper [2], the authors comprehensively summarized the different methods for detection of GB pre- and post-grinding process, out of which, they are broadly classified under destructive- and non-destructive testing (NDT) methods. Under destructive methods, the method of Nital etching revealing the microstructure of the GZ, and microhardness testing revealing the microstructural change as a function of the hardness profile of the GZ can be used. Furthermore, the optical inspection (with sophisticated CCD cameras and visual inspection), residual stress inspection using x-ray diffractometer (XRD), spectroscopic measurement of elemental study in the GZ, electromagnetic methods involving conventional eddy-current sensors, magnetoelastic methods involving the Barkhausen noise analysis and acoustic emission sensors can be used to detect the occurrence of GB on the ground WP post grinding process [2]. Few of the reported methods like XRD and spectroscopy cannot be integrated in the grinding machine, while the others such as electromagnetic and acoustic emission sensors still rely on the conventional sensor systems; wherein, on one side the system is quite bulky (which hinders its integration within the grinding machine), and on the other side the detection resolution is limited due to the size of the conventional probes used.

Furthermore, since the past two decades there have been a constant quest to realize realistic intelligent grinding manufacturing processes, which has further gained momentum in terms of research and development, owing to the emergence of industry 4.0 and big data with the concept of smart interconnected real-time condition monitored manufacturing processes [8–14]. However, among other challenges associated with the implementation of the intelligent grinding process, the availability of sophisticated approaches for the integrated process monitoring (i.e., the sensor systems), considering the difficulty to model process dynamics of grinding processes, is a fundamental requirement for the development of intelligent grinding. In order to fulfill this requirement, a sensor fusion system comprising of more than one sensing principle would enormously increase the detection capability, the sensitivity of the parameters of interest, and also the reliability of the test results. The authors in [15] reported a sensor fusion concept with integration of accelerometers and power-cell sensors for the detection of the roughness of the WP, however, the reported system is not capable of detecting GB. Therefore, in this paper, which is an extended version of a previous research [16], we report the detailed analysis of the sensor fusion concept with the integration of an optical sensor (OS) and a microfabricated eddy-current sensor (μEC) working complementary to each other for NDT of GB using a 42CrMo4 cylinder with varying degrees of reference GB. The proof-of-concept for the NDT of GB using the OS and μEC sensors has been already reported along with the individual development and characterization in [17–19], therefore, this paper deals with the experimental analysis of the complementary sensor functionality of combined OS and μEC sensors for NDT of GB.

2. Sensor fusion concept for non-destructive testing of grinding burn

The occurrence of the GB results in hard to model modifications on and within the GZ of the ground WP. Depending on the severity of the GB, it can get explicitly visible by discoloration (i.e., so-called severe oxidizing GB) on the surface of the GZ (as shown in the schematic in Figure 1 with orange coloration of the ground WP surface), or during the concluding spark-out period of grinding process, however, the discolored oxidized layer might be removed. In this case, the irreversible microstructural changes withing the GZ, which can propagate up to a few hundred micrometers within the GZ, still exist (depicted as hidden GB in Figure 1). Furthermore, the microstructural changes can result in either softening of the GZ or might result in a re-hardened layer (due to quenching effect) on top of the softened layer [4–6]. Therefore, the proposed sensor fusion system in this paper combines two sensor principles, namely the OS and μEC sensors (as shown in schematic in Figure 1) capable of detecting the surface morphological changes and the in-depth microstructural changes (as function of varying driving frequency which influences the depth of penetration of the induced eddy-currents within the WP), respectively. An interesting effect of the surface morphological changes on the GZ surface was shown in our previous paper [17], wherein it was found that the reference GB with varying degrees showed a direct correlation to the average surface roughness values (R_a) of the ground WP surface. The average surface roughness increased with increasing degree of GB. The proposed OS works on the principle of illuminating

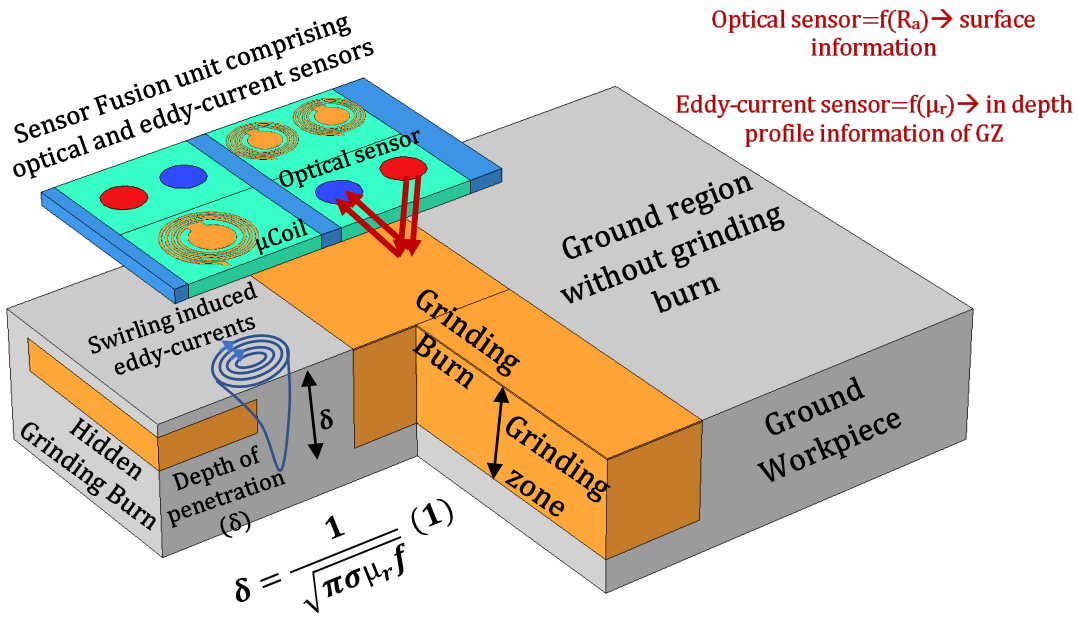


Figure 1: Schematic showing the sensor fusion concept integrating the optical sensor and the microfabricated eddy-current sensor for the NDT of GB. The optical sensor detects the superficial surface modifications for e.g., the change in R_a as function of GB, and the microfabricated eddy-current sensors records the in-depth profile information within the GZ (i.e. by varying the driving frequency). Therefore, making it possible to even detect the so-called hidden and severe degree of GB resulting in significant microstructural changes in the GB.

the WP surface and recording the reflected light, which is transduced into sensor output voltage (detailed explanation of the sensor electronics can be found in [17]). The voltage drop (ΔV) increased with an increase in R_a, which in turn increased with the degree of GB [17].

Among several applications exploiting the use of eddy-currents such as contact-less breaks in this work, µEC sensors are used for the NDT of microstructural modification within the GZ (i.e., an in-depth profile information of GZ). The microfabrication steps are described in [16, 18]and the characterization of the µEC sensors for NDT of GB is shown in [18, 19]. The µEC sensors exploit the fact of the predominant change in the relative permeability (μ_r) within the GZ in case of WP out of hardened ferromagnetic steels, which are inevitably reliant on the grinding technology for many precision applications [3]. As shown in Figure 1, swirling eddy-currents are induced on and within the conductive WP, when it is in the vicinity of the µEC sensors, whose depth of penetration (δ) as shown in the Equation 1 depends on frequency (f) of the driving current, the conductivity (σ) and μ_r.

$$\delta = \frac{1}{\sqrt{\pi\sigma\mu_{rf}}} \quad (1)$$

Furthermore, in case of occurrence of GB, the microstructural changes within GZ affects the nature of the induced eddy-current in WP [20], which is then recorded either as the impedance change (ΔZ) in case of single µEC sensor (absolute sensor) or the voltage change in case of sender-receiver µEC sensor combination. The main advantage of the use of µEC sensors is the possibility of varying δ of the induced eddy-current by varying f, wherein δ is inversely proportional to f. Therefore, by varying f of the µEC sensors, the depth profile within the GZ can be quantitively recorded and compared with the reference material profile. Thereby, making the µEC sensor suitable for the NDT of not only superficial GB, but also the so-called hidden GB within the GZ. Hence, with this proposed approach of sensor fusion concept, the NDT of GB can be enhanced for the quality testing of ground workpieces. Furthermore, the proposed fused sensor system can be integrated within the grinding machine to facilitate the in-process detection of GB and to increase the material removal efficiency.

3. Results and discussion

For the experimental characterization of the sensor fusion concept, reference GB with varying degrees were fabricated on 42CrMo4 steel cylinder as shown in Figure 2(a) (the grinding parameters and conditions are described in [16]). Each region named as S2-S8 in Figure 2(a) belongs to a different degree of GB, except region S3, which is considered as the reference ground region without GB as a calibration for the sensor measurements. Figure 2(b) shows the micrographs recorded using an optical microscope, wherein the discoloration i.e., the oxidative GB along with the undulation pattern corresponding to the abrasive grains in the grinding wheel can be seen. Furthermore, the surface morphology was further characterized using a white light

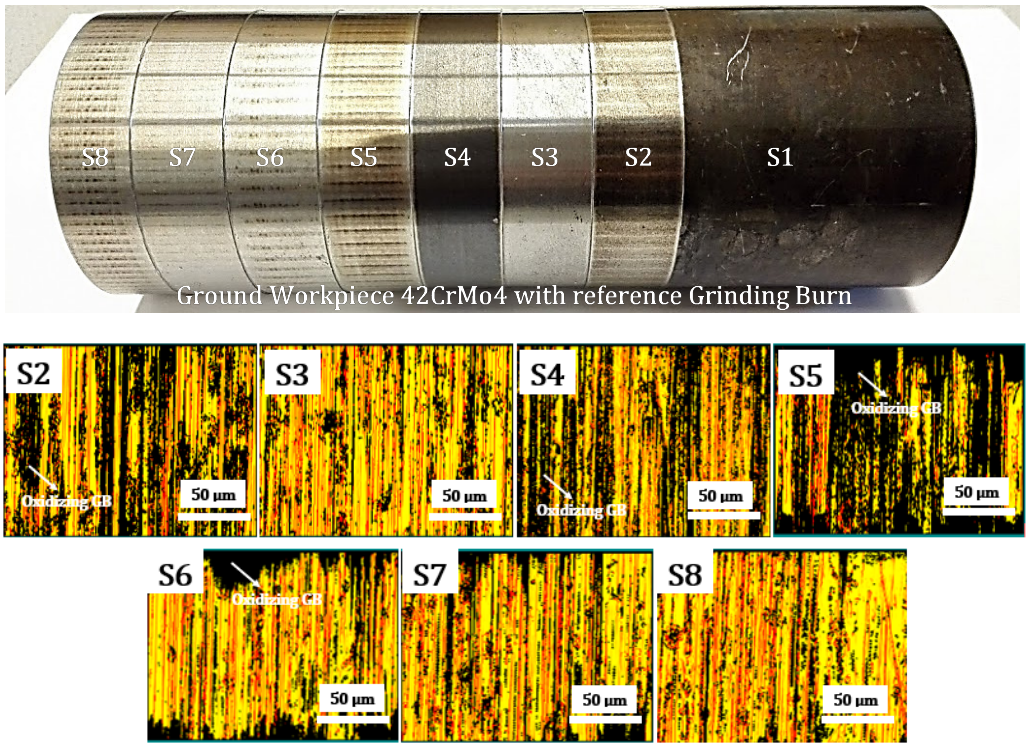


Figure 2: (a) 42CrMo4 cylindrical workpiece ground with reference grinding burn with varying degrees from S2, S4-S8, wherein region S3 is the reference ground region without GB (Region S1 is original unground state of the workpiece), (b) micrographs recorded using an optical microscope for each individual region S2-S8 showing the undulating pattern corresponding to the abrasive grains of the grinding wheel along with the dark discoloration corresponding to the oxidative grinding burn.

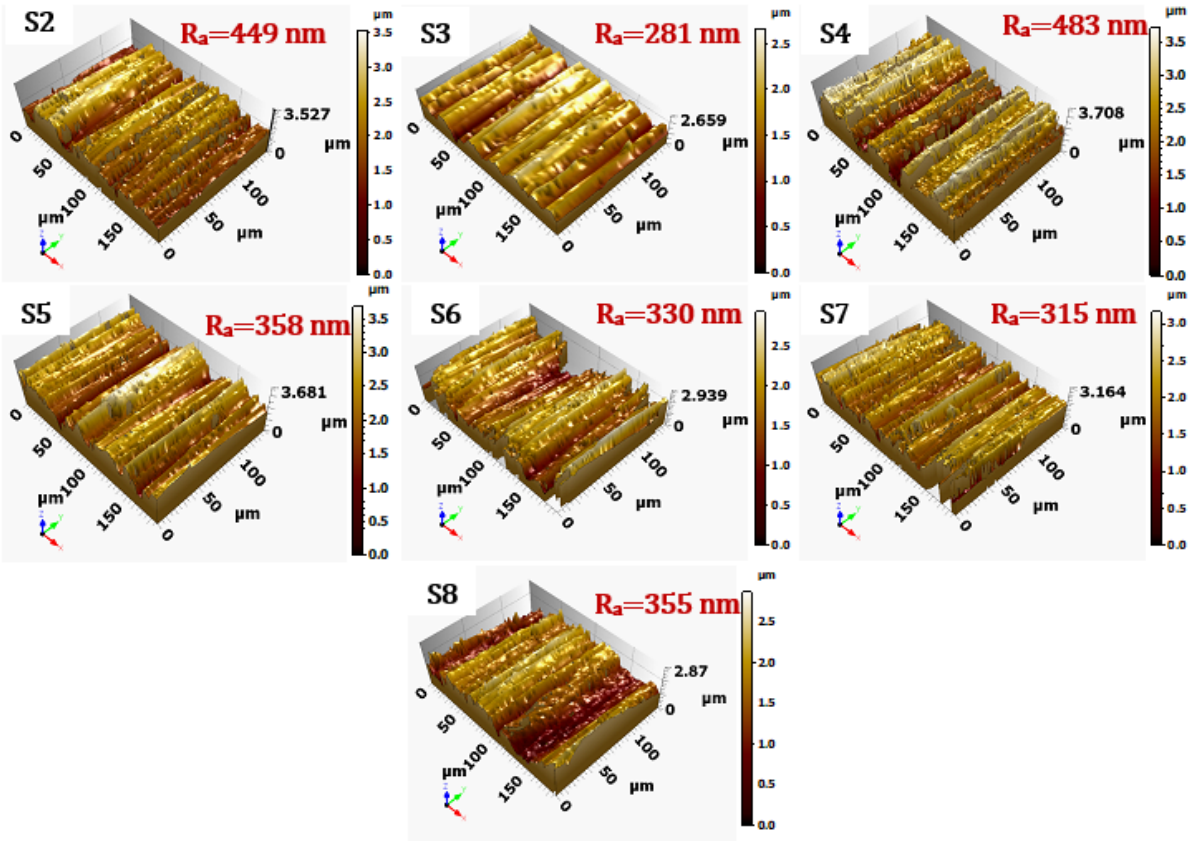


Figure 3: Surface morphology micrographs recorded using a white light interferometer (Zygo Newview) with 50x magnification resulting in a scan area of 188 µm x 140 µm. The average surface roughness values are shown in the top right corner of each micrograph.

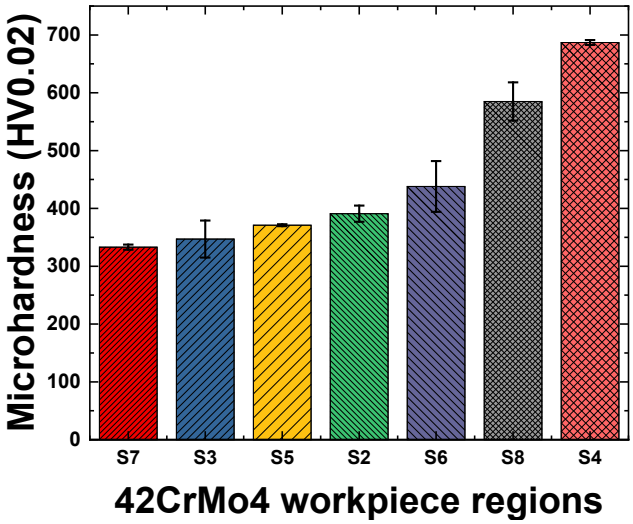


Figure 4: Measured Vickers microhardness on regions S2-S8 on the reference workpiece 42CrMo4 with varying degrees of GB. Each measurement is mean of 5 indentations on each region.

interferometer (Zygo Newview) in order to quantitatively compute the average surface roughness (R_a) values. Figure 3 shows the recorded surface topology micrographs using 50x objective within a scan area of 188 µm x 140 µm for the regions S2-S8 along with the computed R_a values (mean of 10 measurements). The R_a values varied from 281 nm to 483 nm for the reference region S3 and the strongest GB region S4, respectively. This observation can be attributed to the surface morphological modifications occurring on the surface of GZ with the occurrence of GB, which appear not only in the form of discoloration i.e., oxidation GB, but also results in the modification of the surface topology.

Moreover, considering the microstructural modifications as a consequence of GB, which might either soften or re-harden the

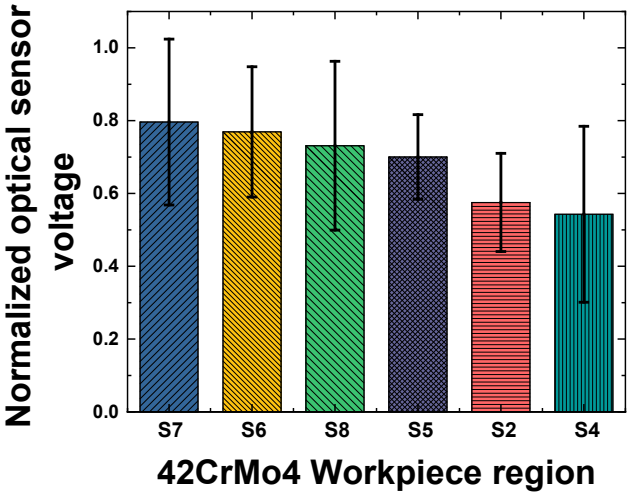


Figure 5: Normalized optical sensor output voltage experimentally measured as a function of varying degrees of GB on regions S2-S8, with a sensor to workpiece distance of 700 µm under quasi-static conditions.

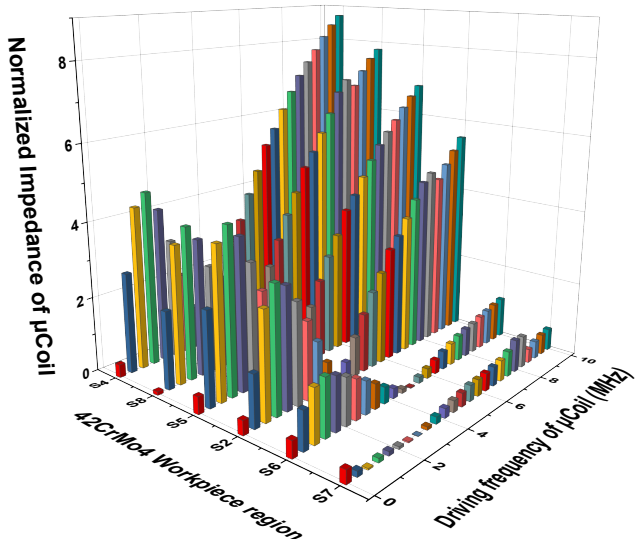


Figure 6: Normalized impedance of 20 turns circular spiral geometry µcoil experimentally measured as a function of varying degree of GB on regions S2, S4-S8, with a µcoil to workpiece distance of 700 µm under quasi-static measurements (the measurements from region S3 are considered as reference).

GZ, the microhardness of each region S2-S8 was measured using Vickers Microhardness device from Innovatest Falcon 501 device with a 20 g load. Furthermore, the indentation depth with the used 20 g load was 1 µm to 2 µm, for the measured microhardness range, and hence the measured microhardness values are more superficial in nature. Figure 4 shows the computed microhardness values (mean of 5 measurements) for the regions S2-S8. The microhardness increased from 333 HV0.02 for region S7 to 685 HV0.02 for region S4. This effect can be attributed to complex microstructural changes, which result in either softening, re-hardening and combination of re-hardened and softened profile within the GZ [6]. From Figure 4, it can be seen that ground region S7 shows the presence of softened layer with reduced microhardness compared to reference region S3, while all other regions showed a re-hardened layer with S8 and S4 showing an almost twofold increment in the microhardness compared to the reference region S3.

The experimental measurements using OS and the µEC sensor were done on a lathe machine to emulate the conditions within a grinding machine. The details of the experimental setup along with the information related to the optical sensor can be found in [17], and the microfabrication steps of µEC sensor with circular spiral geometry (20 turns µcoil used in this work) is reported in [18]. The distance between the sensors and WP was kept constant at 700 µm for all the measurements and the WP was rotated per hand for each new measurement; thereby, resulting in a quasi-static measurement. Figure 5 shows the normalized output voltage of the optical sensor as a function of varying degree of GB on the WP, by taking the voltage recorded from region S3 as reference value. The optical sensor takes into account the surface morphological changes, i.e., either the discoloration as a result of oxidation GB or the modified surface roughness profile on the GZ,

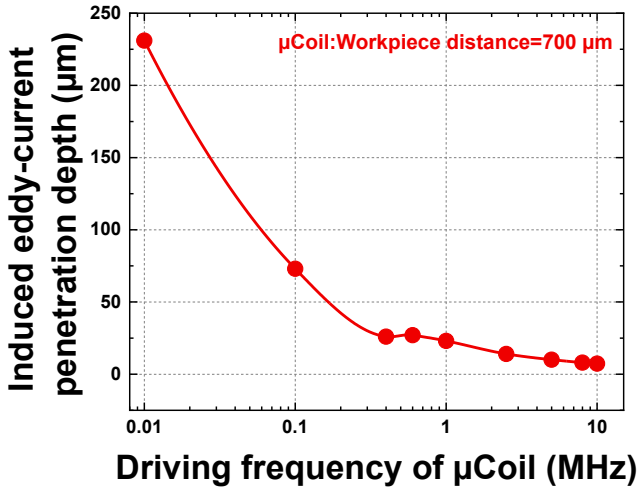


Figure 7: Theoretical computed depth of penetration of induced eddy-current within the 42CrMo4 workpiece at a μEC sensor to workpiece distance of 700 μm and considering the material properties as reported in [21].

both of which would adversely affect the intensity of the reflected light (i.e., in response to the incident light form the optical sensor), which is transduced into voltage by the OS. Therefore, the presence of the GB is detected as the function of voltage change ($\Delta V_{OS} = (V_{ref} - V_{wp})/V_{ref}$) with respect to the reference voltage (V_{ref}) from the reference region S3 and the voltage recorded from ground surface with GB (V_{wp}). As seen in Figure 5, the OS is capable for the NDT of GB, wherein the output voltage changes as a function of the different degree of GB on regions S2 and S4-S8. Furthermore, the recorded ΔV_{OS} values are directly proportional to the R_a values, i.e., the higher the surface roughness, the larger are the ΔV_{OS} values as shown in [17]. As seen in Figure 5, the ΔV_{OS} values dropped by ca. 20 % and 50 % for the regions S5-S8 and regions S2 and S4, respectively. Thereby, depending on the intensity of drop in ΔV_{OS} values, the GB can be roughly classified under moderate (i.e., for ΔV_{OS}=20 %) and strong (i.e., for ΔV_{OS}=50 %) degree of GB. Complementary to the OS measurements, the μEC measurements on the same WP were carried out under the similar working conditions as described above for OS measurements. However, in case of μEC sensors, the impedance response (Z_{wp}) of the μEC sensor as function of the different regions S2-S8 was recorded. Similar to the OS measurements, Z_{wp} for the region S3 was considered as the reference and the normalized impedance change ($\Delta Z = (Z_{ref} - Z_{wp})/Z_{ref}$) for the regions S2, S4-S8 as a function of f is computed and shown in Figure 6. Primary, the ability of the μEC sensors for the NDT of GB is observed for all the ground regions with varying degree of GB, and secondary, on contrary to the OS measurements, where only the surface morphological information was revealed, in case of the μEC sensors, an in-depth profile information is revealed as function of varying f as seen in Figure 6. A very interesting trend of variation in ΔZ as function of f is observed, wherein ΔZ initially increased on increasing f from 1 kHz to 1.2 MHz and decreased further with increase f; however, it increased again from 4 MHz to 10 MHz (the f limit of the used LCR meter was 10 MHz). This effect can be attributed to the

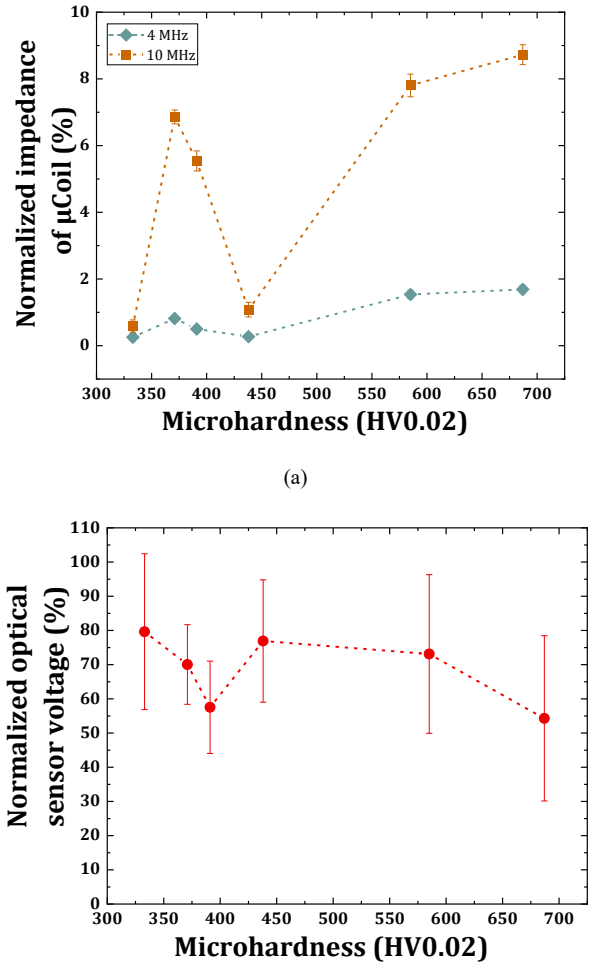


Figure 8: Comparison of sensor fusion concept using the OS and μEC sensors as function of the measured Vickers microhardness values (i.e. destructive approach) (a) normalized impedance of μEC sensor as function of measured Vickers microhardness for the reference GB regions on 42 CrMo4 for 4 MHz (δ = 11 μm) and 10 MHz (δ = 7.5 μm), and (b) normalized optical sensor voltage as function of measured Vickers microhardness values.

microstructural changes occurring within the GZ few hundred micrometers deep in case of the occurrence of GB. At this point it comes to the complementary sensing nature and measurement capability of μEC sensors compared to the OS, where it can quantitatively record the microstructural variation in ΔZ as seen in Figure 6. This effect is because of the variation in δ for the induced eddy-currents as function of f as shown in Equation 1. For e.g., Figure 7 shows the theoretical δ for the induced eddy-currents within the used WP (i.e., 42CrMo4, taking the material properties from [21]), wherein the induced eddy-current tends to be more superficial with increasing driving frequency, i.e., it decreased from 230 μm to ca. 7 μm on increasing the frequency from 0.01 MHz to 10 MHz. Therefore, from Figure 6 it is evident that the microstructural changes within the GZ consist of two zones, namely submerged (>15 μm within GZ) zone and superficial zone (i.e., near to the surface within <15 μm). Furthermore, the ΔZ values increased at higher f, showing that

GB caused severe microstructural changes in the superficial zone in case of ground regions S2, S4, S5 and S8. Additionally, the order of arrangement of the WP ground regions in Figure 6 is with respect to descending order in the ΔZ values (i.e., at higher f), is comparable with the measured microhardness as shown in Figure4. For e.g., the ground region S4 followed by the region S8 showed the maximum microhardness values and hence, a larger ΔZ variation was observed i.e., 8.7 % and 7.8 % at 10 MHz for S4 and S8, respectively, which is evident from Figure 6. Meanwhile, the region S7 showed minimum microhardness and hence, it also showed minimum variation in ΔZ values as seen in Figure 6. Another interesting observation can be noticed for region S6, wherein the ΔZ is larger in the submerged region as compared to the superficial region and an opposite trend is then recorded for region S7. Thereby, showing the complex in-depth nature of the occurrence of the GB. However, in respect to the very similar impedance dependence on frequency of regions S4 and S8, the OS provides a significant signal difference between these surfaces. Moreover, apart from the complementary information related to superficial and in-depth profile obtained using the combination of OS and µEC sensors, the comparison of both sensors merits for the superficial measurement as function of destructive approach i.e., the Vickers microhardness (Figure 4) is shown in Figure 8(a-b). Here, the measured microhardness HV is plotted on the x-axis considering HV as reference for microstructural changes by grinding related GB which is correlated to the µEC (Figure 8a) and OC (Figure 8b) measurements. For the used load of 20 gf an indentation depth of 1 µm – 2 µm is obtained for Vickers microhardness measurements. The normalized impedance ΔZ in Figure 8a is evaluated for 4 MHz (δ = 11 µm) at a frequency where a pronounced minimum in the normalized impedance is found (Figure 6) and for 10 MHz (δ = 7.5 µm) to reveal more the superficial information. It is found that ΔZ for f = 10 MHz initially increased as function of microhardness, however it decreased steeply within the microhardness range of 375 HV0.02 – 450 HV0.02, and subsequently increased as function of microhardness values. At 4 MHz, the minimum of ΔZ occurs at around 450 HV0.02 and the changes of ΔZ towards lower and higher microhardness values are less pronounced. The observed trend can be attributed to the complex microstructural changes occurring within the GZ. ΔV_{OS} showed a different dependence on microhardness and therefore, can provide complementary information about GB compared to µEC especially in the range of 400 HV0.02 - 450 HV0.02. At higher microhardness values, ΔV_{OS} showed a gradual decreasing trend as function of the microhardness values.

Therefore, the sensor fusion by integrating the OS and µEC sensors makes it possible the NDT of GB, by not only considering the surface morphological information on the GZ, but also by revealing the in-depth profile information up to few hundred micrometers within the GZ on occurrence of GB. The presented sensor fusion concept is preliminary approach towards the proof-of-concept for the complementary sensing nature for the NDT of GB. However, there is enough room for the further development and optimization to successfully integrate the developed sensor fusion unit in the grinding machine to realize and fulfill the quest towards the intelligent grinding process.

4. Conclusion

The sensor fusion concept integrating an optical sensor and a microfabricated eddy-current sensor is reported for the non-destructive testing of the grinding burn. The complementary sensing nature of the reported sensor combination results in not only the superficial non-destructive testing of grinding burn, but also an in-depth profile within the grinding zone. The optical sensor voltage dropped as function of varying degree of grinding burn, which was observed to be directly correlated to the average surface roughness of the grinding burn. The voltage dropped by 20 % for moderate degree of grinding burn, and by 50 % for strong degree of grinding burn. The in-depth quantitative profile of the microstructural changes within the grinding zone is recorded with the use of the microfabricated eddy-current sensors (circular spiral µcoils with 20 turns) as a function of the varying driving frequency. A very interesting nature of the grinding burn was observed as function of the impedance variation of the microcoils with two distinct zones, i.e., the sub-merged zone (>15 µm within grinding zone) and superficial zone (from WP surface to 15 µm within grinding zone). Impedance changes of ca. 8 % and 4 % were obtained for stronger degree of grinding burn (i.e., regions S4 and S8) for the superficial and sub-merged grinding burn at 10 MHz and 2.5 MHz, respectively. Utilizing the impedance information of the grinding zone as function of varying driving frequency, it could also be possible to build qualitatively the profile of the microstructural changes for quality testing, and hence ease the observation for the severity of the grinding burn. The reported sensor fusion concept can be further optimized to increase the detection sensitivity and robust packaging immune to the harsh environmental conditions occurring within the grinding machine to realize intelligent grinding processes and in-process monitoring capabilities corresponding to the requirements of industry 4.0. Furthermore, to model the surface morphology of grinding burn, an in-depth surface roughness analysis can be done (considering the parameters beyond root mean squared and average surface roughness values) to unveil the surface modifications at the atomistic level.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank the company ibg Prüfcomputer, Ebermannstadt, Germany, for providing the industrial support with their expertise and know-how for the conceptualization and development of the microfabricated eddy-current sensors for the non-destructive detection of the grinding burn. The authors would like to express their gratitude to the State Government of Baden-Württemberg, Germany, Landesbank Baden-Wuerttemberg in Germany and European Union for the financial support withing the framework of European Regional Development Fund (EFRE)

and HAW-KMU (***32-7545.20/27/57***). The authors highly appreciate the student work done by Mr. Pranav Dhumal and Mr. Jerrine Jacob Mathews for the characterization of the optical sensor.

References

[1] W. B. Rowe, Principles of Modern Grinding Technology, 2nd edition Burlington: Elsevier Science, 2013.

[2] B. He, C. Wei, S. Ding, and Z. Shi, "A survey of methods for detecting metallic grinding burn," Measurement, **134**, 426–439, 2019, doi: 10.1016/j.measurement.2018.10.093.

[3] R. Ito, T. Azuma, T. Kasuga, S. Soma, S. Murakami, and T. Kuriyagawa, "Development of Non-Destructive Inspection System for Grinding Burn - An Application of the Grinding Burn Detecting Technique to Evaluate Residual Stress," AMR, **797**, 307–312, 2013, doi: 10.4028/www.scientific.net/AMR.797.307.

[4] M. W. Seidel, Schleifbrand und dessen Prüfung: Leitfaden für die Praxis. München: Hanser, Carl, 2020.

[5] M. W. Seidel, A. Zösch, and K. Härtel, "Grinding burn inspection," Forschung im Ingenieurwesen, **82**(3), 253–259, 2018, doi: 10.1007/s10010-018-0270-4.

[6] R. Ito, N. Mukaide, T. Azuma, S. Soma, S. Murakami, and T. Kuriyagawa, "Development of Non-Destructive Inspection System for Grinding Burn-in-Process Detection of Grinding Burn," AMR, **1017**, 135–140, 2014, doi: 10.4028/www.scientific.net/AMR.1017.135.

[7] A. Jeffrey Badger and Andrew Torrance, "Understanding the causes of grinding burn helps alleviate the problem," Cutting Tool Engineering, **52**(12), 1-3, 2000.

[8] T. Kaufmann, S. Sahay, P. Niemietz, D. Trauth, W. Maas, and T. Bergs, "AI-based Framework for Deep Learning Applications in Grinding," in 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Herlany, Slovakia, 195–200, 2020.

[9] L. Lv, Z. Deng, T. Liu, Z. Li, and W. Liu, "Intelligent technology in grinding process driven by data: A review," Journal of Manufacturing Processes, **58**, 1039–1051, 2020, doi: 10.1016/j.jmapro.2020.09.018.

[10] W. Guo, B. Li, and Q. Zhou, "An intelligent monitoring system of grinding wheel wear based on two-stage feature selection and Long Short-Term Memory network," Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, **233**(13), 2436–2446, 2019, doi: 10.1177/0954405419840556.

[11] H. B. Hübner, M. A. V. Duarte, and R. B. Da Silva, "Automatic grinding burn recognition based on time-frequency analysis and convolutional neural networks," The International Journal of Advanced Manufacturing Technology, **110(**7-8), 1833–1849, 2020, doi: 10.1007/s00170-020-05902-w.

[12] F. Junejo, "The Application of Artificial Intelligence in Grinding Operation using Sensor Fusion," Geomate, **12**(30), 2017, doi: 10.21660/2017.30.160503.

[13] L. Wang and R. X. Gao, Condition monitoring and control for intelligent manufacturing. London: Springer, 2006.

[14] F. Godoy Neto R., Marchi M., Martins C., R. Aguiar P. and Bianchi E, "Monitoring of Grinding Burn by AE and Vibration Signals," in Proceedings of the 6th International Conference on Agents and Artificial Intelligence, 272–279, 2014.

[15] B. Botcha, V. Rajagopal, R. Babu N, and S. T. Bukkapatnam, "Process-machine interactions and a multi-sensor fusion approach to predict surface roughness in cylindrical plunge grinding process," Procedia Manufacturing, **26,** 700–711, 2018, doi: 10.1016/j.promfg.2018.07.080.

[16] Isman Khazi, Andras Kovacs, Ali Zahedi, Christian Reser, Ulrich Mescheder, Bahman Azarhoushang and Christoph Reich, "Real time In-Situ Quality Monitoring of Grinding Process using Microtechnology based Sensor Fusion," in 2020 IEEE International Conference on Semiconductor Electronics (ICSE), Kuala Lumpur, Malaysia, 180-184, 2020, doi: 10.1109/ICSE49846.2020.9166898

[17] A. Kovacs, I. Khazi, A. Zahedi, U. Mescheder and B. Azarhoushang, "Development of an Optical Sensor for the Non-Destructive Testing of Grinding Burn," in 2020 IEEE SENSORS, Rotterdam, Netherlands, 1-4, 2020. doi: 10.1109/SENSORS47125.2020.9278837

[18] Isman Khazi, Andras Kovacs, Vaibhav Kumar, Pranav Dhumal, Ulrich Mescheder, "Microfabricated 2D planar eddy-current microcoils for the non-destructive testing of grinding burn marks," Nondestructive Testing and Diagnostics, **3**, 29–35, 2019, doi: 10.26357/BNiD.2019.012

[19] I. Khazi, A. Kovacs, A. Zahedi, U. Mescheder and B. Azarhoushang, "Microfabricated Eddy-Current Sensors for Non-Destructive Testing of the Micro Grinding Burn," 2020 IEEE SENSORS, Rotterdam, Netherlands, 1-4, 2020, doi: 10.1109/SENSORS47125.2020.9278595

[20] N. Bowler, Eddy-Current Nondestructive Evaluation. New York, NY: Springer New York, 2019.

[21] T. Bulin, E. Svabenska, M. Hapla, P. Roupcova, C. Ondrusek, and O. Schneeweiss, "Magnetic properties of 42CrMo4 steel," IOP Conference Series: Materials Science and Engineering, **179**, 1-6, 2017, doi: 10.1088/1757-899X/179/1/012010.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1458-1465 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Comparative Study Between Three Methods for Optimizing the Power Produced from Photovoltaic Generator

El hadji Mbaye Ndiaye[1,*], Mactar Faye[1,2], Alphousseyni Ndiaye[1,2]

[1]*Research team Efficiency and Energy System, Alioune Diop University of Bambey, BP 30, Senegal*

[2]*Laboratory of Water, Energy, Environment and Industrial Processes, ESP, Dakar-Fann, S-10700, Senegal*

A R T I C L E I N F O

Article history:
Received: 17 October, 2020
Accepted: 26 November, 2020
Online: 21 December, 2020

Keywords:
MPPT
PSO
ANFIS
Photovoltaic Generator
Artificial Intelligence

A B S T R A C T

The technics of maximum power point tracking is widely used in solar photovoltaic energy and electric power system applications. Traditionally, these technics are based on conventional methods like perturb and observe and incremental conductance. In this work, three methods based on particle swarms optimization, incremental conductance and adaptive neuro fuzzy inference system are presented. A comparative study is carried out. The study of this paper shows that there is a limitation in the incremental conductance method. To overcome the shortage of this last method, particle swarms and adaptive neuro fuzzy optimization methods are used. The behaviors of the three methods are compared and evaluated in simulation under matlab/simulink. Results demonstrate that the adaptive neuro fuzzy inference system is effective for photovoltaic power optimization even for non-uniform climatic conditions. It has the best performances followed by the particle swarms method.

NOMENCLATURE

ANFIS	Adaptive Neuro Fuzzy Inference System	Ipv [A]	PV current
PVS	Photovoltaic Solar	RMSE [-]	Root Mean Square Error
MPP	Maximum Power Point	MSE [-]	Mean Square Error
InC	Incrementale Conductance	MAPE [%]	*Mean Absolute Percentage Error*
PI	Proportional Integral	AIC [-]	Akaike Information Criterion
PID	Proportional Integral Derivated	C1 [µF]	Boost input capacity
MsF	Membership	C2 [µF]	Boost output capacity
MPPT	MPP Tracking	f [kHz]	Frequency of switching the Mosfet
P&O	Perturb and Observ	α [-]	Duty cycle
PVM	Photovoltaic Module	L [mH]	Inductance
GA	Genetic Algorithm	Vpv [V]	PV voltage
FIS	Fuzzy Inference System	Pref [W]	Mesured Power
FL	Fuzzy Logic	Imax [A]	Current at Maximum Power
STC	Standard Test Condition	Vmax [V]	Voltage at Maximum Power
PWM	Pulse Width Modulation	Popt [W]	Optimized Power with ANFIS
PSO	Particle Swarm Optimization	m [-]	coefficient of inertia
ANN	Artificial Neural Network	Q_i [-]	acceleration coefficient
ACO	Ant Colony Optimization	r_i [-]	random number
AI	Artificial Intelligence	Ir [W/m²]	Irradiation
DC	Direct Current	T [°C]	Temperature
Ppv [W]	Photovoltaic power	F [W]	Fitness function
		R_{MPPT} [%]	MPPT efficiency

*Corresponding Author: El hadji Mbaye Ndiaye, Université Alioune DIOP de Bambey, elhadjimbaye.ndiaye@uadb.edu.sn

1. Introduction

Today's world sees rapid industry development and this makes us more energy dependent. Fossil sources occupy the largest share of electricity production [1], [2]. The use of conventional energies (around 87% of global energy consumption) has an undesirable impact on the environment in terms of greenhouse gas (GHG) emissions and safety (nuclear accidents). 36.5 billion tons of CO_2 is emitted in 2020 [3]. To overcome these problems, it is necessary to resort to alternative energies [4–9]. Among them is Photovoltaic Modules (PVM) solar energy. Its annual growth rate over the last ten years is estimated at more than 40% [10]. Despite their low efficiency, PVM still have a high price. To get around this efficiency problem, techniques for optimizing the power generated by the PVM are proposed [11], [12–16]. This technique, makes it possible to run after the maximum power that the module is capable of supplying [17]. The generated power depends on weather conditions irradiation and temperature [18–20]. Under these latter conditions, the electrical characteristics of the PVM have only one optimal. So not too much trouble for the technique to converge the system to the MPP. Under non-uniform conditions of irradiation and temperature, the electrical characteristics of the PVM present several optimal points. The algorithm used must therefore be able to distinguish the global optimum from the local optima [21]. It therefore requires a sophisticated algorithm which will be able to make a global exploration of the search space in order to make the system converge towards the global MPP [22].

The most used of the classical methods are the P&O method and the InC method [5], [23–26]. Several researchers have used the P&O technique in their work. This method, based on the voltage disturbance and the observation of the variation of the power, is very widespread in the literature. Originally, it was designed to exceed the limits of other types of deterministic controls such as Hill Climbing, Open Circuit Voltage Fraction, etc. It also presents limits linked mainly to the response time and the numerous oscillations around the MPP. The InC method is proposed [3, 18, 25]. In [27], according to the results, the incremental conductance algorithm performs better than the perturb and observe algorithm. To overcome the problems linked to the limits of the methods mentioned above under variable of climatic conditions, Soft-Computing methods based on Meta-heuristic algorithms and Artificial Intelligence (AI) algorithms are proposed. MPPT techniques are based almost exclusively on these techniques. They are very numerous and diverse. They range from Evolutionary Algorithms (Genetic Algorithm), Meta-heuristic algorithms (Optimization by Particle Swarms) and AI algorithms (Artificial Neural Networks, Fuzzy Logic, ANFIS) [21], [10, 11, 28–31].

Artificial Neural Networks (ANN) have been used in several works to optimize the power delivered by a PVM. Its operating principle is inspired by that of the human brain. With their great generalization capacity, they are used for solving complex optimization problems [32], [33]. This is the case in [34] where MPPT method based on ANN is compared to MPPT method based on P&O. The results show that ANN method is more robust than P&O method, regardless of the operating conditions of the PVM. Its limits lie in its lack of interpretation and the difficulty of determining the appropriate number of Layer/Neurons. In addition, these latter limits represent strong points for the Fuzzy Logic (FL) algorithm. The interpretive inability encountered with ANN is resolved by the use of LF. It uses its linguistic variables to overcome this problem [35], [36]. Using the classical logic process, it has facility for extension and interaction [37], [38].

FL also has limits which can be circumvented. Which makes them two complementary techniques and the combination of which gives the Neuro-Fuzzy technique including ANFIS.

In reference [39], two ANFIS models are proposed for grid-connected PV system current injection and battery control. The results show that the proposed the models offered allow the battery charging/discharging process to be supervised and at the same time injecting good quality energy into the grid.

In [40] two MPPT techniques based on ANFIS and P&O are compared. Comparison results show that the technique based on the ANFIS algorithm is more robust.

In [5] Five MPPT techniques are proposed. A comparative study is carried out between them. The simulation results show that the best performances are obtain with ANFIS with an efficiency of 99.4%, against 98.1% and 97.5% respectively for FL and P&O.

The power optimization technique using the ANFIS algorithm is more robust than other techniques such as FL, ANN, InC and P&O. It overcomes the problems encountered with ANN and FL as it is a complementary technique linking the two [28, 38, 41].

Other types of techniques based on algorithms whose principle is inspired by the evolution of nature are presented in the literature [11, 42]. Among them there are methods based on the Particle Swarms Optimization (PSO) algorithm [43, 44].

In [45], a comparative study between techniques using PSO, InC and P&O is carried out. Results of simulation show that the PSO technique is the more robust with the most low response time compared to the two others techniques (InC and P&O). In [46], authors proposed a comparative study between PSO, P&O and FL techniques to optimize the MPP of the PVM. Results show that the MPPT technique based on PSO outperformed FL and P&O techniques.

The work in this paper is consisting to do a comparative study between MPPT techniques based on ANFIS, PSO and InC algorithms put under the same conditions in order to indicate the most efficient and the most robust for power optimization problems. A validation of these three commands is done using an experimental database. The scientific contribution of this paper is to design a MPPT technique based on ANFIS algorithm using database collected to solar power plant in tropical zone.

The rest of the article is organized as follows. Section II presents the MPPT proposed approaches. Section III the proposed approach for the MPPT techniques. Section IV presents the results of simulations and discussions. Finally, a conclusion is made in section V.

2. Proposed approach

In solar PV system, power delivered by the PVM is not always the maximum. This is due to the phenomenon of intermittence. As a result, the operating point of the PVM is not even the MPP. It so requires a technique which is able to extract the MPP of the PVM. This technique, called MPPT optimizes the power through the generation of a duty cycle for controlling the static converter.

The block diagram of the studied system is given in Figure 1.

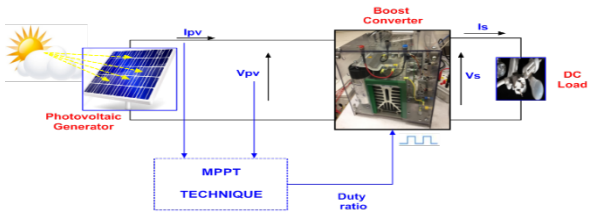


Figure 1: Schematic diagram of the MPPT technique

2.1. Incremental Conductance method based MPPT

The InC algorithm is among the most efficient of the classical algorithms. It is based on the cycle of the change in pressure to the change in voltage of the PVM (equation 1). At MPP, the slope is zero. The tracking is done according to the position of the operational point (or slope) (dPpv / dVpv) relative to the MPP (equation (3)). The latter depends on the value of the conductance (Ipv / Vpv). The sign of the latter indicates whether the MPP is reached or not (equation (2)). It is compared to its increment. This amounts to saying that the MPP depends on the voltage variation and that of the current [18,26],[47,48]. Thus, the algorithm increments or decrements the duty cycle of the static converter to continue the MPP. The flowchart is shown in Figure 2.

$$\frac{dP_{pv}}{dV_{pv}} = I_{pv} + V_{pv}\frac{dI_{pv}}{dV_{pv}} \quad (1)$$

The MPP is found when:

$$\frac{dP_{pv}}{dV_{pv}} = 0 \rightarrow \frac{dI_{pv}}{dV_{pv}} = -\frac{I_{pv}}{V_{pv}} \quad (2)$$

The sign of the slope indicates the direction of evolution of the MPP according to equation (3).

$$\begin{aligned} &\frac{dP_{pv}}{dV_{pv}} > 0, left\ side\ of\ MPP \\ &\frac{dP_{pv}}{dV_{pv}} < 0, for\ V_{pv} > V_{MPP}, right\ side\ of\ MPP \\ &\frac{dP_{pv}}{dV_{pv}} = 0, on\ the\ MPP \end{aligned} \quad (3)$$

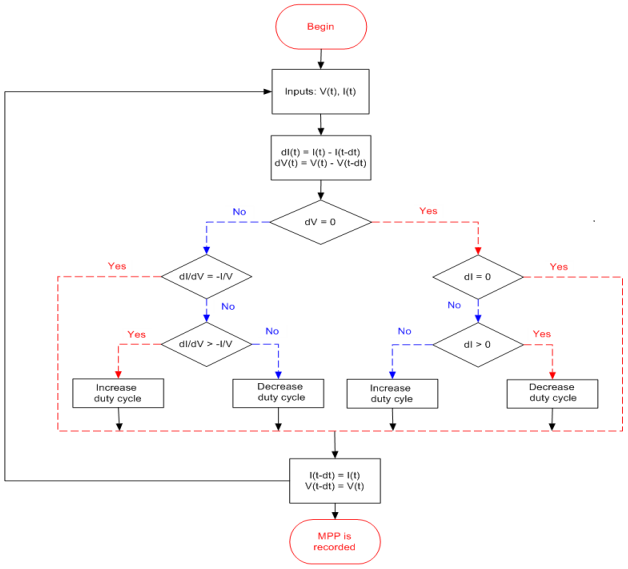


Figure 2: Incremental Conductance flowchart based MPPT

2.2. PSO method based MPPT

Particle Swarm Optimization is an evolutionary meta-heuristic approach. It is used for solving optimization problems. Its principle is based on the behavior of particles (individuals) [49]. In this paper we use a swarm of birds (Figure 3). In this type of swarm, collective intelligence is involved. Particles converge towards those with the best performance [50].

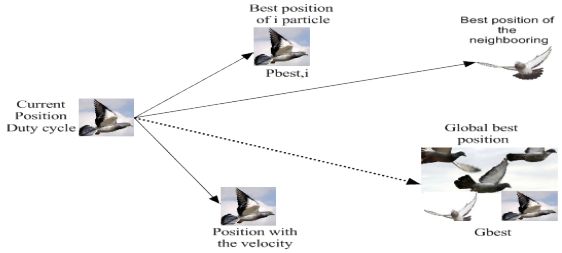


Figure 3: Movement of a particle in the swarm of birds

The duty cycle α of the boost converter represents the particles. Velocity and position are initialized to begin and the movement of the particles are described by equations (4) and (5). For evaluating the best position of the particles, fitness function is calculated according to the equation (6). Vpv and Ipv are calculated for each particle i with a fixed position in the search space [αmin, αmax]. The algorithm converges the system towards the global optimum. For this the duty cycle is initiated. This ratio, depending on P_{besti} and G_{best}, is corrected if it deviates from the best overall duty cycle as in the principle of Figure 3. The particles are then evaluated in terms of position and velocity according to equations (7) and (8). Updates are made to re-evaluate the optimum duty cycle for controlling the boost converter. Table 1 gives the parameters of the PSO. These parameters are defined for the PSO simulations. The flowchart is shown in Figure 4.

Table 1: PSO implementation parameters

Parameters	Values
Q_1	1,2000
Q_2	2
m	0,4000
Iterations	25

$$\alpha_i^{t+1} = \alpha_i^t + \left(\frac{d\alpha}{dt}\right)_i^{t+1} \quad (4)$$

$$\left(\frac{d\alpha}{dt}\right)_i^{t+1} = w\left(\frac{d\alpha}{dt}\right)_i^t + Q_1 r_1 (P_{besti} - \alpha_i^t) + Q_2 r_2 (G_{best} - \alpha_i^t) \quad (5)$$

$$F_i(\alpha^t) = P_{pv,i}^t \quad (6)$$

$$P_{besti} = \alpha_i^t \quad if \quad F_i(\alpha^t) \geq F_i(P_{besti}) \quad (7)$$

$$G_{best} = \max(P_{besti}) \quad (8)$$

F is the fitness function,

α is the duty cycle of the boost converter,

t is the time of simulation

r_1 and r_2: uniformly pull in [0,1].

w: coefficient of inertia.

Qi: acceleration coefficient

P_{besti} is the personal best position of particle *i*

G_{best} is the best position of the particles in the entire population.The MPPT technique based on PSO algorithm is described by equations (4) to (8).

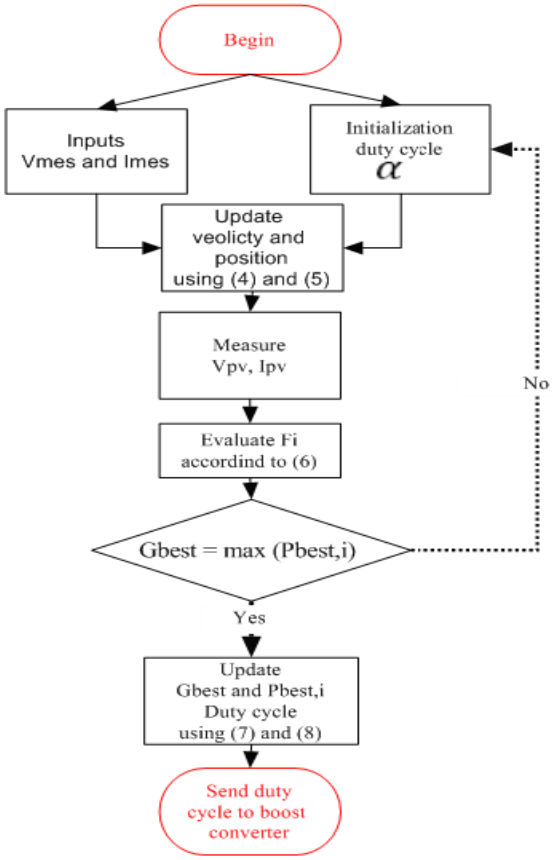


Figure 4: PSO flowchart based MPPT

2.3. ANFIS method based MPPT

ANFIS is an adaptive neuro-fuzzy inference system that has five layers in its structure. These layers refine the fuzzy rules already established by human experts and readjust the overlap between the different fuzzy subsets [51–54].

The neural structure replaces the hidden layers with fuzzy rules. This further simplifies learning and interpreting the results obtained. The structure proposed in this work receives the voltage and current from the PVM as inputs and supplies the output with an optimal power comparable to that of the PVM.

The architecture of the MPPT ANFIS technique is given in Figure 5.

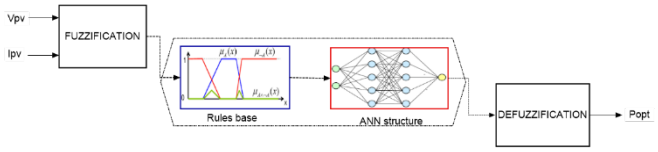


Figure 5. Architecture of the MPPT ANFIS technique

Layer 1: Each Neuron calculates the degree of truth of a fuzzy subset by its transfer function. It is called fuzzification layer. Numeric input values are converted to linguistic variables (equations 9 and 10) for high interpretability.

$$O_i^1 = \begin{cases} \mu_{Ai}(V_{pv}) & for \quad i = 1,2 \\ \mu_{B(i-2)}(I_{pv}) & for\ i = 3,4 \end{cases} \tag{9}$$

With:

$$\mu_{Ai}(V_{pv}) = exp\left[-\left(\frac{Vpv-a_i}{b_i}\right)^2\right] \tag{10}$$

Layer 2: It calculates the degree of activation of antecedents (premises). This is the layer of fuzzy rules (equation 11).

$$w_i = \mu_{Ai}(V_{pv}) * \mu_{Bi}(I_{pv}) \tag{11}$$

Layer 3: It normalizes the degree of activation of the rules: it is the normalization layer (equation 12).

$$G_i = \frac{w_i}{w1+w2} \tag{12}$$

Layer 4: It determines the parameters of the consequence of the fuzzy rules. Previous linguistic variables will be translated again into numeric values before being sent to the last layer. It is defuzzification (equation13).

$$O_i^4 = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{13}$$

Layer 5: It calculates the overall output of the system: it is the output layer. Equation (14) gives the expression of the optimal power generated by the PVM with the ANFIS technique for three fuzzy subsets.

$$P_{opt} = \sum_{i=1}^{nit}\left\{P_{opt,i}.\frac{e^{-\left[\left(\frac{Vpv-a_i}{b_i}\right)^2+\left(\frac{Ipv-a_{i-2}}{b_{i-2}}\right)^2\right]}}{\sum_{i=1}^{3}\left(e^{-\left[\left(\frac{Vpv-a_i}{b_i}\right)^2+\left(\frac{Ipv-a_{i-2}}{b_{i-2}}\right)^2\right]}\right)}\right\} \tag{14}$$

where ai and bi are the parameters of the premise of the MsF, and Popt is the optimal power delivered by the ANFIS controller.

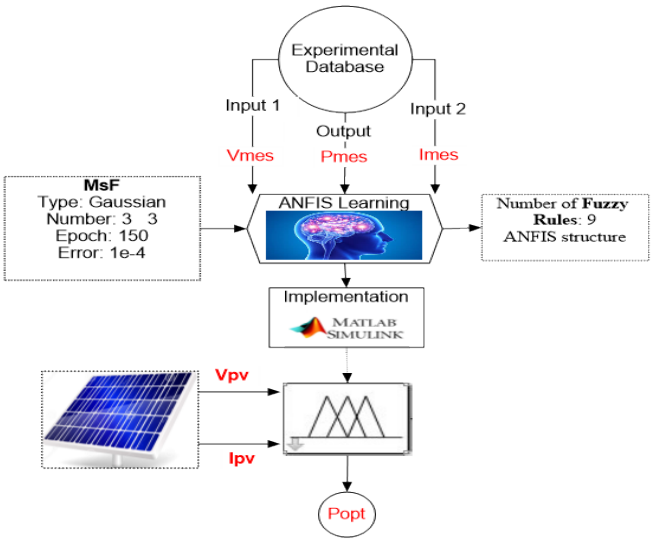
With nit the number of iterations when learning the ANFIS algorithm. Figure 6 gives the flowchart of the technique implemented in simulink and table 2 the learning parameters. The matlab "anfisedit" interface is used for learning the MPPT ANFIS command with a real database made up of two inputs (Vpv and Ipv) and one output. The number and type of MsF are fixed as well as the number of iterations. The hybrid learning algorithm is used and an error tolerance of 1e-4 is arbitrarily set.

Table 2 : ANFIS learning parameters

Parameters	Values
FIS	Takagui-Sugeno
MsF (Type)	Gaussian
MsF (Number)	3 3
Fuzzy rules	9
Epochs	10
RMSE (Training)	0.000123
RMSE (Checking)	0.002012

3. Simulation Results and Discussions

In this section, the results of simulations under matlab/simulink are presented. These simulations are performed with a real database. The latter first allowed to digitally characterize the PVM before being used for the validation of the three techniques presented in previous sections. The PVM consists of two PVs in series. The photovoltaic platform shown in Figure 7 is used in this study. The characteristics of the photovoltaic system being identified (Sharp Module) are given in Table 3. This platform is located at the Polytechnic high school of Cheikh Anta Diop University, Dakar, Senegal. This country is a tropical zone with an adequate rate of sunshine (5.7 kWh/m²/day) for the installation of solar PV plant.



ANFIS flowchart based MPPT

Table 3: Characteristics of the PV

Parameters	Values
V_{CO}	20 V
V_{MPP}	16 V
I_{SC}	2.5600 A
I_{MPP}	2.4300 A
P_{MPP}	38.3800 W



Experimental bench

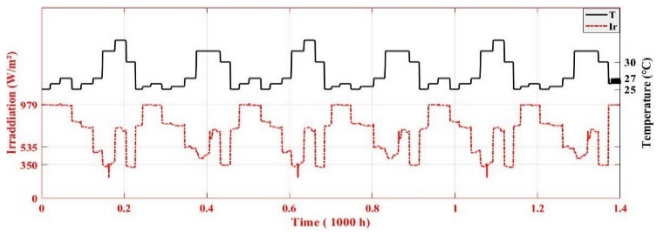


Figure 8: Experimental database of irradiation and temperature

In this subsection, the results obtained by the three MPPT methods are visualized. A comparative study is then carried out in order to detect the best order. Table 4 shows the simulation parameters and the Figure 9 gives the Simulink model.

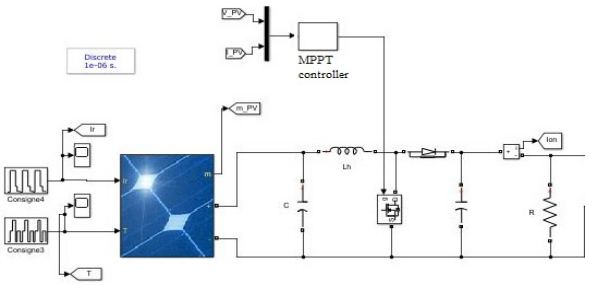


Figure 9: System simulated under matlab/simulink

Table 4: System electrical parameters

Parameters	Values
Input capacity	200.6230 µF
Inductance	1 mH
Output capacity	480 µF
Load	34.8000 Ω
Switching frequence	15 kHz

The RMSE and MAPE criteria are evaluated at the level of equations (17) and (18). They characterize the difference between the power generated by the control and the real power (Figure 14). The efficiency of the MPPT technique is given by equation (19).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_{ref} - P_{pv_i}\right)^2} \quad (17)$$

$$MAPE = 100 * \frac{1}{n}\sum_{i=1}^{n}\left|\frac{P_{ref} - P_{pv_i}}{P_{ref}}\right| \quad (18)$$

$$E_{MPPT}(\%) = \frac{P_{PV,MPPT}}{P_{Mesure}} \quad (19)$$

Figure 12 and Figure 13 respectively show the currents and voltages of the PVM obtained with the different MPPT techniques. They follow changes in weather conditions (Figure 8) which have a considerable influence on the point of operation of the PVM. In Figure 12, we see that the currents obtained with the PSO and ANFIS techniques are almost identical while that obtained with the InC technique is much lower. The opposite phenomenon is observed on the voltage curves in Figure 13. Indeed, the MPPT InC technique pursues the MPP first by comparing the voltage and the current of the PVM. Then, as the principle is based on conductance, the control performs compensation to achieve the desired MPP. Only, as it evaluates each time the variation of the power compared to the voltage, it often diverges with an enormous overshoot of the voltage. In addition, being also a static control, it takes a relatively long time to adapt to a climatic disturbance. This induces a small and slow variation in its duty cycle (Figure 10). As a result, it has difficulty extracting maximum power for these non-uniform irradiation and temperature conditions. Figure 11 represents the power curves of the three controls with respect to the measured power (reference). It reveals overruns for the PSO

and InC techniques while the ANFIS technique follows the set-point with an accuracy of 93.46%.

Furthermore, the performance criteria presented in Figure 14 show that the ANFIS technique is better with an extremely low RMSE (0.0194), no overshoot (D=0) and a higher efficiency of around 99.9984%. It is followed by the PSO technique with a RMSE of 1.7235 and an efficiency of 94.8748%.

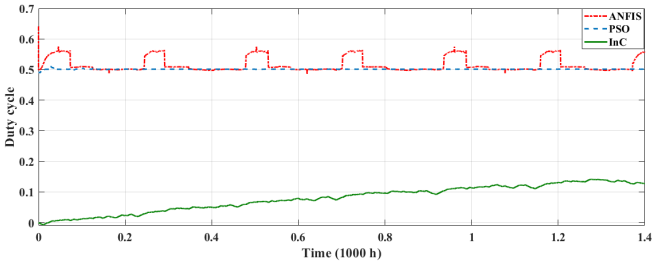


Figure 10: Variation in the duty cycle

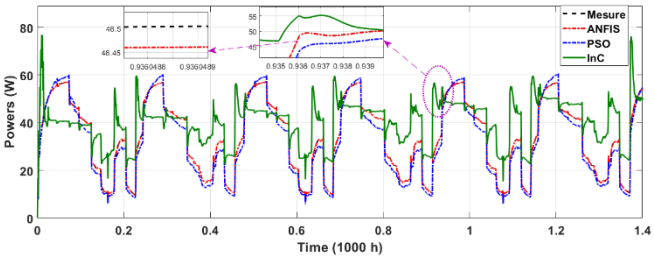


Figure 11: PVM powers

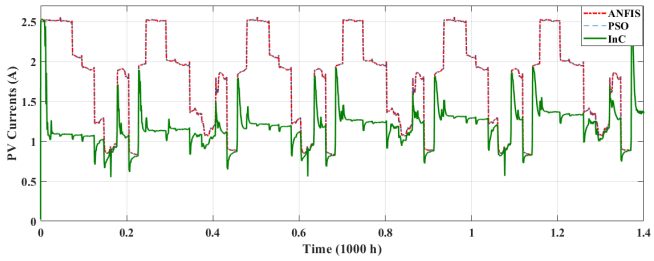


Figure 12: Currents of the PVM

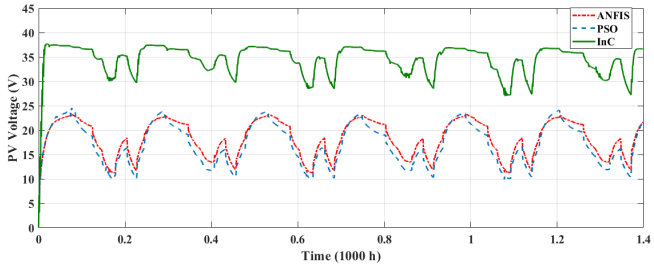


Figure 13: PVM voltages

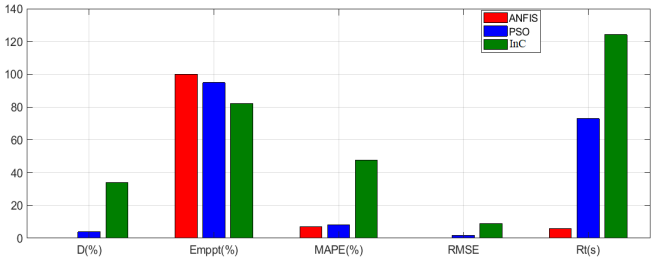


Figure 14: MPPT controller's performance parameters

These results corroborate those found in the literature which highlight the oscillating nature of the power obtained with the InC technique due to their inability to detect with precision the overall maximum in a situation of non-uniform climatic conditions [3]. They also confirm the results presented in [38] where the authors made a comparison between ANFIS and InC.

4. Conclusion

MPPT techniques are used to optimize power generated by PVM. In this work, a comparative studied between PSO, ANFIS and InC is done. An experimental validation of these MPPT techniques is also done using real database. The simulation results show that the MPPT technique based on ANFIS algorithm has the best performances. However, if we have to choose between these methods, the two parameters to consider are the complexity of the implementation and the performance of the method. For the first criterion, the choice will be relatively focused on the InC because of its great ease of implementation. For the second criterion, ANFIS will be chosen because of its best performance in terms of response time and accuracy. But we can conclude that the ANFIS and PSO methods give good results compared to the InC method.

Conflict of Interest

The authors declared that there is no conflict of interest.

Acknowledgments

The authors thank the efficiency and energy system research team of Alioune Diop University of Bambey (Senegal) and the Laboratory of Water, Energy, Environment and Industrial Processes, ESP (Dakar-Senegal).

References

[1] W.H. Ali, P. Cofie, J.H. Fuller, S. Lokesh, E.S. Kolawole, "Performance and Efficiency Simulation Study of a Smart-Grid Connected Photovoltaic System," Energy and Power Engineering, 71–85, 2017, doi:10.4236/epe.2017.92006.

[2] E. hadji M. Ndiaye, A. Ndiaye, M. Faye, S. Gueye, "Intelligent Control of a Photovoltaic Generator for Charging and Discharging Battery Using Adaptive Neuro-Fuzzy Inference System," International Journal of Photoenergy, **2020**, 144–156, 2020.

[3] S. Shabaan, M.I.A. El-sebah, P. Bekhit, "Maximum power point tracking for photovoltaic solar pump based on ANFIS tuning system," Journal of Electrical Systems and Information Technology, **5**(1), 11–22, 2018.

[4] N. Vela, F. Chenlo, "Early degradation of silicon PV modules and guaranty conditions," **85**, 2264–2274, 2011, doi:10.1016/j.solener.2011.06.011.

[5] M.A. Enany, M.A. Farahat, A. Nasr, "Modeling and evaluation of main maximum power point tracking algorithms for photovoltaics systems," Renewable and Sustainable Energy Reviews, **58**, 1578–1586, 2016.

[6] F. Chekired, A. Mahrane, M. Chikh, Z. Smara, "Optimization of energy management of a photovoltaic system by the fuzzy logic technique," Energy Procedia, **6**, 513–521, 2011, doi:10.1016/j.egypro.2011.05.059.

[7] V. Indragandhi, R. Logesh, V. Subramaniyaswamy, V. Vijayakumar, P. Siarry, L. Uden, "Multi-objective optimization and energy management in renewable based AC/DC microgrid," Computers and Electrical Engineering, **70**, 179–198, 2018, doi:10.1016/j.compeleceng.2018.01.023.

[8] A. Ndiaye, "Etude et conception de commandes intelligentes dédiées aux systèmes d'injection d'énergie photovoltaïque dans le réseau électrique de distribution," Thèse de Doctorat de Ecole Supérieur Polytechnique de Dakar, 2014.

[9] A.A. Fröhlich, E.A. Bezerra, L.K. Slongo, "Experimental analysis of solar energy harvesting circuits efficiency for low power applications," Computers and Electrical Engineering, **45**, 143–154, 2015, doi:10.1016/j.compeleceng.2014.09.004.

[10] F.K. Slimane Hadji, Jean-Paul Gaubert, "Theoretical and experimental analysis of genetic algorithms based MPPT for PV systems," Energy Procedia, **74**, 772–787, 2015, doi:10.1016/j.egypro.2015.07.813.

[11] M.K. M. Zagrouba, M. Bouaïcha, A. Sellami, "Optimisation par les algorithmes génétiques et modélisation par la méthode LPV d'un système

photovoltaïque," Vème Congrès International Sur Les Energies Renouvelables et l'Environnement, 1–6, 2010.
[12] S. Titri, C. Larbes, K. Youcef, K. Benatchba, "A new MPPT controller based on the Ant colony optimization algorithm for Photovoltaic systems under partial shading conditions," Applied Soft Computing Journal, **58**, 465–479, 2017.
[13] E. hadji M. Ndiaye, A. Ndiaye, M. Faye, "Experimental Validation of PSO and Neuro-Fuzzy Soft-Computing Methods for Power Optimization of PV installations," in 8th IEEE International Conference on Smart Grid, IEEE Xplore, Elhadjimbaye2020: 189–197, 2020, doi:10.1109/icsmartgrid49881.2020.9144790.
[14] B. Long, L. Huang, H. Sun, Y. Chen, F. Victor, K. To, "An intelligent dc current minimization method for transformerless grid-connected photovoltaic inverters," ISA Transactions, (xxxx), 2018, doi:10.1016/j.isatra.2018.12.005.
[15] M. Farhat, "Advanced ANFIS-MPPT Control Algorithm for Sunshine Photovoltaic Pumping Advanced ANFIS-MPPT Control Algorithm for Sunshine Photovoltaic Pumping Systems," 2012 First International Conference on Renewable Energies and Vehicular Technology Advanced, (April 2012), 2014.
[16] H. Bounechba, A. Bouzid, H. Snani, A. Lashab, "Electrical Power and Energy Systems Real time simulation of MPPT algorithms for PV energy system," International Journal of Electrical Power and Energy Systems, **83**, 67–78, 2016, doi:10.1016/j.ijepes.2016.03.041.
[17] S. Hadji, J. Gaubert, F. Krim, "Real-time Genetic Algorithms-based MPPT: Study and comparison (Theoretical an Experimental) with conventional methods," Energies, (Ic), 2018, doi:10.3390/en11020459.
[18] R.O. Nabil Karami, Nazih Moubayed, "General reviews and classification of different MPPT techniques," Renewable and Sustainable Energy Reviews, **68**(September 2016), 1–18, 2017.
[19] P. Hunoor, S.R. Savanur, "Design and Analysis of ANFIS Controller to Control Modulation Index of VSI Connected to PV Array," European Journal of Advances in Engineering and Technology, **2**(5), 12–17, 2015.
[20] P.F. Torres, A.F.P. Costa, V.L. Chaar Junior, W.L. Monteiro, M.A.B. Galhardo, J.T. Pinho, W.N. Macêdo, "A mobile educational tool designed for teaching and dissemination of grid connected photovoltaic systems," Computers and Electrical Engineering, **76**, 168–182, 2019, doi:10.1016/j.compeleceng.2019.03.017.
[21] F. Belhachat, C. Larbes, "Global maximum power point tracking based on ANFIS approach for PV array con fi gurations under partial shading conditions," Renewable and Sustainable Energy Reviews, **77**(February), 875–889, 2017, doi:10.1016/j.rser.2017.02.056.
[22] W. Issaadi, S. Issaadi, A. Khireddine, "Comparative study of photovoltaic system optimization techniques : Contribution to the improvement and development of new approaches," Renewable and Sustainable Energy Reviews, **82**(August 2017), 2112–2127, 2018, doi:10.1016/j.rser.2017.08.041.
[23] E. Fadil, H. El, A. Yahya, "Maximum Power Point Tracking Algorithm for Photovoltaic Systems under Partial Shaded Conditions," IFAC-PapersOnLine, **49**(13), 217–222, 2016.
[24] T.H. Kwan, X. Wu, "High performance P & O based lock-on mechanism MPPT algorithm with smooth tracking," Solar Energy, **155**, 816–828, 2017, doi:10.1016/j.solener.2017.07.026.
[25] H. Bounechba, A. Bouzid, K. Nabti, H. Benalla, "Comparison of perturb & observe and fuzzy logic in maximum power point tracker for PV systems," Energy Procedia, **50**, 677–684, 2014.
[26] D. Gueye, A. Ndiaye, F. Mactar, "Design Methodology of Novel PID for Efficient Integration of PV Power to Electrical Distributed Network," International Journal of Smart Grid, **2**(1), 2018.
[27] I.V. Banu, R.Ä. Beniug, M. Istrate, "Comparative Analysis of the Perturb-and-Observe and Incremental Conductance MPPT Methods," THE 8th INTERNATIONAL SYMPOSIUM ON ADVANCED TOPICS IN ELECTRICAL ENGINEERING May 23-25, 2013 Bucharest, Romania, (July 2014), 2013, doi:10.1109/ATEE.2013.6563483.
[28] Y.K. Semero, J. Zhang, D. Zheng, S. Member, "PV Power Forecasting Using an Integrated GA-PSO-ANFIS Approach and Gaussian Process Regression Based Feature Selection Strategy," CSEE JOURNAL OF POWER AND ENERGY SYSTEMS, **4**(2), 210–218, 2018, doi:10.17775/CSEEJPES.2016.01920.
[29] S. Paul, J. Thomas, "Comparison of MPPT using GA optimized ANN employing PI controller for solar PV system with MPPT using Incremental Conductance," International Conference on Power, Signals, Controls and Computation (EPSCICON), (January), 8–10, 2014.
[30] S. Hadji, F. Krim, J. Gaubert, U. De Poitiers, "Development of an algorithm of maximum power point tracking for photovoltaic systems using Genetic Algorithms," 7th International Workshop on Systems, Signal Processing and Their Applications (WOSSPA), 43–46, 2011.
[31] P. Jood, S.K. Aggarwal, V. Chopra, "Performance assessment of a neuro-fuzzy load frequency controller in the presence of system non-linearities and renewable penetration," Computers and Electrical Engineering, **74**, 362–378, 2019, doi:10.1016/j.compeleceng.2019.02.009.
[32] A. Harrag, H. Bahri, "Novel neural network IC-based variable step size fuel cell MPPT controller Performance , efficiency and lifetime improvement," International Journal of Hydrogen Energy, **42**(5), 3549–3563, 2016, doi:10.1016/j.ijhydene.2016.12.079.
[33] S. Saravanan, R.B. N, "RBFN based MPPT algorithm for PV system with high step up converter," Energy Conversion and Management, **122**, 239–251, 2016.
[34] S. Messalti, A. Harrag, A. Loukriz, "A new variable step size neural networks MPPT controler: Review, simulation and hardware implementation.," Renewable and Sustainable Energy Reviews, **68**(August 2015), 221–233, 2017.
[35] F. Chekired, Z. Smara, A. Mahrane, M. Chikh, S. Berkane, "An energy flow management algorithm for a photovoltaic solar home," Energy Procedia, **111**(September 2016), 934–943, 2017, doi:10.1016/j.egypro.2017.03.256.
[36] M. Nabipour, M. Razaz, S.G.H. Seifossadat, S.S. Mortazavi, "A new MPPT scheme based on a novel fuzzy approach," Renewable and Sustainable Energy Reviews, **74**(February), 1147–1169, 2017, doi:10.1016/j.rser.2017.02.054.
[37] O. Kraa, M. Becherif, M.Y. Ayad, R. Saadi, M. Bahri, A. Aboubou, I. Tegani, "A Novel Adaptive Operation Mode based on Fuzzy Logic Control of Electrical Vehicle," Energy Procedia, **50**(0), 194–201, 2014, doi:10.1016/j.egypro.2014.06.024.
[38] E.M. Ndiaye, A. Ndiaye, M.A. Tankari, G. Lefebvre, "Adaptive Neuro-Fuzzy Inference System Application for The Identification of a Photovoltaic System and The Forecasting of Its Maximum Power Point," 7th International IEEE Conference on Renewable Energy Research and Applications, ICRERA 2018, **5**, 1–7, 2018.
[39] M.M. Ismail, A.F. Bendary, "Smart Battery Controller using ANFIS for Three Phase Grid Connected PV Array System," Mathematics and Computers in Simulation, 2018.
[40] K. Amara, A. Fekik, D. Hocine, M. Lamine, "Improved performance of a PV solar panel with Adaptive Neuro Fuzzy Inference System ANFIS based MPPT," 7th International IEEE Conference on Renewable Energy Research and Applications, ICRERA 2018, **5**, 1098–1101, 2018.
[41] B. Kebe, O. Ba, B. Niang, L. Thiaw, "Etude , synthèse et implémentation d ' un contrôleur neuro- floue pour la commande de groupe turbo-alternateur," **2**(January), 46–53, 2018.
[42] M. Elloumi, R. Kallel, G. Boukettaya, "A comparative study of GA and APSO algorithm for an optimal design of a standalone PV/Battery system," 15th International Multi-Conference on Systems, Signals & Devices (SSD), 1104–1109, 2018.
[43] A. El Dor, "Perfectionnement des algorithmes d ' optimisation par essaim particulaire : applications en segmentation d ' images et en électronique," Thèse de Doctorat de l'Unicersité Paris-Est, 2013.
[44] K. Khezzane, F. Khoucha, "Application de la Technique PSO pour la Poursuite du PPM d ' un Système Photovoltaïque," The 3nd International Seminar on New and Renewable Energies, 1–6, 2014.
[45] E.A. Gouda, M.F. Kotb, D.A. Elalfy, "Modelling and Performance Analysis for a PV System Based MPPT Using Advanced Techniques," EJECE, European Journal of Electrical and Computer Engineering, **3**(1), 1–7, 2019.
[46] O. Ben Belghith, L. Sbita, F. Bettaher, "MPPT Design Using PSO Technique for Photovoltaic System Control Comparing to Fuzzy Logic and P & O Controllers," Energy and Power Engineering, **8**, 349–366, 2016, doi:10.4236/epe.2016.811031.
[47] T.T. Guingane, Z. Koalaga, E. Simonguy, F. Zougmore, D. Bonkoungou, T.T. Guingane, Z. Koalaga, E. Simonguy, F. Zougmore, D.B. Modélisation, "Modélisation et simulation d ' un champ photovoltaïque utilisant un convertisseur élévateur de tension (boost) avec le logiciel MATLAB / SIMULINK," JOURNAL INTERNATIONAL DE TECHNOLOGIE, DE L'INNOVATION, DE LA PHYSIQUE, DE L'ENERGIE ET DE L'ENVIRONNEMENT, **2**(1), 2016.
[48] I.V. Banu, M. Istrate, "Modeling of Maximum Power Point Tracking Algorithm for Photovoltaic Systems," ICEPE, (July 2014), 2012, doi:10.1109/ICEPE.2012.6463577.
[49] S. Motahhir, A. El Hammoumi, A. El Ghzizal, "The Most Used MPPT Algorithms : Review and the Suitable Low-cost Embedded Board," Journal of Cleaner Production, 2019, doi:10.1016/j.jclepro.2019.118983.
[50] S. Krishnamurthy, J.P. Ram, "EL-PSO based MPPT for Solar PV under Partial Shaded Condition," Energy Procedia, **117**, 1047–1053, 2017, doi:10.1016/j.egypro.2017.05.227.
[51] J.R. Jang, "ANFIS : Adaptive-Network-Based Fuzzy Inference System," IEEE, **23**(3), 1993.
[52] A. Fallah, F. Zarei, H. Zarrabi, M.J. Lariche, "ANFIS-GA modeling of dynamic viscosity of N- Alkane in different operational conditions," Petroleum Science and Technology, **0**, 1–7, 2018, doi:10.1080/10916466.2018.1458117.
[53] R.K. Kharb, F. Ansari, S.L. Shimi, "Design and Implementation of ANFIS based MPPT Scheme with Open Loop Boost Converter for Solar PV

Module," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, **3**(1), 6517–6524, 2014.

[54] A. Sharifian, M.J. Ghadi, S. Ghavidel, L. Li, J. Zhang, "A new method based on Type-2 fuzzy neural network for accurate wind power forecasting under uncertain data," Renewable Energy, **120**, 220–230, 2018, doi:10.1016/j.renene.2017.12.023.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1466-1473 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Investigation of the Effect of Optimal Plane Spacing Between Electrode Planes for the EIT Industrial Applications

Yew Lek Chong, Renee Ka Yin Chin*

Faculty of Engineering, University Malaysia Sabah, Kota Kinabalu, Sabah, 88400, Malaysia

ARTICLE INFO

Article history:
Received: 10 November, 2020
Accepted: 10 December, 2020
Online: 21 December, 2020

Keywords:
Electrical Impedance Tomography
Plane Spacing
Image quality

ABSTRACT

In this paper, the effect of plane spacing between electrode planes on Electrical Impedance Tomography (EIT) reconstructed images is investigated. Image properties of models for various plane spacings between electrode planes on EIT imaging were investigated by applying conventional measurement strategies. Sensitivity analysis and spatial resolution analysis were used to study the influence of the different plane spacing between electrode planes on imaging properties. In the sensitivity analyses, the results indicate that there are insignificant differences in sensitivity level for the models with different plane spacings, regardless of measurement strategies applied. From the spatial resolution analyses, the findings are conclusive as there are visible differences in the spatial resolution across the off-electrode plane. A comparative study using reconstructed images was also done. The true distributions with the different number of objects are used as references to assess resulting reconstructed images obtained from models with different plane spacing between electrode planes. Results indicate the model with plane spacing between electrodes planes, which is one quarter to the height of the model, provides the better quality of reconstructed images, in terms of estimations of dimension, and colour contrast of the imaged object.

1. Introduction

Electrical Impedance Tomography (EIT) is useful for visualising the conductivity distributions of an imaged space. Due to its non-intrusiveness, low cost and versatility, EIT has been widely used in medical [1] and industrial [2] applications. Recently, EIT is gaining more attention in industrial applications as it provides intuitions and insights to the reaction and dynamic of the process, as well as useful information of flow characteristic in a pipeline or process vessel. However, EIT imaging suffers from relatively low spatial resolution, especially towards the centre of the imaged space, which is ultimately impacting the quality of the reconstructed images. This is due to the ill-posed nature of the EIT problem [3]. As a result, numerous efforts have been made to further understand this problem and explore ways to improve spatial resolution in EIT imaging.

In an EIT system, the planar arrangement is a commonly employed electrode configuration, where electrodes are mounted in an equally spaced manner, on the perimeter of the process tank or vessel, and the centre of electrodes are aligned on a defined height. In industrial applications, an effective electrode configuration system greatly depends on the vessel or tank size. The criteria of the electrode configuration include the plane spacing between electrode planes and the number and size of electrodes. So far, there is no past work that studied plane spacing of the multiple electrode planes in EIT applications systematically. Instead, the past publications are more focus on the electrode parameters, such as size and number of electrodes, in medical and industrial applications.

In [4], the authors investigated the optimal distance of the multiple sensor plane using a 3D EIT imaging technique based on simulation and experimental study. The investigation showed that the quality of the reconstructed images was accessed based on the correlation coefficient and the relative error. The results showed that 3.5 cm plane spacing between electrode planes, for a model with a diameter of 16 cm and a height of 21.5 cm, produced the highest value in correlation and achieved the lowest values in the relative errors as well as provided optimal performance in the distribution tests. In [5], the authors investigated the performance of drive and measurement electrode patterns using various plane spacing between electrode planes. The results highlighted that the Signal to Error (SNR) decreases as the spacing between two electrode planes decreases, regardless of the transverse and

*Corresponding Author: Renee Ka Yin Chin, Email: reneekychin@ums.edu.my

longitudinal current patterns applied. The transverse current patterns provided a better resolution than longitudinal current patterns but were more susceptible to noise and error, leading to poorer sensitivity in detecting impedance changes in the centre of imaged space. However, in industrial applications, the plane spacing between electrode planes may become a minor concern for the researchers.

Electrode parameters, such as size and number of electrodes, have been investigated in the past few decades. In [5], an experiment was carried out to examine the effect of the height of electrodes on the quality of reconstructed images. The findings indicated that the sensitivity or detectability of the EIT system reduced or remained the same as the height of electrodes was increased to twice the height of the imaged object or target. In [6], the authors investigated the optimum sized electrodes for Electrical Resistance Tomography (ERT). The findings highlighted that the optimum size of electrode coverage is 60% and 80% of the wall surface area of the model. In [7], [8], the authors investigated the electrode properties of ERT. They found that dot or point electrodes could eliminate the effect of contact impedance. Dot or point electrodes are more commonly applied in medical applications. However, it is impractical when implementing dot or point electrodes in industrial applications, as it is more costly to fabricate [9].

Electrodes used in applications are often in rectangular shape. This is not ideal, as narrow angles on the electrode corners result higher current densities in these areas. Circular electrodes are ideal for this purpose due to the uniformity of its shape [10]. However, forward solving requires discretisation of the model into tetrahedral elements in EIT. An accurate modelling of circular electrodes requires a large number of small elements to best represent the shape of electrodes, which results in an increased number of computational elements, and more importantly, these elements do not contribute to the representation of the imaged space. When the coarser mesh structure is used, the inaccurate representation of the shape of the electrode will occur and it leads to errors in forward and inverse solving. As a compromise, rectangular electrodes are often used.

In this paper, the influence of different plane spacing between electrode planes on EIT applications in industrial processes (e.g. pipeline flow and suspension monitoring, and mixing processes) on the quality of reconstructed images is studied. The main purpose of this investigation is to investigate a minimum requirement of space between electrode planes, such that information is not lost. This investigation was carried out through investigating the effect of different electrode configurations on imaging properties and reconstructed images. The rest of the paper consists of five (5) sections. The model setup for various plane spacing between electrode planes is described in Section II. Section III discusses the findings obtained from the comparative studies on image properties. Section IV describes the experimental setup for image reconstruction, as well as its results. Finally, Section V provides the conclusion of this study.

2. Model Setup

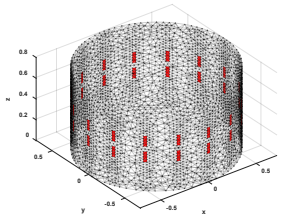
In this work, a cylindrical model with a radius of 71.4 cm and a height of 80 cm is used. The dimensions of the model are chosen to be sufficiently big, such that it can accommodate a variety of different electrode configurations, including spacing between electrode planes, which is the main focus of the work presented in this paper.

The electrode configuration used for this work consists of 32 electrodes arranged in two planes. Each plane consists of 16 electrodes arranged equi-spaced on the periphery of the model. The electrodes are rectangular in shape, with the dimensions of 3 cm in width and 10 cm in height. In this study, the spacing between planes is set as 5 cm (the smallest gap), 10 cm, 15 cm, 20 cm, 25 cm and 30 cm (largest gap). A summary of the plane spacing between electrode planes and the height of the centre of the upper and lower electrode planes is provided in Table I.

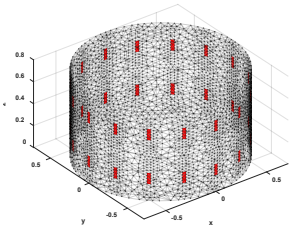
Table 1: Plane Spacing Between Two Electrode Planes and its Center of Upper and Lower Electrode Plane

Number of electrodes per plane	Plane spacing between two electrode planes	Center of upper electrode plane (cm)	Center of lower electrode plane (cm)
16	5 cm	47.5	32.5
	10 cm	50.0	30.0
	15 cm	52.5	27.5
	20 cm	55.0	25.0
	25 cm	57.5	22.5
	30 cm	60.0	20.0

The work presented in this paper is simulated using Electrical Impedance Tomography and Diffuse Optical Tomography Reconstruction Software (EIDORS), which is an open-source toolkit available for MATLAB [11]. The discretisation of the models was done by applying the Finite Element Method (FEM) through Netgen [12]. The FEM models with various spacing between electrodes planes are as depicted in Figure 1. It is should be noted that the number of elements for each of the models is the same as they are based on the same cylindrical model. The work presented in this paper is through simulation as it is a more feasible option due to the requirement of using electrodes with different plane spacing.



(a) 5 cm plane spacing between electrode planes model



(b) 30 cm plane spacing between electrode planes model

Figure 1: Examples of FEM models with varying plane spacing between electrode planes.

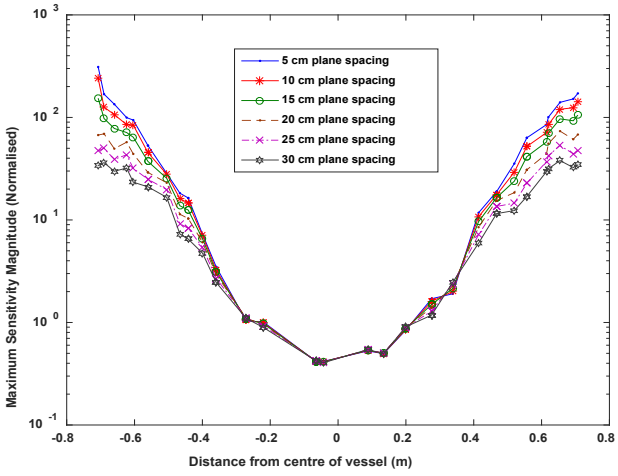
All measurements were simulated using the adjacent current injection protocol, which produces *N*(N-3)* number of

measurements, where *N* is the number of electrodes. In the case of two planes of 16-electrodes per plane model, 416 full-frame measurements were generated. Another set of measurement data were repeated to simulate by using another opposite current injection protocol, which produces *N*(N-4)* number of measurements. There are 384 full-frame measurements generated for a model with two planes of 16-electrodes per plane.

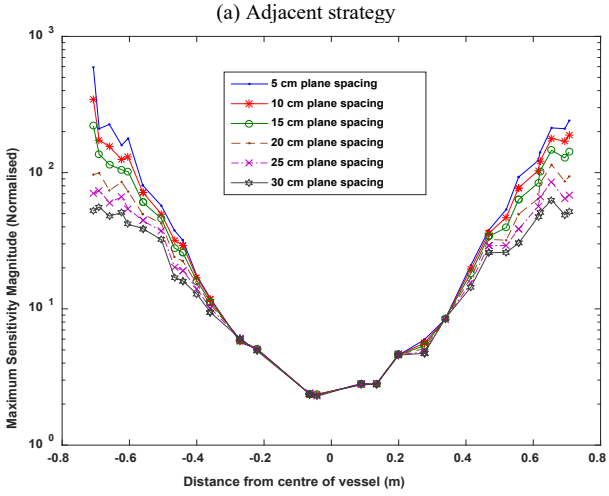
3. Experimental Results

3.1. Sensitivity Analysis

Sensitivity analysis provides information on the potential success of detecting any changes in conductivity in the imaged space. In the discretised model, each element has a sensitivity value for any measurement acquired, which makes up the sensitivity (Jacobian) matrix. For this comparative study, the maximum sensitivity for each discretised element for a full set of measurement, normalised to the volume of the element, is computed and compared, as shown in Figure 4 and Figure 5.



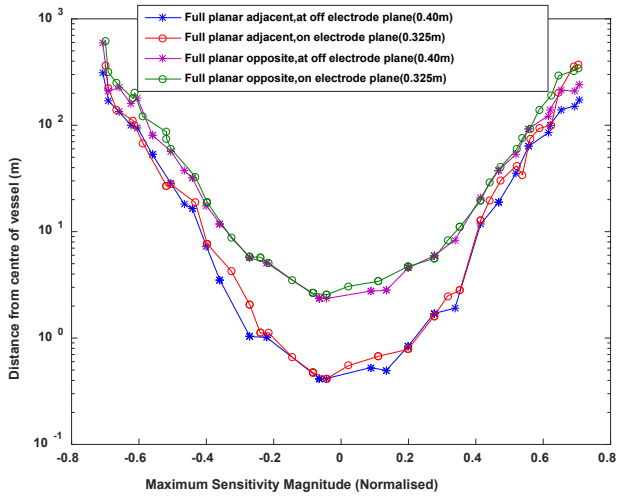
(a) Adjacent strategy



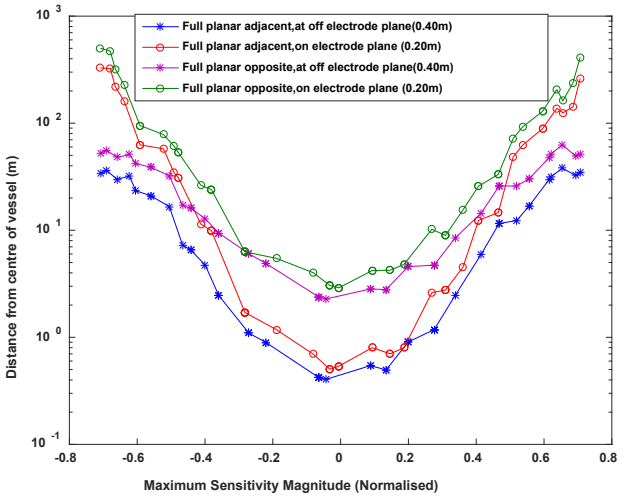
(b) Opposite strategy

Figure 4: Sensitivity comparison for models with different plane spacing electrode planes at the off-electrode plane (the height is taken at 0.40 m from the bottom of the model) by applying the (a) adjacent strategy and (b) opposite strategy.

Figure 4(a) shows the comparison of maximum sensitivity magnitude for the models with different spacing between electrode planes across an off-electrode plane when applying adjacent current injection protocol. From Figure 4(a), it can be seen that there are insignificant differences, as the spacing between electrode planes increases, especially towards the centre of the imaged space. Similar results are observed using the opposite current injection strategy, as reflected in Figure 4(b). This may due to the planar measurement strategy used, whereby measurements are only taken on the plane where the current source and sink pairs are located.



(a) The model with 5 cm plane spacing between electrode planes



(b) The model with 30 cm plane spacing between electrode planes

Figure 5: Sensitivity comparison for the models both (a) 5 cm and (b) 30 cm plane spacing between electrode planes, respectively, at the off-electrode plane and on the electrode plane.

From Figure 5, sensitivity is higher for opposite strategy due to the current penetrating through the imaged space, as opposed to the adjacent strategy, where current flowing typically tends to be more confined nearer to the periphery of the model [13]. The other obvious observation is that the sensitivity is higher on the electrode plane in comparison to those obtained off-electrode plane. This is expected, as the current signal is stronger on the electrode plane compared to the off-electrode plane [14]. This observation is similar to a previous study [15], where it indicates that the possibility of detecting a change in conductivity reduces when further away from the electrode plane.

Comparing Figure 5(a) and Figure 5(b), it can be observed that when the space between electrode planes are smaller (5 cm), there are insignificant differences between sensitivity on- and off-electrode planes. This is likely to be due to the closeness of the locations of the two electrode planes. In comparisons, when the space between electrode planes is bigger (30 cm), there are noticeable differences between sensitivity on- and off-electrode planes. Although the EIT image reconstruction problem is often treated as a 2D problem, where images and data are captured on a specific plane, it is essentially a 3D problem.

3.2. Spatial Resolution Analysis

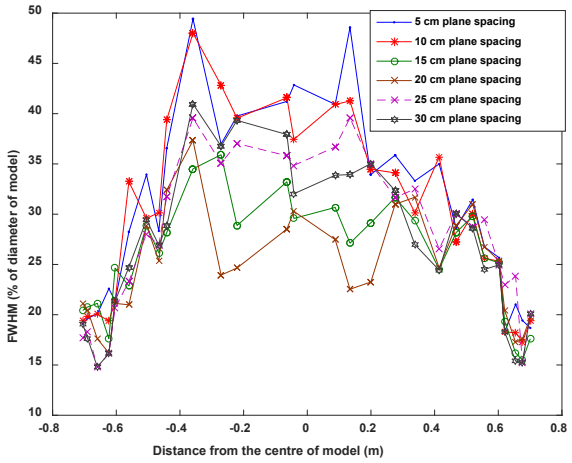
Spatial resolution analysis utilises information given in the resolution matrix, which contains the correlation between elements in a discretised model. In this analysis, the Full-Width Half-Maximum (FWHM) technique [16] is applied to investigate the spatial resolution of models with different spacing between electrode planes. The FWHM method applied in this paper is based on the technique, which was first proposed by [17] and later adopted by [18]. In spatial resolution analysis, the minimum distance between two point sources, which shares the same solution value in an image is measured. The same solution value shared is known as half the peak value. Spatial resolution for each point is computed and compared. The results are as shown in Figure 6 and Figure 7.

Figure 6 shows the spatial resolution comparing models with different spacing between electrode planes using both adjacent and opposite strategies. It is shown that the regions near the wall of the model, in general, have better resolution (low FWHM) in comparison with the area towards the centre of the model, which resulted in higher FWHM. This general trend obtained is consistent with the sensitivity analysis obtained in Figure 4(a) and Figure 4(b), where they indicate that sensitivity is lower in the centre of the model but is higher near the wall region of the model. With a higher possibility of detecting a change in conductivity, and thus it increases the chance of change detected to be reconstructed.

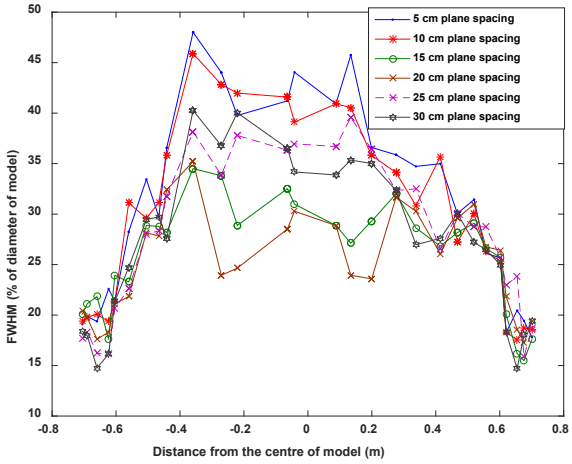
It is worth noting that unlike the sensitivity analysis, spatial resolutions do not show symmetry across the plane, and irregularities (i.e. peaks or valleys) are observed. One of the factors that affect spatial resolution is the size of the discretised element, and unlike sensitivity analysis, the computed spatial resolution for each discretised element is not normalized to the volume of the element, as it will provide misrepresentation on the visibility contrast.

From Figure 6(a), it can be observed that models with 5 and 10 cm plane spacing between electrode planes produced the worst resolution overall, especially in the centre region of the model. However, it is worth noticing that model with 20 cm plane spacing between electrode planes resulted in better spatial resolution in general. This finding is not anticipated as the sensitivity analyses are shown in Figure 4, where it indicates that there are insignificant differences in maximum sensitivity magnitude across the off-electrode plane, for both current injection strategies.

Referring to Figure 7, comparing spatial resolution for the on- and off-electrode planes, it is observed that there are insignificant differences in the resolution for an on-electrode plane, especially with the gap between the two electrode planes is small (5 cm). This result is consistent with the results observed in its corresponding comparative analysis, as shown in Figure 5. The spatial resolution analysis in Figure 7 also shows that the overall resolution is evenly matched for usable information which is available for image reconstruction.



(a) Adjacent strategy



(b) Opposite strategy

Figure 6: Spatial resolution comparison for models with different plane spacing by applying adjacent strategy in (a) and opposite strategy in (b) at 0.40 m height from the base of the model.

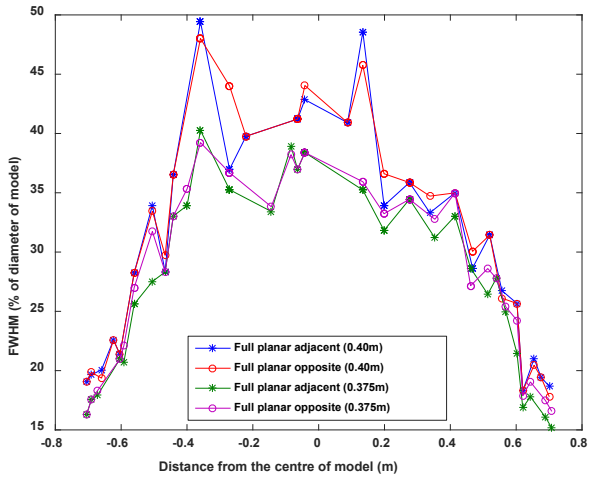
It is interesting to see in Figure 7(a) that there is very little difference in spatial resolution when comparing spatial resolutions for the adjacent and opposite strategies. This is unexpected, as shown in Figure 5, where there is a distinguishable difference in sensitivity between the two different measurement strategies. One of the reasons is likely to be due to both strategies producing a similar amount of usable (stable) information for reconstruction, which resulted in the spatial resolution for these two strategies to be evenly matched.

4. Experimental Set-up

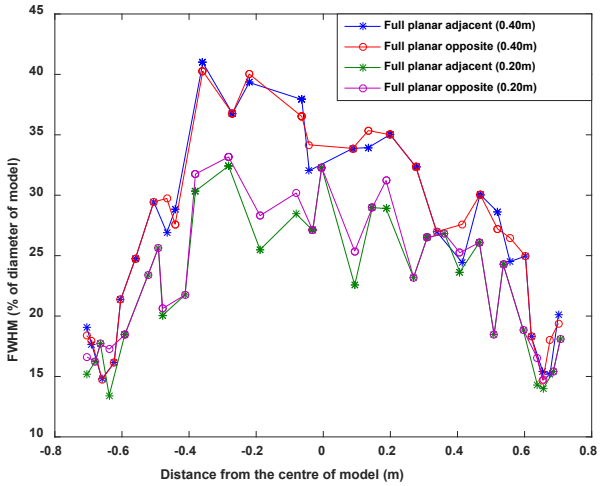
4.1. Experimental Set-Up

Measurements for image reconstruction are simulated using the EIDORS package [11], which is MATLAB-based. As

mentioned earlier, this study is more feasible to be done in simulation, as laboratory validation requires a process tank that allows multiple settings for electrodes configurations. For validation, two test distributions are used for experiment and comparison.



(a) The model with 5 cm plane spacing between electrode planes

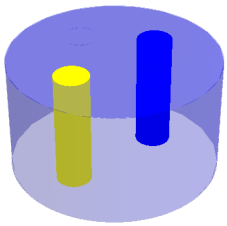


(b) The model with 30 cm plane spacing between electrode planes

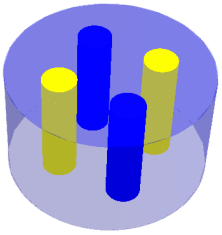
Figure 7: Spatial resolution comparison for the models both (a) 5 cm and (b) 30 cm plane spacing between electrode planes, respectively, at the off-electrode plane and on the electrode plane.

Measurements are simulated using a finely discretised FEM model. The level of discretization is chosen to be significantly higher than the discretized model used for forward and inverse solving, as it is intended to mimic a physical model. Two measurements are taken for each test distribution, for each electrode configuration, and each measurement strategy.

In this work, two test distributions are used for image reconstruction, as shown in Figure 8. The first test distribution consists of two cylinders, each with a radius of 12 cm. The conductivity of the background is 0.012 S/m (approximately the conductivity value of tap water). One of the cylinders has a higher conductivity than background conductivity, and the other is non-conducting. The arrangement of test objects is as shown in Figure 8(a). The second test distribution is similar to the first test distribution, with the addition of one conducting cylinder. The arrangement of test objects is as shown in Figure 8(b). All reconstructed images were obtained by applying the Non-Linear Gauss-Newton (NLGN) algorithm in inverse solving.



(a) True distribution with two cylindrical objects



(b) True distribution with four cylindrical objects

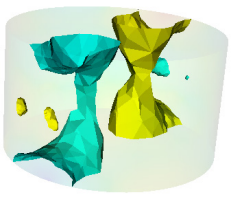
Figure 8: True distribution models with (a) two cylindrical objects, (b) four cylindrical objects and (c) four floating objects. The colour scale is applied, where yellow colour indicates the conducting object and blue colour indicates the non-conducting object.

4.2. Results and Discussions

The image reconstructions obtained from models with different spacing between electrode planes are shown in Figure 9 to Figure 12. The colour scale is maximised to show the best contrast. The reconstructed images are assessed and discussed based on the dimensions of the reconstructed objects and colour contrast of the reconstructed conductivity distribution.

Although sensitivity analyses indicated there are insignificant differences in the maximum sensitivity magnitude as the gap between electrode planes increases, this is not reflected in spatial resolution analysis and reconstructed images. When comparing the reconstructed images from Figure 9-12, it can be seen that there are visible differences in terms of shape and estimation of dimensions of the objects.

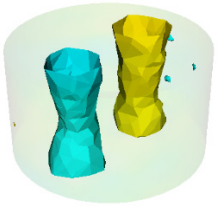
In Figures 9-12, it can be seen that the models with 5cm and 10 cm plane spacing between electrode planes generally produced the worst reconstruction images. The shape and the dimension of the cylindrical test objects are not reconstructed properly, especially towards the top and the bottom of the imaged space. There are also noticeable artefacts visible in the imaged space. The results are similar for experiments done using both the adjacent and opposite strategies.



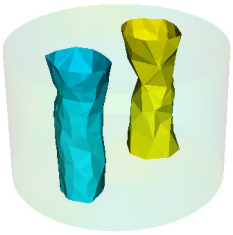
(a) 5 cm plane spacing between electrode planes



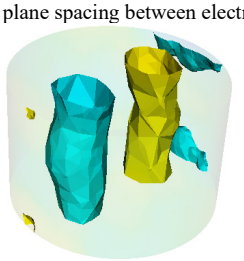
(b) 10 cm plane spacing between electrode planes



(c) 15 cm plane spacing between electrode planes



(d) 20 cm plane spacing between electrode planes

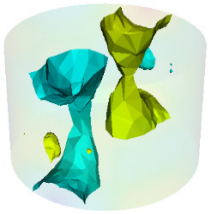


(e) 25 cm plane spacing between electrode planes



(f) 30 cm plane spacing between electrode planes

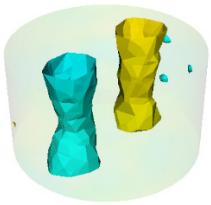
Figure 9: Reconstructed images comparison for models with various plane spacing (a)-(f) by applying the adjacent current protocol.



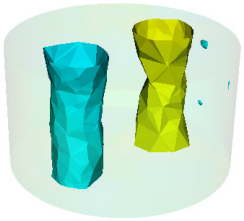
(a) 5 cm plane spacing between electrode planes



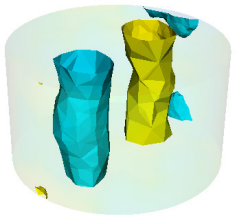
(b) 10 cm plane spacing between electrode planes



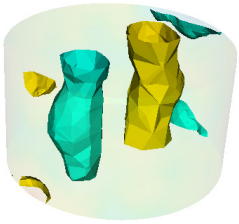
(c) 15 cm plane spacing between electrode planes



(d) 20 cm plane spacing between electrode planes

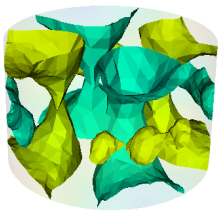


(e) 25 cm plane spacing between electrode planes

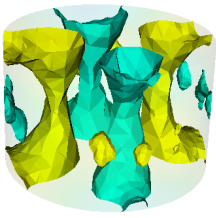


(f) 30 cm plane spacing between electrode planes

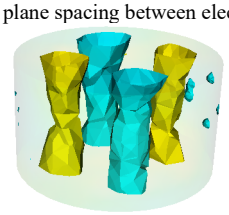
Figure 10: Reconstructed images comparison for models with various plane spacing (a)-(f) by applying the opposite current protocol.



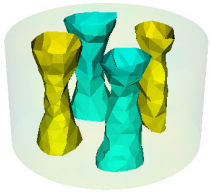
(a) 5 cm plane spacing between electrode planes



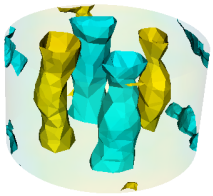
(b) 10 cm plane spacing between electrode planes



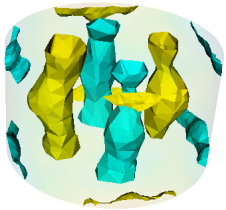
(c) 15 cm plane spacing between electrode planes



(d) 20 cm plane spacing between electrode planes

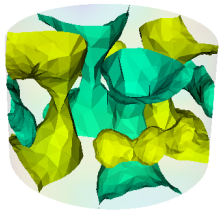


(e) 25 cm plane spacing between electrode planes

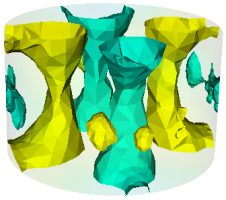


(f) 30 cm plane spacing between electrode planes

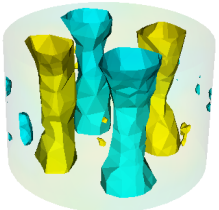
Figure 11: Reconstructed images comparison for models with various plane spacing (a)-(f) by applying the adjacent current protocol.



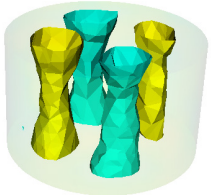
(a) 5 cm plane spacing between electrode planes



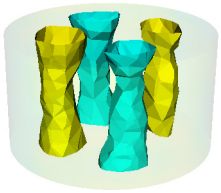
(b) 10 cm plane spacing between electrode planes



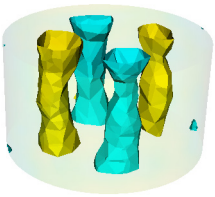
(c) 15 cm plane spacing between electrode planes



(d) 20 cm plane spacing between electrode planes



(e) 25 cm plane spacing between electrode planes



(f) 30 cm plane spacing between electrode planes

Figure 12: Reconstructed images comparison for models with various plane spacing (a)-(f) by applying the opposite current protocol.

Models with 15 cm and 20 cm gaps generally produced better reconstructed images. With the exception of the reconstructed image using a test distribution with four cylindrical objects, the application of the adjacent strategy for the model with 20 cm gap, it can be seen in Figures 9-12 that the cylindrical test objects are reconstructed in the right configuration, with similar dimensions to the true distribution. There are also little to no artefacts visible in the imaged space.

For the widest gaps, namely the 25 cm and 30 cm electrode spacing models, the images reconstructed using the opposite strategy are better compared to those obtained using the adjacent strategy. The images obtained using the adjacent strategy has a significant amount of artefacts, and it is noticeable that the shape of the test objects are not reconstructed correctly where the gap of the electrodes are (i.e. space between electrode planes). It can be seen in Figures 9(e-f) and 11 (e-f) that the dimensions are overestimated in that space. In contrast, reconstructed images obtained using the opposite strategy turn out better, with the test objects shown to be cylindrical objects, and with less artefacts. This result is consistent with the results observed in the sensitivity analyses.

5. Conclusion

A summary of the results is tabulated in Table 2. This study is conducted through two methods: statistical method, through comparing sensitivity and spatial resolution of different models used; and through reconstructed images.

Table 2: Summarization of the results on image properties

Analysis	Best model	Worst model
Sensitivity Analysis	Results are inconclusive; Opposite strategy is better than adjacent strategy	
Spatial Resolution Analysis	15 and 20 cm gap (both strategies)	5 and 10 cm gap (both strategies)
Reconstructed Images	15 and 20 cm gap (opposite strategy)	5 and 10 cm gap (both strategies)

From Table 2, it can be seen that the results obtained through spatial resolution analysis are reflected in the reconstructed images, where the best quality and worst quality images are

reflective of the results obtained through spatial analysis. In this case, however, the sensitivity analysis results were inconclusive, as the sensitivity level in the centre region of the imaged space is indistinguishable.

It is worth noting that the results given in Table 2 are relative to the dimension of the model used for this study. When the model is scaled up or scaled down, an investigation similar to that carried out in this study should be conducted to ensure that the scaling effect does not deteriorate the quality of the reconstructed images.

Through this work, it is shown that it is important to investigate the spacing between electrode plane, to ensure that the information is not lost due to the inability to capture any changes that occur. It is impractical to have electrode planes close to each other, as it will result in a waste of resources. Therefore, a minimal distance ought to be set in order to ensure that any changes in conductivity can be measured and the information can be used to reconstruct and visualize the changes.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to express their gratitude to the financial support provided by the KPT-RACE (RACE0010-TK-2013) grant, KPT-FRGS (FRGS0425-TK-1/2015) grant and UMS-SGA (SGA0007-2019) grant for this work.

References

[1] H. X. Wang, C. Wang, W. L. Yin, "Optimum design of the structure of the electrode for a medical EIT system," Measurement Science and Technology, **12**(8), 1020–1023, 2001. https://doi.org/10.1088/0957-0233/12/8/305

[2] J. Nowakowski, E. Hammer, D. Sankowski, D. Styra, R. Wajman, R. Banasiak and A.Romanowski, "New concept of ECT/ERT/GRT tomography for multiphase flow measurements," Automatyka (pó³- rocznik AGH), **14**, 741-748, 2010.

[3] W. Lionheart, N. Polydorides and A. Borsic, Part 1: The Reconstruction Problem, In Holder D. (ed) Electrical Impedance Tomography: Methods, History and Applications, Bristol, 3-64, 2005.

[4] Z. Hao, S. Yue, B. Sun, and H. Wang, "Optimal distance of multi-plane sensor in three-dimensional electrical impedance tomography," Computer Assisted Surgery, **22**, 326–338, 2017.

[5] J. Hope, F. Vanholsbeeck, and A. McDaid, "Drive and measurement electrode patterns for electrode impedance tomography (EIT) imaging of neural patterns activity in peripheral nerve," Biomedical Physics and Engineering Express, **4**(6), 2018. https://doi.org/10.1088/2057-1976/aadff3

[6] J. C. Newell, Y. Peng, P. M. Edic, R. S. Blue, H. Jain, and R. T. Newell, "Effect of electrode size on impedance images of two-and three-dimensional objects," in 1998 IEEE Transactions on Biomedical Engineering, **45**(4), 531–534, 1998. https://doi.org/10.1109/10.664209

[7] P. A. T. Pinheiro, W. W. Loh, F. J. Dickin, "Optimal sized electrodes for electrical resistance tomography," Electronics Letters, **34**(1), 69-70, 1998. https://doi.org/10.1049/el:19980092

[8] F. Dickin, M. Wang, "Electrical resistance tomography for process applications," Measurement Science and Technology, **7**(3), 247–260, 1996. https://doi.org/ 10.1109/10.664209

[9] L. D. Suits, T. Sheahan, J. Lee, J. Santamarina, "Electrical Resistivity Tomography in cylindrical Cells—Guidelines for hardware pre-design," Geotechnical Testing Journal, **33**, 2010. https://doi.org/10.1520/GTJ102366.

[10] M. Hanke, B. Harrach and N. Hyvonen," Justification of Point Electrode Models in Electrical Impedance Tomography," Mathematical Models and Methods in Applied Sciences, **21**(06), 1395–1413, 2011. https://doi.org/10.1142/S0218202511005362

[11] K. S. Cheng, D. Isaacson, J.C. Newell, D.G. Gisser," Electrode models for electric current computed tomography," in 1989 IEEE Trans. Biomed Engr., **36**, 918–924, 1989.

[12] N. Polydorides, W.R.B. Lionheart, "A Matlab toolkit for three-dimensional electrical impedance tomography: a contribution to the Electrical Impedance and Diffuse Optical Reconstruction Software project," Meas. Sci. Technol., **13**, 1871–1883, 2002.

[13] J. Schoberl, Netgen: Automatic Mesh Generator, 2008.

[14] P. O. Gaggero, "Miniaturization and distinguishability limits of electrical impedance tomography for biomedical application", Ph. D Thesis, University of Neuchatel, 2011.

[15] P. Metheral, D. C. Barber, R. H. Smallwood, "Three-dimensional electrical impedance tomography," **380**, 509-5012, 1996. https://doi.org/https://doi.org/10.1038/380509a0

[16] J.L. Wheeler, W. Wang, M. Tang, "Á comparison of methods for measurement of spatial resolution in two-dimensional circular EIT images," Physiol. Meas., **23**, 169-175, 2002. https://doi.org/10.1088/0967-3334/23/1/316

[17] S.C. Murphy, "Ad hoc electrode arrangement for electrical tomography," Ph. D Thesis, University of Manchester, United Kingdom, 2008.

[18] W. Menke, Geophysical Data Analysis: Discrete Inverse Theory (Rev. Ed), Academic Press: San Diego, 1989.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1554-1559 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Application of Deep Belief Network in Forest Type Identification using Hyperspectral Data

Xianxian Luo[1,2], Songya Xu[*,3], Hong Yan[4]

[1]*Faculty of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou, 362000, China*

[2]*Fujian Provincial Key Laboratory of Data Intensive Computing, Quanzhou, 362000, China*

[3]*Faculty of Educational Science, Quanzhou Normal University, Quanzhou, 362000, China*

[4]*Fujian Forest Inventory and Planning Institute, Fuzhou, 350000, China*

ARTICLE INFO

Article history:
Received: 23 August, 2020
Accepted: 08 December, 2020
Online: 25 December, 2020

Keywords:
Deep Learning
Deep Belief Network
Forest Mapping
Hyperspectral Data

ABSTRACT

Forest mapping by remote sensing is a hot topics in forestry. At present, many researchers focus on the research of forest type classification or tree species identification using different machine learning methods and try to improve the accuracy of classification of satellite image. However, forest type classification using deep belief network (DBN) is still limited in previous literatures. Our research focuses on forest mapping in the western part of Dehua county in southern China. Most important objective was to assess the feasibility of forest mapping from hyperspectral data using deep learning. The HJ-1A hyperspectral data was adopted in this paper. We applied deep belief network and got a thematic map of four forest types, such as coniferous forest, broad-leaved forest, mixed forest and non-forest. Our finding shows that optimal network depth of DBN model is 3 and best node in each layer is 256 in our experiment. Overall accuracy is 85.8% and kappa coefficient is 0.785 with best-fit parameters in DBN model, while for SVM is 73% and 0.6447 respectively. DBN obtain better performance compared with support vector machine. Furthermore, network depth and number of nodes in each hidden layer in DBN model has a significant effect on overall accuracy and Kappa coefficient. In general, DBN is promised to be dominant method of forest mapping by hyperspectral data.

1. Introduction

It is very important to obtain the forest information timely and correctly for forest management and land cover mapping in forest ecosystems. For example, forest mapping plays an important role in the issues of forest restoration and forest degradation assessment [1]-[4]. Nowadays, it is an effective way to obtain forest structure from satellite image [5], [6]. Hyperspectral image is a typical data of satellite image and widely used in forest mapping because there are lots of spectral bands of this image and beneficial for forest type recognition. What's more, A series of achievements have been made based on hyperspectral remote sensing technology [7], [8]. However, it leads to the Hughes phenomenon because of redundant spectrum of hyperspectral image [9]. Normally, it needed to be reduced by different methods before image classification [10]. After dimension reduction, traditional classifiers were adopted, such as random forests [6] and support vector machines [7], SVM is the most popular and nonparametric supervised learning method. It is widely used in classifying forest ecosystems [11]. Fortunately, deep learning can extract feature automatically, past studies have showed that deep learning can achieve higher accuracy of image processing of remote sensing avoiding human intervention [12]. Deep belief network is typical model of deep learning. There is an input layer, one output layer and several restricted in Boltzmann machines [13]. Pre-training and fine-tuning are two important steps in DBN model. Pre-training is an unsupervised learning from bottom to up and output result with a classifier. Backpropagation was coined by Rumbelhart [14] and used to update the parameter of model for better result.

There are lots of references about image classification using DBN. It seems that DBN was first adopted by Lü et al. in 2014 year [15]. The experiments showed that the optimal network depth was three hidden layers and best nodes of each hidden layer were

*Corresponding Author: Songya Xu, 29974817@qq.com

64. In year 2015, Chen et al. created DBN model combing with principal component analysis and logistic regression. They confirmed better effect of DBN than that of SVM [12] adopting two Hypersperctral datasets, such as Indian Pines and University of Pavia. On the whole, there are a lot of application situation of deep learning [16], but few literatures is reported about forest mapping using deep learning from Hyperspectral image.

Therefore, one of our research objectives is to confirm the performance of forest mapping by DBN using hyperspectral data. Another objective of this study is to compare the classification performance of DBN model and SVM algorithm. Third objective is to investigate how optimal parameters of DBN affect overall accuracy and Kappa coefficient of forest type identification. The major contribution is to determine the optimal parameters of DBN for forest mapping from Hyperspectral data.

2. Methods

2.1. Restricted Boltzmann Machines (RBM)

RBM is a directionless, full connective model [17]. A RBM consists of two layers. One of them is an input layer, or visible units *v*. Output layer, or hidden units *h* is another layer of RBM. There are no connections within each layer. From Figure 1, we know that the units between visible and hidden are full connected by weight W. The energy of a joint configuration with *v* visible units and *h* hidden units is

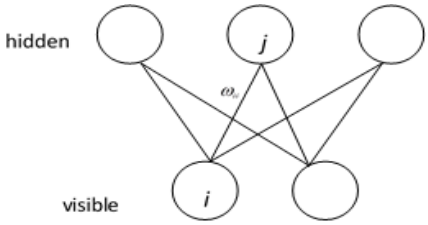


Figure 1: The model of an RBM

$$E(v,h|\theta)=-\sum_{i=1}^{I}a_i v_i-\sum_{j=1}^{J}b_j h_j-\sum_{i=1}^{I}\sum_{j=1}^{J}v_i h_j \omega_{ij} \quad (1)$$

Where $\theta=\{a_i,b_j,\omega_{ij}\}$ represents parameters of RBM. ω_{ij} is the weight between input layers *i* and hidden layers *j*. a_i is the bias of v_i, b_j is the bias of h_j. *I* and *J* are total numbers of input and output layers respectively. The probability of the joint configuration over both layers is

$$P(v,h|\theta)=\frac{e^{-E(v,h|\theta)}}{Z(\theta)} \quad (2)$$

where $Z(\theta)$ is the normalized factor, equals to $\sum_{v}\sum_{h}e^{-E(v,h|\theta)}$.

In the binary case where $h_j\in\{0,1\}$, given input layer, a conditional probability of an output layer being 1 is given as:

$$P(h_j=1|v,\theta)=\sigma\left(b_j+\sum_{i=1}^{I}v_i\omega_{ij}\right) \quad (3)$$

In the same way, the conditional probability of input layer given the output layer:

$$P(v_i=1|h,\theta)=\sigma\left(a_i+\sum_{j=1}^{J}h_j\omega_{ij}\right) \quad (4)$$

where $\sigma(x)$ is a function, such as a sigmoid function,

$$\sigma(x)=\frac{1}{1+e^{-x}}.$$

Normally, total error *E* is minimized to acquire the most favorable parameters θ, RBM.

$$E=\frac{1}{2}\sum_{k}^{K}(d_k-o_k)^2 \quad (5)$$

where the d_k stands for the true outputs, o_k corresponds real outputs. In general, there are *K* outputs. Gradient descent is a good method to minimize total error, it is ecessary to compute the partial derivative of *E* with respect to each weight in the network. The parameters, such as weights and biases will be iterated by stochastic gradient ascent which can be formulated as [18]:

$$\Delta\omega_{ij}=\varepsilon\left(\langle v_i h_j\rangle_{data}-\langle v_i h_j\rangle_{recon}\right) \quad (6)$$

$$\Delta a_i=\varepsilon\left(\langle v_i\rangle_{data}-\langle v_i\rangle_{recon}\right) \quad (7)$$

$$\Delta b_j=\varepsilon\left(\langle h_j\rangle_{data}-\langle h_j\rangle_{recon}\right) \quad (8)$$

where ε is a key parameter, that is learning rate. $\langle\bullet\rangle_{data}$ is expectation for data distribution and $\langle\bullet\rangle_{recon}$ is the expected value of the model distribution.

2.2. Deep Belief Network

Several RBMs stacked into a classical model of deep belief network. There are two important steps in the DBN [19]. One is pre-training and the other is fine-tuning, shown in the Figure 2. Initial parameters are produced in the pre-treatment by an unsupervised method, such as gradient descent method to avoid local optimum. Back-propagation is normally used to update model parameters according to error, such as squared error or cross entropy error.

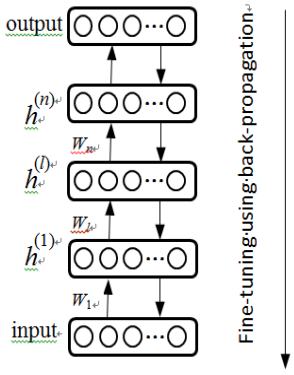


Figure 2: A generative model of deep belief network

2.2.1. Pre-training

Pre-training of DBN model is an unsupervised procedure. At the beginning, the input layer is initialized to a training vector. Then the subsequent hidden layer of RBM is trained through the output of the previous layer. Moreover, this whole process can be repeated until the final hidden layer. By doing so can the DBN extract more and deep feature from the original input data.

2.2.2. Fine-Tuning using Backpropagation algorithm

Backpropagation was coined by Rumbelhart [20]. Backpropagation update model parameters using cross entropy error because there is four kinds of outputs in this paper.

$$\delta_k = o_k(1-o_k)(d_k - o_k) \tag{9}$$

For the *l*th hidden layer,

$$\delta_i^l = (1-o_i)o_i \sum_j^n \omega_{ij}^l \delta_j^{l+1} \tag{10}$$

The batch update rules are

$$\omega_{ij}^l = \omega_{ij}^l + \varepsilon_{fine-tuning} o_i \delta_j^{l+1} \tag{11}$$

$$b_j^l = b_j^l + \varepsilon_{fine-tuning} \delta_j^{l+1} \tag{12}$$

2.3. Indicators of Evaluation of Classification

- *Overall accuracy*

The overall accuracy is total assessment of classified quality, which equals the total pixels sum divided by correctly classified pixels by total pixels. The formula for calculating the overall accuracy based on the confusion matrix can be listed as following:

$$p = \sum_{i=1}^{c} m_{ii} / N \tag{13}$$

here, *C* represents the number of categories, and m_{ii} represents the elements on the diagonal of the confusion matrix, and *N* represents the total number of test samples.

- *Kappa coefficient*

Kappa coefficient adopts a multivariate discrete analysis technique to reflect the consistency between classification results and reference data. It considers all factors of the confusion matrix, and it is a more objective evaluation index, which is defined as:

$$k = \frac{N\sum_{i=1}^{c} m_{ii} - \sum_{i=1}^{c}(m_{i+}m_{+i})}{N^2 - \sum_{i=1}^{c}(m_{i+}m_{+i})} \tag{14}$$

Among them, m_{i+} and m_{+i} represent the sum of the line *i* of the confusion matrix and the sum of the column *i* respectively. The higher the Kappa coefficient, the higher the classification precision.

3. Data and Process

3.1. Study site and hyperspectral data

Hyper Spectral Imager (HSI) is the first Chinese space-borne hyperspectral sensor aboard the HJ-1A satellite. In September 6, 2008, China launched the environmental and disaster monitoring and forecasting small satellite constellation A (HJ-1A), and the HJ-1A satellite was equipped with CCD camera and hyperspectral imager (HSI). With HSI interference imaging spectroscopy, the cutting width is 50km, the ground pixel resolution is 100 meters, 110-128 spectral bands, the spectral resolution of up to 2.08 nm, in the continuous spectral image of earth observation and available features to achieve the direct identification from the space object surface. HSI often produces certain stripe noise on the image. Satellite data was HJ-1A star HSI data 2 level products, imaging time is August 24, 2011, a total of 115 bands with working spectrum range 459 ~ 956 nm.

The research area has 8 downtowns in the west of Dehua county, Quanzhou City, Fujian Province. The administrative area of study site and its pseudo color synthetic imaging (105th, 7th, 40th band for pseudo colour synthesis) are shown in Figure 3. The image data of the product is geometrically corrected by previously corrected image with polynomial transform, and then orthorectified and outputted onto a fixed uniform spatial grid using nearest neighbour resampling. The correction error is not less than one pixel. The corrected image is unified into the specified map projection coordinate system (Xi'an 1980 coordinate system). There is obvious stripe noise in the part of the HSI image data, mainly in bands 1-29. Therefore, these 29 bands are eliminated from the HSI image. The remaining 86 bands are used for research within the wavelength range (529.6350-951.54 nm).

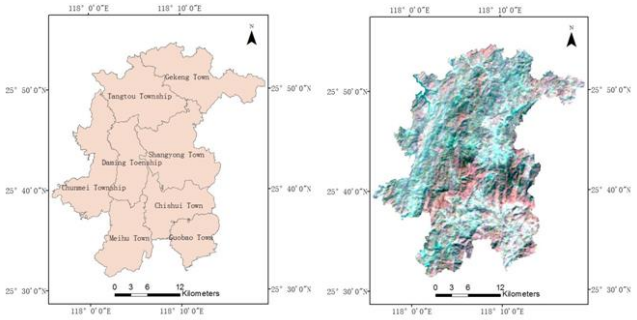


Figure 3: Study area and pseudo colour synthetic image

3.2. Experimental processing flow

According to the field data, the labelled samples were selected in the experiment. Then 86 bands of data were used as input of DBN. Before the training, these data were normalized by min-max normalization, which was mapped into the value between 0 and 1. Then, these data were shuffled and split into training and testing data randomly. After pre-training and fine tuning, categorical data was represented by one-hot coding, as shown in Table 1, 1 was Coniferous forest, 2 was Broad-leaved forest, 3 was Mixed forest, 4 was Non-forest.

The experiment is based on Windows 10 with 64-bit operating system. The processor is Intel (R) Core (TM) i5-8250U CPU @1.60GHz. The experiment is carried out in PyCharm 2018.3x64

editor to code and tune parameters. TensorFlow 1.11.0 framework was first set up as development environment. Also, The Python Extension Library is loaded, including Numpy, Pandas, Matplotlib, etc.

Then the selected samples were converted into a CSV file that is easy to be processed by Python program. All these samples were shuffled into training and testing data. Moreover, training samples were thrown into the DBN model. The trained parameters of model were stored into Tensorboar. Finally, the classified map of was outputted.

3.3. Distribution of samples

There are total 97258 pixels in the pseudo color synthetic image, shown in Table 1. The samples of four types in study site are selected manually. Different allocation of training samples and test samples are selected by experiments again and again. At last, 28000 pixels of known categories were selected as total samples. Among them, 51989 pixels pertain to the coniferous forest, 6142 pixels are broad-leaved forest, 16283 pixels are mixed forest, and the remaining 28986 pixels are non-forest. In the training process, there are 10,000 training samples for coniferous forest and 6,000 training samples for other forest type.

Table 1: Class code of dataset

Code of label	Forest type	Number of training samples	Number of testing samples	Number of all samples
1	Coniferous forest	10000	4000	45848
2	Broad-leaved forest	6000	4000	6142
3	Mixed forest	6000	4000	16283
4	Non-forest	6000	4000	28985
	total	28000	16000	97258

4. Experimental Results

4.1. Selection of super parameters

In the experiment, the dimension of the input data is 86. At the moment, network structure of DBN model is only designed by experience without theoretical basis. For simplicity, we assume that each hidden layer has the same number of nodes. The number of layers of the DBN is selected from {2, 3, 4, 5, 6, 7}, and the number of hidden layer nodes is selected from {16, 32, 64, 128, 256, 512}. According to the pre-experiments and references, several parameters are set as following: pre-training and fine-tuning of the learning rate is 0.001, the size of mini-batch is 100. Note that 10000 iteration times are run due to stability. Next, we discuss the influence of different network depth and the number of hidden layer nodes on the classification effect by fixing all these super parameters.

4.2. Network depth

Overall accuracy (OA) and Kappa coefficient were normally selected as the evaluation of image classification of remote sensing. The higher OA and Kappa coefficient, the better effect of classification precision. As we all know, network depth is key parameter of DBN. It is the number of hidden layers. Reference 21 summarized the range of optimal network depth is from 2 to 3 based on previous researches.

In experiment, we keep super parameters fixed and set the number of nodes in each hidden layer 256. Six different numbers (2, 3, 4, 5, 6, 7) of network depth were tested. Figure 4 shows that DBN with 3 hidden layers performs best, which both overall accuracy and Kappa coefficient are the largest. There is no obvious feature that how network depth affects the OA and Kappa coefficient. Reference [21] pointed out that there is no perfect theoretical basis for network structure selection. The optimal parameters should be given by experiments in different applications.

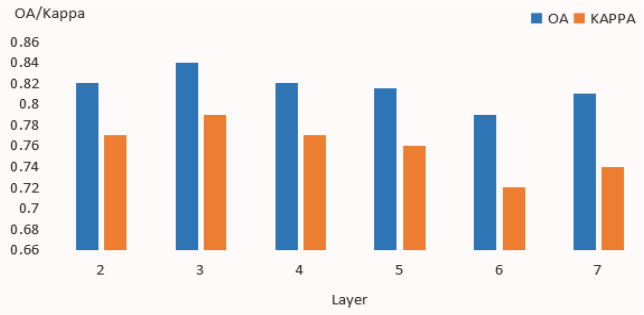


Figure 4: Effect of network depth on OA and Kappa coefficient

4.3. Number of nodes in each hidden layer

Number of nodes in the hidden layer is also import parameter in DBN model. When nodes are too large, it may cause the over-fitting problem. Oppositely, when nodes are too small, perhaps it will fail to extract deep information and gain high classification accuracy. Summarized by reference 21, the range of number of nodes in the hidden layer is from 50 to 500 according to previous research.

In experiment, we set super parameters unchanged and keep the network depth 3. Six various nodes (16, 32, 64, 128, 256, 512) were selected one by one and tested. Effect of number of nodes in each hidden layer on OA and Kappa coefficient is showed in Figure 5. It demonstrates that DBN with 256 nodes performs best, because it got the biggest overall accuracy and Kappa coefficient.

There is no obvious feature that how network depth affects the OA and Kappa coefficient. Reference [21-24] pointed out that there is no perfect theoretical basis for network structure selection. The optimal parameters should be given by experiments in different applications. Meanwhile, OA and Kappa coefficient keep relatively stable with changing nodes.

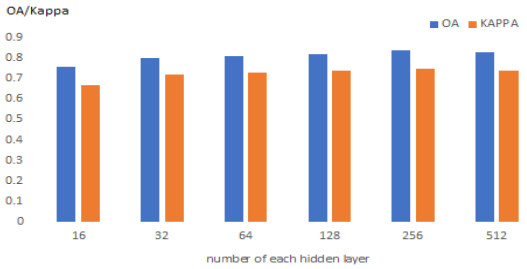


Figure 5: Effect of number of nodes in each hidden layer on OA and Kappa coefficient

4.4. Comparative analysis

SVM is a supervised and non-parametric machine learning algorithm. There is different type of kernels. A radial basis function (RBF) was adopted and has been used in some former works concerning forest type classification. The penalty factor C is selected from [1, 0.1, 0.001]. In order to keep the equit, the same samples are adopted between SVM and DBN method. Furthermore, to optimize of SVM method, five-fold cross validation and network search are adopted. When the C value is 1, the maximum overall accuracy is 73%, and the Kappa coefficient is 0.6447. For deep belief network, the optimal number of network depth is 3 and number of each node is 256.The classification process resulted in a thematic map of four kinds of forest types in Dehua county, as shown in Figure 6. As can be shown, the coniferous forest is green, broad-leaved forest is yellow, mixed forest is pink, and non-forest is blue.

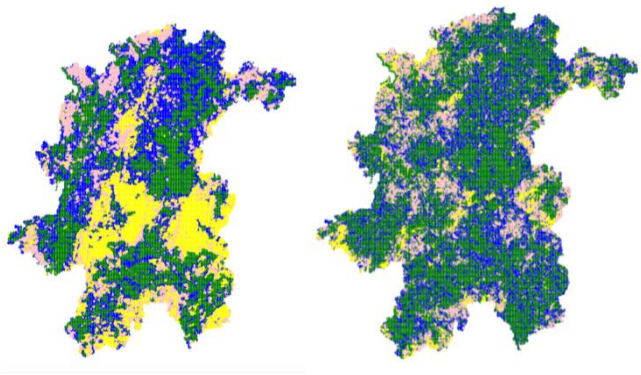
Table 2 reports the kappa values and the accuracy of forest types for both DBN and SVM classification algorithms. The accuracies of four kinds of forest types for SVM oscillated between 0.64 and 0.74 while DBN method varied from 0.82 to 0.89. Hyperspectral image classification using DBN algorithm yields the higher accuracy than that of SVM classifier for each forest type. Broad-leaved forest was generally better recognized than coniferous forest. Broad-leaved forest is classified with highest accuracy for both methods, 0.92 for DBN, while for SVM it is 0.84. The second top-classified forest type tends to be non-forest for DBN, while for SVM it is coniferous forest. Slightly lower result is obtained for mixed forest by DBN, while for non-forest by SVM. Furthermore, the worst effect performed by DBN algorithm is coniferous forest, compared to mixed forest for SVM. As the result, DBN model tends to attain better capability than SVM method from accuracy. There are two reasons due to it. One is that all feasible spectral features are thrown into input of DBN model, yet three bands or several components are selected as input for SVM method. The other is that there is large amount of data when all hyperspectal bands are considered. This is good for training of DBN model.

5. Conclusion and Discussion

This paper examined the capability DBN model in forest type classification with HJ/1A hyperspectral image. Many experiments are tested to obtain optimal parameter in DBN model. Conclusions can be drawing out as followed. At the beginning, the results showed that DBN model outperform SVM algorithm. Then, optimal network depth is 3 and best node in each layer is 256 in our experiment. Overall accuracy is 85.8% and kappa coefficient is 0.785 with best-fit parameters in DBN model.

Table 2: Comparison with SVM

Forest type	*SVM*	*DBN*
coniferous forest	0.74	0.82
broad-leaved forest	0.84	0.92
mixed forest	0.64	0.87
non-forest	0.70	0.89
OA	0.73	0.875
Kappa	0.6447	0.835



(a) SVM (b) DBN

Figure 6: Comparison of classification results based on different machine learning techniques

What makes our study unique is to make use of deep belief network to classify the forest type. Regardless of the spatial resolution of 100 m of HJ/1A hyperspectral image, our results provide satisfied recognition of four kinds of forest types. There are several open questions and future research directions remain worthy of investigation in the future. Firstly, in this paper, it is novative to adopt deep belief network to identify forest types. The classification accuracy of this method are better than that of support vector machine. However, it is still unclear that the mechanism of how this method can solve the traditional problem of "same object with different spectra" and "same spectra with different object ". Secondly, our results confirm the capability to map forest by deep belief network. Our experience in the study reveals that network depth and width affect accuracy of classification. The optimal structure needed to be adjusted according to different remote sensing data and further identification, such as tree species mapping. Finally, combining with field survey data, nearly one third of the samples are selected as training and testing samples to improve the classification effect. It is uncertain that how can we obtain the best result with so high number of total samples. It is a good idea to expand the sample by Generative Adversarial Network. Last but not least, optimal network structure is only for forest type recognition. The optimal network structure of specific forest type recognition needs further study. At the same time, it is urgent to establish the criteria and norms of forest classification using deep learning and remote sensing image.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The work was supported by the Project of Natural Science Foundation of Fujian Province under Grant No. 2020J01785, the project from the department of Quanzhou Science and Technology under Grant No. 2016N057 and the project from the department of education under Grant No. JAT170475. The author gratefully acknowledges the support of K.C. Wong Education and DAAD (No. 91551268). The authors also acknowledge the support by Fujian Provincial Key Laboratory of Data-Intensive Computing, Fujian University Laboratory of Intelligent Computing and Information Processing, and Fujian Provincial Big Data Research Institute of Intelligent Manufacturing.

References

[1] G. Modica, A. Merlino, F. Solano, R. Mercurio, "An index for the assessment of degraded Mediterranean forest ecosystems", Forest Systems, **24**, 1-3, 2015, doi : 10.5424/fs/2015243-07855.

[2] X.G. Kang, "Forest Management". Beijing: China Forestry Publishing House, 10-28, 2011.

[3] I.D. Thompson, M.R. Guariguata, K. Okabe, et al, "An operational framework for defining and monitoring forest degradation", Ecology and Society, **18**(20), 2013, doi: 10.5751/ES-05443-180220.

[4] L. Kurvonen, M.T. Hallikainen, "Textural information of multi-temporal ERS-1 and JERS-1 SAR images with applications to land and forest type classification in boreal zone", IEEE Transactions on Geoscience and Remote Sensing, **37**, 680-689, 1999, doi:10.1109/36.752185.

[5] Mei, A. X., Peng, W. L., Qin, Q.M. et al. "Introduction to remote sensing". Beijing: Higher education press, 225-286, 2006.

[6] A. Ghosh, F.E. Fassnacht, P.K. Joshi, B. Koch, "A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales". International Journal of Applied Earth Observation and Geoinformation, **26**, 49-63, 2014, doi:10.1016/j.jag.2013.05.017.

[7] Y.Y. Dian, Z.Y. Li, Y. Pang, "Spectral and Texture Features Combined for Forest Tree species Classification with Airborne Hyperspectral Imagery". J Indian Soc Remote Sens, **43**(1), 101-107, 2015, doi:10.1007/s12524-014-0392-6..

[8] N. Puletti, N. Camarretta, P. Corona, "Evaluating EO1-Hyperion capability for mapping conifer and broadleaved forests". European journal of remote sensing, 157-169, 2016doi:10.5721/eujrs20164909.

[9] G. Hughes, "On the mean accuracy of statistical pattern recognizers". IEEE transactions on information theory, **14**(1), 55-63, 1968, doi:10.1109/TIT.1968.1054102.

[10] R.E. Faßnacht, "Assessing the potential of imaging spectroscopy data to map tree species composition and bark beetle-related tree mortality". Freiburg: Freiburg university, 1-20, 2013, doi: 10.1016/j.rse.2013.09.014.

[11] M. Mustafa, A.R. James, S.J. Karl, R. David, M. Charles, "Remote Distinction of A Noxious Weed (Musk Thistle: CarduusNutans) Using Airborne Hyperspectral Imagery and the Support Vector Machine Classifier". Remote Sensing, **5**(2), 612-630, 2013, doi: 10.3390/rs5020612.

[12] Y.S. Chen, X. Zhao, X.P. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network". IEEE Journal of selected topics in applied earth observations and remote sensing. 1-12, 2015, doi:10.1109/JSTARS.2015.2388577.

[13] E. Alpaydin, Introduction to machine learning(third edition), The MIT press, 267-311, 2014.

[14] L. Deng, S.S. Fu, R.X. Zhang, "Application of deep belief network in polarimetric SAR image classification". Journal of image and graphics, **21**, 933-941, 2016, doi: 10.1109/IGARSS.2015.7326284.

[15] Q. Lü, Y. Dou, X. Niu, J.Q. Xu, F. Xia, "Remote sensing image classification based on DBN model". Journal of computer research and development. 51, 1911-1918, 2014, doi:10.7544/issn1000-1239.2014.20140199.

[16] P. Zhong, Z.Q. Gong, S.T. Li, et al., "Learning to Diversify Deep Belief Networks for Hyperspectral Image Classification". IEEE Transactions on Geoscience and Remote Sensing, 1-15, 2017, doi:10.1109/TGRS.2017.2675902.

[17] G.E. Hinton, J.S. Terrance, Learning and relearning in Boltzmann machines. *MIT Press, Cambridge*, 282-317, 1986.

[18] H.M. Xie, S. Wang, K. Liu, S.P. Lin, B. Hou. "Multilayer feature learning for polarimetric synthetic radar data classification". in IEEE International Geosicence and Remote Sensing Symposium. 2818-2821, 2014, doi: 10.1109/IGARSS.2014.6947062.

[19] G.E. Hinton, R.R. Salakhutdinov. "Reducing the dimensionality data with neural networks", Science, **313**, 504-507, 2006, doi: 10.1126/science.1127647.

[20] L. Deng, S.S. Fu, R.X. Zhang. "Application of deep belief network in polarimetric SAR image classification". Journal of image and graphics, **21**, 933-941, 2016, doi:10.11834/jig.20160711.

[21] C.C. Yao, X.X. Luo, Y.D. Zhao, et al, "A review on image classification of remote sensing using deep learning" in Proceedings of 2017 3rd IEEE Internaitonal Conference on Computer and Commmocaotpms, 1947-1955, 2017, doi: 10.1109/CompComm.2017.8322878.

[22] S.J. Ge, J. Lu, H. Gu, et al., "Polarimetrie SAR image classification based on deep belief network and superpixel segmentation." in 2017 3rd International Conference on Frontiers of Signal Processing, 115-119, 2017, doi: 10.1109/ICFSP.2017.8097153.

[23] M.M. Lau, J.T.S. Phang, K.H. Lim, "Convolutional Deep Feedforward Network for Image Classification." in 2019 7th International Conference on Smart Computing & Communications, 2019, doi:

[24] A. Mughees, L.Tao, "Multiple Deep-Belief-Network-Based Spectral-Spatial Classification of Hyperspectral Images." Tsinghua Science and Technology, **24**(2), 183-194, 2019, doi: 10.26599/TST.2018.9010043.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1560-1568 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Computational Modelling and Algorithmic Design Approach of Digital Watermarking in Deep Neural Networks

Revanna Sidamma Kavitha[1 *], Uppara Eranna[2], Mahendra Nanjappa Giriprasad[1]

[1]*Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University Anantapur, Ananthapuramu, 515002, India*

[2]*Department of Electronics and Communication Engineering, Ballari Institute of Technology and Management, Ballari-583104, India*

ARTICLE INFO

Article history:
Received: 21 October, 2020
Accepted: 12 December, 2020
Online: 25 December, 2020

Keywords:
Digital Water Marking
Neural Networks
Algorithms
Convolution Neural Network

ABSTRACT

In this paper we propose an algorithmic approach for Convolutional Neural Network (CNN) for digital watermarking which outperforms the existing frequency domain techniques in all aspects including security along with the criteria in the neural networks such as conditions embedded, and types of watermarking attack. This research addresses digital watermarking in deep neural networks and with comprehensive experiments through computational modeling and algorithm design, we examine the performance of the built system to demonstrate the potential of watermarking neural networks. The inability of intruder towards the retrieval of data without the knowledge of architecture and keys is also discussed and results of the proposed method are compared with the state of the art methods at different noises and attacks.

1. Introduction

The research presented here is the extension of the work originally presented at International Conference on Artificial Intelligence and Signal Processing (AISP),2020 [1]. The digital revolution and the internet have paved a way to the creation of massive digital information containing images, videos, transactions, intellectual properties. The ease with which this digital data can be copied and reproduced has created avenues for copyright infringements. The massive explosion of digital multimedia devices has resulted in the creation of a large chunk of data and increased demand (and role) of data hiding techniques. Digital watermarking is employed for various applications such as copyright protection (ownership assertion), broadcast monitoring (really broadcasted or not), tamper detection (persistent item identification, forgery detection), data authentication and verification (integrity verification), fingerprinting (transaction tracking and privacy control), content description (labelling and captioning), publication monitoring and copy control (unauthorized distribution) covert communication (data hiding), Medical applications (Annotation and privacy control), and Legacy system enhancement (backward compatibility) [2]. All information hiding techniques revolve around 3 parameters: Imperceptibility, robustness and payload capacity [3]. While payload capacity is important in steganography, in digital watermarking the trade-off between imperceptibility and robustness is needed for the excellent quality of data hiding [4]. The ability of a data hiding technique to remain unchanged to human perception is called imperceptibility i.e. an unintended user should not be able to make out whether the image has undergone watermarking [5]. Robustness is the ability of watermarked data to be immune to attacks and threats. Other goals of digital watermarking are Security (watermark should be secret and undetectable by an unauthorized parties), effectiveness (ease of detection immediately after embedding), uniqueness (multiple watermarks to coexist), cost effectiveness (in terms of hardware and computational speed), and scalability [6]. Several Artificial Intelligent techniques such as neural network, evolutionary computation, fuzzy logic, swarm intelligence, Probabilistic reasoning and multi-agent systems were proposed to secure watermark in digital media [7]. These techniques will be employed at either the transmitter side during embedding the watermark or during the extraction stage at receiver. Recent approaches have also attempted to incorporate AI methods in the pre-processing stage [8].

With such diverse approaches and their unparalleled capabilities, the hiding of data in the digital image will be far more superior in terms of imperceptibility and robustness as compared to the existing ones [9].

*Corresponding Author: R.S Kavitha, drkavithakavana@gmail.com

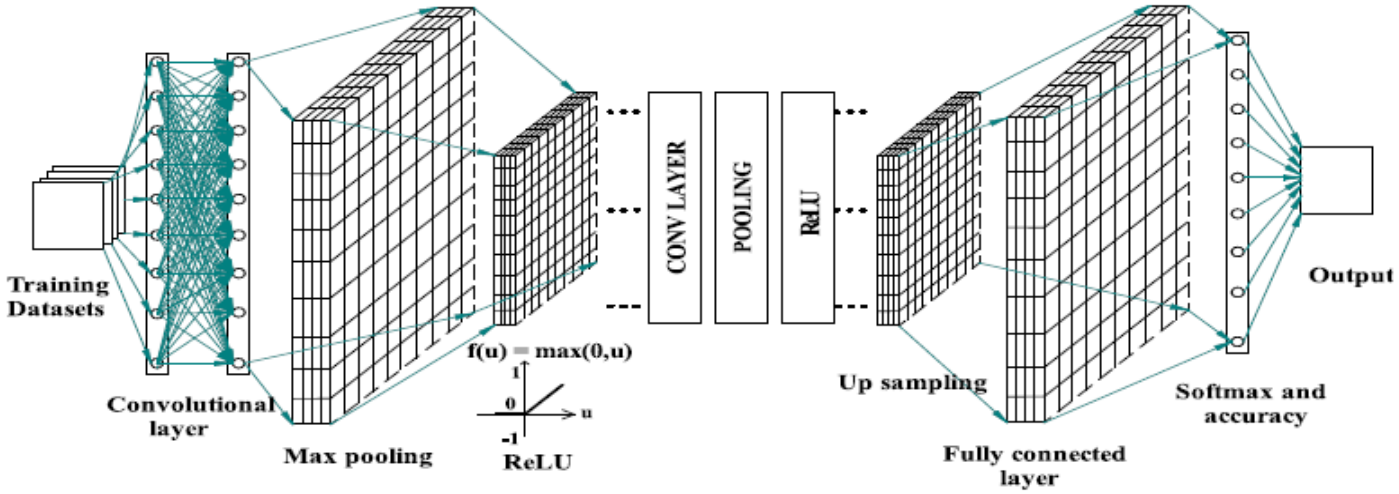


Figure 1: A typical digital watermarking process with Deep learning and its associated architectures

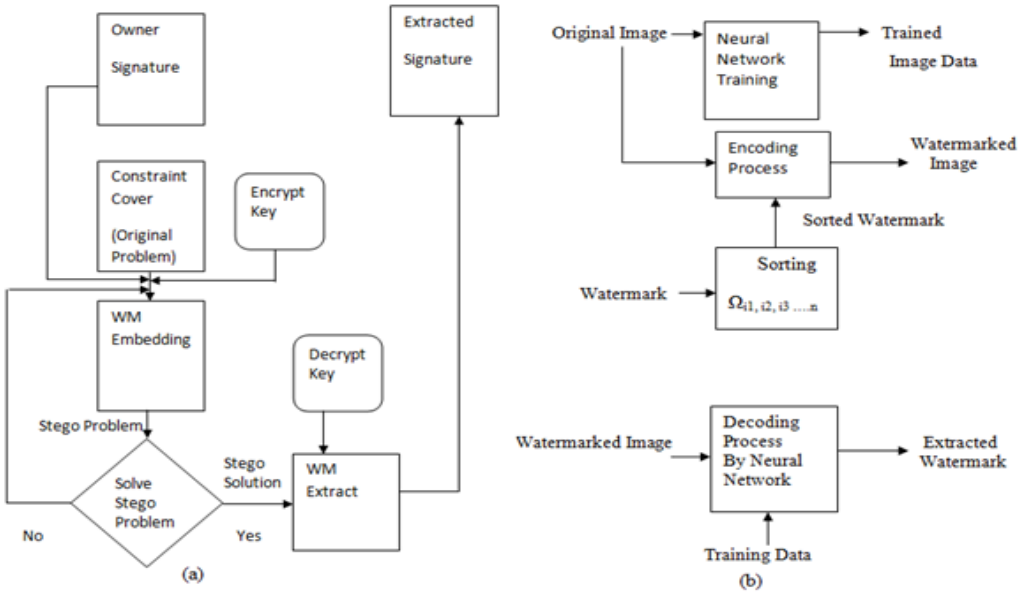


Figure 2: Experimental setup block diagram for digital watermarking process

A typical digital watermarking mechanism with embedding and extracting stages is depicted in Figure 1. It enables owners of the digital documents to embedded their copyright information for information security.

Digital watermarking can be done on text, image, audio, video and graphics in spatial or frequency domain. The watermark can be of noise type (pseudo noise, Gaussian random and chaotic sequences) or image type (binary image, stamp, logo and label) [10]. Based on the deployment conditions various watermarking techniques can be used. For public use, visible watermarks are preferred while for private applications and to arrest unauthorized copying invisible watermarking can be used [11]. Fragile watermarks are used in tamper-proof applications whereas robust watermarking is used in applications where the watermark should remain intact even after modification or tampered with [12]. According to the detection stage, visual watermarking is more robust which needs the original media and the embedded watermark for detection while blind watermarking does not require any of these. This is the most demanding type of watermarking as the watermarking is generated and embedded at the transmitter, while detection and extraction will happen at receiver as illustrated in Figure 2. Watermarking is also done even in prepossessing stage and the approach presented here does not impact the efficiency of the network in which a watermark is inserted as the embedded watermark while it is training the host network [13, 14].

The robustness of hard and soft decision detectors can be measured using Receiver Operating Characteristic graphs while Bit Error Rate is used for the detector response with bit sequence.

2. Literature Review

Many standard quantitative measures and metrics have been proposed to evaluate digital watermarking while comparing with their counterparts. Here we report all the standard methods found in the literature as illustrated in Table 1 and those used in this work to measure imperceptibility and robustness (along with capacity and computational cost) [34]. We use the methods for performance evaluation by ensuring

Table 1: Summary of literature review on Algorithm Transforms and Classifications

Reference		Survey/Tutorial	Parameter(s)
Frequency Domain + Back Propagation NN			
[15]		Discrete Wavelet Transform	Robustness
[16]		Multiwavelet Transform	Imperceptibility
[17]		Discrete Wavelet Transform	Robustness
[18]		Fourier Transform	Fidelity
[19]		Discrete wavelet Transform	Imperceptibility
[20]		Discrete wavelet Transform	Imperceptibility
Frequency Domain + Radial Basis NN			
[21]		Discrete Cosine Transform	Robustness
[22]		Discrete wavelet Transform	Imperceptible
[23]		Discrete wavelet Transform	Invisible and Robust
[24]		Discrete wavelet Transform	Robustness
Frequency Domain + Hopfield NN			
[25]			Capacity
[26]			Imperceptibility
[27]			Image Quality
[28]		Discrete Cosine Transform	Invisible and Robust
Frequency Domain + Full Counter Propagation NN			
[29]			Robustness, Imperceptibility
[30]		Discrete Cosine Transform	Robustness
[31]		Discrete Cosine Transform	Complexity, Capacity, PSNR
[32]		Discrete Cosine Transform	Imperceptibility, and robustness
Frequency Domain + Synergetic NN			
[33]		Discrete Wavelet Transform	Robustness and Imperceptibility

i) model and sources of distortion remain uniform
ii) All test images are 8-bit grey-scale images and are defined in same color space.

Quality assessment can also be done by comparing the original watermark and extracted watermark. This alternate metric is called a Normalized Correlation which exploits the correlation between the original watermark and the extracted one. The value lies between [0 1], any value nearer to 1 assures better quality. Another way of defining the similarity between the original watermark and the extracted one is accuracy ratio. It is the ratio of correct bits to the total bits. The architecture of CNN is different, unlike neural network where all layers are fully connected, here the layers are recognized in 3D: height, width, and depth. Further neurons in one layer are connected to only a small region of the next layer. Finally, the output is a single vector of probability scores, organized along the depth dimension [35]. The CNN consists of series of convolutional, pooling/sub-sampling layers followed by a fully connected layer. To implement the act of recognition in machines we need to show an algorithm of millions of images before it makes a pattern by generalizing the input and start making predictions for images it has never seen before [36]. The aim is to evolve a more robust and imperceptible watermarking scheme that can cater to the needs of content protection [37].

3. Algorithm Design

Watermarking can be achieved in any of the following two methods: either by changing the pixel values (least significant) of the image or by changing the coefficient values [38]. The quality of watermark depends on the method used. The first method where bits (representing the pixel values) are manipulated refers to a spatial domain which is very simple but not robust and can be easily perceived. In simple terms, spatial domain techniques refer to replacing pixels of the original image by watermark image [39]. In a given image first, the target pixels are identified and are replaced by pixels of watermark image. The spatial domain techniques are simple, fast and with less computational complexity [40]. They are immune to cropping and noising but are sensitive to signal processing attacks. In pursuit of enhancing robustness, if more pixels are manipulated we may end up with visible watermarks. These algorithms should carefully achieve a trade-off between robustness and imperceptibility [41]. In this section the experimental setup of watermarking a digital image is described with Algorithm 1 and Algorithm 2. At the various stages of digital watermarking from encoding of original image to decoding process and finally up to extracting the watermark and obtaining the high-resolution output decoded image. The aim is to evolve a more robust and imperceptible watermarking scheme that can cater to the needs of content protection and piracy prevention [42]. All the transform (and few hybrid approaches with NN) methods are good for watermarking but lack learning and adaptability [43]. We propose a digital watermarking method using deep learning methods which exploit the expressiveness of deep NN to securely embed invisible, imperceptible, attack-resilient binary signatures into the cover images. Coming to the

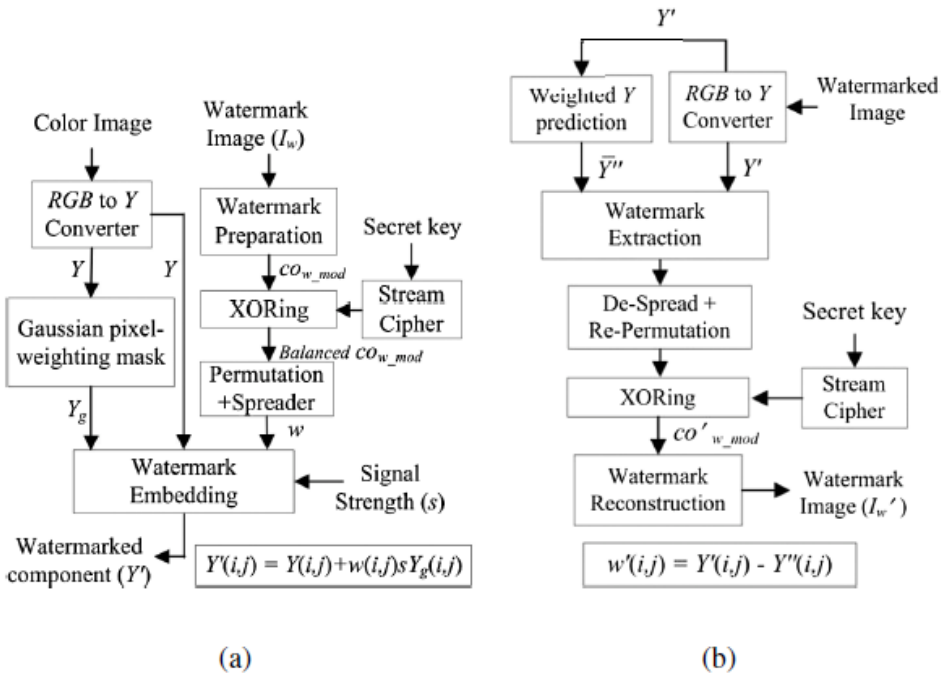


Figure 3: Block diagram representation of a typical digital watermarking process (a)Watermark Embedding and (b)Watermark Extraction.

decoder, we adopt adversarial model techniques to cause disorders to decode the desired signature [44]. We perform extended gradient descent under the Expectation over Transformation framework. In training the decoder network an Expectation-Maximization (EM) framework is employed to learn feature transformations that are more resilient to the attacks [45]. Experimental results indicate that our model achieves robustness across different transformations (all transformations, including scaling, rotation, adding noise, blurring, random cropping, and more) [46]. The aim is to evolve a more robust and imperceptible watermarking scheme that can cater to the needs of content protection and piracy prevention [47]. All the transform (and few hybrid approaches with NN) methods are good for watermarking but lack learning and adaptability [48]. We propose a digital watermarking method using deep learning methods which exploit the expressiveness of deep NN to securely embed an invisible, imperceptible, attack resilient binary signature into the cover images [49]. Coming to the decoder, we adopt adversarial model techniques to cause disorders to decode the desired signature.

Algorithm 1: Watermarking using Deep NN(Part-I)

Result: Encoding
Initialization;
Input_IMG: Image for Watermarking
Compute α transparency;
Construct an overlay;
Add Watermark to the overlay
while *setup* **do**
  Start encoding model;
  Image Encoding Process Starts;
  return encoded images;
  if *Given the model and targets* **then**
    compute the cross entropy loss;
  Else
    Given number of Iterations;
    Encode the set of images with specified binary targets;
    Image Encoding Process Ends;
  End Encoding Model
End

We perform extended gradient descent under the Expectation over Transformation framework. In training the decoder network an Expectation-Maximization (EM) framework is employed to learn feature transformations that are more resilient to the attacks. To implement the act of recognition in machines we need to show an algorithm of millions of images before it makes a pattern by generalizing the input and start making predictions for images it has never seen before [50]. The architecture of CNN is different, unlike NN where all layers are fully connected, here the layers are recognized in 3D: height, width, and depth. Further neurons in one layer are connected to only a small region of the next layer. Finally, the output is a single vector of probability scores, organized along the depth dimension. The CNN consists of series of convolutional, pooling/sub-sampling layers followed by a fully connected layer. The Figure 3 elucidates the block diagram representation of a typical digital watermarking process with Figure 3(a) as Watermark embedding process and Figure 3(b) watermark extraction phase. In this the input of the deep learning algorithm is the colour image and watermark to embed inside the colour image to watermark the deep Neural Network (DNN). The Algorithm 1 and Algorithm 2 are formulated based on the Figure 3(a) and 3(b).

Algorithm 2: Watermarking using Deep NN(Part-II)

Result: Decoding
Initialization;
While *setup* **do**
  Start the Decoding model;
  if *decoding network* **then**
    Initialize Decoding;

Extract Watermark;
Obtain High Resolution
Decoded Image;
Else
End of Decoded Network;
Return predictions;
Return decoding model;
Image Decoding Process Ends;
End Decoding Model
End

The Algorithm 1 illustrates the encoding process of digital watermarking using deep neural network and the Algorithm 2 presents the decoding process of digital watermarking using deep neural network.

4. Performance Evaluation

Fair performance evaluation of any system is fundamental for its acceptance and accreditation. Many standard quantitative measures and metrics have been proposed to evaluate digital watermarking while comparing with their counterparts. Here we report all the standard methods found in the literature and those used in this work to measure imperceptibility and robustness (along with capacity and computational cost). We use the methods for performance evaluation by ensuring i) model and sources of distortion remain uniform ii) all test images are 8-bit gray-scale images and are defined in same color space. For an image size of M x N pixels, with a pixel value of O for original image (without watermark) and W for watermarked image, the performance metrics can be calculated as below:

MSRE:

$$MSRE(O,W) = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{J=0}^{N-1}[\mathrm{O(i,j)} - W(i,j)]e^2$$

PSNR:

$PSNR(O,W) = 10\log 10[\{255\}\,e^2/\mathrm{MSRE\ (O,W)}$

4.1. Robustness Metrics

Robustness Metrics: Robustness of a watermark is a measure of resistance to attacks. The detectors dictate the evaluation method to measure robustness. The three types of detector responses are i) hard decisions (true or false for the presence and absence of watermark respectively), ii) soft decisions (correlation or similarity coefficients in terms of real numbers) iii) bit sequence (if the embedded watermark is in form of a message). The robustness of hard and soft decision detectors can be measured using Receiver Operating Characteristic (ROC) graphs while Bit Error Rate (BER) is used for the detector response with bit sequence.

ROC:

True Positive Fraction (TPF): $\frac{Number\ of\ True\ Positive\ Results}{True\ Positive+False\ Negative\ Results}$

False Positive Fraction (FPF): $\frac{Number\ of\ False\ Positive\ Results}{Fasle\ \ Positive+True\ Negative\ Results}$

Quality assessment can also be done by comparing the original watermark and extracted watermark. This alternate metric is called a Normalized Correlation which exploits the correlation between the original watermark and the extracted one. The value lies between [0 1], any value nearer to 1 assures better quality. Another way of defining the similarity between the original watermark and the extracted one is accuracy ratio. It is the ratio of correct bits to the total bits. When the attack is not going to affect the commercial value then it is better to consider only the watermark-to-noise ratio. This ratio signifies the power of watermark signal against the noise introduced by such attacks. For a watermark of size m x n pixels, the performance metrics can be calculated as below:

$$NC(OW,EW) = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{J=0}^{n-1}[\delta[\mathrm{OW(i,j)} - EW(i,j)]$$

where,

δ (X,Y) = 1 if X=Y and 0 otherwise;

5. Results Obtained

In this section we discuss the watermarking in the presence of various attacks at different noise levels. We present the results in terms of robustness of the watermark (BER and NC). The CNN is trained using standard descent back propagation algorithm. The performance consistency of the proposed method is verified by considering 2 different cover images: Lena and Camera man as presented in Figure 4 shows the watermarking process with different transforms (scaling, rotation, adding noise, blurring, random cropping, and more)and also confirms that that the proposed method is applicable to any cover image and with any watermark. BER is a measure of noise injected when the

signal is received after transmission channel. The BER shows the signal loss and fading in a wireless channel. The BER for various noise levels is shown in Figure 5(a). Normalized (cross) correlation is a template-matching method in digital watermarking. The template will be an image that shows a critical feature; by repeatedly computing a statistic between the watermarked image and corresponding pixels of a subset of an original image presents the noise correlation for various noise levels of 2,5,10 and 15. The important parameter in watermarking is the loss of original information and the accuracy with which the watermark is hidden in the cover image. There is a fine balance between the robustness and imperceptibility. The trade-off is to achieve high robustness without showing any trace of watermark. The accuracy of our model is gradually increasing as the epochs increases, on the contrary loss is gradually decreasing. The models training and testing results are shown in Figure 6 with the over fit, under fit and good fit are shown in Figures 6(a), 6(b) and 6(c) respectively. Our model Adam performs well and is consistent as compared to other models as shown in Figure 7. The Figure 8 shows the noise injected/present and the model performance. In the Figure 8(a) noise in hidden layer is shown while Figure 8(b) depicts the noise in input layer.



Figure 4: Typical digital watermarking process for Lena and Cameraman Images (a)original image (b)watermarked image (c)initial decoded image with watermark (d) extracted watermark (e)high resolution output image

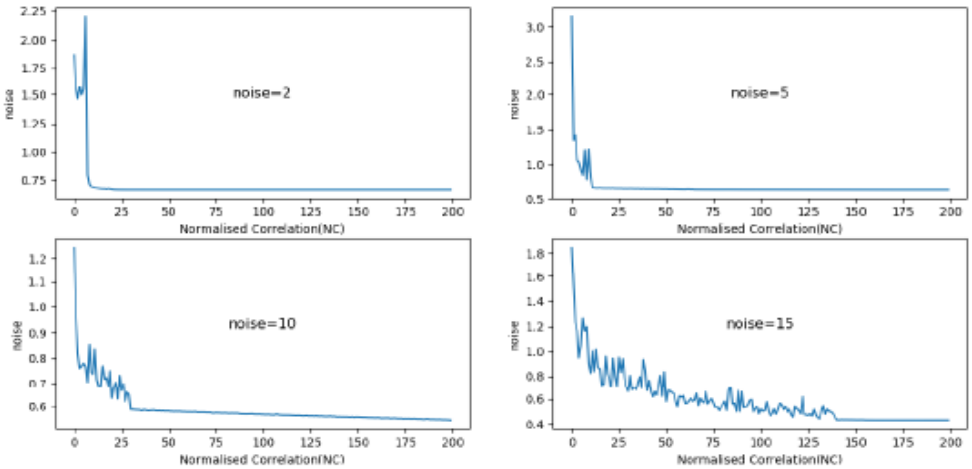


Figure 5: Verification of Robustness of the digital watermark against various type of transformations, shows the BER of watermark for various noise levels 2, 5, 10 and 15

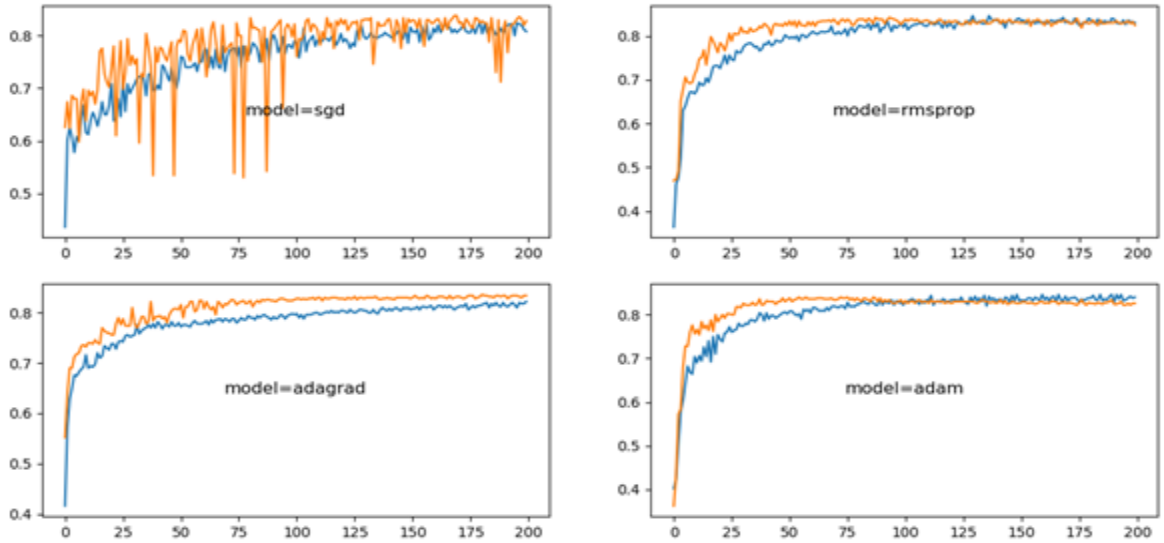


Figure 6: Validation of Neural Network Model

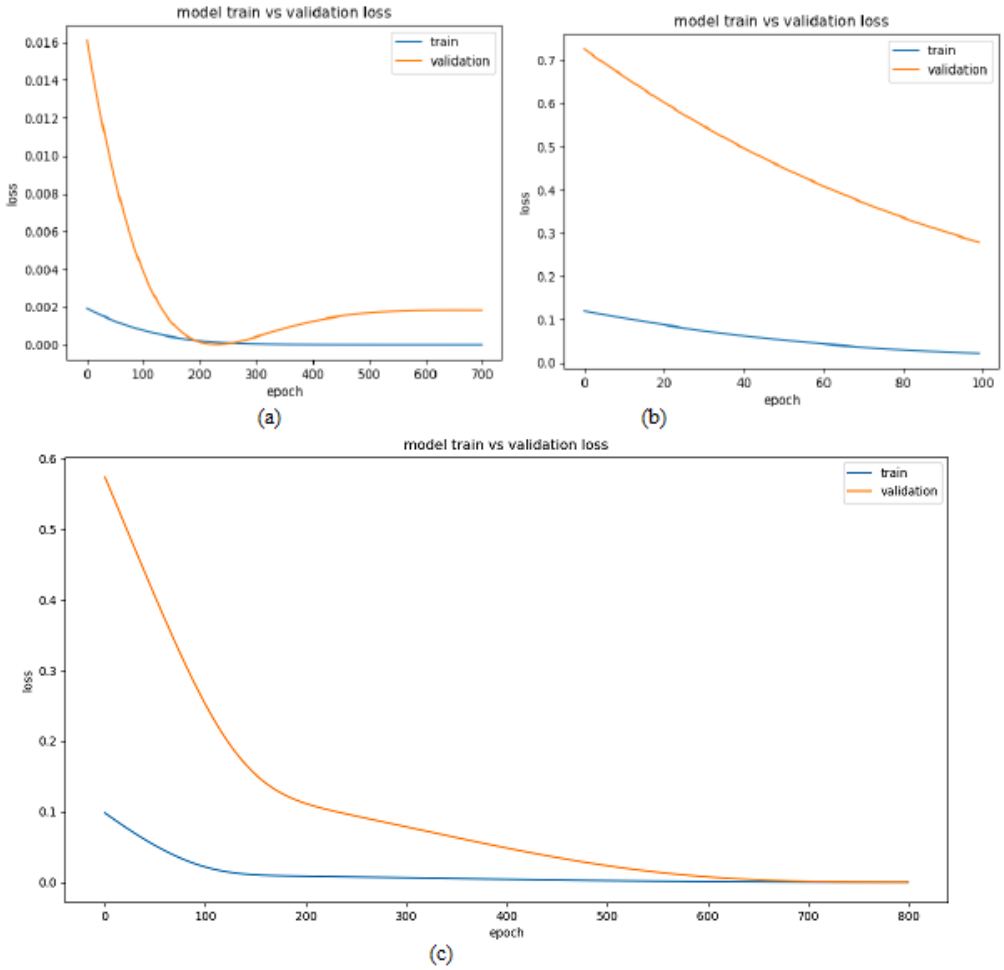


Figure 7: Fitting the model for validation (a)over fit (b) under fit and (c) good fit.

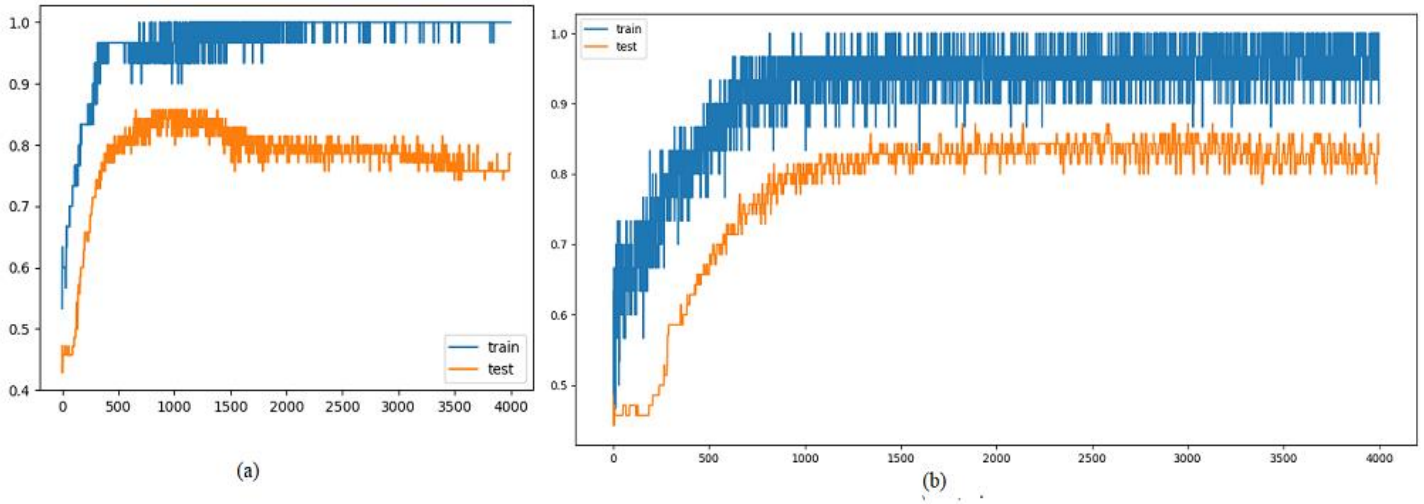


Figure 8: Injected/present noises in(a)hidden layers and (b)input layer of the proposed convolutional neural network.

Table 2: Comparison of Results

S.No	Noise/Attack	Parameter	[51]	[52]	[53]	[54]	[55]	[56]	[57]	[58]	Our Work
1	No Noise	PSNR(db)	58.91	48.29	48.47	51.45	34.33	56.47	35.6	47	**62.3**
		NC	1	1	1	1	1	1	0.99	0.98	**1.00**
2	With Noise	PSNR(db)	55.10	43.04	32.74	31.66	22.43	40.18	29.1	41.72	**58.78**
		NC	1	0.91	0.81	0.92	0.91	0.93	0.95	0.91	**1.00**

The graph shows that training and test results are in good accordance and validate the model used for training and optimization in digital watermarking. The proposed model is suitable for video too as the training is done off-line hence can be used in real-time applications. Hence the use of deep neural networks will enable us to realize complex features like learning weight sharing and update mechanism, noise-resilience and immunity towards attacks, and scaling. The proposed method has low loss and the accuracy goes on increasing as the number of epochs increases. The accuracy is 85% for 15 epochs and will slightly increase for more epochs. The level of high accuracy shows that the original image and watermarked image are indistinguishable.

This high performance of the model can be owed to the learning feature embedded in the watermarking. A digital watermarking task comprises embedding a signal into an image in accordance with robustness and quality constraints, it can be said that it is in essence a multi-objective optimization problem. The faster convergence of the algorithm (more accuracy with less epochs) can be achieved by introducing reinforcement learning or transfer learning method which has got huge attention in recent times. The current digital watermarking can also be validated by applying it to protect the copyrights of trained neural networks where ownership protection and piracy prevention is of utmost priority. The traditional approaches in digital watermarking are not fit for the digital data that can be stored efficiently and with very high quality and manipulated easily using computers [59, 60]. The aim is to evolve a secure digital communication that remarkably pushes forward the limits of legacy digital watermarking schemes across all dimensions of performance metrics. As this research space is ever increasing and innovations have spurred at all levels of communication, considerable progress is required in understanding the deep learning approaches for digital watermarking. The convergence of technological, economic and environmental forces is driving the digital watermarking and deep learning simultaneously, then each drives the other forward. Be it the deep neural networks driving digital watermarking or vice-versa, the continued expansion of each is good for the other. The complete suitability of digital watermarking for securing deep neural networks dataset is yet to be conducted. The trained models can be viewed as intellectual property, and it is a worthy challenge to provide copyright protection for trained models. We emphasis on how the copyrights of trained models can be protected computationally and propose for neural networks a digital watermarking technology. We propose a conceptual framework for integrating a watermark into models of deep neural networks to safeguard copyrights and identify violation of trained models of intellectual property. The Table 2 illustrates the comparison of obtained results with the previous works and it is noteworthy that the proposed work has shown an improvement of 5.75% in PSNR without noise and 6.68% in PSNR with noise = 5.

6. Conclusions

In this paper, we proposed a learning framework for robust digital image watermarking technique based on deep neural network. As observed, previous efforts in this space focused on optimization of embedding parameters with use of evolutionary computing. Demonstration of the watermarking under various noise and attacks is performed. The detailed experiments were carried out and we analyzed the performance of designed system. We have shown that our model could embed a watermark without impairing a deep neural network's efficiency. In future the research would continue in the direction towards digital watermarking based on intelligence.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] R.S. Kavitha, U. Eranna, M.N. Giriprasad, “DCT-DWT Based Digital Watermarking and Extraction using Neural Networks,” in 2020 International Conference on Artificial Intelligence and Signal Processing, AISP 2020, 2020, doi:10.1109/AISP48273.2020.9073104.

[2] N. Memon, P.W. Wong, “A buyer-seller watermarking protocol,” IEEE Transactions on Image Proc., **10**(4), 643-649, 2001, doi:10.1109/83.913598.

[3] I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S. U Khan, "The rise of “big data” on cloud computing: Review open research issues." Information systems, **47**, 98-115, 2015, doi:10.1016/j.is.2014.07.006.

[4] N.F. Johnson, Z. Duric, S. Jajodia, N. Memon, “Information Hiding: Steganography and Watermarking—Attacks and Countermeasures,” Journal of Electronic Imaging, 2001, doi:10.1117/1.1388610.

[5] J.C. Ingemar, M.L. Miller, A.B. Jeffrey, J. Fridrich, T. Kalker, Digital Watermarking and Steganography, 2008, doi:10.1016/B978-0-12-372585-1.X5001-3.

[6] Y.S. Singh, B.P. Devi, K.M. Singh, “Image comp using Multilayer Feed Forward Artificial Neural Network with Conjugate Gradient,” in Proceedings of the 2012 World Congress on Information and CommunicationTechn,976-980,2012, doi:10.1109/WICT.2012.6409216.

[7] A.Z. Tirkel, G.A. Rankin, R.M. Van Schyndel, W.J. Ho, N.R.A. Mee, C.F. Osborne, “Electronic Water Mark,” Digital Image Computing, Technology and Applications (DICTA’93), 1993.

[8] S. Wu, S. Zhong, Y. Liu, “Deep residual learning for image steganalysis,” Multimedia Tools Apps, **77**(9), 10437-10453, 2017, doi:10.1007/s11042-017-4440-4.

[9] S.J. Lee, S.H. Jung, “A survey of watermarking techniques applied to multimedia,” in IEEE International Symposium on Industrial Electronics,1,272-277, 2001, doi:10.1109/isie.2001.931796.

[10] R. Warkar, P. More, D. Waghole, “Digital audio watermarking and image watermarking for information security,” in 2015 International Conference on Pervasive Computing: Advance Comm. Tech. and Application for Society, ICPC2015, doi:10.1109/PERVASIVE.2015.7086980.

[11] A.K. Singh, B. Kumar, M. Dave, A. Mohan, “Robust and Imperceptible Spread-Spectrum Watermarking for Telemedicine Applications,” Proceedings of the National Academy of Sciences India Section A - Physical Sciences, **85**(2),295-301, 2015, doi:10.1007/s40010-014-0197-6.

[12] M. Barni, F. Bartolini, V. Cappellini, A. Piva, “A DCT-domain system for robust image watermarking,” Signal Processing, 1998, doi:10.1016/S0165-1684(98)00015-2.

[13] T. Ashkan et al.,"A modified dual watermarking scheme for digital images with tamper localization/detection and recovery capabilities." 2012 9th International ISC Conference on Information Security and Cryptology. IEEE, 2012.

[14] P.K. Dhar, J.M. Kim, “Digital watermarking scheme based on fast Fourier transformation for audio copyright protection,” International Journal of Security and Its Applications, **5**(2),33-48, 2011. doi: 10.1016/j.future.2018.07.029.

[15] I.J. Cox, J. Kilian, F.T. Leighton, T. Shamoon, “Secure spread spectrum watermarking for multimedia,” IEEE Transactions on Image Processing, 1997, doi:10.1109/83.650120.

[16] D. Kundur, D. Hatzinakos, “Digital watermarking for telltale tamper proofing and authentication,” Proceedings of the IEEE, 1999, doi:10.1109/5.771070.

[17] A. Tiwari, M. Sharma, R.K. Tamrakar, “Watermarking based image authentication and tamper detection algorithm using vector quantization approach,” AEU - International Journal of Electronics and Communications, **78(5)** ,114-123, 2017, doi: 10.1016/j.aeue.2017.05.027.

[18] Z. Ma, M. Jiang, H. Gao, Z. Wang, “Block chain for digital rights management,” Future Generation Computer Systems, **89**(7), 746-764, 2018, doi: 10.1016/j.future.2018.07.029.

[19] J. Kishigami, S. Fujimura, H. Watanabe, A. Nakadaira, A. Akutsu, "The Block Chain-Based Digital Content Distribution System," in Proceedings - 2015 IEEE 5th International Conference on Big Data and Cloud Computing, B D Cloud 2015, 2015, doi:10.1109/BDCloud.2015.60.

[20] Stefik, Mark J., et al. "System for controlling the distribution and use of rendered digital works through watermaking." U.S. Patent No. 6,233,684. 15 May 2001.

[21] T. Jitha Raj, E.T. Sivadasan, "A survey paper on various reversible data hiding techniques in encrypted images," in Souvenir of the 2015 IEEE International Advance Computing Conference, IACC 2015, 2015, doi:10.1109/IADCC.2015.7154881.

[22] A.A. Philip, P.R. Geetharanjin, "Fingerprint Encryption and Dual Watermarking to Verify the Security Issues in Teleradiology," in Proceedings of the 3rd International Conference on Communication and Electronics Systems, ICCES 2018, 2018, doi:10.1109/CESYS.2018.8724112.

[23] I.J. Cox, M.L. Miller, "The first 50 years of electronic watermarking," EURASIP Journal on Applied Signal Processing, **2**(2)**,** 820-936, 2002, doi:10.1155/s1110865702000525.

[24] J. Bajpai, A. Kaur, "A literature survey - Various audio watermarking techniques and their challenges," in Proceedings of the 2016 6th International Conference - Cloud System and Big Data Engineering, Confluence 2016, 2016, doi:10.1109/CONFLUENCE.2016.7508162.

[25] C. Kumar, A.K. Singh, P. Kumar, "A recent survey on image watermarking techniques and its application in e-governance," Multimedia Tools and Applications, **77(3)**, 3597-3622, 2018, doi:10.1007/s11042-017-5222-8.

[26] M.H. Alkawaz, G. Sulong, T. Saba, A.S. Almazyad, A. Rehman, "Concise analysis of current text automation and watermarking approaches," Security and Communication Networks, **9**(18), 6365-6378, 2016, doi:10.1002/sec.1738.

[27] A. Świerkosz, "Digital watermarking in telemedicine an example from ECG - Review of challenges, methods and applications," Advances in Intelligent Systems and Computing, 248-255, 2017, doi:10.1007/978-3-319-47154-929.

[28] S. Mirghasemi, P. Andreae, M. Zhang, "Domain-independent severely noisy image segmentation via adaptive wavelet shrinkage using particle swarm optimization and fuzzy C-means," Expert Systems with Applications, **133**(4), 126-150, 2019, doi: 10.1016/j.eswa.2019.04.050.

[29] Q.D. La, M. V. Ngo, T.Q. Dinh, T.Q.S. Quek, H. Shin, "Enabling intelligence in fog computing to achieve energy and latency reduction," Digital Communications Networks, 2019, doi: 10.1016/j.dcan.2018.10.008.

[30] H. Kandi, D. Mishra, S.R.K.S. Gorthi, "Exploring the learning capabilities of convolutional neural networks for robust image watermarking," Computers Security ,**65**(3), 247-268, 2017, doi: 10.1016/j.cose.2016.11.016.

[31] M.S. Hsieh, D.C. Tseng, Y.H. Huang, "Hiding digital watermarks using multiresolution wavelet transform," IEEE Transactions on Industrial Electronics, **48**(5) ,875-882, 2001, doi:10.1109/41.954550.

[32] N.M. Makbol, B.E. Khoo, "Robust blind image watermarking scheme based on Redundant Discrete Wavelet Transform and Singular Value Decomposition," AEU - International Journal of Electronics and Communications,**67(2)**,102-112, 2013, doi: 10.1016/j.aeue.2012.06.008.

[33] H.T. Hu, L.Y. Hsu, J. Garcia-Alfaro, "Exploring DWT-SVD-DCT feature parameters for robust multiple watermarking against JPEG and JPEG2000 compression," Computers and Electrical Engineering, **41**(8), 52-63, 2015, doi: 10.1016/j.compeleceng.2014.08.001.

[34] T.T. Takore, P.R. Kumar, G.L. Devi, "A modified blind image watermarking scheme based on DWT, DCT and SVD domain using GA to optimize robustness," in International Conference on Electrical, Electronics, and Optimization Techniques, 2016, doi:10.1109/ICEEOT.2016.7755190.

[35] A.K. Singh, "Improved hybrid algorithm for robust and imperceptible multiple watermarking using digital images," Multimedia Tools and Applications, **76(6)**,8881-8900, 2017, doi:10.1007/s11042-016-3514-z.

[36] B.J. Saha, Arun, K.K. Kabi, C. Pradhan, "Non-blind watermarking technique using enhanced one-time pad in DWT domain," in 5th International Conference on Computing Communication and Networking Technologies, ICCCNT 2014, 2014, doi:10.1109/ICCCNT.2014.6963061.

[37] N.M. Kumar, T. Manikandan, V. Sapthagirivasan, "Non-blind image watermarking based on similarity in contourlet domain," in International Conference on Recent Trends in Information Technology, ICRTIT 2011, 2011, doi:10.1109/ICRTIT.2011.5972280.

[38] G.H. Yann LeCun, Yoshua Bengio, "Deep learning,", Nature, **521**(7553)**,** 436–444, 2015.

[39] R. Namba, J. Sakuma, "Robust watermarking of neural network with exponential weighting," in Asia CCS 2019 - Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security, 2019, doi:10.1145/3321705.3329808.

[40] E. Le Merrer, P. Pérez, G. Trédan, "Adversarial frontier stitching for remote neural network watermarking," Neural Computing and Applications, **32**(13), 9233-9244, 2020, doi:10.1007/s00521-019-04434-z.

[41] B. Prashanth, "Design and Implementation of Radar Cross-Section Models on a Virtex-6 FPGA," Journal of Engineering (United Kingdom), 2014, doi:10.1155/2014/489765.

[42] L. Pang, S. Zhu, C.W. Ngo, "Deep Multimodal Learning for Affective Analysis and Retrieval," IEEE Transactions on Multimedia, **17**(11), 2008-2020, 2015, doi:10.1109/TMM.2015.2482228.

[43] B. Prashanth, P.A. Kumar, G. Sreenivasulu, " Design and Implementation of floating point ALU on a FPGA processor," in 2012 International Conference on Computing, Electronics and Electrical Technologies, ICCEET 2012, 2012, doi:10.1109/ICCEET.2012.6203790.

[44] B. Vijay Krishna, B. Venkata Prashanth, P. Sujatha, "Design and implementation of DPFC for multi-bus power system," International Journal of Engineering and Technology (UAE), **7**(2), 2018, doi:10.14419/ijet.v7i2.8.10318.

[45] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, doi:10.1109/CVPR.2015.7298594.

[46] S.R. Qasim, H. Mahmood, F. Shafait, " Rethinking table recognition using graph neural networks," in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2019, doi:10.1109/ICDAR.2019.00031.

[47] J. Li, P. He, J. Zhu, M.R. Lyu, "Software defect prediction via convolutional neural network," in Proceedings-IEEE International Conference on Software Quality, Reliability and Security, 2017, doi:10.1109/QRS.2017.42.

[48] Y. Uchida, Y. Nagai, S. Sakazawa, S. Satoh, "Embedding watermarks into deep neural networks," in ICMR 2017 - Proceedings of the 2017 ACM International Conference on Multimedia Retrieval, 2017, doi:10.1145/3078971.3078974.

[49] A. Van Den Oord, S. Dieleman, B. Schrauwen, "Deep content-based music recommendation," in Advances in Neural Information Processing Systems, **26(1)**, 2643-2651, 2013, doi:10.1145/2500098.2500109.

[50] A. Karatzoglou, B. Hidasi, "Deep learning for recommender systems," in RecSys 2017 - Proceedings of the 11th ACM Conference on Recommender Systems, 2017, doi:10.1145/3109859.3109933.

[51] A. Karpathy, L. Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, doi:10.1109/TPAMI.2016.2598339.

[52] H.O. Song, Y. Xiang, S. Jegelka, S. Savarese, "Deep Metric Learning via Lifted Structured Feature Embedding," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, doi:10.1109/CVPR.2016.434.

[53] B. Kim, H. Kim, K. Kim, S. Kim, J. Kim, "Learning not to learn: Training deep neural networks with biased data," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019, doi:10.1109/CVPR.2019.00922.

[54] L. Shao, D. Wu, X. Li, "Learning deep and wide: A spectral method for learning deep networks," IEEE Transactions on Neural Networks and Learning , **25**(12), 2303-2308, 2014, doi:10.1109/TNNLS.2014.2308519.

[55] A.B. Duque, L.L.J. Santos, D. Macêdo, C. Zanchettin, "Squeezed Very Deep Convolutional Neural Networks for Text Classification," in Lecture Notes in Computer Science, 2019, doi:10.1007/978-3-030-30487-4_16.

[56] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Computer Society Conference on Computer Vision Pattern Recognition, 2016, doi:10.1109/CVPR.2016.90.

[57] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," IEEE Transactions on Image Processing, **26**(7), 3142-3155, 2017, doi:10.1109/TIP.2017.2662206.

[58] B.D. Rouhani, H. Chen, F. Koushanfar, "Deep Signs: An End-to-End Watermarking Framework for Ownership Protection of Deep Neural Networks," in International Conference on Architectural Support for Programming Languages and Operating Systems - ASPLOS, 2019, doi:10.1145/3297858.3304051.

[59] A.I. Orhean, F. Pop, I. Raicu, "New scheduling approach using reinforcement learning for heterogeneous distributed systems," Journal of Parallel and Distributed Computing, **117**(1)**,** 292-302, 2018, doi:10.1016/j.jpdc.2017.05.001.

[60] Q. Guo, X. Wang, Y. Wu, Z. Yu, D. Liang, X. Hu, P. Luo, "Online knowledge distillation via collaborative learning," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020, doi:10.1109/CVPR42600.2020.01103.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1569-1576 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

ESP2: Embedded Smart Parking Prototype

Tarek Frikha[1,*], Hedi Choura[1], Najmeddine Abdennour[1], Oussama Ghorbel[1,2], Mohamed Abid[1]

[1]*CES-Laboratory, Sfax University, Sfax, 3029, Tunisia*

[2]*Computer Sciences Department, Jouf University, Skaka, 2014, Kingdom of Saudi Arabia*

ARTICLE INFO

Article history:
Received: 16 October, 2020
Accepted: 16 December, 2020
Online: 25 December, 2020

Keywords:
Embedded System
Smart City
Smart Parking
Wireless Sensor Network
Actuator
Sensor
Mobile application
Web Application

ABSTRACT

The technological evolution and the frequent use of the Internet as a technology within the framework of the Internet of Things and Robotics has affected various fields such as Industry 4.0, smart agriculture, smart cities, home automation and autonomous cars. These advances applied to everyday life have facilitated the implementation of new concepts to minimize pollution, traffic jams and accidents while making life easier. In this article, we propose an intelligent parking concept based on embedded systems. In order to solve the problem of permanent traffic jams and their repercussions on large cities we have implemented this system. It is a complete platform allowing remote reservation from different platforms. It will also collect the various data obtained by parking sensors. These data will be processed and saved if necessary. This concept was first developed to highlight the feasibility and adaptability of this system in the context of smart cities. It has been realized using wireless sensor networks connected to embedded platforms. These platforms are not only connected to motors, sensors and actuators but also to servers to save data and deploy reservations. These reservations are made via web and mobile application.

1. Introduction

The robotics automation, wireless sensors, artificial intelligence and embedded systems are famous scientific research fields with a wide margin for innovation and growth. Although these are areas that are literally different and divergent. However, they remain compatible. One can thus join these different axes in order to carry out applications that make life easier for mankind.

In this framework, we can mention several types of uses such as smart houses [1], smart agriculture [2] and smart parking [3], which are smart city [4] applications. We can also talk about e-health [5] and Industry 4.0 [6] which are applications in constant development.

In this work, we're going to touch on smart parking. The automated systems for vehicle parking management [7] are widely available in the recent years. However, it is still an exclusive product that is not largely used for a couple of reasons including being very cost expensive and complex. These difficulties also contribute to other problems such as hard maintenance and resource management issues [8]. Nowadays, android applications and Arduino based technologies [9] are the mainstream and the most popular solution in numerous domains due to being open sourced. This attribute makes them easy to implement and heavily supported with the majority of hardware and software. For these reasons we used android based application for our embedded smart parking and management system that outperformed the other available options in resource optimization and performance.

In our previous work [10] we presented a fully embedded architecture for an automated intelligent parking system that communicates through an Android application, which delivered an intelligent parking management and a top-notch costumer service for smart parking.

In this paper, we describe in details our previous work [10] while we expand on the practical implementation and the communication interface. With the objective of having a functional prototype that develops the research work and can be integrated in the field of smart cities, we will present two platforms:

- an embedded platform that has the role of data acquisition and processing in our parking lot. The use and architecture will be described later on.
- a client/server platform with Android and Web applications: these will describe the reservation, confirmation and payment parts of the proposed service.

*Corresponding Author: Tarek Frikha, tarek.frikha@enis.tn

Our paper will be presented as follows. We start by an overview on the smart parking state of the art. In the theoretical analysis section, we will introduce the architectural design of the car parks on one hand and the developed computer application to be realized on the other. Then, we present the practical implementation section where we demonstrate the developed applications as well as the implemented embedded architecture. Finally, we compare between the implemented Intelligent Parking results before we declare the conclusion and perspectives.

2. State of Art

Several works have presented smart parking. In [11], they present a flexible and real time information system. This parking is easy to use for the client via web connexion. A CMOS cam is used to control the parking. The disadvantages of this system are the need to inquire for the available slot, the high implementation cost. The most important disadvantage is the bottlenecks due to an overload of incoming informations.

In [12], they use RFID technology. It permits to have not only secured but also an easy and fast checking out of the vehicles in less time. The data security is provided by ZigBee Wireless Technology. The data of the arrival and departure of the vehicle are stored in database. To inform the driver informations are displayed on the LCD proposed screens in the parking. Despite all the advantages, the system has many disadvantages. It allows one by one. That's why it consumes a lot of time and prevent multiple check out and check in at the same time. The drive can't make a reservation via mobile device. Another disadvantage of RFID is the increase of cost and of the node to node implementation time.

In [13], if the most important disadvantage is that the parking does not alarm the user about the car theft, the proposed smart parking use ultrasonic sensors and RFID. Needing for deployment ZigBee and 3G (or 4G, 5G…), the most important advantages of the proposed work are the use of the Car Parking Framework that gives two performance classes in term of financial and consumption cost deployment. In addition, the heterogeneous wireless sensor network cause energy fluctuations. To remedy they use LIBP gathering protocol.

In [14], they use Zigbee for deployment and ferromagnetic sensor. Suryady and al introduced another concept of smart parking based on cloud-based platforms as a service (PaaS). This technology gives the smart parking an interesting quality. The smart-parking system is faster to use. It reduces time, cost and system maintenance. The platform has the advantage of being flexible to collaborate with other platforms. This system has also disadvantaged. It does not provide neither guidance for the drive nor payment due to the application insecurity to protect user privacy. The user has no possibility to display spaces for user on smartphone or parking LCD.

In [15], the smart parking is more complete. It has many qualities like his flexibility and security increase. This security appears for the user because of less of human interaction for booking slots. The drive can find the availability of parking space and then books slots. The IoT's devices provide data storage. The data are processed and collected to be saved. The drive has another security aspect to preserve the user from steals by theft management.

As we already declared, there are numerous automated parking management systems, yet only three methods which are Park Assist [16], Hectronic [17] and Parking plus [18] that stand out and achieve respectable results. These intelligent car parks are available on the market but have some missing features that need to be addressed.

2.1. Park Assist smart parking system

the TKH group developed this smart parking system that have various traits including guiding the client to the empty parking space while providing access to a mobile application with the vehicle geolocation and the reliance on machine learning analysis for the parking spaces to insure the right exploitation.

2.2. Hectronic intelligent parking

Contrary to Park assist, this intelligent parking system delivered by the German international company Hectronic has a separate entrance and exit terminals. The main feature of this system is their powerful and well built security alongside other adequately constructed services.

2.3. Parking Plus mobile application

This parking system is based on a mobile application that delivers a central parking service by gathering information from a cluster of parking zones to provide the driver with the details about the nearest empty parking place and the possibility to reserve it and pay for it by the phone.

When we analyze these various methods, we found that some of them have security shortcomings both in capturing information for reservation and payment or in securing vehicles in case of theft. Therefore, in this work we present an intelligent car park allowing the capture of information on the different vehicles when they enter, exit and during their presence in our car park. An offline backup is made throughout the day before being transmitted to the cloud in order to maintain a trace if necessary.

Reservation, payment and placement of the car in the car park are done in a secure way via not only web but also mobile applications compatible on different types of terminals.

Given the lack of space proposed for parking lots and the permanent increase in the number of cars in large cities generally and particularly in Tunisia, our goal is to set up an intelligent system to give the user the choice, if possible, to choose his parking space, if not to allocate an automatic space to optimize the space.

Indeed, thanks to the automatic parking system chosen, we will have the possibility:

- to secure the parking lot by minimizing human intervention: cars will be closed and parked at the entrance of the parking lot. Their placement will be carried out automatically afterwards.
- to optimize space and minimize accidents: the cars will be parked automatically and therefore no space or accidents due to bad maneuvers will occur.
- to minimize pollution and traffic jams: by allowing the reservation system to not search for parking spaces in a

random way and therefore optimize fuel consumption and minimize pollution.

The innovation proposed in our paper is to create a fully automated parking system that solves the limitations described above.

This design will present not only the architecture of the realized prototype but also the technologies based on the IoT to allow the automatic placement of cars in the proposed concept.

There will be a design part, an IoT part and a hardware implementation part which will be described in parts 3 and 4.

3. Theoretical Analysis

Looking at the systems available on the market, we can initially assume that depending on their points of strength we can choose from them a tolerable approach [19]. However, these methods have many apparent concerns such as the inefficient space exploitation with Park Assist and Hectronic by the need for a driving pavement to connect the vehicle to the driving spot and the capacity problem caused by the lack of long term parking strategy [20]. We can also point out the high service cost and the absence of security plan for Parking Plus. In this section we will describe the different axis of our own parking architecture that fix these issues.

We begin with our proposed parking architecture for car parks that is set up in narrow plots of land. We present a cylinder shape architecture as shown in figure 1. The second part of this theoretical analysis introduce the proposed embedded system to collect, process and update the information before transmission step. The last part describes the frontend and the backend of the application architecture. It will describe different actors and secured reservation and payment service to be used by the driver.

3.1. Smart parking architecture

3.1.1. Proposed smart parking architecture

In order to solve the parking management concerns that the previous methods failed to answer, we conceptualized a structure with a cylinder-shaped architecture for vehicle parking as revealed in figure 1.

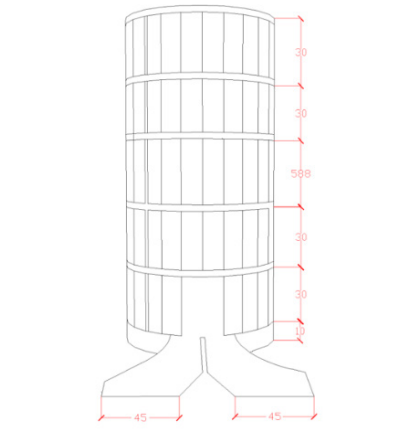


Figure 1: The front face display of the cylinder-shaped architecture for our parking

We choose using this architectural shape in an attempt to reduce the space spent on pavement allowing the vehicle movement within the structure and switching to a unique vehicle area, which is the mechanical component that act as an elevator and the selector for the parking place simultaneously.

3.1.2. Detailed units of the architecture

To simplify and optimize the driver's access to the proposed architecture, the parking place selection is done via the android mobile application. Afterwards, the driver has to station his vehicle on the entrance elevator panel that has the same size as standard parking place as shown in figure 2.

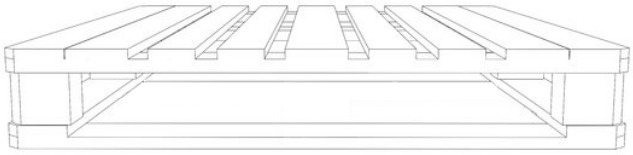


Figure 2: A side view of the entrance elevator panel

Having reserved a parking space via web or mobile application, the customer will provide the information for the car registration number. This number will be the reference to allow the car to enter the parking place.

Once the vehicle is checked to the entrance panel, a two-armed cart raise the panel as an elevator. Figure 3 present an example of this activity.

Afterwards, the vehicle will be directed directly to the parking space relying on the elevator's ability to rotate 360 degrees. This feature allows the system to set the vehicle in the appropriate location selected in advance by the customer. The exit process is also similar to the entry procedure but backwards. The customer has to do his check out and pay for the service via the mobile application and once he does the elevator selects the vehicle corresponding to the car registration number and bring it down to the exit of the parking lot.

The full automation of elevator with the minimization of manpower and human intervention in the parking process will make the process secure by avoiding human errors and effective by efficient exploitation of the parking facility.

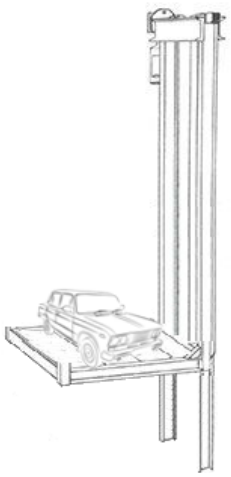


Figure 3: A sketch example of a vehicle elevator

3.2. Hardware embedded architecture

Having chosen a suitable architectural form for our intelligent car park, we will present the embedded implementation of our system in this section. It will, in fact, consist of three parts. Firstly, we have chosen a set of sensors and actuators that will be installed in order to obtain information about the parking spaces if they are available or occupied. Second, a system will collect this information from the installed sensors to update and save the occupancy in the parking lot. This data will be saved locally and subsequently updated in the cloud. In part C, we will present the software backup platform implemented.

3.2.1. Used set of sensors and actuators (SSA)

As described in the state-of-the-art section, our intelligent car park will not only make the driver's life easier by offering him an available space but will also provide him with a secure space for both him and his vehicle.

Thus, a surveillance camera at the entrance but also on each floor of the car park will detect the license plate of the car in order to validate its access to the car park but also to the reserved space.

As soon as the car reaches its space a sensor will transfer the information and cancel the space for any other reservation. At the exit of the car park, the information is transmitted directly to the driver and the space will be directly released for another reservation.

The billing for this operation will be proportional to the time occupied in our car park.

To sum up, we'll need:

- Surveillance cameras at the entrance and exit of the car park but also at the entrance and exit of each floor of our car park.

- Barrier level detection sensors to validate entry and exit to the car park.

- Infrared sensors to transmit information on the presence and absence of the car from the reserved space.

In parallel with the sensors, the need for actuators is equally important. Indeed, these will allow us to manage the different engines of our car park remotely. We can therefore mention the following:

- The motors of the car park barriers
- The engines of the car lift
- Vehicle placement jacks as shown in Figure 3.

Due to the importance of precision in the realization of our system. Different electrical voltage values must be received from the sensors and sent to the different motors. The accuracy of these data requires a system that can directly manipulate the analog data by minimizing the use of digital-to-analog and analog-to-digital converters. This will be detailed in the section sensor and actuators control and collection platform.

3.2.2. Sensor and actuators control and collection platform (SAC2P)

In order to control the various sensors and actuators present in our system, we will use an on-board platform that can receive and transmit analog data. In the prototype that will be described in section IV, part B, we will use an embedded Arduino platform that will allow us to validate the proposed approach.

This platform will be connected via Bluetooth to both actuators and sensors. It will thus control entry to the car park, the placement of the car and the sending of information to the information backup and transfer platform. Figure 4 describes sensor and actuators control and collection platform.

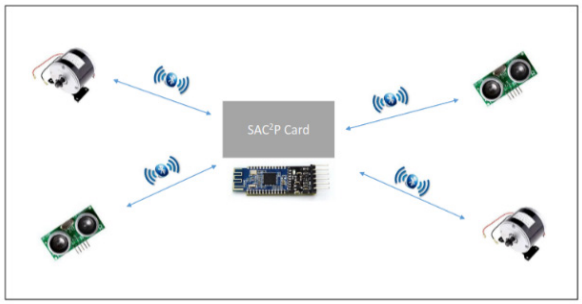


Figure 4: Sensor and actuators control and collection platform

3.2.3. Information backup and transfer platform (IBTP)

The Information backup and transfer platform part offers a digital embedded platform.

Several types of platforms can be proposed such as GPUs, FPGAs, ASICs. This platform will have four main tasks with the following priorities:

a) Acquisition and processing of information from surveillance cameras.

b) The administration of the Sensor and actuators control and collection platform via Bluetooth.

c) The offline backup and management of the various data recovered.

d) Updating information and saving it on the cloud server via wifi or 4G

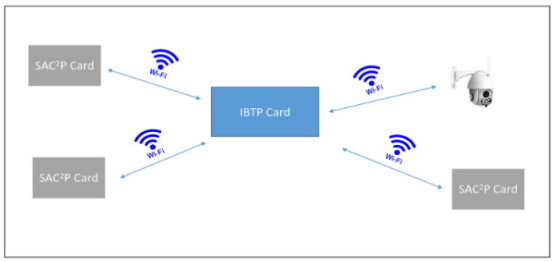


Figure 5: Information backup and transfer platform

Since, for our application, obtaining a real-time result is not paramount, it is not necessary to use GPUs that are quite expensive in terms of hardware and energy costs. Thus, a simple multiprocessor architecture based on 4 processors working in parallel is largely enough.

For our application, we decided to use a Raspberry Pi as a platform to meet the needs of our Smart Parking prototype. In part 4, we will describe in more detail the system implemented.

In figure 5, there is a description of information backup and transfer platform.

3.3. Software reservation platform architecture

After having described and set up the architecture part 3.1 and the hardware part 3.2, we will highlight in part 3.3 the software part and particularly the online booking platform.

In the reservation part we will have several interfaces.

Firstly, the administrator will access the different functionalities via the web. He will be able to modify the terms of payment, accept complaints, validate possible subscriptions from members and deal with disputes.

It is the administrator who will be able to manipulate the number of spaces, the availability as well as the occupation of the car park, the rates per hourly session, penalties for late payment.

The customer will access the online booking platform either via mobile application or via web.

Each customer must first register on the system either via mobile application or via the web. He will indicate his personal data (surname, first name, telephone number, e-mail, license plate, make and colour of the car, ...). He will then choose the method of payment either by credit card or in cash.

As a nominal scenario, a customer will be able to reserve a place on one or the other of the applications. He will indicate the duration of his parking period. At the time of reservation, the user will indicate the method of payment. At 15 minutes from the end of the indicated parking period, a SMS and an e-mail will be sent to inform the customer. He will be able to extend his reservation by reply or maintain the duration indicated at the time of reservation.

On leaving the car park, the parking space is released and is therefore open for a new reservation.

If the customer leaves the car park on time and has paid by credit card, they will leave the car park directly: the camera at the car park entrance will identify the car and open the entrance barrier. If he is going to pay in cash, a machine will take the coins or notes and give him the rest of the amount.

If the customer is late without replying to the SMS received, he will have to pay a penalty. The amount charged for leaving the car park will be the equivalent of a full day's parking.

If the customer is late and this delay has been indicated by message, the customer will pay the difference when leaving the car park.

4. Practical implementation

In this part we will present the parts describing the implementation and setup of our demonstrator. We will start with the modeling part including the class diagram and the use case diagram.

Then we propose the mobile SW part (Arduino application) and finally the implementation of the embedded part implemented on an Arduino platform.

4.1. Uses case diagram

In this section we will present the actors of our platform. The administrator, the client as a physical actor and the system as a logical actor.

0As shown in Figure 6, the administrator has the possibility to access to the different functionalities via the web. He will be able to accept claims, check the number of seats and set rates per hourly session.

The administrator can also be able to manipulate the penalties for late payment, validate the possible subscription of the member and change the terms of payment, etc.

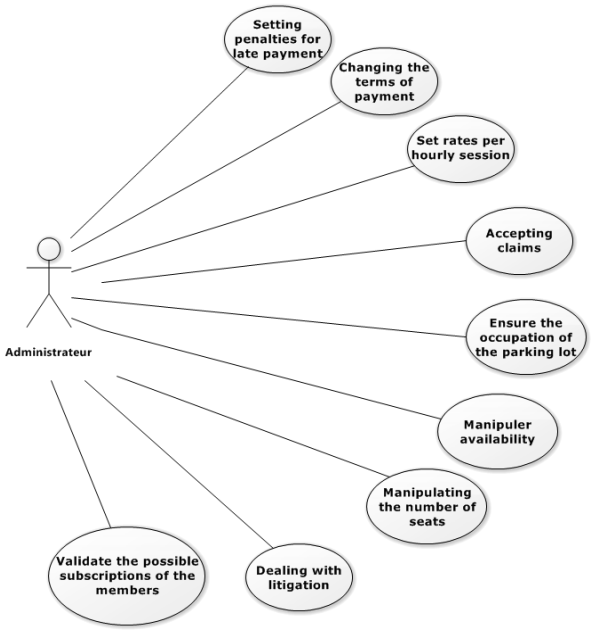


Figure 6: Admin Use Class Diagram

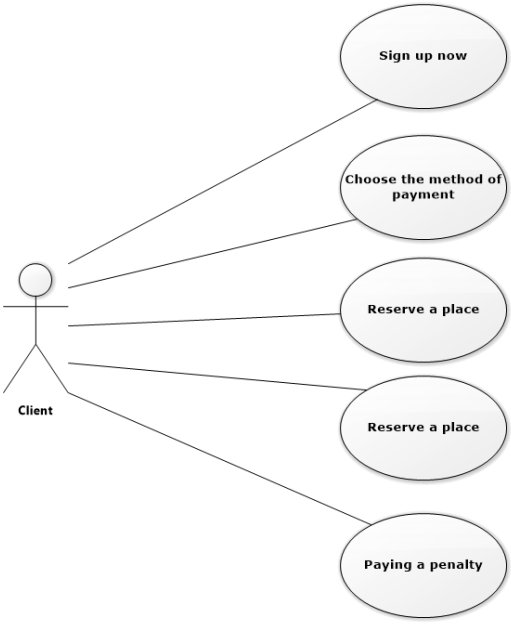


Figure 7: Client Use Class Diagram

The client, as mentioned in Figure 7, must start by registering on the system either via mobile application or via the web. He I indicate his personal data such as surname, first name, phone number, e-mail, license plate, make and color of the car, etc. After that, he chooses the method of payment either by credit card or cash. He will be able to reserve a place by indicate the duration of his parking period.

At the time of his reservation, the client will indicate the method of payment. He will be able to extend his reservation by reply or maintain the duration indicated at the time of reservation.

The last actor is the system, as represented in Figure 8, it must be able to identity the car, take coins or bills, remit the rest of the amount, open the entry barrier and before 15 minutes of the end of the period subscription, it sends a SMS to client to inform him; if he wants to extend the period or he free the space.

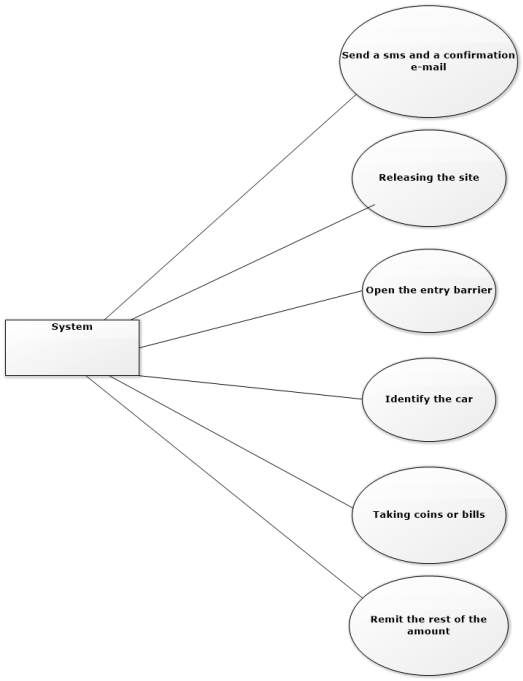


Figure 8: System Use Cas Diagram

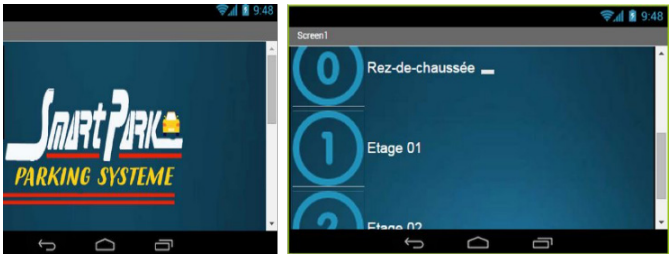


Figure 9: Our smart park application main menu

4.2. Android application

The Android operating system is by far the largest used operating system for smartphones in the world for the recent years. To communicate with our automated smart system, we developed an Android application to exploit the ease of implementation, compatibility and support [21]. We used the App Inventor development tool generate by Google to create our Android web application for smartphones [22]. Our application contains a main menu for the user to select the desired parking floor from the parking structure as shown in figure 9.

From this menu the driver first selects the floor and then searches for the empty parking location. Finally, the application creates a security code for the specific location to be used by the user when he exits the parking facility. It is also possible for the user to let the parking system to choose automatically which place to allocate.

4.3. Arduino technology embedded system:

To control the various parts of our system we choose to work with a simple Arduino board [23] and we have decided on the Arduino mega board. This embedded bored is responsible for the communication with the client commands coming from the Android application. It is also responsible for checking and monitoring the sensors that provide the information about empty locations and vehicle presence, while managing the mechanical elements of the elevator and the rotating panel.

For the communication part, the Arduino mega board needs the Arduino WIFI shield in order to connect to the network via WIFI [24]. As displayed in figure 10, this wifi shield is pin by pin compatible with the Arduino mega board, which makes it easy to work with.

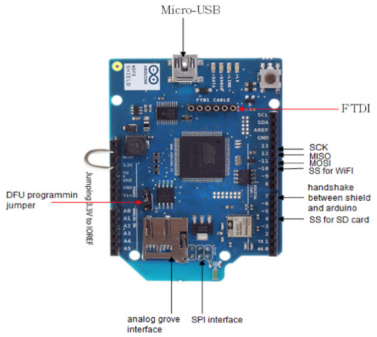


Figure 10: The Arduino WIFI shield

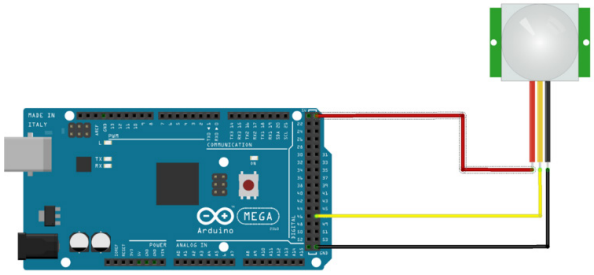


Figure 11: Arduino mega board and the PIR sensor wiring diagram

In order to collect the information about the availability of the parking locations and the elevator panel a cluster of spread out sensors is crucial for our system. Therefore, we decided to lay out a collection of the PIR (Passive Infrared Sensor) in every parking location, the entrance, the exit and in the elevator panel. As shown in figure 11, connecting the PIR sensor to our Arduino mega board is a very simple operation that demands the use of only three pins. We also considered using the HC-SR04 ultrasonic sensor for the same operation mainly for compatibility reasons since it is also an

Arduino technologies component. However, considering the low expensive PIR sensor was the optimal choice since it does not compromise the proficiency of our system [25].

To manage the mechanical part of our system, which are the elevator panel motor and the rotation motor, we also relied on our Arduino mega board. For our prototype , we sufficed with the use of simple direct current motor driver to emulate a real life situation where our Arduino mega board can also control the motor of the elevator using two relays and terminal levers for the 230V alternative current motor [26]. Our choice for a simple DC motor driver is the L298N circuit that grants our Arduino board full authority over the speed and the rotation direction of the direct current motor [27]. For the panel rotation motor, we used a servomotor connected to the Arduino mega. The servomotor delivers an immense management and control to every detail of the panel rotation and angular rotation manipulation [28]. With the use of the servomotor, we guarantee that our smart parking system is capable to reach all parking positions throughout the facility thanks to the 360 degrees of freedom available to the rotating panel. In figure 12 we present the wiring diagram for both the DC motor and the servomotor.

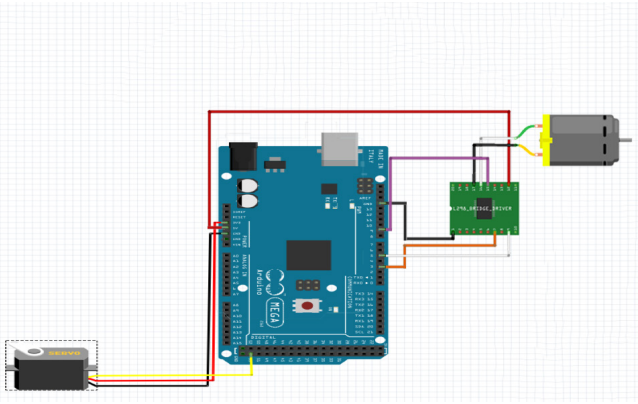


Figure 12: The Arduino mega board wiring diagram with the L298N motor driver, the DC motor and the servo motor.

4.4. Result comparison

Comparing the system we have set up with the different intelligent parking lots existing in the literature we notice that our system has several points in common with them but has advantages in terms of space optimization and automation of the system.

Table 1 illustrates this comparison.

Table 1: ESP2 parking comparison

	Mobile App	Automatic	Secured	Smart
Park Assist	Yes	No	*	No
Hectronic	Yes	No	***	No
Parking +	Yes	No	***	Yes
ESP2	Yes	Yes	*****	Yes

As shown in Table 1, it can be seen that all the systems described in the literature offer a mobile application for booking.

However, the car parking requires the intervention of the man (driver or parking manager) to park the car unlike our system which is fully automated based on sensors, IoT and embedded systems.

The minimization of human intervention makes it possible to secure the system against theft, minimize human errors and save energy consumption.

Finally, as Parking+, our system has an intelligent part to update the status of the parking lot and therefore make it easier to park the car while minimizing the waiting time for parking.

5. Conclusion

Throughout this paper we introduced a smart vehicle parking system that effectively manage the storage and automates vehicle parking using an Android application as the front-end user communication platform with our embedded control system. Our embedded control unit includes the main Arduino mega embedded board and the Arduino WIFI shield that ensures WIFI connection. The information for the presence of vehicles and the availability of empty parking location is acquired by the collection of PIR sensors and provided to the Arduino mega board to take the appropriate decision. This main control unit is also responsible for the management and control of the elevator panel and the rotation engine in charge of accessing the empty locations and stored vehicles.

Our novel smart parking system with the Android customer communication unit has delivered a better storage management capability than the previously mentioned commercially available competitors. The dominance of our developed system over Park Assist and Hectronic parking system is clearly manifested in the long term cost gained from the smart management of the allocated space that our competitors sacrifice in order to provide the driving pavement needed to connect the vehicle to the driving spot. The Parking Plus application also miss the security quality that our solution provides compared to their nonexistent answer to this problem.

Although our system is a great solution for smart parking, we do not deny the fact that it has a few shortcomings such as the need for more flexibility when it comes to the architectural structure of the facility and the power management of the mechanical parts (elevator panel and rotation engine). These limitations can be the subject of future research that can include introducing a sustainable energy resources and other architectural designs. we can also ensure a better security by implementing a decentralized Blockchain based security control system in case of an expansion to a multiple units of smart parking systems.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. Li, W. Gu, W. Chen, Y. He, Yannian Wu, Y. Zhang, “Smart Home: Architecture, Technologies and Systems”, Procedia Computer Science, 131, 393-400, ISSN 1877-0509, 2018, doi:10.1016/j.procs.2018.04.219.

[2] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, P. Nillaor, "IoT and agriculture data analysis for smart farm", "Computers and Electronics in Agriculture", 156, 467-474, ISSN 0168-1699, 2019, doi: 10.1016/j.compag.2018.12.011.

[3] F. Al-Turjman, A. Malekloo, "Smart parking in IoT-enabled cities: A survey", Sustainable Cities and Society, 49, 101608, ISSN 2210-6707, 2019

[4] A. Mathijssen and A. J. Pretorius, "Specification, Analysis and Verification of an Automated Parking Garage",. 40, 2005.

[5] K. Ghidouche Aït-Yahia, F. Ghidouche, G. N'Goala, "Smart city of Algiers: defining its context", Leonidas Anthopoulos, Smart City Emergence, Elsevier, 391-405, 2019. https://doi.org/10.1016/B978-0-12-816169-2.00019-5

[6] I. Chiuchisan and O. Geman, "Trends in embedded systems for e-Health and biomedical applications," 2016 International Conference and Exposition on Electrical and Power Engineering, 304-308, 2016, doi: 10.1109/ICEPE.2016.7781352

[7] R. Y. Zhong, X. Xu, E. Klotz, S. T. Newman, "Intelligent Manufacturing in the Context of Industry 4.0: A Review," Engineering, **3**(5), 616-630, ISSN 2095-8099, 2017, doi: 10.1016/J.ENG.2017.05.015.

[8] Howard, K. Charles, K. Cayetano, and O. Omojola. "Parking management systems." U.S. Patent No. 6,885,311. 26 Apr. 2005.

[9] A. D'Ausilio "Arduino: A low-cost multipurpose lab equipment." Behavior research methods, **44**(2), 305-313, 2012. doi:10.1016/2012.04.219.

[10] https://www.parkassist.com

[11] A. Najmeddine and T. Frikha, "A fully automated intelligent system for vehicle storage and parking management using android and arduino hardware and applications," 2019 IEEE International Conference on Design & Test of Integrated Micro & Nano-Systems (DTS), Gammarth-Tunis, Tunisia, 2019, 1-4, doi: 10.1109/DTSS.2019.8915278.

[12] P. Dharma Reddy, A. Rajeshwar Rao, A.S Musthak. "An Intelligent Parking Guidance and Information System by using image processing technique". International Journal of Advanced Research in Computer and Communication Engineering. **2**(10), 2013.

[13] M. Patil, V.N Bhonge, "Wireless sensor network and RFID for smart parking system". International Journal of Emerging Technology and Advanced Engineering, **3**(4), 188-192, 2013

[14] E. Karbab, D. Djenouri, S. Boulkaboul, & A. Bagula, "Car park management with networked wireless sensors and active RFID" In 2015 IEEE International Conference on Electro/Information Technology (EIT), 373-378, 2015.

[15] Z. Suryady, G.R Sinniah, S. Haseeb, M.T Siddique, M.F.M Ezani, "Rapid development of smart parking system with cloud-based platforms." In Information and Communication Technology for The Muslim World (ICT4M), 2014, 1-6

[16] B.M.K Gandhi, M.K Rao. "A Prototype for IoT based Car Parking Management system for Smart cities", Indian Journal of Science and Technology, **9**(17), 2016.

[17] https://hectronic.com/en/solutions/parking/

[18] https://www.parkingplus.ro

[19] T. Miroslav, S. Gutta, and V. Philomin. "Smart parking advisor." U.S. Patent No. 6, 426, 708. 30 Jul. 2002.

[20] J. Yatin, & al. "Understanding smart and automated parking technology." Int. J. u-e-Serv. Sci. Technol, **8**, 251-262, 2015.

[21] M. Goadrich, M. P. Rogers. "Smart smartphone development: iOS versus Android." Proceedings of the 42nd ACM technical symposium on Computer science education. ACM, 2011.

[22] D. Wolber, &al. App Inventor. " O'Reilly Media, Inc.", 2011.

[23] Arduino, S. A. "Arduino." Arduino LLC 2015.

[24] A. Javed, "Building Arduino Projects for the Internet of Things." Experiments with Real-World Applications." United States of America: Apress Media, LLC,177- 193, 2016.

[25] J. Xin, et al. "Target detection and classification using seismic and PIR sensors." IEEE Sensors Journal **12**(6), 1709-1718, 2012.

[26] https://www.instructables.com/id/3-floor-elevator-arduino/

[27] J.D Warren, J. Adams, and H. Molle. "Arduino for robotics." Arduino robotics. Apress, Berkeley, CA, 51-82, 2011.

[28] J. Blum, "Exploring Arduino: tools and techniques for engineering wizardry" John Wiley & Sons, 80-85, 2013

[29] http://www.ee.ic.ac.uk/pcheung/teaching/DE1_EE/stores/sg90_datasheet.pdf

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 6, 1577-1592 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Framework for Adoption and Diffusion of Mobile Applications in Africa

Chinedu Wilfred Okonkwo[*,1,2], Magda Huisman[2], Estelle Taylor[2]

[1]*School of Consumer Intelligence and Information Systems, University of Johannesburg, Johannesburg, 2001, South Africa*

[2]*School of Computer, Statistical and Mathematical Sciences, North-West University, Potchefstroom Campus, 2520, South Africa*

ARTICLE INFO

Article history:
Received: 11 October, 2020
Accepted: 28 November, 2020
Online: 25 December, 2020

Keywords:
Mobile Applications
Adoption
Diffusion
Framework
Africa

ABSTRACT

The adoption and diffusion of mobile applications (mobile apps) has become the base of modern activities in Africa owing to the services and values that are obtained through mobile apps innovations. More emphasis has been on the development of mobile apps whereas the adoption and diffusion process as well as their predictors are ignored or given less attention. This study explores and develops a framework for the adoption and diffusion of mobile apps in Africa. A survey was conducted on the basis of the diffusion of innovation framework in five-selected African nations. A total of 1285 of the 2300 distributed questionnaires were returned, giving a response rate of 55.87%. The results indicated that many factors/predictors drive the acceptance and use of mobile apps and these factors were structured to develop a common framework for mobile apps adoption and diffusion in Africa.

1 Introduction

The advancement in mobile technology has had a major impact on the manner in which business and personal activities are performed. It could be said, in particular, that a radical evolution started with the advent of handheld devices such as mobile phones that caused the creation of distinct kinds of mobile apps that are used in separate activities [1, 2]. A number of mobile technology start-ups have been created in Africa to encourage mobile app innovation [2, 3]. Advances in digital technology have enhanced communication systems in Africa, leading to a growing amount of African communities linked to digital communications [4, 5]. The use of mobile apps are changing the African society, socially, economically and politically [1], [6, 7]. Despite the penetration and growth of mobile services, some mobile apps are rejected and discontinued to be used. Moreover, there are some non-users of mobile apps in Africa [8]. The process of acquiring and disseminating mobile app services in Africa must therefore be improved, particularly for those with local content.

It is imperative to information and communication technology (ICT) experts and scholars to know the flow of information technologies [3]. This can be achieved by conducting a research on adoption and diffusion, which implies accepting to use [2, 9, 10] and continuing the use of a new innovation [11, 12]. Various models have been proposed to aid this process including frameworks for mobile health [13], mobile banking [14] and e-government [15]. These models dealt on a particular mobile app in a specific country and cannot be generalised.

However, African mobile consumers are growing interest in home-grown mobile apps. Several mobile apps have originated in Africa, especially in South Africa, Nigeria and Kenya including social networking mobile apps like Mxit, 2go and How Far, entertainment mobile apps like Simfy Africa, Spinlet, iRokoTV and SOLO, and financial mobile apps like M-pesa [2]. These mobile apps are mostly used in the country where they are created whereas in the other countries of the continent, there is little or no adoption of the mobile apps. In other words, there is a significant difference between the rate of adoption of each respective mobile app between the country where it is created and the other countries in Africa.

This study proposes a general framework for mobile apps adoption and diffusion in African social system. Drawing from the literature on the adoption of mobile app innovation and deliveries, this research intends to address the following questions:

1. What are the factors that influences mobile apps adoption in

[*]Corresponding Author: Chinedu Wilfred Okonkwo, School of Consumer Intelligence and Information Systems, University of Johannesburg, Johannesburg, 2001, South Africa., +27 83 579 6557 & Email: chinedung@yahoo.com

Africa and

2. How can we structure these factors to form a general model to enhance the acceptance and use of mobile app innovation in Africa.

In order to achieve these goals, we conducted a survey in the five countries representing the five geographical locations in Africa to assess the user's view of mobile app use in order to explore the factors predicting the acceptance and use of mobile apps in Africa. These factors were structured to develop a framework. With leverage on the report that there is a need to know the diffusion of mobile technologies and the important of mobile technological services on developing nations such as African countries, this study would therefore make the following contributions. We have:

1. Examined and discovered the acceptance and use of mobile apps in Africa,
2. determined the factors/predictors that influence users' intention to adopt and use mobile apps innovations in Africa, and
3. developed a general framework that can improve the adoption and use of different kinds of mobile apps across African nations.

The remainder of this research work is organised as follows: Section 2 discusses the background information of the research work. Section 3 describes the related works and the gap. Section 4 shows the research conceptual model and Section 5 explains the research methodology. Section 6 describes the methods of analysis and Section 7 zooms into the survey results. Section 8 presents the regression analysis and Section 9 discusses analysis of variance (ANOVA) results. Section 10 interprets the results to validate the proposed framework. Section 11 concludes the study.

2 Background information

A brief theoretical overview of the study is presented in this part.

2.1 What are mobile apps?

Mobile app creativity has recently attracted scholars from different fields of study to carry out studies in their respective fields. As a consequence, mobile apps have been described by several writers, write-ups and web pages in many perspectives including Technology [16, 17], Educational [18], Context-aware Computing [19], Internet of Things [20], Business [21], Cloud Computing [22] and Social Services [23]. In this study, mobile apps are considered to be mobile technological innovations of a mobile computer software that allows users to carry out various activities including political, social, economic, and educational regardless of the user's location and background.

2.2 Features of mobile apps

There are five main characteristics of mobile apps, according to Rogers [9]. They are relative advantage,compatibility, complexity, observability and trialability. Such characteristics help to decrease the level of misunderstanding and increase the adoption rate of mobile apps. Technology reliability and cost are conceptualized in this study as important features of mobile apps [8].

2.3 Innovation decision process

This is a process of decision-making which includes an information-seeking and information-processing practices in which mobile app adopters are driven to reduce the confusion surrounding the effect of acceptance and use of a new mobile app innovation. Rogers's diffusion framework consists of five stages of the innovation decision process including knowledge, persuasion, decision-making, implementation and confirmation [9]. For this study implementation and confirmation are merged because they describe diffusion (putting into use and continue to use) and referred to as the diffusion stage. Thus, the innovation decision process of this research entails four phases including knowledge, persuasion, decision-making and diffusion. These stages are briefly discussed below.

Knowledge stage:The mobile apps adoption and diffusion begins with the knowledge about mobile apps. This stage involves acquisition, method and sharing of information about mobile apps. Having the right knowledge about mobile apps requires information sources, learning methods and information sharing. In Africa, there are two distinct sources of information about mobile apps including external initiated actions and personal initiated actions, learning methods are interpersonal communications, the Internet, operational manual and informal training, whereas information sharing involves organizations and events, and individual activities [8].

Persuasion stag: is shaping the minds of adopters towards adopting and using mobile apps. The main factor under this stage is the social influence [9]. Decision-making stage: This includes the adoption and rejection of mobile apps. In this stage, an individual/adopter of mobile apps chooses to adopt or reject the mobile apps. The adoption covers the adopter's category whereas rejection covers the reasons for non-adoption [9]. There are five categories of adopters in Africa including late majority and laggards, non-adopters, innovators, early adopters and early majority [8]. Also, the reasons for non-adoption of mobile apps in Africa are using mobile apps is time-consuming, using mobile apps takes a lot of phone memory, using mobile apps distracts attention from other activities and using mobile is expensive [8]. Diffusion stage: This involves the implementation and confirmation. That is putting mobile apps into use and continuing to use them. It covers the use of mobile apps, ways of using mobile apps and reasons for discontinuance of the use of mobile apps. There are two distinct ways of using mobile apps in Africa including business use and individual use, and the reasons for discontinuance include a better replacement, unsatisfactory performance, High complexity and lack of support [8].

2.4 Adopting Unit

The adopting unit consists of features of individual adoption, comprising personality variables, socioeconomic and communication behavior. [9].

Personality variables: Personality variables entail the individual characteristics (demographic) that are part of the factors influenc-

ing the outcome of the innovation decision process [9]. Adopters' attitudes, anxiety (degree of nervousness) and self-efficacy are conceptualized as part the personality variables. Attitude is measured by the feelings and reactions of the participants on the importance and contributions of mobile apps in Africa [24].

Socioeconomic characteristics: Socioeconomic characteristics involve the external factors that may influence individuals' adoption and diffusion of mobile apps. These characteristics include the contributions of mobile apps to the development of Africa [25], facilitating conditions and employment status[8] .

Communication behaviour: This covers the primary purpose of using and disseminating the information about a newly developed mobile app [9]. It is one of the primary goals that encourages a person to embrace and disseminate mobile apps. In this study, It deals with the function of information sharing as specified in the abstraction model of the research. This research deals with two key reasons of spreading the information about a newly created mobile app including: [8].

1. to facilitates both individual and business activities, and
2. to improve and make communications easier.

2.5 Adoption

According to Rogers [9], adoption is the absolute willingness to embrace and use new advanced technologies (mobile apps). GSMA [2] explained that adoption is embracing and using a new technology. It is the acceptance [10] and use [11] of an emerging product such as mobile apps by the consumer. Innovation is new when it is first introduced to a person. Therefore, we describe the adoption as the users' acceptance and use of a new invention.

2.6 Diffusion

Diffusion is a means of spreading the information and use of an innovation among the individuals of a social structure across certain networks [9, 3]. It is a mechanism by which a commodity is spread from one customer to another [11]. For this research, with regards to Rogers framework, we combine the implementation and confirmation phases to form diffusion. Thus, we defined diffusion as the method of sharing, uses and discontinued the uses of a product.

2.7 Usage of mobile applications in Africa

GSMA [26] recorded that, there are more than five billion mobile subscribers across the globe and the number is expected to rise to 5.9 billion by the year 2025. The spread of mobile technology is wider than any other technology [27]. Mobile service penetration is expanding at an extraordinary rate in Africa, and mobile apps are part of the main foundations of all mobile services. The introduction of mobile services in Africa is rapidly growing such that the mobile app is the driving force behind some of the emerging developments in African nations.[16, 7]. The African continent currently ranks among the world's largest adopters of mobile technology services, introducing new mobile services to the population [28, 29].

The beginning of mobile app development can be linked to the program required to operate the emerged mobile devices [30]. Over the years, telecommunications in Africa was primarily through a fixed-line system and the networks were regulated by the respective countries' central governments [31, 32]. In the late 1990s, the advent of mobile telecommunications brought significant changes to the African continent, so that countries like Nigeria, Cote d'Ivoire, and Uganda had more mobile phones than fixed lines [31]. In Africa, the mobile telecommunications system has a faster and broader penetration and deployment than fixed lines, and some nations, such as the South Africa, Comoros, Uganda, Kenya, Mauritius, Malawi, and the Seychelles, have increased their rural mobile network coverage as of 2010 [33]. The advancement of mobile technology has contributed to the creation, acceptance and dissemination of mobile apps to promote mobile services, leading to the production and rise in use of different kinds of mobile devices such as smartphones.

The procurement and service of a fixed-line telephone system is highly expensive. As a result, much of the African population in particular rural dwellers and low-income earners cannot afford or have access to it [4, 32]. A broad range of network coverage is provided by mobile technology enabling quick access to the African population to telecommunications services [34, 35]. Communication services are currently accessible and affordable and more than 70% of the African people is linked to the digital telecommunications network[27]. Thus, mobile technology services enable digital inclusion.

The development, delivery and use of mobile technology services included a group of interconnected network entities called the mobile ecosystem [16, 36, 37]. The African mobile ecosystem comprises of various entities that are involve in the development and delivery of mobile innovation services including mobile network operators, sales agencies, government and other stakeholders [16, 7]. Several facets of the mobile ecosystem in conjunction with a defined system development method are involved in the process of creating, delivering, implementing and disseminating mobile apps [38]. With regard to the stakeholder's priorities, aspirations, operational services, and some predicting factors [39] mobile apps are developed for different services. There are several types of mobile apps that offer unique services to their users, including: Mobile Learning [40], Mobile Commerce [41], Mobile Health [42], Mobile Agriculture [43], Mobile Banking [44], and Mobile Governance [45] used to offer different online services respectively. In this study, these various kinds of mobile apps are categorised as the classification of mobile apps. The use of mobile apps in Africa has added tremendously to the growth of Africa [1]. That is, mobile services are rising the living standards of the people of Africa, resulting in the promotion of the economic, social, political, and educational development of all African nations. [46, 47]. Thus, the ecosystem promotes mobile apps innovations which accelerate economic growth, though there are some issues to be considered.

While mobile app developments are growing in Africa, some areas, especially rural locations, are still experiencing a slow penetration of this advanced technology due to certain factors or issues of influence [29]. Around 60% of Africans resides in remote areas

where a large number of potential adopters are limited by the poor network connectivity, low income, unstable economic situations, and social and political instability [29]. In addition, some scholars have established numerous important factors that need to be addressed, including relative advantage, complexity, compatibility, trialability, observability, [9], [48]–[49], social influence, user perception, perceived technology reliability [49, 50], facilitating conditions [51], demographic, mobile technology self-efficacy, ICT anxiety, perceived financial cost [49, 52, 53] and culture [54], for greater acceptance and distribution of mobile apps in Africa.

However, in the African background, it is important to understand the whole procedures of adopting and disseminating mobile apps and to create a well-defined structure for this process. The goal of this study is to address these issues by developing a well-defined framework that can encourage the wider and more efficient use of home-made mobile apps, not only in the country of origin but also in other African nations. Through a survey conducted in five selected nations representing five geographical regions in Africa the process of mobile apps adoption and diffusion was examined. These countries include south (South Africa), west (Nigeria), north (Morocco), east (Kenya) and central (Democratic Republic of Congo).
These countries were selected because:

1. Each country is situated in one of Africa's five regions.
2. There is a significant population in each of the countries chosen.
3. More references were made to them in the literature review than to those in their respective regions.
4. Each region is presumed to have a similar lifestyle type.

2.8 What is a Framework?

A framework is a blueprint or guide consisting of principles, constructs and concept of a theory [55]. According to Fulton, Krainovich-Miller [55, 56], conceptual framework is a structure of concepts that offers the foundation for the development of an abstract idea. For this study, a framework is an abstraction for a successful mobile apps adoption and diffusion among entire African social system. This study seeks to develop a conceptual framework for mobile apps adoption and diffusion in Africa. Applying regression analysis to the research model, the relationships between the research constructs can be revealed.

3 Related Works

Although there is no general framework for the adoption and diffusion of mobile apps in all African geographical areas but some related works have been carried out. The existing frameworks apply to a particular category of mobile app and to a unique environment or country. A model for the adoption of mobile health in developing countries was proposed by Kiongo [13] using Kenya as a case study. In the framework, social influence, facilitating conditions, reliability, cost and simplicity are considered as determining factors of the intention to adopt and use mobile health. In a model for mobile banking adoption in Ghana, Crabbe, Standing [14] highlighted that usefulness, reliability and culture are factors shaping the attitudes and behavioural intention to use mobile banking. Bwalya and Healy [15] proposed a conceptual model for e-government adoption in the SADC region with consideration to the following factors; culture, user support, usefulness, ease of use and appropriate ICT infrastructure. Social influence, facilitating conditions, cost, culture and reliability are the predicting factors considered in a model for the adoption of e-commerce in Nigeria [57]. According to Moore and Benbasat [58], relative advantage and complexity are equivalent to usefulness and ease of use. In the light of the above-proposed frameworks, there is a need for a general model that can be applied to different domains and different regions/countries in Africa to increase the use of African-based mobile apps across all African nations.

3.1 The Gap

The development of mobile apps in Africa is growing rapidly [7]. A variety of mobile app start-ups have been introduced across several African countries to promote mobile app technologies. [2]. The African mobile ecosystem has promoted the creation of various kinds of mobile apps that are available on the African technology market. As a result, a large number of mobile apps originated in Africa [2]. From the literature, all the existing suggested frameworks are for specific kinds of mobile apps [13]–[15], [57] and there is no framework for the general mobile apps adoption and diffusion in Africa. In practice, most of the mobile apps created in Africa are actively used in the country where they are originated, while there is low or no acceptance of these mobile apps in other nations of the continent. It is necessary to have a common structure that will enable a wider adoption of mobile apps (especially local mobile apps) across African nations. This research will therefore assist to close the gap by suggesting an overall framework for mobile apps adoption and diffusion in Africa.

4 Research Conceptual Model (RCM)

Drawing from the diffusion of innovation framework, previous related studies on Africa context regarding the framework for adoption of mobile apps and their findings, a conceptual model for this study is, therefore, proposed as displayed in figure 1. The RCM demonstrates the methodological approach to accomplish the research aim. The model consists of three components including mobile apps innovations, innovation decision process and adopting unit. The mobile apps innovations component describes mobile apps and their features; innovation decision processes component deals with the acceptance, rejection, use, discontinuance and continues the use of mobile apps whereas adopting unit contains the influencing factors of mobile apps adoption and diffusion. Achieving the primary goals of this research work would validate this model.

The Mobile apps section describes that: as an innovation, a newly created mobile app should be considered to be better than the previous similar apps that it replaces, fits well with the intended potential users, easy to understand and use, tested before implemen-

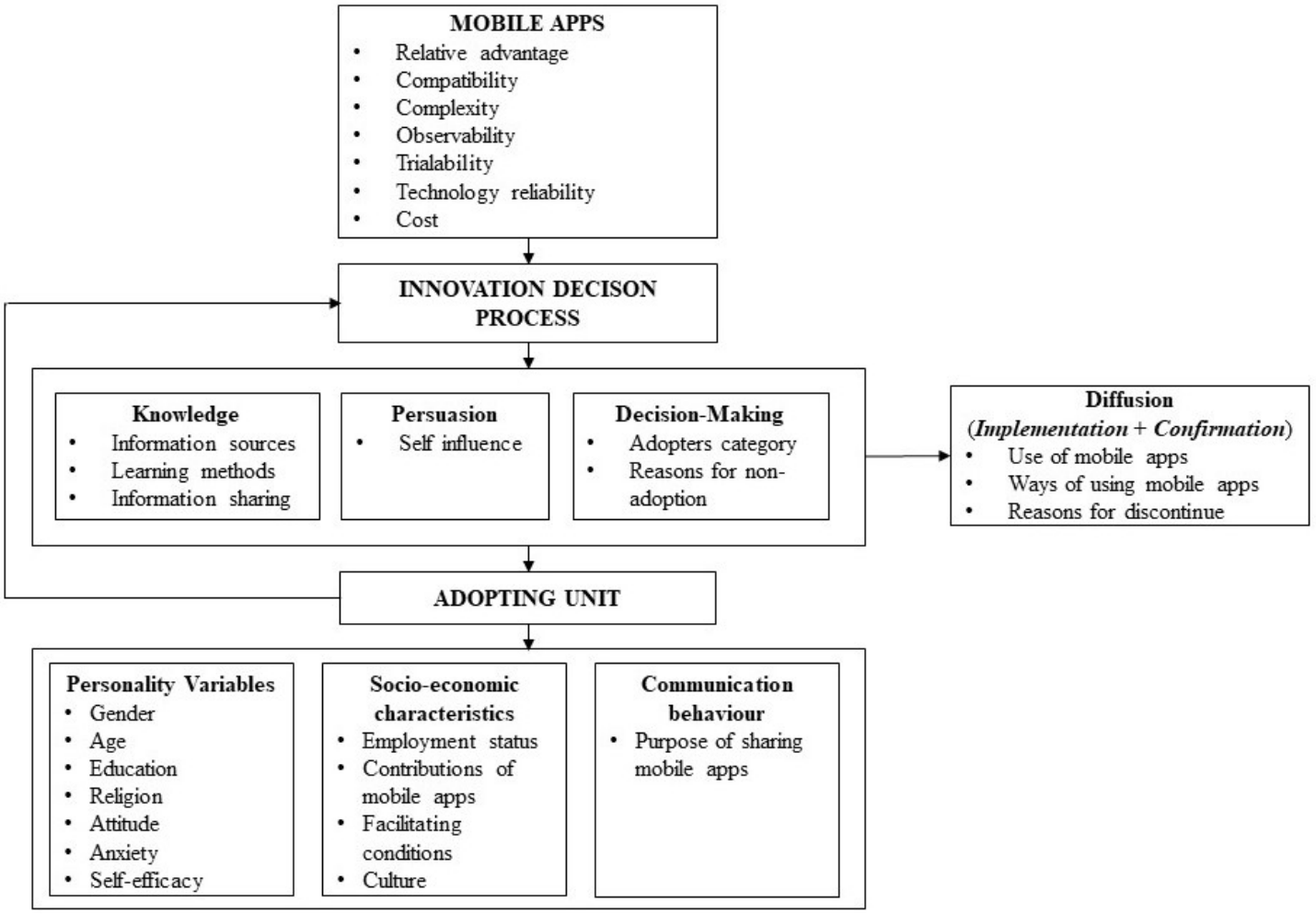


Figure 1: Research conceptual model

tation, accessible to people, and be credible. The innovation decision process explains the need for the public to be aware about the nature of a new mobile app, including source of information, the method of using the mobile app, and how to exchange information about it. Given these will help to influence the decision-making process of the adopters, either to accept or reject the mobile app. The adopting unit deals directly with the individual adopters, whereby certain personalized factors, such as personality variables, socio-economic characteristics, and communication behaviour, can affect the process of adoption. Overall, with the influence of good innovation characteristics, personalised factors and good decision-making process, a newly developed mobile app may gain widespread acceptance within the African social system.

5 Research Methodology

This research requires quantitative data from mobile apps users residing in five selected countries representing various African geographical regions. Hence, the positivism paradigm was adopted whereby a scientific method was used to gather self-determined data from the participants without the researcher's influence. The non-probabilistic technique of sampling was introduced since it is not easy in Africa to get the correct amount of mobile app users. In choosing the sample population for data gathering, purposive and convenience methods were considered. Purposive sampling involves the selection of a target audience based on a specific purpose. The study aims at the mobile apps adoption and diffusion and the population was selected from the area where there are large number of users of mobile apps for various activities. Convenience sampling is a particular type of data collection approach that entails the gathering of data from the participants that can be conveniently reached within the chosen population.

As the entire likely sample could not be covered by the researcher, therefore it was appropriate to pick a sample population containing all the desired participants. Thus, the purposive method was needed. The researcher was unable to reach everyone in the selected sample except those who were easily accessible; the convenience technique became essential at this juncture and was also applied. Educational institutions (universities) were selected with the purposeful method, and data was obtained from the participants with the convenience method. In order to achieve the research goals, the RCM formed the basis of the survey performed in the selected

nations.

5.1 The Research Instrument – The Questionnaire

The questionnaire was used as a data collection tool. It consists of seven sections containing close-ended questions with a 5-point scale rating of 1–5 (1 strongly disagreed and 5 strongly agreed). These sections are:

1. Background information – collects individuals (participants) details.
2. Mobile applications – examines the extent of participants' awareness of mobile app technologies.
3. mobile apps innovation Adoption and diffusion – examine Africa's acceptance and use of mobile apps.
4. Innovation features – examine the participants' perceptions of the importance of the characteristics of mobile apps.
5. Impact of mobile apps – examine the contributions of mobile apps to African growth.
6. Predicting factors – examine the factors influencing the acceptance and use of mobile apps in Africa.
7. Adopter's category – to determine the types of adopters of mobile apps existing in Africa.

A pilot or experimental test was conducted to make sure that the questionnaire can resloved the important parts of the study. All the recommendations were properly considered and effected. Also, the questionnaire was vetted and approved by a Statistical Consultant. The target population is the mobile apps users from the chosen nations in Africa. In the delivery of the study questionnaires, drop off and electronic/online approaches were applied. Participants were drawn from various age groups and job status to ensure that general views are covered by the collected data. A contact person was hired to help us in collecting responses from the respondents in each of the respective counties. The person was informed of the research process and the aim of the study was clarified. A total of 1285 of the 2300 distributed questionnaires were returned, providing a response rate of 55.87%. The responses comprise of 957 drop off and 328 electronically (online google forms) completed questionnaires. The responses obtained from the participants of each country was shown in Table 1.

Table 1: Participants' responses from each country

Country	**Amount disseminated**	**Amount received**
South Africa	500	335
Nigeria	500	370
Morocco	400	116
Kenya	400	212
DRC	500	252
Total	2300	1,285

Table 2: Reliability

Research variables	Mean	Std dev.	Cronbach Alpha Coefficient
Business interactions.	3.19	0.99	0.76
Individual interaction.	3.83	0.86	0.69
External initiated actions.	3.21	0.98	0.66
Personal initiated actions.	3.61	0.99	0.60
Organizations and events.	3.36	0.99	0.61
Individual activities.	3.68	0.86	0.60
Relative advantage and compatibility	3.59	0.84	0.79
Observability	3.44	0.87	0.60
Trialability	3.48	0.79	0.69
Perceived technology reliability	3.38	0.83	0.61
Self-efficacy	3.58	0.79	0.64
Social influence	3.56	0.96	0.64
Facilitating conditions	3.36	0.99	0.65
Financial cost	3.44	0.98	0.64
Cultural effect	3.31	0.88	0.63
Contributions to the healthcare system.	3.52	0.96	0.75
Contributions to education and social activities.	3.65	0.75	0.61
Contributions to the economy.	3.62	0.85	0.61
Contributions to politics.	3.60	0.86	0.60
Contributions to finance.	3.79	0.87	0.61
Contributions to agriculture.	3.39	0.95	0.69
Late Majority and Laggards	3.25	0.95	0.63
Non-Adopters	3.17	1.02	0.78
Innovators	3.19	0.88	0.80
Early adopters	3.33	0.89	0.72
Early majority	3.34	0.88	0.71

6 Analysis

To explain the frequency and response percentage of the participants, statistical analysis was performed on all the study variables. Cronbach alpha reliability analysis was used to measure reliability coefficients of the investigating variables (criteria 0.6). In order to assess the difference in the groups of dependent variables, analysis of variance (ANOVA) was conducted on the study variables, involving individual profile data (criteria, $p < 0.05$). Finally, regression analysis was performed to measure the relationship among the dependent and the independent variables. The results obtained from the data analysis will be interpreted on the basis of the RCM components. See section 7

6.1 Reliability

Using Cronbach alpha coefficient analysis, the reliability of the study variables was checked and it was found that the factors were internally consistent ($\alpha = 0.67$). Also, the correlation scores of all the elements are the same (r=1), meaning that the factor elements

associate with each other. The reliability results are displayed in Table 2.

7 Survey Results

The main objective of this research is to develop a framework for mobile app adoption and diffusion in Africa. The quantitative data was obtained from 1,285 participants from five chosen nations representing the geographical regions of Africa. The data collected was carefully extracted, statistically analysed and the results were obtained. The outcomes of the survey are presented according to the mean values of the responses obtained from the participants with respect to scale rating except the participants' profiles which is presented on percentages.

7.1 Participants' profiles

The research sample population consists of 703 (54.71%) of males and 582 (45.29%) of females. This suggests a reasonably equal distribution of the gender categories (male and female). Mobile apps are designed to be used equally by all, regardless of gender [56]. It shows a good representation of the African population. Participants were of various age groups, from less than 20 to 70 years of age and above. The results revealed that the majority of participants were between the age groups of 20-30 and 31-40 with the proportions of 409 (31.8%) and 379 (29.5%) respectively. That is, >60% of the participating population was between the ages of 20 and 40 years. All levels of education, including Matrics, Diploma, University, Masters, Doctorate and others were involved in this study. 'Others' means the set of people that are deemed to be uneducated and could not fall into these classifications. 93.3% of respondents are educated, while 3.7% are uneducated. This shows that the responses were collected from individuals who had acquired some knowledge of mobile apps innovations and services. A large proportion of the participants are working class and students with 615 (47.9%) and 402 (31.3%) respectively. Such people have the capacity of embracing and using mobile apps, even though some financial costs are involved. In addition, students and workers are also more involved in many things than elderly and unemployed people, allowing more use of mobile apps for different tasks.

7.2 Mobile apps Component

This section investigated the participants' perceptions of the features of mobile apps.

Relative advantage - is the degree to which users of a specific mobile app considered it to be more improved than another equivalent mobile app. The mean value of the participants' responses is 3.74 meaning that relative advantage is an influencing feature of mobile apps.

Compatibility – is the extent to which mobile apps suit with the exiting values, practices and demands of the prospective adopters. The obtained mean value of the participants' responses is 3.78 suggesting that compatibility is a significant feature of mobile apps.

Complexity – the degree to which participants believe that the use of mobile apps is affected by how simple it is to use the mobile app. The obtained mean value of the participants' responses is 3.78 suggesting that complexity is an influencing characteristic of mobile apps.

Observability – is the degree to which participants believe that seeing the impact or what can be achieved with the mobile app will convince people to accept and use the mobile app. The obtained mean value is 3.52 suggesting that observability is an influencing feature of mobile apps.

Trialability – is the degree to which the participants believe that testing a mobile app before adopting can influence the adoption of the mobile app. The obtained mean value of the participants' responses is 3.72 suggesting that trialability is an influencing feature of mobile apps.

Perceived technology reliability - deals with the belief, danger, and credibility involved in using a new mobile app. The obtained mean value of the participants' responses is 3.38 suggesting that perceived technology reliability is an influencing characteristics of mobile apps.

Cost – are the financial costs involved with the adoption and usage of mobile app services. The obtained mean value of the participants' responses is 3.44 suggesting that cost is an influencing feature of mobile apps.

7.3 Innovation Decision Process Component

This section investigated the innovation decision process of mobile apps including the knowledge, persuasion, decision-making, and how the information about mobile apps are being spread in Africa.

Information sources – sources of disseminating information about the existence of mobile apps. The obtained mean values are external initiated actions, 3.21 and personal initiated actions, 3.61 indicating that the sources are important to the acceptance and use of mobile apps.

Learning methods – are the methods of learning how to use mobile apps correctly. The obtained mean values from the participants' responses are: interpersonal communications, 3.6; the internet, 3.5; operational manual, 3.3; informal training, 3.3 and formal training, 2.8. This implies that all the methods are significant except formal training method.

Information sharing – are the communication channels through which the participants share the information about mobile apps with one another. The obtained mean values from the participant's responses are: organizations and events, 3.36 and personal activities, 3.68 indicating that these are significant communication channels of mobile apps adoption diffusion.

Social influence – is the level to which the adopters of mobile apps are influence by others (family, friends, government, etc.). The obtained mean value is 3.56 indicating that social influence has a significant effect on the consumers intention to adopt and use mobile apps.

Late majority and Laggards category – are the most skeptical (wait and see) and most localized set of adopters of mobile apps. The obtained mean value is 3.25 indicating a significant group of mobile apps adopters in Africa.

Non-adopters category – are the set of participants that do not adopt and use mobile apps at all. The obtained mean value from the participants' responses is 3.18 indicating that there are non-users of

mobile services in Africa.

Innovators category – involve a group of users eager to experience or practice new innovation, regardless of the advantages or losses associated with it. The obtained mean value from the participants' responses is 3.19 indicating a significant group of mobile apps adopters in Africa.

Early adopters category – they are social system leaders and advisors and their acts are confined within the limits of the social system. The obtained mean value from the participants' responses is 3.34 indicating a significant group of mobile apps adopters in Africa.

Early majority category – they are non-leaders with strong social interactions who actively promote invention and they are neither the first nor the last to adopt the innovation. The obtained mean value from the participants' responses is 3.35 indicating a significant group of mobile apps adopters in Africa.

Reasons for non-adoption – are the factors that cause the rejection of mobile apps. The obtained mean values from the responses of the participants are: mobile apps occupies a large memory space in a phone, 3.4; use of mobile apps takes time, 3.3; use of mobile apps distracts focus from other tasks, 3.3; use of mobile apps is costly, 3.3. The results indicate that these reasons may cause the rejection of mobile apps.

Use of mobile apps – Ask to decide whether participants are mobile app adopters/users. The obtained mean value from the participants' responses is 4.0 indicating that most of the sample population are mobile apps users.

Ways of using mobile apps – determine the different aims of using mobile apps. From the responses of the participants, the obtained mean values are: business use, 3.20 and individual use, 3.83 indicating that these are the important ways of using mobile apps in Africa.

Reasons for discontinuance – are the causes of stopping the use mobile apps. The obtained mean values from the participants' answers are: a new improved one to replace the existing one, 3.5; unsatisfactory results, 3.5; not easy to use, 3.3; lack of support, 3.2. The results imply that these are significant motives that may cause users to stop using a particular mobile app.

7.4 Adopting Unit Component

Attitude – is the reaction (positive or negative) of adopters towards new mobile apps. The obtained mean value from the the answers of the participants is 3.39 suggesting that attitude has significant effect on the acceptance and use of mobile apps.

Nervousness/anxiety – is the degree of nervousness whenever a new mobile app is introduced. The obtained mean value from the participants' responses is 3.20 suggesting that anxiety has significant effect on the acceptance and use of mobile apps.

Predicting factors – are the factors that influence mobile apps adoption and diffusion. The obtained mean values from the participants' responses are: self-efficacy, 3.58; facilitating conditions, 3.56; culture, 3.31. The results imply that these are important factors that may impact mobile apps adoption and diffusion.

Contribution of mobile apps – are the effect of the use of mobile apps services on the growth of Africa. The obtained mean values from the participants' responses are: contributions to healthcare system, 3.52; contributions to education and social activities, 3.65; contributions to economy, 3.62; contributions to politics, 3.60; contributions to finance, 3.80; contributions to agriculture, 3.40.

Purpose of information sharing – are the motives for spreading the news about mobile apps. The obtained mean values from the participants' responses are: to enable easy communications, 3.88 and to facilitate ones activities, 3.55 indicating that these are significant purposes of information sharing regarding mobile apps.

8 Regression Analysis

To analyse the association among the dependent and independent variables, multiple regression analysis was conducted. Two regression techniques were used to ensure that more accurate results are produced, including Enter and Stepwise techniques. Because the proposed framework measures the adoption and diffusion of mobile apps, the variables in the adoption and diffusion sections of the RCM are the dependent variables, while the remaining research variables are the independent variables. Regression analysis determines how the variations in the independent variables are associated to the variations in the dependents variables. These variations are represented by the coefficients while the p-values show the significant level of the variations. P <0.05 is the appropriate significant level. It suggests a 95% confidence interval (the possibility of obtaining the accurate effect). Any of the variables (predictor) that has p <0.05 is potential to be included into the model. The R^2 shows the proportion of the variance accounted for by the independent variables in the dependent variable. In Table 3, the effect sizes and the significant level of R^2 are shown, according to Ellis and Steyn [59].

Table 3: Effect sizes and significant levels of R^2

Effect size	Effect	Values of R^2	Conclusions of R^2
<0.15	Small	<0.13	Non-significant
0.15-0.35	Medium	0.13-0.25	Significant
>0.35	Large	>0.25	Practically important

The dependent variables are; the Use of mobile apps, Business uses, Individual uses, and the five categories of adopters.

The independent variables are; External initiated actions, Personal initiated actions, Organizations and events, Individual activities, relative advantage and compatibility, Observability, Trialability, Perceived technology reliability, Self-efficacy, Social influence, Facilitating conditions, Cost, Cultural effect, Contributions to the healthcare system, Contributions to education and social activities, Contributions to the economy, Contributions to politics, Contributions to finance and Contributions to agriculture. Also included to the independent variables are Informal training, Operational manual, The Internet, Interpersonal communications, Good innovations and Add to development. The significant results obtained from the regression analysis performed are shown in Tables 4 – 11.

8.1 General Use of Mobile App

This model indicates that the R^2 is 0.291, thus, the important independent variables accounted for 29.1% of the dependent variable variance (*General use of mobile apps*). R^2 >0.25 which is practically significant. Eight variables/predictors entered into the model which accounted for 27.3% of the entire dependent variable variance. These variables are Add to development, Good innovations, Relative advantage and compatibility, Perceived technology reliability, Contributions to healthcare, Occupy phone memory, Satisfy user expectations and Social influence. Thus, R^2 = 28.2% suggesting that the variables added in this model are practically significant for the proposed framework. See Table 4

8.2 Business Use

It indicates that the R^2 is 0.380, thus, the important independent variables accounted for 38.0% of the dependent variable variance (*Business use*). R^2 >0.25 which is practically significant. Eighteen variables that accounted for 37.0% of the overall dependent variable variance entered the model. These variables are Enable work activities, Relative advantage and compatibility, Culture, Add to development, Anxiety, Informal training, Internet, Contributions to finance. Satisfy User expectations, Working manual, Interpersonal communications, Advantageous, Contributions to agriculture, Contributions to politics, Observability, Social influence, Contributions to educations and social activities and Enable easy communications. Thus, R^2 >25.0% showing that the variables added in this model are practically significant for the proposed framework. See Table 5

8.3 Individual use

This model indicates that the R^2 is 0.447, so the significant independent variables accounted for 44.7% of the dependent variable variance (*Individual use*). R^2 >0.25 which is practically significant. Also, 15 variables that accounted for 43.8% of the overall dependent variable were added into the model. These variables are Add to development, Self-efficacy, Good innovations, Trialability, Internet, Interpersonal communications, Time consuming, Contributions to finance, Tested correctly, Observability, Contributions to agriculture, Contributions to education and social activities, Individual activities, Fits well, and Contributions to healthcare. Thus, R^2 >25.0% suggesting that the variables added in this model are practically significant for the intended framework. See Table 6

8.4 Late Majority and Laggards

It indicates that the R^2 is 0.256, so the important independent variables accounted for 25.6% of the dependent variable variance (*Late majority and laggard*). R^2 >0.25, which is practically significant. Also, 12 variables that accounted for 26.7% of the dependent variables variance were added into the model. These variables are Contributions to agriculture, Contributions to politics, Formal training, Used by others, Self-efficacy, Contributions to economy, Satisfy user expectations, Organizations and events, High complexity, Social influence, Perceived technology reliability, A better replacement, Occupy phone memory, and Informal training. Thus, R^2 >25.0% suggesting that the variables added in this model are practically significant for the intended framework. See Table 7.

8.5 Non-adopters Category

This model indicates that the R^2 is 0.312, so the significant independent variables accounted for 31.2% of the dependent variable variance (*Non-adopters*). R^2 >0.25, which is practically significant. Also, 15 variables that accounted for 30.1% of the dependent variable variance were added into the model. These variables are Financial cost, Formal training, High complexity, Use by others, Culture, Satisfy user expectations, Facilitating conditions, Lack of support, Distraction of attention, Contributions to educations and social activities, Contributions to economy, Trialability, Observability, Good innovations, and Perceived technology reliability. Thus, R^2 >25.0% suggesting that the variables added in this model are practically significant for the intended framework. See Table 8.

8.6 Innovators Category

It suggest that the R^2 is 0.326, thus, the important independent variables accounted for 32.6% of the dependent variable variance (*Innovators*). R^2 >0.25, indicating practically significant. Also, 14 variables that accounted for 29.7% of the overall dependent variable variance were added into the model. These variables are Relative advantage and compatibility, Informal training, Contributions to healthcare, Perceived technology reliability, Formal training, Fits well, Add to development, Facilitating conditions, High complexity, Contributions to economy, Culture, Unsatisfactory performance, Tested correctly and Satisfy user expectations. Thus, R^2 >25.0% demonstrating that the variables added in this model are practically significant for the intended framework. See Table 9

8.7 Early Adopters Category

It indicates that the R^2 is 0.316, thus, the important independent variables accounted for 31.6% of the dependent variable variance (*Early Adopters*). R^2 >0.25, indicating practically significant. Also, 9 variables that accounted for 29.7% of the overall dependent variable variance were added into the model. These variables are Relative advantage and compatibility, Self-efficacy, Contributions to agriculture, Contributions to politics, Culture, Fits well, Contributions to education and social activities, Satisfy user expectations, and Facilitating conditions. Thus, R^2 >25.0% suggesting that the variables added in this model are practically significant for the suggested framework. See Table 10.

8.8 Early Majority Category

This model indicates that the R^2 is 0.271, thus, the important independent variables accounted for 27.1% of the dependent variable variance (*Early majority*). R^2 >0.25, which is practically significant. Also, 14 variables entered into the model which accounted for 25.7% of the total dependent variable variance. These variables are Self-efficacy, Financial cost, Contributions to education and social activities, Occupy phone memory, Informal training, Tested correctly, A better replacement, Advantageous, Contributions to

Table 4: Model summary of the General use of mobile apps

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.539a	0,291	0,270	0,967	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,291	13,746	37	1241	0,000

Table 5: Model summary of Business use

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.617a	0,380	0,362	0,79478	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,380	20,574	37	1241	0,000

Table 6: Model summary of Individual use

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.669a	0,447	0,431	0,64418	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,447	27,126	37	1241	0,000

Table 7: Model summary for the Late majority and Laggard category

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.506a	0,256	0,234	0,83069	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,256	11,555	37	1241	0,000

Table 8: Model summary for Non-adopters category

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.559a	0,312	0,292	0,86144	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,312	15,234	37	1241	0,000

Table 9: Model summary for Innovators category

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.571a	0,326	0,306		**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
				0,73808	0,326	16,203	37	1241	0,000

Table 10: Model summary of Early adopters category

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.562a	0,316	0,296	0,75119	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,316	15,521	37	1241	0,000

Table 11: Model summary of Early majority category

Model Summary									
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Change Statistics**				
1	.509a	0,271	0,237	0,76663	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**
					0,259	11,741	37	1241	0,000

politics, Internet, External initiated actions, Culture, Satisfy user expectations and Trialability. Thus, R^2 >25.0% suggesting that the variables added in this model are practically significant for the proposed framework. See Table 11.

9 ANOVA Test

In order to examine the correlation among the profiles of the participants and the dependent variables, the ANOVA test was used. It helps to assess the impact or statistical importance of the participants' age, educational level, employment position and religion on the acceptance and distribution of mobile apps in Africa. The ANOVA results obtained are depicted in the Table 12 with regard to the five selected countries. This suggests that with the exception of innovators and early adopters, age has relevant impact on all the dependent variables. Education level indicated an important effect on Business use and Individual use. Employment status showed an important influence on business use and non-adopter's category while religion has a substantial impact on business use, individual use, non-adopter's and innovators categories concerning the acceptance and use of mobile apps within the African social system.

10 Interpretations and Validation of the Proposed Framework

The RCM was developed and used in this research as the basic principles. The model entails three parts or components, which include:

1. Mobile apps Innovations,
2. Innovation decision process, and
3. Adopting unit.

Each of these part contains a variety of factors that constituted the study variables. Various statistical analyses were conducted to measure the importance of the study variables from the perspectives of the participants. The obtained mean values from the responses of the participants show that the research variables are essential for mobile apps to be adopted and disseminated in Africa. With regard to the proposed framework, several analyses were also, carried out to confirm if the variable fits well in the model.

To examine the relationship between the model variables, multiple linear regression analysis was carried out. For dependent variables, eight factors were considered, including the Use of Mobile apps, Business uses, and Individual uses, Late majority and Laggards, Non-adopters, Innovators, Early adopters, and Early majority. These variables are derived form the adoption and diffusion component of the model. Regarding independent variables, the remaining factors were considered. The association among the variables was tested by regression analysis. After the analysis, the eight dependent variables resulted in the eight models. These models accounted over 25.0% of the variance (R^2 >25.0% <50.0%). With regards to the criteria for significant association namely the p-values, R^2, normality, tolerance and VIF, the output from regression analysis on the test of the relationships between the dependent and independent variables are interpreted below.

Model for the general use of mobile apps: R^2 = 0.273 suggests that the model is practically significant to the general use of mobile apps. This means that the model explained 27.3% of the variance concerning the general use of mobile apps in Africa. The model included a total of eight positive predictors: Add to development, Good innovations, Relative advantage and compatibility, Perceived technology reliability, Contributions to health care, Occupies phone memory, Satisfies user expectations and Social influence. All these variables have a positive impact on Africa's overall use of mobile apps.

The model therefore suggests that the dependent variable (general use of mobile apps) increases as predictive variables increase. this implies that that the adopters' attitude towards adoption and use of mobile apps can be shaped by good mobile apps that can contribute to Africa's development. Competitive benefit and usability increase the opportunity for mobile apps to be adopted and used. Safety assurance and decreased risk of using mobile apps, as well as the conviction of close friends and family increase the general use of mobile apps, even though the use of mobile apps takes up a great portion of phone storage.

Model for Business use: R^2 = 0.370 suggests that the model is practically significant to the use of mobile apps for business purposes. This means that the model explained 37.0% of the variance in the use of mobile apps in Africa for business purposes. In the model, 14 predictors, namely Enables work activities, Relative advantage and compatibility, Culture, Adds to development, Anxiety, Informal training, Internet, Satisfies user expectations, Working manual, Interpersonal communication, Contributions to agriculture, Observability, Social influence and Enables easy communication have a positive relationship with the use of mobile apps for business purposes, while four predictors, Contributions to finance, Advantageous, Contributions to politics, and Contributions to education and social activities have a negative relationship with the use of mobile apps for business purposes.

This means that the use of mobile apps for business purposes requires mobile apps to be good, beneficial and business-friendly in order to improve their meaningful business activities. Enables work activities and easy communications improve the use of mobile apps for business interactions. Organizations, such as banks are sharing the information about their mobile banking apps and other service-related mobile apps, such as FNB eWallet with their clients to improve their services. Increasing these predictors will increase the rate of adoption and diffusion of these mobile apps meant for business purposes.

African countries are developing nations, and developing mobile apps with local content containing social behaviour and norms will attract more adopters, especially rural dwellers and the less educated. Individuals can easily observe mobile apps created with local content, such as local language, and be convinced to use them. M-commerce apps in a local language can easily be adopted and used by people in remote areas to improve their lives in that area. Interestingly, anxiety about the use of mobile apps for business purposes has been positive. It is possible because most mobile app users were nervous when using mobile apps, especially for business transactions. Using the banking apps, as well as M-commerce apps

Table 12: ANOVA on Participants' profiles

Variables	p-values (ANOVA)			
	Age	Education level	Employment status	Religion
Use of mobile apps	0.023	0.598	0.442	0.066
Business use	0.000	0.000	0.000	0.000
Individual use	0.000	0.000	0.227	0.022
Late majority and laggards	0.000	0.746	0.990	0.264
Non-adopters	0.000	0.080	0.004	0.012
Innovators	0.176	0.835	0.777	0.023
Early adopters	0.063	0.607	0.197	0.065
Early majority	0.010	0.110	0.572	0.313

attracts some degree of fear and nervousness, but they facilitate easy, quick and comfortable transactions. Although people are anxious about using such apps, the number of users increases rapidly. Users need to learn how to use mobile apps adequately for business. Increasing the learning methods of how to use mobile apps will improve the adoption and diffusion of mobile apps. Unfortunately, the use of mobile apps for business is not profitable, and it attracts some costs. The more banking apps you use to carry out banking transactions, the more fees you pay. The use of mobile apps for business is also not affected by political, educational and social advancement.

Model for Individual use: $R^2 = 0.438$ suggests that the model is practically significant to the individual use of mobile apps. This shows that the model explained 43.8% of variance for individual use of mobile apps in Africa. The model included a total of 15 predictors, including 13 positive predictors (Adds to development, Self-efficacy, Good innovations, Trialability, Internet, Interpersonal communication, Time-consuming, Contributions to finance, Tested correctly, Observability, Contributions to education and social activities, Fits well, and Individual activities) and two negative predictors (Contributions to agriculture and Contributions to health care). Positive predictors have a positive relationship with the individual use of mobile apps, while negative predictors have a negative relationship with the use of mobile apps for individual purposes.

The model suggests that the attitude of mobile app adopters and their use thereof can be shaped by good mobile apps that can contribute to the development in Africa. The advantages of using mobile apps, including financial, educational and social benefits, affect the adoption and diffusion of mobile apps for individual purposes. The benefits of mobile social networking apps, for example have improved the adoption and use of mobile apps by individuals. Individual use of mobile apps involves the technical knowledge of how to use mobile technology, and the trial of new technology, as well as seeing others using it increases the adoption and diffusion of mobile apps. Users need to learn how to use mobile apps. The availability of learning methods, such as the Internet and interpersonal communication increase the rate of adoption and diffusion of mobile apps for individual use. The model also shows that, due to time and costs involved, there is less interest in using mobile apps to perform agricultural and health activities. In other words, the time and cost involved in using mobile apps reduces mobile app adoption and dissemination in Africa.

Model for Late majority and Laggards category: $R^2 = 0.265$ indicates that the model is practically significant to the Late and Laggards category of mobile app adopters. It implies that 26.5% of the variance can be explained by the adopters of mobile apps in Africa's Late majority and Laggards category model. The model included a total of 12 predictors, including nine positive (Contributions to agriculture, Contributions to politics, Formal training, Used by others, Self-efficacy, Contributions to the economy, Organizations and events, High complexity, Perceived technology reliability, Occupies phone memory, and Informal training) and three negative predictors (A better replacement, Satisfies user expectations and Social influence). Any increase in positive predictors leads to an increase in the Late majority and Laggards category, while any increase in negative predictors leads to a reduction in the category of the Late majority and Laggards adopters in Africa.

The Late majority and Laggards category of adopters is the set of adopters that is sceptical about adopting new mobile apps and they adopt a "wait and see" attitude towards adopting new technology. For this category to adopt and use mobile apps, they need proper information sources, such as organization and events, as well as seeing others using them and knowing the benefits of using mobile apps. In addition, they require training and reliability assurance on the adoption and diffusion of mobile apps. Since they are the local set in the social system, they become more discouraged by their close peers and are more afraid to adopt new mobile apps.

Model for Non-adopters category: $R^2 = 0.301$ indicates that the model is practically significant to the non-adopters' category of mobile apps adopters. This means that the model explained 30.1% of the variance. The model included a total of 15 significant predictors, containing 11 positive (Financial cost, Formal training, High complexity, Use by others, Culture, Facilitating conditions, Lack of support, Distraction of attention, Contributions to the economy, Observability, and Perceived technology reliability) and four negative predictors (Good innovations, Trialability, Satisfies user expectations and Contributions to education and social activities). The 11 predictors have a positive relationship with the category of non-adopters, while the four negative predictors have a negative relationship with the category of mobile app users in Africa who are not adopters.

The model explained that the high complexity, lack of support and the financial cost involved in the adoption and diffusion of mobile apps cause some adopters in the social system of Africa not to adopt and use mobile apps. The cultural beliefs, as well as the extent of support from government and other stakeholders, and dis-

traction from other activities increase the existence of non-adopters in Africa. It is expected that the perception of mobile apps as good innovations will change the attitudes of non-adopters. Providing a trial opportunity, satisfying users' expectations and the benefits of the use of mobile apps, especially in education and social activities, will reduce the number of non-adopters in Africa.

Model for Innovators category: $R^2 = 0.317$ indicates that the model is practically significant to the innovators' category of mobile apps adopters. It implies that the innovator category model explained 31.7% of the variance. The model is related to 14 important variables/predictors, including 12 positive variables (Relative advantage and compatibility, Informal training, Contributions to health care, Perceived technology reliability, Formal training, Fits well, Facilitating conditions, Contributions to the economy, Culture, Unsatisfactory performance, Tested correctly and Satisfies user expectations) and two negative variables (Adds to development and High complexity). The positive variables increase the category of innovators as they increase, while the negative variables reduce the category as they increase.

The model explained that the challenge of competitive advantage, availability of necessary environment from government and other stakeholders, as well as trying to satisfy the consumers' demands results in the increase of the innovators' category of adopters in Africa. The advancement of technology increases users' experience and requirements, thus there is a need for innovators to create mobile apps that will suit the new improvements. For purposes of improving the economy, poor performance of mobile apps and social behaviour, more improved mobile apps are needed. This encourages or raises more innovators in the social system. On the other hand, the higher the development progress and higher complication of mobile apps in Africa, the lesser innovators are required.

Model for Early adopters category: R2 = 0.297 suggests that the model is practically significant to the early adopters' category of mobile app adopters. It means that the model described 29.7% of the variance. The model is connected to nine predictors, including eight positive predictors (Relative advantage and compatibility, Self-efficacy, Contributions to agriculture, Contributions to politics, Culture, Fits well, Contributions to education and social activities and Satisfies user expectations) and one negative predictor (Facilitating conditions). Any rise in positive predictors leads to an increase in the early adopt category, while any increase in negative predictors leads to a decrease in the early adopt category.

The model shows that the ability to use a mobile app with advanced features will be of good interest and increase the potential adopters in the early adopter's category. In addition, the ability to use mobile apps to conduct educational, agricultural and political activities, as well as to fulfill consumer needs, is increasing the category of early adopters. Conversely, the availability of a suitable setting decreases the size of the group of early adopters.

Model for the Early majority category: $R^2 = 0.257$ shows that the model is practically important for mobile apps adopters in the early majority category. This means that 25.7% of the variance was explained by the model. The model is associated with 14 predictors, including 11 positive (Self-efficacy, Financial cost, Contributions to education and social activities, Occupies phone memory, Informal training, Tested correctly, Advantageous, Contributions to politics, Internet, Culture, and Trialability) and three negative predictors (A better replacement, External initiated actions, and Satisfies user expectations). A rise in positive predictors results to an increase to the size of early majority adopter category while an increase in negative predictors reduces the size of the category.

This shows that testing a new mobile app and finding it useful will increase the interest of consumers to embrace and use the mobile app without considering the impact on the phone's memory. The availability of methods of learning, such as the Internet and informal training, increases the user's skill and knowledge. The creation of better mobile apps and the acquisition of information from external sources, such as TV and radio, on the other hand increases consumer demand, thereby decreasing the acceptance of prior mobile apps.

Regarding the personality variables (age, education level, employment status and religion), ANOVA was conducted to examine the association among personality variables and the dependable variables. The findings showed that age has a statistically significant impact on all dependent variables, with the exception of the categories of Innovators and Early Adopters. The level of education showed a major effect on business and individual uses. A significant impact from the employment status was noticed on the business use and the non-adopters categories. Concerning religion, the impact was much on business use, individual use, non-adopters' and innovators' categories. Therefore, age, level of education, employment status and religion have an important impact on the acceptance and diffusion of mobile apps in Africa in various ways.

To explore the influence of gender on the acceptance and use of mobile apps, a t-test study was carried out. No noticeable effect was shown by the result. There are no variances among males and females in the adoption and delivery of mobile apps. Mobile apps are created for everyone regardless of the gender.

Overall, various substantial portions of the variance were explained by the eight models and all R^2 scores were over 25.0 points, (R^2 >25.0% <50.0%). This demonstrates that these models are practically important for the adoption and diffusion of mobile applications in Africa, and all the factors/predictors of the RCM can be applied to the proposed system. Some personality profiles have also been shown to have an effect on the adoption and distribution of mobile apps in Africa, including age, education, employment status and religion. As a consequence, a correlation occurs among them and the dependent variables. The association among the study variables typically shows that they are valuable predictors of mobile apps adoption and diffusion in Africa.

10.1 The Framework

Based on the findings obtained from various statistical studies and recommendations, Figure 2 presents the final proposed framework for the effective adoption and diffusion of mobile apps in Africa.

11 Conclusions and Future Work

11.1 Conclusion

This research aims to create a general framework for mobile apps adoption and diffusion in Africa. A three components model was conceptualized consisting of the research variables. The positivism

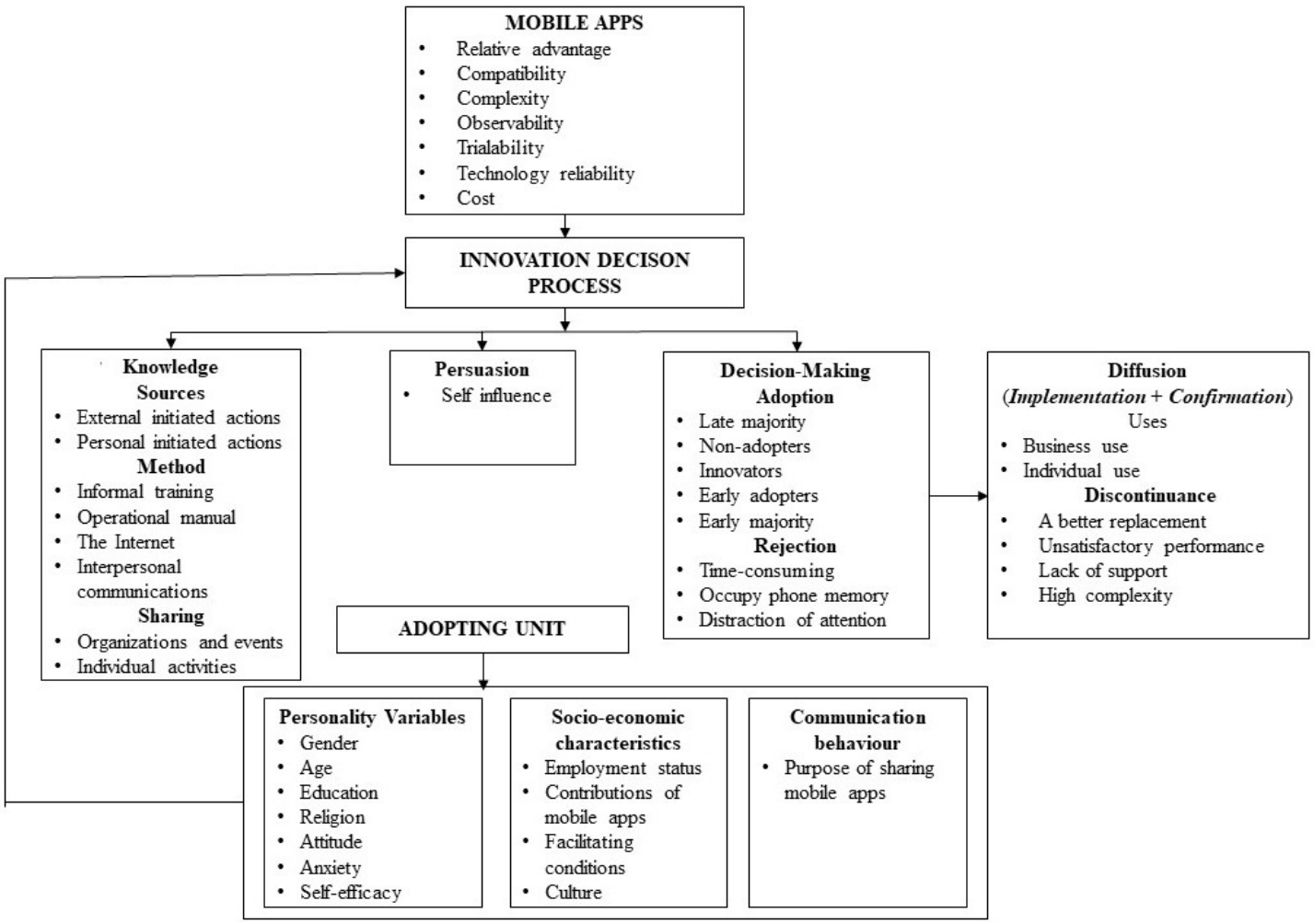


Figure 2: Validated Proposed Framework

research approach was applied. Through a survey, quantitative data were collected from the mobile apps users living in the five distinct nations representing the various regions in Africa. Various statistics analyses were carried out and the results show that the participants perceived these variables to be significant factors for mobile apps adoption and diffusion in Africa. Furthermore, a regression analysis was conducted to measure the relationships among these variables in association with mobile apps adoption and diffusion. The variable in the adoption and diffusion component of the RCM are the dependable variables whereas the remaining variables are the independent variables. The findings indicated that variables (predictors) explained more than 25.0% of the variance (R^2 >25.0% <50.0%) with appropriate level of significance ($p < 0.05$). These results suggest that the research variables are practically significant and they are potential or meaningful predictors of mobile apps adoption and diffusion in Africa. Thus, the variables fits well in the proposed framework. Also, the study revealed that some personality profiles such as age, education level, employment status , and religion have certain degree of influences on mobile apps adoption and diffusion in Africa. Therefore, there is a relationship among the personality variables and the dependent variables. Overall, the relationships existing among the research variables and their significant impact on mobile apps adoption and diffusion suggest that they are potential predictors of mobile apps adoption and diffusion in Africa. Consequently, the study proposed a three components framework for mobile apps adoption and diffusion in Africa.

11.2 Limitations

Africa is a region made up of multilingual nations and a variety of political challenges. Language gaps, cultural disparities and political uncertainty in most nations posed some serious limitations to this research work especially in Kenya, Morocco, and DRC. Also, due insufficient funding, it was not possible to reach all target audiences, particularly some rural dwellers, instead the convenient participants were used.

11.3 Future work

African nations are growing with potential population that are ready to embrace and use mobile technology innovations such as mobile apps. In order to further encourage mobile technology adoption, further research is needed on information awareness, the expansion of mobile services to unconnected areas, and how to bridge the gap

between educated and uneducated users of mobile apps in Africa.

References

[1] J. Hellström, P.-E. Tröften, The innovative use of mobile applications in East Africa, Swedish international development cooperation agency (Sida), 2010. https://www.sida.se/contentassets/b55f10e43c874c9da41203d991a7eaa0/14993.pdf.

[2] GSMA, The mobile economy Sub-Saharan Africa 2015., GSMA Intelligence. Regional report. https://gsmaintelligence. com/research. Accessed on 23/03/2016., 2015.

[3] R. Nickerson, M. Austreich, J. Eng, "Mobile technology and smartphone apps: A Diffusion of innovations analysis," in 20th Americas Conference on Information Systems, AMCIS 2014, 2014, 1–12.

[4] G. Van Zyl, Why Kenya's fixed line market is shrinking., itweb. https: //itweb.africa/content/KjlyrvwdnK57k6am. Accessed on 24/06/2016., 2013.

[5] S. Lee, M. Marcu, S. Lee, "An empirical analysis of fixed and mobile broadband diffusion," Information Economics and Policy, **23**(3-4), 227–233, 2011, doi:10.1016/j.infoecopol.2011.05.001.

[6] A. A. Shaikh, H. Karjaluoto, "Mobile banking adoption: A literature review," Telematics and informatics, **32**(1), 129–142, 2015, doi:10.1016/j.tele.2014.05.003.

[7] S. Murugesan, "Mobile apps in Africa," IT Professional, **15**(5), 8–11, 2013, doi:10.1109/MITP.2013.83.

[8] C. W. Okonkwo, M. Huisman, E. Taylor, "An investigation on the adoption and diffusion of mobile apps in Africa." Thesis submitted in fulfilment of the requirements for the degree Doctor of Philosophy in Computer Science at the North-West University, Potchefstroom campus, South Africa., 2019., 2019.

[9] E. Rogers, Diffusion of Innovations (5th ed)., New York: Free Press., 2003.

[10] I. Lule, T. K. Omwansa, T. M. Waema, "Application of technology acceptance model (TAM) in m-banking adoption in Kenya." International Journal of Computing & ICT Research, **6**(1), 2012.

[11] D. G. Taylor, T. A. Voelker, I. Pentina, "Mobile application adoption by young adults: A social network perspective," 2011. http://digitalcommons.sacredheart.edu/wcob_fac/1.

[12] D. W. Stacks, M. B. Salwen, An integrated approach to communication theory and research, Routledge, 2014.

[13] N. Kiongo, "A framework for mobile health adoption in developing countries: Case study Kenya," A framework for mobile health adoption in developing countries: Case study Kenya, 2014. http://hdl.handle.net/11295/74353.

[14] M. Crabbe, C. Standing, S. Standing, H. Karjaluoto, "An adoption model for mobile banking in Ghana," International Journal of Mobile Communications, **7**(5), 515–543, 2009. DOI: https://doi.org/10.1504/IJMC.2009.024391.

[15] K. J. Bwalya, M. Healy, "Harnessing e-government adoption in the SADC region: a conceptual underpinning," Electronic Journal of E-government, **8**(1), 23, 2010.

[16] C. Z. Qiang, S. C. Kuek, A. Dymond, S. Esselaar, "Mobile applications for agriculture and rural development," 2012. http://hdl.handle.net/10986/21892.

[17] K. S. McNamara, "Mobile applications in agriculture and rural development: Framing the Topic, and Learning from Experience," in World Bank Workshop on Mobile Innovations for Social and Economic Transformation, 2009.

[18] M. O. M. El-Hussein, J. C. Cronje, "Defining mobile learning in the higher education landscape," Journal of Educational Technology & Society, **13**(3), 12–21, 2010. https://www.jstor.org/stable/jeductechsoci.13.3.12.

[19] C. Perera, A. Zaslavsky, P. Christen, D. Georgakopoulos, "Context aware computing for the internet of things: A survey," IEEE communications surveys & tutorials, **16**(1), 414–454, 2013, doi:10.1109/SURV.2013.042313.00197.

[20] E. Oriwoh, P. Sant, G. Epiphaniou, "Guidelines for internet of things deployment approaches–the thing commandments," Procedia Computer Science, **21**, 122–131, 2013, doi:10.1016/j.procs.2013.09.018.

[21] A. Giessmann, K. Stanoevska-Slabeva, B. De Visser, "Mobile enterprise applications–current state and future directions," in 2012 45th Hawaii International Conference on System Sciences, 1363–1372, IEEE, 2012. https://doi.org/10.1109/HICSS.2012.435.

[22] H. T. Dinh, C. Lee, D. Niyato, P. Wang, "A survey of mobile cloud computing: architecture, applications, and approaches," Wireless communications and mobile computing, **13**(18), 1587–1611, 2013, doi:10.1002/wcm.1203.

[23] J. Donner, C. A. Tellez, "Mobile banking and economic development: Linking adoption, impact, and use," Asian journal of communication, **18**(4), 318–332, 2008, doi:10.1080/01292980802344190.

[24] C. W. Okonkwo, M. Huisman, E. Taylor, "The adoption of m-commerce applications: rural dwellers perspectives," in 12th, IADIS, International conference. Information systems., pages:99–106, 2019. DOI: 10.33965/is2019_201905L013.

[25] C. W. Okonkwo, M. Huisman, E. Taylor, "Socio-Economic Contributions of Mobile Applications in Africa: Impact of Local Mobile Applications," in 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC), 1–8, IEEE, 2019. DOI: https://doi.org/10.1109/IMITEC45504.2019.9015912.

[26] G. Intelligence, "The Mobile Economy. 2020," 2020.

[27] G. Association, et al., "The mobile economy 2018," London: GSM Association, 2018.

[28] L. Hosman, E. Fife, "The use of mobile phones for development in Africa: Top-down-meets-bottom-up partnering," The Journal of Community Informatics, **8**(3), 2012.

[29] G. Association, et al., "The Mobile Economy Sub-Saharan Africa," London: GSM Association, 2017.

[30] B. Simon, "A History of Mobile Application Development," https://manifesto.co.uk/history-mobile-application-development/. Accesses on 23/06/2016. 2014., 2017.

[31] I. T. Union, World Telecommunication/ICT Development Report, International Telecommunication Union, 2006.

[32] H. K. Chavula, "Telecommunications development and economic growth in Africa," Information Technology for Development, **19**(1), 5–23, 2013. https://doi.org/10.1080/02681102.2012.694794.

[33] M. E. Batuo, "The role of telecommunications infrastructure in the regional economic growth of Africa," The Journal of Developing Areas, **49**(1), 313–330, 2015. https://www.jstor.org/stable/24241299.

[34] S. Etzo, G. Collender, "The mobile phone 'revolution'in Africa: rhetoric or reality?" African affairs, **109**(437), 659–668, 2010. DOI: https://doi.org/10.1093/afraf/adq045.

[35] S. H. Lee, J. Levendis, L. Gutierrez, "Telecommunications and economic growth: an empirical analysis of sub-Saharan Africa," Applied Economics, **44**(4), 461–469, 2012. DOI: https://doi.org/10.1080/00036846.2010.508730.

[36] J. Zhang, X.-J. Liang, "Business ecosystem strategies of mobile network operators in the 3G era: The case of China Mobile," Telecommunications policy, **35**(2), 156–171, 2011. DOI: https://doi.org/10.1016/j.telpol.2010.12.009.

[37] R. Lindgren, O. Eriksson, K. Lyytinen, "Managing identity tensions during mobile ecosystem evolution," Journal of Information Technology, **30**(3), 229–244, 2015. DOI: https://doi.org/10.1057/jit.2015.8.

[38] W. C. Okonkwo, M. Huisman, "The use of system development methodologies in the development of mobile applications: Are they worthy of use?" in 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 2, 278–283, IEEE, 2018. DOI: https://doi.org/10.1109/COMPSAC.2018.10243.

[39] C. W. Okonkwo, M. Huisman, "The predicting success factors of mobile applications development: organizational perspectives," in 12th, IADIS, International conference. Information systems., 11–18, 2019. DOI: 10.33965/is2019_201905L002.

[40] W.-H. Wu, Y.-C. J. Wu, C.-Y. Chen, H.-Y. Kao, C.-H. Lin, S.-H. Huang, "Review of trends from mobile learning studies: A meta-analysis," Computers & Education, **59**(2), 817–827, 2012. DOI: https://doi.org/10.1016/j.compedu.2012.03.016.

[41] S. Okazaki, F. Mendez, "Exploring convenience in mobile commerce: Moderating effects of gender," Computers in Human Behavior, **29**(3), 1234–1242, 2013. DOI: https://doi.org/10.1016/j.chb.2012.10.019.

[42] C. Free, G. Phillips, L. Watson, L. Galli, L. Felix, P. Edwards, V. Pa-tel, A. Haines, "The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis," PLoS medicine, **10**(1), e1001363, 2013. DOI: https://doi.org/10.1371/journal.pmed.1001363.

[43] A. Gichamba, I. A. Lukandu, "A model for designing M-agriculture applications for dairy farming," The African Journal of Information Systems, **4**(4), 1, 2012. https://digitalcommons.kennesaw.edu/ajis/vol4/iss4/1.

[44] P. Van der Boor, P. Oliveira, F. Veloso, "Users as innovators in develop-ing countries: The global sources of innovation and diffusion in mobile banking services," Research Policy, **43**(9), 1594–1607, 2014. DOI: https://doi.org/10.1016/j.respol.2014.05.003.

[45] J. Hellström, "Mobile Governance: Applications, Challenges and Scaling-up in Poblet, M.(Ed.),"Mobile Technologies for Conflict Management: Online Dispute Resolution, Governance, Participation"," 2012. DOI: https://doi.org/10.1007/978-94-007-1384-0_12.

[46] D. Baelden, L. Van Audenhove, "Participative ICT4D and living lab research: The case study of a mobile social media application in a rural Tanzanian University setting," Telematics and Informatics, **32**(4), 842–852, 2015. DOI: https://doi.org/10.1016/j.tele.2015.04.012.

[47] J. C. Aker, I. M. Mbiti, "Mobile phones and economic development in Africa," Journal of economic Perspectives, **24**(3), 207–32, 2010. DOI: 10.1257/jep.24.3.207.

[48] S. Tunmibi, A. Aregbesola, E. Asani, "Factors influencing the adoption of smart phones by university students–A cross-border approach," African Journal of Computing & ICT, **8**(1), 1–16, 2015.

[49] B. M. Achieng, B. K. Ingari, "Factors influencing the adoption of mobile banking in Kenya's commercial banks: A case of Kenya Commercial Bank (KCB) Kilindini branch," International Journal of Scientific and Research Publications, **5**(10), 1–14, 2015.

[50] R. A. Oluoch, Factor affecting adoption of mobile banking technology in Kenya: A case of bank customers within Nakuru Municipality, Ph.D. thesis, 2012.

[51] J. Mtebe, R. Raisamo, "Investigating students' behavioural intention to adopt and use mobile learning in higher education in East Africa," International Journal of Education and Development using ICT, **10**(3), 2014. https://www.learntechlib.org/p/148476/.

[52] W. P. Mtega, R. Bernard, A. C. Msungu, R. Sanare, "Using mobile phones for teaching and learning purposes in higher learning institutions: The case of Sokoine University of Agriculture in Tanzania," 2012. https://repository.ubuntunet.net/handle/10.20374/130.

[53] F. Iddris, "Barriers to adoption of mobile banking: Evidence from Ghana," International Journal of Academic Research in Business and Social Sciences, **3**(7), 356–370, 2013. DOI: https://doi.org/10.6007/IJARBSS/v3-i7/59.

[54] F. O. Bankole, O. O. Bankole, I. Brown, "Mobile banking adoption in Nigeria," The Electronic Journal of Information Systems in Developing Countries, **47**(1), 1–23, 2011. DOI: https://doi.org/10.1002/j.1681-4835.2011.tb00330.x.

[55] A. Osanloo, C. Grant, "Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house"," Administrative issues journal: connecting education, practice, and research, **4**(2), 7, 2016. DOI: 10.5929/2014.4.2.9.

[56] S. Fulton, B. Krainovich-Miller, C. Cameron, "Finding and appraising the literature," Nursing Research in Canada-E-Book: Methods, Critical Appraisal, and Utilization, 90, 2014.

[57] S. Chiemeke, A. Evwiekpaefe, "A conceptual framework of a modified unified theory of acceptance and use of technology (UTAUT) Model with Nigerian factors in E-commerce adoption," Educational Research, **2**(12), 1719–1726, 2011. http://www.interesjournals.org/ER.

[58] G. C. Moore, I. Benbasat, "Development of an instrument to measure the perceptions of adopting an information technology innovation," Information systems research, **2**(3), 192–222, 1991. DOI: https://doi.org/10.1287/isre.2.3.192.

[59] S. Ellis, H. Steyn, "Practical significance (effect sizes) versus or in combination with statistical significance (p-values): research note," Management dynamics: journal of the southern african institute for management scientists, **12**(4), 51–53, 2003.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1643-1650 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

FPGA-Based Homogeneous and Heterogeneous Digital Quantum Coprocessors

Valeriy Hlukhov*

Computers Department, Lviv Polytechnic National University, Lviv, 79013, Ukraine

ARTICLE INFO

Article history:
Received: 15 September, 2020
Accepted: 19 December, 2020
Online: 25 December, 2020

Keywords:
Digital quantum computer
Digital quantum coprocessor
Digital qubit
Quantum Fourier transform
FPGA

ABSTRACT

Quantum computers are heterogeneous device. It consists of a main CPU and a quantum accelerator. True quantum accelerator (or coprocessor) is analog and probabilistic device. Qubits are the basic building blocks of quantum computers. But qubits can be digital. A digital qubit is similar to RISC processor pipeline and is an unique chain of digital gates.

In this work, it is proposed to execute quantum routines in quantum computer not on the quantum chip but on the chip of a digital FPGA. This paper presents the architecture of such FPGA – an architecture of digital quantum coprocessor. The paper presents two types of digital quantum coprocessors - heterogeneous and homogeneous. The advantage of a homogeneous coprocessor is shown.

The IP Core generator was developed to create VHDL descriptions of digital quantum elements and digital quantum coprocessors in general.

In this paper heterogeneous quantum computer which consists of a main CPU and a FPGA-based quantum accelerator (coprocessor) has been proposed. And these FPGA-based digital quantum coprocessors can have a homogenous or heterogeneous structure. Quantum coprocessors have up to 1024 qubits in one FPGA. A homogeneous quantum coprocessor performs better than a heterogeneous one. Also, its implementation is easier.

The measured ratio of correct results for a 1024-qubit homogenous coprocessors is more then 50 %.

1. Introduction

This paper is an extension of work originally presented in IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT) [1].

Quantum algorithms are a mixture of classical logic and quantum routines which can be executed on the quantum chip. In [2] a heterogeneous quantum computer architecture was presented. According to [2] a quantum computer consists of both a classical and quantum computing part.

In this work, it is proposed to execute quantum routines in quantum computer not on the quantum chip but on the chip of a digital field programmable gate array, on a FPGA. This paper presents the architecture of such FPGA – an architecture of digital quantum coprocessor.

As an extension of the results described in [1] heterogeneous and homogeneous digital quantum coprocessors are considered. It is important to note that a heterogeneous quantum computer and proposed heterogeneous quantum coprocessor are completely different concepts.

Also, the number of digital qubits in the coprocessors under study was increased to 1024.

Von Neumann architecture had for a long time a single processor. Then homogeneous multi-core processor dominated the processor development. In the era of microprocessors, the understanding came that heterogeneity is the best way forward to improve the compute power. System architecture with heterogeneous accelerators includes the main CPU and heterogeneous coprocessors such as floating-point math coprocessors, graphics and neural accelerators, FPGAs. In [2] and [3] heterogeneous quantum computer which consists of a main CPU and a quantum accelerator has been proposed. In this paper heterogeneous quantum computer which consists of a main CPU and a FPGA-based quantum accelerator (coprocessor) has been proposed. And these FPGA-based digital quantum coprocessors can have a homogenous or heterogeneous structure and can be used at the lowest level of full-stack quantum accelerators along with quantum chips and quantum simulators [3].

A quantum computer contains N qubits and a digital quantum coprocessor contains N digital qubits. In both cases as a result of any calculations they can produce any of 2^N results. But the same

*Corresponding Author: Valeriy Hlukhov, Lviv Polytechnic National University, Computers Dept., Lviv, 79013, Ukraine, Glukhov@polynet.lviv.ua

calculations can lead to different results. And every i-th result will be produced with probability p_i.

A qubit can be thought of as a device that has a group of inputs for data and instructions that control its behavior. Unlike a real qubit, a digital qubit can have an additional group of outputs. The exact qubit state code is generated on this group of outputs. An important element of the digital qubit is a pseudo random number generator (PRNG). Together with the qubit state code, the pseudo-random code is used to generate a probabilistic result at the one-bit output of the qubit (Figure 1) [4].

Quantum volume VQ is a metric that measures the performance of a quantum computer's capabilities and error rates [5]. In the simplest case, the quantum volume is d·N. Now the quantum volume of a real quantum computers is very small (at the level of milliseconds), but for digital quantum coprocessors it is practically unlimited.

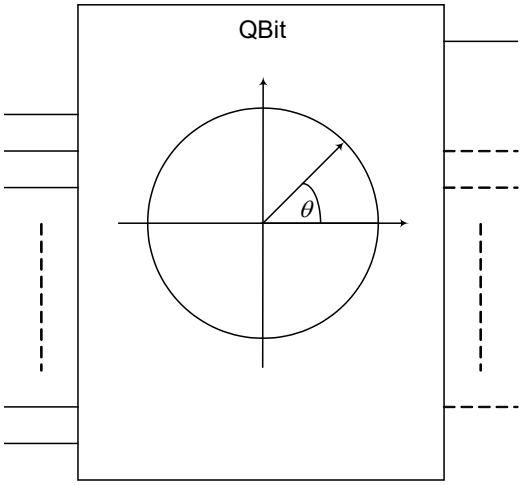


Figure 1: Digital qubit symbol

The DiVincenzo [6] criteria are conditions necessary for constructing quantum computer. Not a single word about the physical nature of a quantum computer is included in this criterion. Therefore, you can try to create digital quantum computer. It can be created either as a software model or as a hardware device.

"You can run a simulator on your phone that can run a 20-qubit system. But once you get up close to 50, the amount of memory you need to run a simulator gets into the petabytes [of memory]. Over 50 and there is no way you can pretend to be a quantum computer. At that point you really need a true quantum computer" [7].

The number of qubits required for solving practical problems is now estimated at several thousand [8]. Creating a true quantum computer with so many qubits is a very difficult task, and a digital quantum coprocessor can already be created on one FPGA.

The purpose of the article is to demonstrate the advantages of digital quantum computers over true ones. For this, it is necessary to develop a circuit for a digital qubit and a digital quantum coprocessor as part of a digital quantum computer. To compare different coprocessors, it is necessary to develop methods for their comparison. In this works digital quantum computers with 2^n qubits (n = 5, 6, …, 10) have been used for research.

Also, the purpose of the work is to determine the timing and hardware characteristics of the developed digital quantum coprocessors.

An important task is the problem of the reliability of the results of modeling the operation of digital quantum coprocessors. In this work, it was successfully solved, it was shown that simulation results coincide with the results obtained during prototyping.

2. Theoretical Foundations of Quantum Computing

$$|0\rangle, |1\rangle, ..., |N-1\rangle$$

$$|j\rangle$$

Along with measurement, the quantum superposition can be changed under external influence.

$$|\psi\rangle = \sum_{j=0}^{N-1} \lambda_i |j\rangle \quad p_j = \lambda_j^2 \quad P = \sum_{j=0}^{N-1} \lambda_i^2 = 1$$

$$p_0 = \cos^2\theta \qquad p_1 = \sin^2\theta$$

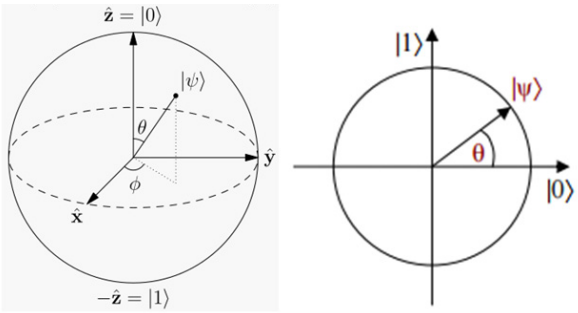


Figure 2: A Bloch sphere (left) and a unit circle (right)

A true quantum computer is an analog device and it has no memory. It has only gates. Therefore, there are no quantum programs. In the drawing of a true analog quantum computer, a sequence of quantum gates shows only the time sequence of qubit state changes. Software tool [14] allows to create circuits from quantum gates, simulate their work and manage it with C-like language instructions. This is very similar to the FPGA design.

In the schema of a digital quantum computer, a sequence of digital quantum gates shows both the time sequence of qubit state changes and the relative physical position of the gates themselves in the space.

A good illustration of a qubit can be an electron spin. An electron can change its spin in time from 1 to 20 nanoseconds [15]. This time can serve as a base operation time for compare the performance of true and digital quantum coprocessors.

To compare the capabilities of true and digital quantum computers, one can calculate the quantum volume as a metric that measures the performance of a quantum computer's capabilities and probability that the qubit will work for some time t without failure:

$$QV = N*p(t),$$

$p(t) = e^{-t/MTBF}$ is probability that the qubit will work for some time t without failure where (MTBF - Mean Time Between Failures);

N is qubits number.

For true qubits MTBF is microseconds and millisecond and p(t) run to 0 when t > 1 ms, QV << N.

For digital qubits MTBF is practically unlimited (MTBF of modern FPGA is at level 50 years) and p(t) run to 1. So, for digital quantum coprocessor QV = N.

The results of comparison of analog data processing methods, which are used in true quantum computers, and digital methods are summarized in the Table 1. All this speaks about the prospects of creating hardware digital quantum computers.

3. The Structure of FPGA-based Digital Quantum Coprocessor

A classical computer controls the operation of a quantum coprocessor (Figure 3) provides it with an input data, instructions and checks the result of its work [4]. This interaction is well described as a full-stack of the layers of an accelerator at [3].

A generalized functional diagram of full-stack hardware resources of quantum computer with realized in FPGA quantum coprocessor is given in Figure 4 [4].

Table 1: Comparison of analog and digital data processing methods

Characteristic	Analog processing methods	Digital processing methods
Speed	+	-
Versatility	-	+
Microminiaturization	-	+
Accuracy	-	+
Zoom	-	+
Transmission in space	-	+
Transmission in time (memory)	-	+
Immunity	-	+
Reliability	-	+
Testing, debugging, diagnostics	-	+
Quantum volume	-	+

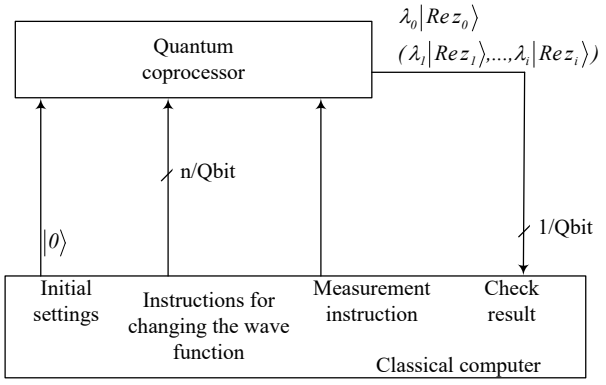


Figure 3: A classical computer with quantum coprocessor.

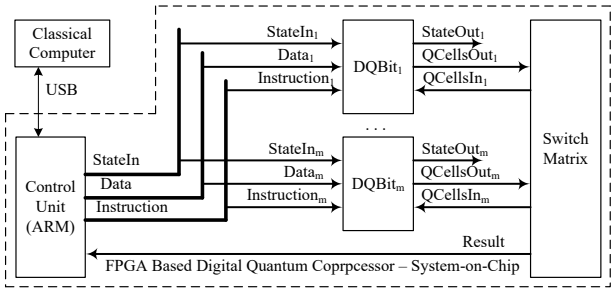


Figure 4: A digital quantum coprocessor for classical computer.

The top-level functionality of this stack is provided by the classic host computer. Functioning at the microarchitecture level is provided by embedded in FPGA microprocessor (AWP, control unit).

And directly quantum computing is provided by a set of digital qubits and a switch matrix which connects the qubits to each other and transmits the final state code of all of them or only those required at the moment to the control unit.

The connections between qubits can be static or dynamic. In this work static connections have been used. They do not change while the computer is running.

A digital qubit consists of j series-connected digital quantum cells (Figure 5).

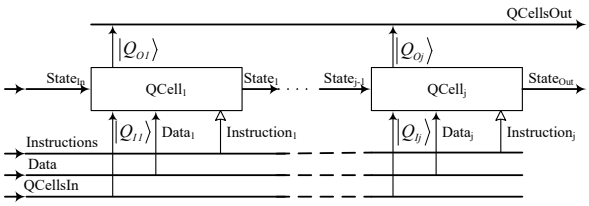


Figure 5: RISC processor-like structure of a digital qubit DQBit.

$|Q_{Oj}\rangle$

$|Q_{Ij}\rangle$

$D = arcsin\sqrt{A}$

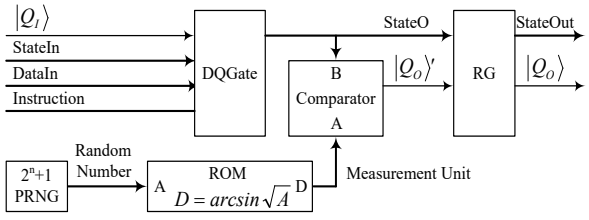
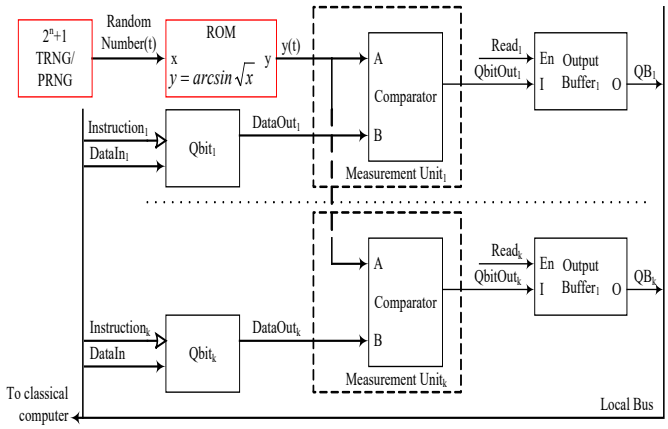


Figure 6: A digital quantum cell QCell.



A digital quantum gate transforms the input qubit state code into the output code under the influence of the instructions and their data, as well as under the influence of measured states of another qubits.

Inverse functional transformation allows the creation and research of two types of digital quantum coprocessors:

each quantum cell of the heterogeneous coprocessor has its own pseudo-random number generator and its own functional transformer;

a homogeneous coprocessor contains only one pseudo-random number generator and only one functional transformer for all quantum cells, for all quantum qubits (Figure 7).

A simplified topology of a part of a FPGA chip with resources available to the user for design is shown in the Figure 8. These resources are sets of configurable logic blocks CLB, programmable switching machines PSW and I/O blocks IOB. The PSM and the wires laid on the chip between them and CLB make it possible to connect the CLB to each other. CLB can be programmed to act as a qubit (QB) or quantum cell (QC). This allows many digital qubits to be placed on one chip. The interaction of qubits and the control of qubits is carried out by digital methods using PSM. The use of COTS FPGAs makes it possible to create digital quantum coprocessors with a large number of qubits, which can operate for a long time under normal conditions.

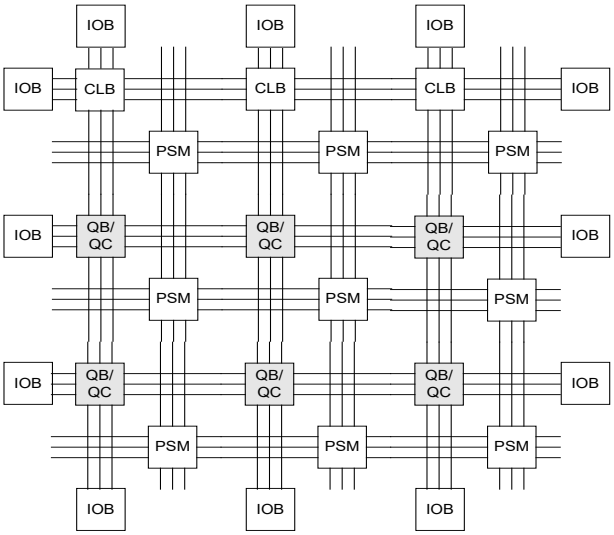


Figure 8: Simplified topology of digital quantum coprocessor FPGA chip

Recently, other developers have also tried to connect digital technologies with true quantum computers:

place true qubits on the crystal (similar to the Figure 8 shown) and organize interaction between them using digital methods [16]. It uses a custom crystal that can operate at ultra-low temperatures (near 4° K);

carry out optical control of true qubits embedded in the chip [17];

use a pseudo-random number generator in true quantum computers [18].

In this study digital quantum coprocessors were implemented for the case of wave function real amplitudes (Figure 2) and for polar coordinate system to represent the movement of the vector (Figure 13) [12].

Algorithm design steps for digital quantum computer design on FPGA are standard for FPGA design:

to create or to find an algorithm for solving the problem;

to find or to create a mathematical description of the solution to the problem;

for FPGA-based circuits to create graphic symbols of library elements and their descriptions in hardware description language;

- to create a schema from library elements;
- to simulate the created schema;
- to implement the project;
- to verify the project;
- to make a prototype project.

The IP Core generator was developed to create VHDL descriptions of digital quantum elements and schema of digital quantum coprocessors in general.

4. Testing Digital Quantum Coprocessors

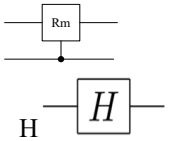
The quantum Fourier transform (QFT) as a part of Shor's algorithm [19] (Figure 9) for factorization [11] was chosen to compare different variants of quantum coprocessors.

The QFT is defined as

$$\sum_j \alpha_j |j\rangle \rightarrow \sum_k \tilde{\alpha}_k |k\rangle$$

$$\tilde{\alpha}_k \equiv \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{2\pi i j k / N_{\alpha_j}}$$

Only 2 types of digital quantum gates are required during QFT (Figure 10, Figure 11):



- Hadamard transform H

$$R_m = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{2\pi i / 2^m} \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

When implementing a digital qubit in the form of a chain of digital quantum gates and implementing only a QFT circuit on FPGA, there is no need to change the functions of the quantum gates. Therefore, the functions of each gate are determined only by their circuit. And there is no need for instructions that change the

function of the gates. Therefore, they are not shown in the Figure 12, but are shown in the general diagram Figure 4. And the data path in the Figure 12 is a chain of links named Q*, along which the changing qubit state code goes from one gate to another. The measured states of a qubit, which are called q*, are transferred from one gate to another as required by the algorithm for solving the problem. QFT determines spectrum of qubits states.

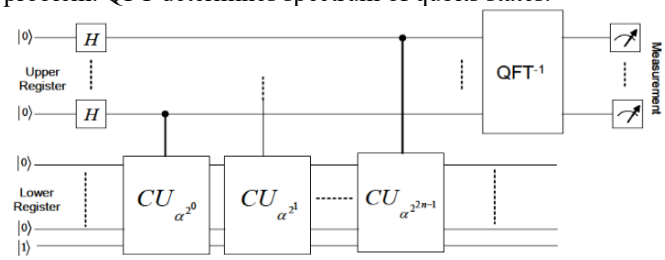


Figure 9: Quantum factorization by Shor's algorithm

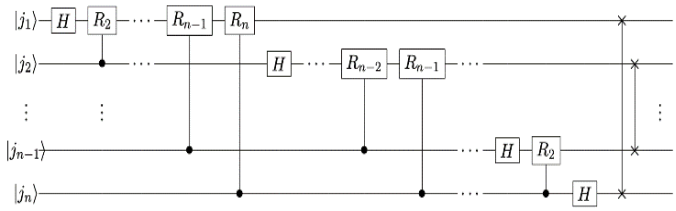


Figure 10: QFT drawing

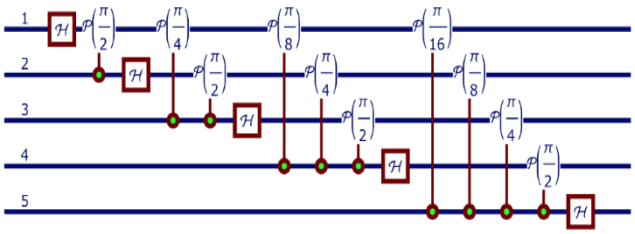


Figure 11: Simplified drawing of QFT (5-qubits)

The diagram for FPGAs corresponding to the Figure 11 is shown in the Figure 12.

The QFT results were obtained as described in [12] and [13].

5. Simulation and Verification of FPGA-based Digital Quantum Coprocessor

For analysis the input state of qubits, which can conditionally be described as |XXX..X0>, where X corresponds to the neutral position of the vector in the unit circle - at angle of π/4, was selected. The probability of measuring the input state with odd code is p_{odd}=0, and with even codes is $p_{even} = 100/2^{n-1}$ %, where n is qubits quantity and the spectrum of qubit states at the QFT input will look like in Figure 14 (for 4 qubits).

QFT simulation results for this case are shown in Figure 15. The 4 qubit digital quantum coprocessor determines the period of the input states spectrum with probability 38,5 %, so it correctly executes a QFT.

Also digital quantum coprocessors with up to 1024 qubits, have been investigated [13], [20]. Similar studies were performed for other periods in input states spectrum (for 0, 1, 2, 4, …, 2^{n-1} periods) [13].

Implementation results are summarized in Table 2.

Table 2: Implementation Results of Heterogonous Digital Quantum Co-Processors (2^{n-1} periods)

Qubit number, n	CLK period (ns)	LUT number	RAM number	FPGA resoirces, % R	FPGA 7z type [21].
128	2.353	32352		14	045
64	4.289	16043		15	010
32	3.874	7983			010
16	3.606	4000			010
10	3.976	2522			010
10	4.180	4307		24	010
10	5.137	3754			010
10	5.281	3650	20	16	010

32-qubit digital quantum coprocessor (Figure 18) was implemented on ZedBoard Zynq-7000 ARM_FPGA SoC Development Board (Digilent, Figure 17) for verification.

Verification and simulation results are almost the same (Table 3). This indicates the creation of both a high-quality model of a digital quantum coprocessor and its prototype, which can be used in further research. Each experiment was repeated 4000 times to determine the probability of getting the correct result (Figure 19). The average level of correct results during the verification of a heterogeneous coprocessor is 9%, during the simulation of a homogeneous processor is 61%, during the verification of a homogeneous processor is 58%.

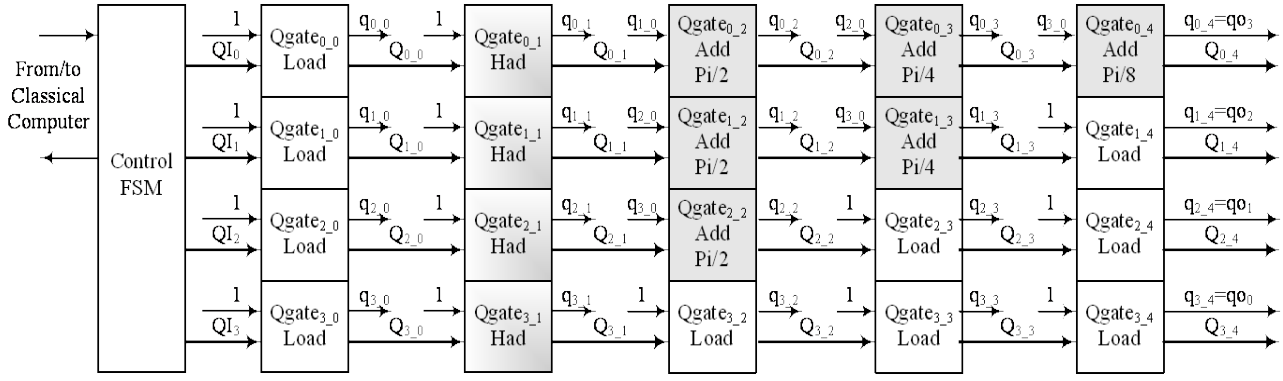


Figure 12: FPGA schema of QFT (4-qubits)

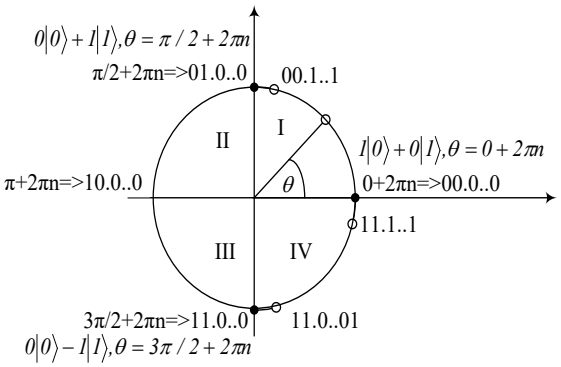


Figure 13: A polar coordinate system (angle θ codes)

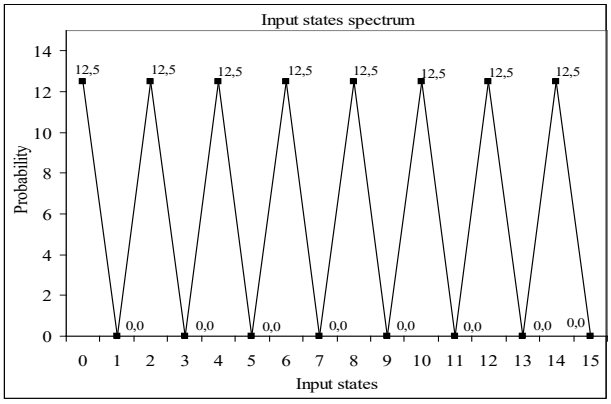
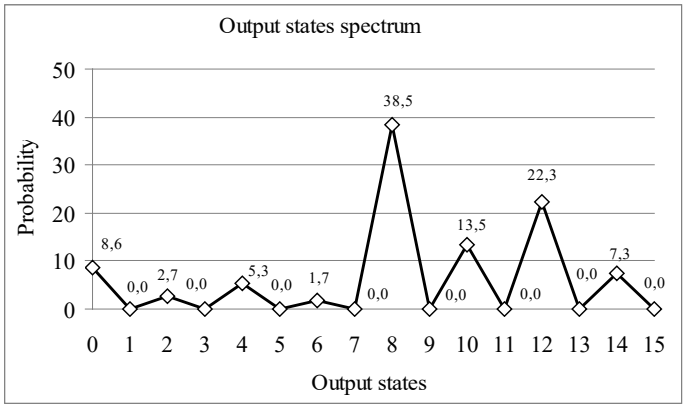


Figure 14: Input state spectrum

$$\theta_3 = \theta_2 = \theta_1 = 45°, \theta_0 = 0°$$

A homogeneous digital quantum coprocessor has shown better results than a heterogeneous one. Therefore, further research was carried out on a homogeneous digital quantum coprocessor.

The probabilities of obtaining correct results during the verification for the operation of multi-qubit homogeneous digital quantum coprocessors are shown in the Figure 21. The work of coprocessors with 32, 64, 128, 256, 512 and 1024 qubits were investigated.



$$\theta_3 = \theta_2 = \theta_1 = 45°, \theta_0 = 0°$$



Figure 16: Digital quantum computer

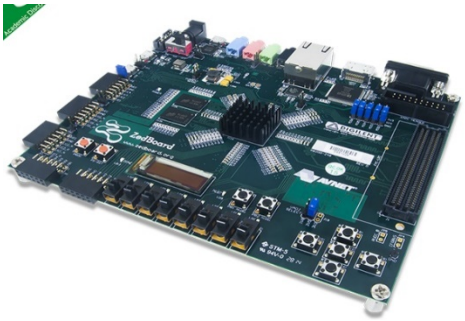


Figure 17: FPGA Development Board

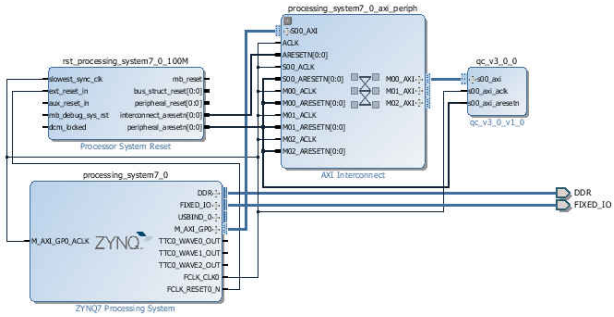


Figure 18: Schema of digital quantum coprocessor for FPGA

Table 3: Simulation vs verification

Qubits number N	Qubits width W, bits	True results probability % H	Cycle T, ns	Process
32	3	61	3.874	Simulation
32	3	58	8.0	Verification

FPGA resources for 32-qubit heterogeneous and homogeneous coprocessors are presented in the Table 4.

FPGA topology of 32-qubit digital quantum coprocessor is presented in Figure 20.

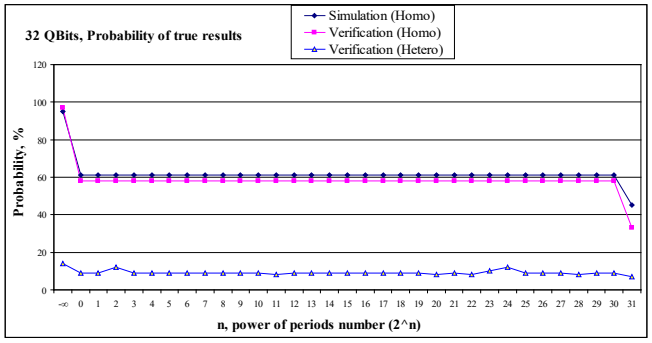


Figure 19: Probability of true results in digital quantum coprocessors

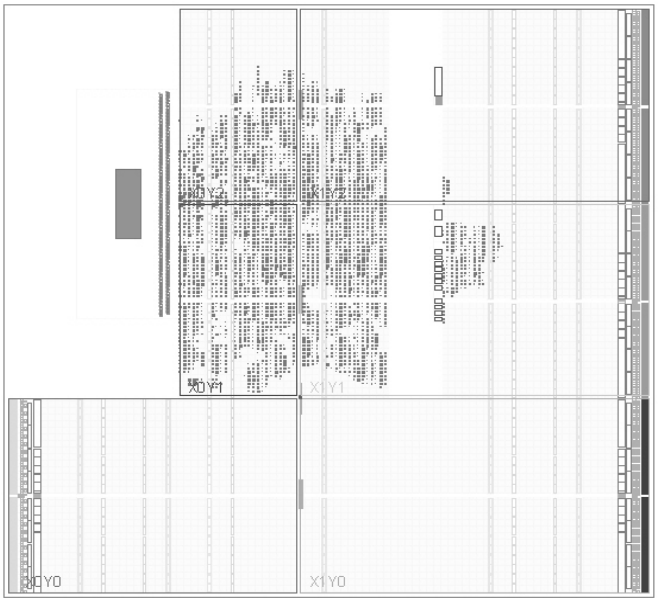


Figure 20: 32-qubit digital quantum coprocessor topology

This result indicates the creation of both a high-quality model of a digital quantum coprocessor and its prototype, which both can be used in further research.

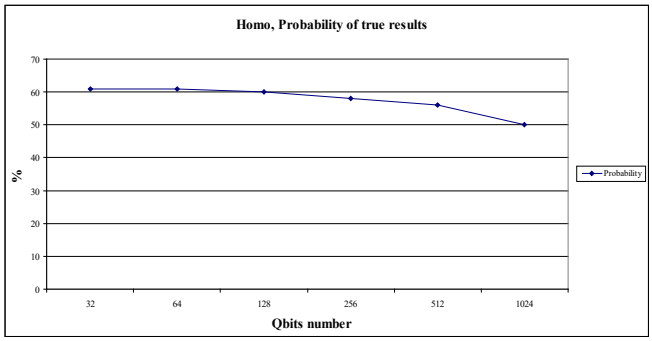


Figure 21: Probability of true results in homogenous digital quantum coprocessors

Table 4: FPGA resources for 32 qubit coprocessors

	Homo	Homo	Hetero	Hetero
	Used	Util %	Used	Util %
Slice LUTs	5831	10.96	7669	14.42
# of logical nets	21575		28147	

6. Conclusion

The paper shows the advantage of digital quantum computers over true computers in terms of quantum volume. The paper presents heterogeneous quantum computer which consists of a main CPU and a FPGA-based quantum accelerator (coprocessor). This FPGA-based digital quantum coprocessors can have a homogenous or heterogeneous structure. Structures of digital quantum qubits and cells are also presented.

The IP Core generator was developed to create VHDL descriptions of digital quantum elements and digital quantum coprocessors in general.

32 qubits homogeneous digital quantum coprocessor generates correct quantum Fourier transform results with a probability of 61% during simulation and 58% during prototyping.

1024 qubits homogeneous digital quantum coprocessor generates correct quantum Fourier transform results with a probability more than 50 % during simulation.

Homogeneous digital quantum computer generates the correct results of the quantum Fourier transform 7 times more often than heterogeneous one.

The hardware cost of homogeneous digital quantum coprocessors is 20% less than that of heterogeneous coprocessors.

The number of logical nets in homogeneous digital quantum coprocessors is 20% less than in heterogeneous coprocessors.

Digital quantum coprocessors verification results are practically the same as simulation results. This indicates the creation of both a high-quality model of a digital quantum coprocessor and its prototype, which both can be used in further research.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] V. Hlukhov, "FPGA Based Digital Quantum Computer Verification," in 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DeSSerT), 178-182, 2020, doi: 10.1109/DESSERT50317.2020.9125077.

[2] X. Fu et al., "A heterogeneous quantum computer architecture," in 2016 ACM International Conference on Computing Frontiers (CF'16), 323–330, 2016, doi: 10.1145/2903150.2906827.

[3] K. Bertels, A. Sarkar, T. Hubregtsen, M. Serrao, A. Mouedenne, A. Yadav, A. Krol, I. Ashraf, "Quantum Computer Architecture: Towards Full-Stack Quantum Accelerators," in 2020 EDAA Design, Automation & Test in Europe Conference & Exhibition (DATE), 1–6, 2020, doi: 10.23919/DATE48585.2020.9116502.

[4] V. Hlukhov, B. Havano, "FPGA-based Digital Quantum Coprocessor," Advances in Cyber-Physical Systems, **3**(2), 12-31, 2018, doi: 10.23939/acps2018.02.067.

[5] N. Moll, P.Kl. Barkoutsos, L.S. Bishop and others, "Quantum optimization using variational algorithms on near-term quantum devices," Quantum Science and Technology, **3**(3), 030503, 2017, doi:10.1088/2058-9565/aab822.

[6] D. DiVincenzo, "The physical implementation of quantum computation," Fortschritte der Physik: Progress of Physics, **48**(9–11), 771–783, 2000.

[7] E. Scannell, "IBM quantum computers' usefulness in sight - using binoculars," in TechTarget News. available at: https://searchdatacenter.techtarget.com/feature/IBM-quantum-computers-usefulness-in-sight-using-binoculars, accessed 25 Nov. 2020.

[8] "Applying Moore's Law to Quantum Qubits," Quantum Computing Report, avaliable at: https://quantumcomputingreport.com/our-take/applying-moores-law-to-quantum-qubits, 2019, accessed 25 Nov. 2020.

[9] E. Grumbling, M. Horowitz (eds.), Quantum computing: progress and prospects, The National Academies of Sciences, Engineering, and Medicine, Washington, DC: National Academies Press, 2019.

[10] M. Khalil-Hani, Y. Lee, M. Marsono, "An accurate FPGA-based hardware emulation on quantum Fourier transformm," in 13th Australasian Symposium on Parallel and Distributed Computing (AusPDC 2015), 23–30, 2015.
[11] S. Gushanskiy, V. Pereverzev, "Simulation of quantum computing using hardware cores," (in Russian), Scientific Journal of KubSAU, 123(09), 545–557, 2016, avaliable at: http://ej.kubagro.ru/2016/09/pdf/37.pdf, accessed 02 Dec. 2020.
[12] V. Hlukhov, "Implementing quantum Fourier transform in a digital quantum coprocessor," Advances in Cyber-Physical Systems, **4**(1), 6–13, 2019, doi: 10.23939/acps2019.01.006.
[13] V. Hlukhov, "FPGA-based K-qubit digital quantum coprocessor," Electrotechnic and Computer Systems, **31**(107), 104–117, 2019, doi: 10.15276/eltecs.31.107.2019.10.
[14] "Microsoft Quantum Documentation," 2020, available at: https://docs.microsoft.com, accessed 02 Dec. 2020.
[15] J. Smith, P. Dalgarno, R. Warburton, A. Govorov, K. Karrai, B. Gerardot, P. Petroff, "Voltage control of the spin dynamics of an exciton in a semiconductor quantum dot," Physics Rev. Lett., **94**, 197402, 2005.
[16] S. Moore, "What Intel Is Planning for The Future of Quantum Computing: Hot Qubits, Cold Control Chips, and Rapid Testing", IEEE Spectrum, 2020, available at: https://spectrum.ieee.org/tech-talk/computing/hardware/intels-quantum-computing-plans-hot-qubits-cold-control-chips-and-rapid-testing, accessed 24 Aug 2020.
[17] Ch.P. Anderson, A. Bourassa, K.C. Miao, G. Wolfowicz, P.J. Mintun, A.L. Crook, H. Abe, J.U. Hassan, N.T. Son, T. Ohshima, D.D. Awschalom, "Electrical and optical control of single spins integrated in scalable semiconductor devices", Science, **366**(6470), 1225–1230. 2019. doi: 10.1126/science. aax9406.
[18] K. Miyamoto, K. Shiohara, "Reduction of Qubits in Quantum Algorithm for Monte Carlo Simulation by Pseudo-random Number Generator," Physical Review A., 022424, 2020, doi: 10.1103/physreva.102.022424.
[19] P. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," in 35th Annual Symposium on Foundations of Computer Science, 124–134, 1994, available at: https://www.jstor.org/stable/2653075?seq=1, accessed 25 Nov. 2020.
[20] V. Hlukhov, "Comparison of Homogeneous and Heterogeneous Digital Quantum Coprocessors," in International Workshop on Computational Methods and Information Transformation Systems, a satellite of IEEE 2020 XV International Scientific and Technical Conference on Computer Science and Information Technologies, 70–73, 2020.
[21] "DS190 (v1.11.1) Zynq-7000 SoC Data Sheet: Overview," 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1651-1656 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Type 2 Diabetes Risk and Physical Activity in outpatients treated in Health Centers in a District of North Lima, 2020

Deisy Chipana-Collahua[1], Mariluz Chipana-Collahua[1], Rosa Villegas-Ortiz[1], Brian Meneses-Claudio[2,*], Hernan Matta-Solis[1]

[1]*Faculty of Health Sciences, Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

[2]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, 15314, Lima-Perú*

ARTICLE INFO

Article history:
Received: 31 August, 2020
Accepted: 09 December, 2020
Online: 25 December, 2020

Keywords:
Diabetes risk
Lifestyle
Outpatients
Physical activity
Risk factors

ABSTRACT

Diabetes Mellitus Type 2 is a chronic disorder that affects the way the body metabolizes sugar (glucose) in the blood and depends on a combination of risk factors, such as genes and lifestyle. Although certain risk factors such as family history, age, or ethnicity cannot be changed, those related to diet, physical activity, and weight can be changed. These lifestyle changes can affect the probability of developing type 2 diabetes and its complications. In this research work, the prevention of the patient to develop type 2 diabetes with the option of leading a healthy lifestyle is proposed and reinforcing the population of each establishment with health education. It is a quantitative approach, a non-experimental, descriptive, and correlational study, with a population of 300 outpatients from a district of North Lima, who answered a questionnaire with sociodemographic data and the instruments of the level of physical activity IPAQ and risk of FINDRISK type 2 diabetes mellitus. In the results with respect to the level of physical activity, low 125 (41.7%) predominated and at risk of type 2 diabetes mellitus, 113 (27.7%) predominated. In conclusion, the population should be educated to have an adequate lifestyle by improving physical activity to prevent the risk of contracting diabetes.

1. Introduction

Type 2 diabetes is currently considered one of the non-communicable diseases with the greatest negative impact and projection on the quality of life of the world population and represents a world health problem [1]; it is diagnosed in 85-95% of all those who suffer from it, with a higher percentage in less developed regions (one in 14 adults will have diabetes mellitus). It is estimated that nearly four million deaths a year are directly related to this condition, based on changes in the lifestyle of the population, associated with sedentary lifestyle, obesity, high blood pressure and other cardiovascular risk factors [2]. For these reasons, it is the fourth cause of death in the entire universe [3].

Globally, the number of patients diagnosed with type 2 diabetes mellitus has exceeded the references by more than 100% increase in the last 20 years [4]. In its development, it is related to lifestyle, age, sex, ethnicity/race and family history of Diabetes Mellitus Type 2 (DMT2) [5]. The Body Mass Index (BMI) increased means the probability of developing DMT2, to a greater extent in individuals with a Normal BMI compared to the obese, identified physical activity as a fundamental strategy in its prevention [6].

In Peru, the number of people with diabetes mellitus is increasing very rapidly and the main cause of its rapid increase is the important change in the lifestyle of the Peruvian population, whose current characteristics represent an excessive intake of high-calorie foods (junk food and sugary drinks), and what is equally worrisome: a reduction in physical activity that directly leads to a high morbidity of overweight and obesity [7].

Regular physical activity is essential for disease prevention, particularly chronic non-communicable diseases; Health promotion; quality of life; and reduction of mortality [8].

In [9], the study prevalence and incidence of type 2 diabetes mellitus in the general adult population of Peru, they define the prevalence of type 2 diabetes mellitus nationwide according to the National Survey of Nutritional, Biochemical, Socioeconomic and Cultural Indicators study was 5.1% in 2004-2005, and it increased to 7.0% in 2010-12, Perú Diabetes (PERUDIAB, national study).

* Corresponding Author: Brian Meneses-Claudio, Sr., +51 1 950159924 & bmeneses@uch.edu.pe

The Risk Factors of Non-communicable Diseases study in Lima and Callao reported a prevalence of 3.9% in 2006, and in 2010-2012 the Centre of Excellence in Chronic Diseases (CRONICAS) study also in Lima found this indicator at 5.5%. Regarding the incidence, three of the nine studies analyzed PERUDIAB estimated the incidence of diabetes in 19.5 new cases per 1,000 person-years. The CRONICAS study, in selected populations of Lima, Puno and Tumbes, found the same result. The Perú's Rural to Urban Migrants (PERU MIGRANT) study estimated a 50% lower incidence than the previous two studies, although more than half of the population in this study lived in rural areas or were migrants from rural to urban areas.

In [10], the author studied physical activity in outpatients with diabetes mellitus type 2 concluded that 88% of diabetic patients had inactivity or minimal physical activity, estimating that patients with diabetes have a higher level of inactivity than the general population.

In [11], the author indicates that leading a healthy lifestyle pattern could lead to a lower risk of diabetes mellitus type 2, which identifies 142 associations related to the risk of developing DMT2. In which, they pointed out that Eleven associations (8%) presented evidence for the risk of DMT2 (low consumption of whole grains, metabolically healthy obesity, increased sedentary time). Thirty-four associations (24%) were associated with an increased risk of DMT2 and presented highly suggestive evidence, they were the following: high BMI, low educational level, gestational diabetes, increased consumption of processed meat, high level of total physical activity. Twenty-nine associations had suggestive evidence (20%), and 42 associations had weak evidence (30%) of DMT2 risk. All but six associations with convincing or highly suggestive evidence were based exclusively on prospective cohort studies.

The main objective of the study is to determine the risk of developing diabetes mellitus type 2 in relation to its level of physical activity in adults treated in health centers in a district of North Lima. It is important the need to modify the lifestyles of the population, associated with sedentary lifestyle, obesity, high blood pressure and other cardiovascular risk factors and thus identify physical activity as a fundamental strategy in its prevention.

Diabetes mellitus risk is understood as the probability of presenting or suffering from type 2 diabetes in adults attended in District-Rímac health centers, which can be evaluated by the presence of indicators such as age, body mass index, anthropometry waist, physical activity, consumption of vegetables or fruits, blood pressure medication, diagnosis of high glucose values and a history of diabetes mellitus [12].

The technique to be used in the research work will be the survey, which a significant amount of data will be obtained optimally and effectively. To measure the risk of type 2 diabetes mellitus, the data collection instrument will be used, the Finnish Diabetes Risk Score (FINDRISK) test that considers eight items grouped into eight dimensions and the level of physical activity will also be evaluated using the International Physical Activity Questionnaire (IPAQ), which considers 27 items grouped into three dimensions.

The data collection process will be carried out during the month of June through surveys, after the data will initially be entered into a data matrix designed in the Microsoft Excel 2013 program, then will be exported to the IBM SPSS Statistics version 24.0 program. It will apply descriptive statistical tests such as absolute frequencies, relative frequencies, and measures of central tendency, determining the risk of developing diabetes and the level of physical activity.

The following research work is structured as follows: In section II, the methodology that will determine the research work will be presented, with emphasis on the instrument, techniques to be used, scope of the study, valid inclusion, and exclusion criteria for the estimated population, among others. In section III, the results will show the tables and graphs obtained after data processing with an emphasis on determining the risk of developing type 2 diabetes mellitus and the level of physical activity of the studied population. In section IV, it will present the discussions of the research work and finally in section V, the conclusions obtained through the present research work, as well as the future work that leads to the timely detection of type 2 diabetes mellitus and the adoption of healthy styles.

2. Methods

The steps to carry out the survey as the population and sample were obtained, type of study, inclusion and exclusion criteria, analysis of the variables, data collection, location and application of the IPAQ and Finnish Diabetes Risk Score (FINDRISK) instruments. The paper before its application went through an ethics committee in which it was approved.

2.1. Population and Sample

It was applied to 300 outpatients at risk of developing diabetes, in the nine Health Centers of the Rimac district, considering the total distribution of patients who participated in the study, the sample corresponds to 35 participants from each center: H.C. Rímac, H.C. Caquetá and H.C. Flor de Amancaes. Likewise, 34 participants from Leoncio Prado, San Juan de Amancaes, Mariscal Castilla, Ciudad y Campo and Villa Los Ángeles.

2.2. Inclusion and Exclusion Criteria

Inclusion criterion is that all patients who receive regular care and wish to collaborate voluntarily in being evaluated to obtain a risk diagnosis participate and exclusion criteria; Patients who did not wish to participate in the study are excluded.

2.3. Type of study

The present work, due to its data collection characteristics and the measurement of the variables, has a quantitative approach, a non-experimental, descriptive and correlational study.

2.4. Analysis of the variables

The present study presents a main variable that is the risk of diabetes, the data collection instrument was used, the FINDRISK test that comprises eight items grouped into eight dimensions and also can be evaluated through the presence of indicators such as age, body mass index, anthropometry of the waist, physical activity, consumption of vegetables or fruits, medication for blood

pressure, diagnosis of high glucose values and a history of diabetes mellitus [13].

The level of physical activity will be evaluated using the IPAQ questionnaire. The short version evaluates three specific characteristics of activity: intensity (mild, moderate or vigorous), frequency (measured in days per week) and duration (time per day). Activity of moderate intensity is considered as which produces a moderate increase in respiration, heart rate and sweating for at least 10 continuous minutes, and vigorous activity, such as which produces a greater increase in the same variables, for 10 minutes or more [14].

2.5. Instruments

The data collection technique was the survey and the selected instruments were the FINDRISK test and the short version IPAQ questionnaire to measure the level of physical activity.

The data collection process will be carried out during the month of June through the surveys, later the data will be initially entered into a data matrix designed in the Microsoft Excel 2013 program, and then they will be exported to the IBM SPSS Statistics version 24.0 program. And the data analysis will apply descriptive statistical tests such as absolute frequencies, relative frequencies and measures of central tendency, determining the risk of developing diabetes and the level of physical activity.

The instrument was applied in the nine establishments of the Rímac district. The respective permission was requested from Directorate of Integrated Health Networks (DIRIS) Lima Norte, to have access to health centers, for this purpose an enrollment was made in the ethics committee to approve the instrument to apply; After approval, it was presented to the head of each establishment and permission was given to the triage health personnel in charge of the area to develop the questionnaire.



Figure 1: The patient is triage taking anthropometric measurements, before going to the external office to be evaluated

As shown in Figure 1, the instructions on how to carry out the respective filling in of the questionnaires are explained to the patients, they were informed to mark a single alternative in the questions avoiding smearing, and the patients were also consulted to ask before any doubt that they would not understand, to obtain real data.

Figure 2 considers the distribution according to sex of the total number of people who participated in the study, segmented by each health center in the Rímac district, an analysis that considers 35 HC respondents: Rímac, Caquetá and Flor de Amancaes; the second block considers the HC with 34 respondents: Leoncio Prado, San Juan de Amancaes, Mariscal Castilla, Ciudad y Campo and Villa Los Angeles; and finally the block of 25 respondents the HC Piedra Liza, an activity that was carried out according to schedule and considered the minimum established quota.



Figure 2: Filling out the survey of outpatients treated in the health center of a district of North Lima - 2019 (N = 300)

2.6. Developing the surveys

The patients were interested in reading the questionnaire; some patients were hesitant to answer because of their advanced age they did not understand the questions, they did not interpret it in the health field.

Based on good data collection in most of the patients with good attitude and with consent, they collaborated in the survey. In some patients, they did not give their consent because of the time or they had another meeting, the discomfort of the patient and/or relatives when waiting for care was also observed, so they arrived before the mentioned time and were not treated, since that this has an impact on answering the survey.

It is important in the data collection, the technical staff of the triage service since their function is to take the anthropometric measurements (weight, height) and the abdominal perimeter of all the patients before entering the medical office and thus be diagnosed and evaluated.

3. Result

After collecting the data, graphic tables were generated where it will synthesize the information and graphics so that the results obtained are easier to understand.

Table 1: Sociodemographic data on the risk of type 2 diabetes and physical activity in outpatients treated at health centers in a district of North Lima, 2019 (n=300)

Sex of hospitalized patient	**Frequency (f)**	**Percentage (%)**
Female	214	71.3
Male	86	28.7
Health Center	**Frequency (f)**	**Percentage (%)**
CMI Rímac	35	11.7
Caquetá	35	11.7
Flor de Amancaes	35	11.7
Leoncio Prado	34	11.3

San Juan de Amancaes	34	11.3
Mariscal Castilla	34	11.3
Ciudad y Campo	34	11.3
Villa los Ángeles	34	11.3
Piedra Liza	25	8.3
Degree of Study	**Frequency (f)**	**Percentage (%)**
No studies	2	0.7
Primary Incomplete	2	0.7
Primary Complete	4	1.3
High School Incomplete	18	6.0
High School Complete	91	30.3
University study Incomplete	41	13.7
University study Complete	142	47.3
Marital Status	**Frequency (f)**	**Percentage (%)**
Single	121	40.3
Cohabitant	87	29.0
Married	65	21.7
Separated	7	2.3
Divorced	6	2.0
Widowed	14	4.7
Type of Family	**Frequency (f)**	**Percentage (%)**
Nuclear	111	37.0
Single parent	36	12.0
Extended	71	23.7
Expanded	40	13.3
Reconstituted	7	2.3
Equivalent	9	3.0
Single Person	26	8.7
Mental health disorder	**Frequency (f)**	**Percentage (%)**
Depression	3	1.0
Family violence	3	1.0
Not have	294	98.0

Table 1 considers the socio-demographic data regarding the sex of the participants, most of the surveyed sample, equivalent to 71.3%, is female. There are nine health centers in North Lima that were part of the study, which between 34 and 35 patients (11.7% and 11.3% respectively) for each health center made up the study sample, except for Piedra Liza Center, which had 25 patients (8.3%). Likewise, the data of the surveyed sample (47.3%) has completed higher education, followed as a percentage of those with complete secondary education (30.3%). Regarding marital status, most respondents (40.3%) were "single". In a similar way, 37% of the surveyed patients denied having a mental health disorder.

Figure 3 considers the history of diabetes, which the risk is considered considering direct affinity, in which 108 patients are observed, representing 36% have no history, that is, relatives with a diagnosis of diabetes, followed by 104 patients representing 34.7% consider that there is a history of diabetes (grandparents, aunt, uncle, cousin, brother) and lastly 88 patients representing 29.3% consider that there is a history of diabetes (parents, siblings, children), they consider the line Transmissible with higher prevalence.

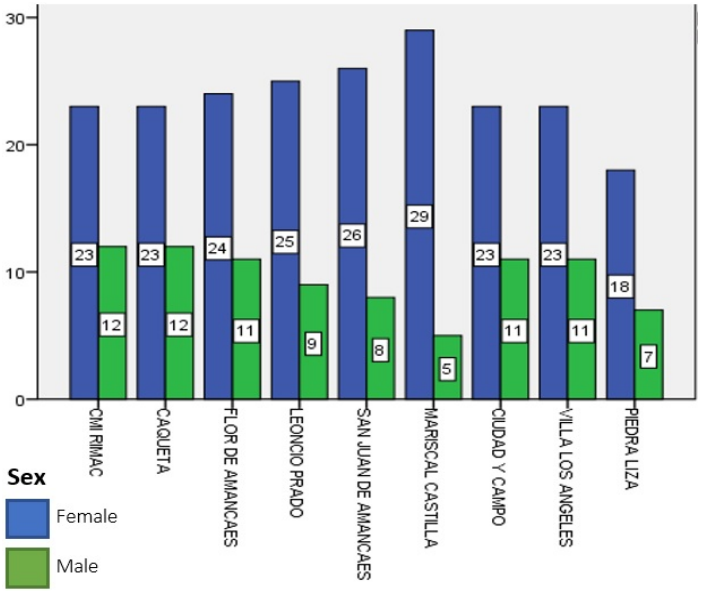


Figure 3: History of diabetes mellitus in outpatients attended in different health centers of a district of North Lima - 2019 (N = 300)

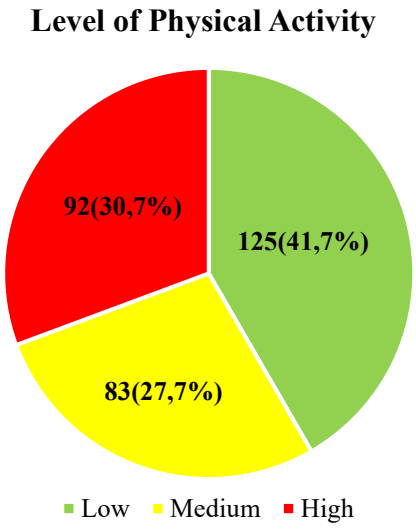


Figure 4: Patients surveyed who perform physical activity, cared for in the establishments of a District of North Lima - 2019 (n=300)

In Figure 4 regarding the level of physical activity, it is observed that 92 (30,7%) of the patients present a high level of physical activity, 83 (27,7%) a medium level of physical activity and 125 (41,7%) low level of physical activity.

In Figure 5 regarding the risk of type 2 diabetes mellitus, it is observed that 79 (26,3%) of the patients present a low risk, 113 (27,7%) present a slightly high risk, 46 (15,3%) present a moderate risk, 52 (17,3%) present a high risk and 10 (3,3%) of the patients present an extremely high risk.

Table 2: Relationship between the Risk of type 2 diabetes mellitus and Level of physical activity in outpatients attended in health centers of a district of North Lima, 2019 (n = 300) based on the Pearson's Chi-square statistical test (X^2)

			Level of physical activity			Total
			Low	Moderate	High	
Risk of type 2 diabetes mellitus	Low	Count	30	27	22	79
		% of total	10,0%	9,0%	7,3%	26,3%
	Slightly high	Count	55	28	30	113
		% of total	18,3%	9,3%	10,0%	37,7%
	Moderate	Count	17	10	19	46
		% of total	5,7%	3,3%	6,3%	15,3%
	High	Count	19	17	16	52
		% of total	6,3%	5,7%	5,3%	17,3%
	Very High	Count	4	1	5	10
		% of total	1,3%	0,3%	1,7%	3,3%
Total		Count	125	83	92	300
		% of total	41,7%	27,7%	30,7%	100,0%

Chi-square tests	Value	gl	Sig. (p)
	9,653	8	0,290

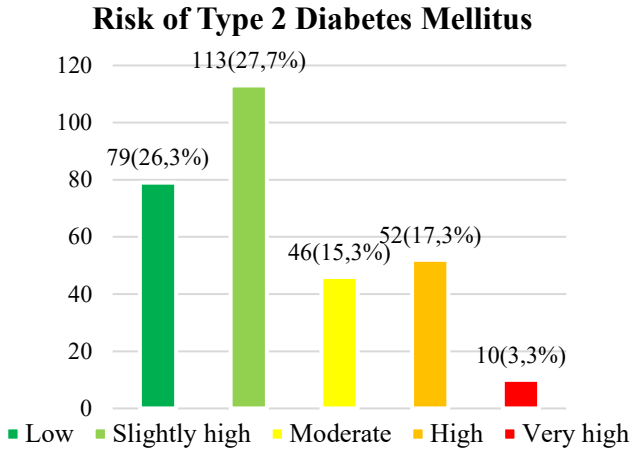


Figure 5: Risk of type 2 diabetes mellitus in patients treated in establishments in a District of North Lima - 2019 (n = 300)

Table 2 shows the relationship between risk of type 2 diabetes mellitus and level of physical activity in outpatients treated in health centers in a district of North Lima, where the significance value of the chi-square test is of 0,290 (p> 0.05) (X^2 = 9,653; g.l. = 8), therefore, there is insufficient evidence to deny the dissociation hypothesis. Given the above, it is concluded that there is no significant relationship between the risk of type 2 diabetes and physical activity in outpatients in health facilities in a district of northern Lima, 2019 with a confidence level of 0,05 (95%).

The importance of obtaining data allows us to identify the relationship between physical activity and the risk of T2DM and which factors are risk factors, which we can carry out prevention strategies that allow maintaining the well-being of the person, also serves for future research with studies that allow evaluating the relationship between physical activity and risk of T2DM in which they can look for the effectiveness of the strategies applied in this research work.

4. Discussion

For the present research work, the focus was given to the patient on their physical health, although we can affirm that the risk of having type 2 diabetes mellitus today is one of the problems in our country due to the sedentary lifestyle that people present due to physical inactivity.

In the results on the level of physical activity, the patients present a low level of physical activity 125 (41,7%), this is since the patients have an inappropriate lifestyle, where sedentary lifestyle is the first factor present when physical activities are not performed, therefore patients tend to have an increased risk of non-communicable diseases that can compromise their health throughout their lives. In [15], the author argues that sedentary lifestyle is a predisposing factor for metabolic diseases where long hours spent at a computer, being on the phone and the eating of fast-food foods that provides a high caloric index, that is not waste by not performing any physical activity, compromises the health of the person generating health problems.

In the risk results for T2DM, patients present a slight risk of having this disease 113 (27,7%), this is since the prevalence of morbidity for contracting T2DM is increasingly considered higher, because people nowadays they manage an inadequate lifestyle, a poor diet, diminished family activities, also family history, where the person who has a relative with this disease if it does not have

an adequate lifestyle can develop the disease. Therefore, adopting prevention and control strategies will reduce the risk of contracting T2DM. In [16], the author argue that family history of diabetes risk tends to increase the degree of susceptibility to the disease and that it is not possible to modify it, but also changes in the lifestyle of the person as a preventive measure by modifying food consumption, increasing physical activity and reducing body weight, are actions that allow the person to increase the benefits in their own health, ensuring well-being and trying to maintain a healthy life. Likewise, in [17], they argue that the self-care that patients with diabetes mellitus should exercise in themselves is important because it is related to the actions that each person take for their own benefit of their health including physical activity, proper nutrition, and self-control.

5. Conclusion

It is concluded that the population should be educated about the importance of increasing physical activity and of a healthy lifestyle where the well-being of health in the person is balanced.

It is concluded that nutritional counseling should be provided when making home visits to patients with diabetes mellitus.

It is concluded that the strengthening and control of prevention programs in primary care related to chronic diseases, including type 2 diabetes mellitus, should be considered.

It is concluded that preventive actions should be taken in the patient with diabetes mellitus such as a healthy lifestyle, a diet program that suits the patient and physical activity.

The main limitations in this research is the understanding that patients had about filling out the surveys, also, the time they had to fill out the surveys was also limited because the patients were waiting to be attended in the health centers and some did not want to fill out the survey; most of the patients voluntarily agreed to participate in the research work to find out if they had any risk of contracting T2DM.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] A. Olokoba, O. Obateru, L. Olokoba, "Type 2 diabetes mellitus: A review of current trends.," Oman Medical Journal, **27**(4), 269–273, 2012, doi:10.5001/omj.2012.68.

[2] M. Mazidi, P. Toth, M. Banach, "C-reactive Protein Is Associated With Prevalence of the Metabolic Syndrome, Hypertension, and Diabetes Mellitus in US Adults.," Angiology, **69**(5), 438–442, 2018, doi:10.1177/0003319717729288.

[3] L. Slater, S. Stuckey, R. Chandra, "Tratamiento actual de la diabetes mellitus tipo 2.," Neuroimaging Pharmacopoeia, **20**(1), 289–292, 2016, doi:10.1007/978-3-319-12715-6_40.

[4] A. Koopman, J. Beulens, E. Voerman, P. Simone, A. Heijden, T. McDonald, M. Langendoen, F. Rutters, "The association between GAD65 antibody levels and incident Type 2 Diabetes Mellitus in an adult population: A meta-analysis.," Metabolism: Clinical and Experimental, **95**, 1–7, 2019, doi:10.1016/j.metabol.2019.03.001.

[5] A. Sharma, L. Sikora, S. Bush, "Management of Diabetes Mellitus in Adults at the End of Life: A Review of Recent Literature and Guidelines.," Journal of Palliative Medicine, **22**(9), 1133–1138, 2019, doi:10.1089/jpm.2018.0614.

[6] F. Petermann, X. Díaz, Á. Garrido, A. Leiva, M. Martínez, C. Salas, F. Poblete, C. Celis, "Association between type 2 diabetes and physical activity in individuals with family history of diabetes.," Gaceta Sanitaria, **32**(3), 230–235, 2018, doi:10.1016/j.gaceta.2017.09.008.

[7] A. Pinilla, M. Barrera, "Prevención en diabetes mellitus y riesgo cardiovascular: enfoque médico y nutricional.," Revista Facultad de Medicina, **66**(3), 459–468, 2018, doi:10.15446/revfacmed.v66n3.60060.

[8] S. Seclén, "Diabetes Mellitus en el Perú: hacia dónde vamos.," Revista Medica Herediana, **26**, 3–4, 2015, doi:10.20453/rmh.2015.2340.

[9] R. Carrillo, A. Bernabé, "DIABETES MELLITUS TIPO 2 EN PERÚ: UNA REVISIÓN SISTEMÁTICA SOBRE LA PREVALENCIA E INCIDENCIA EN POBLACIÓN GENERAL.," Rev. Perú Med. Exp. Salud Publica, **36**(1), 26–36, 2019, doi:10.17843/rpmesp.2019.361.4027.26.

[10] A. Manzaneda, M. Lazo, G. Málaga, "Physical activity in outpatients with type 2 diabetes in a national hospital of Perú.," Rev. Perú Med. Exp. Salud Pública, **32**(2), 311–315, 2015.

[11] V. Bellou, L. Belbais, L. Tzoulaki, E. Evangelou, "Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses.," PLoS ONE, **13**(3), 1–27, 2018.

[12] Å. Bjørn, K. Midthjell, S. Krokstad, V. Rangul, A. Bauman, "Prolonged sitting may increase diabetes risk in physically inactive individuals: an 11 year follow-up of the HUNT Study, Norway.," Diabetologia, **60**(5), 830–835, 2017, doi:10.1007/s00125-016-4193-z.

[13] I. Mendiola, I. Urbina, A. Muñoz, G. Juanico, G. López, "Evaluación del desempeño del Finnish Diabetes Risk Score (findrisc) como prueba de tamizaje para diabetes mellitus tipo 2.," Atención Familiar, **25**(1), 22–26, 2017, doi:10.22201/facmed.14058871p.2018.1.62925.

[14] R. Barrera, "Cuestionario Internacional de actividad física (IPAQ)," Revista Enfermería Del Trabajo, **7**(2), 49–54, 2017.

[15] B. Pérez, "Salud: entre la actividad física y el sedentarismo.," An Venez Nutr, **27**(1), 119–128, 2014.

[16] Y. Llorente, P. Soca, D. Rivas, Y. Borrego, "Factores de riesgo asociados con la aparición de diabetes mellitus tipo 2 en personas adultas.," Revista Cubana de Endocrinología, **27**(2), 123–133, 2016.

[17] M. Soler, E. Pérez, M. López, D. Quezada, "Conocimientos y autocuidado en pacientes con diabetes mellitus tipo 2.," Archivo Médico de Camagüey, **20**(3), 244–252, 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1683-1689 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Sentiment Analysis in English Texts

Arwa Alshamsi[1], Reem Bayari[1], Said Salloum[2,3,*]

[1]*Faculty of Engineering & IT, The British University, Dubai, 345015, UAE*

[2]*Research Institute of Sciences & Engineering, University of Sharjah, Sharjah, 27272, UAE*

[3]*School of Science, Engineering, and Environment, University of Salford, Manchester, M5 4WT, UK*

ARTICLE INFO

Article history:
Received: 23 September, 2020
Accepted: 24 December, 2020
Online: 28 December, 2020

Keywords:
Sentiment Analysis
Balanced Dataset
Unbalanced Dataset
Classification

ABSTRACT

The growing popularity of social media sites has generated a massive amount of data that attracted researchers, decision-makers, and companies to investigate people's opinions and thoughts in various fields. Sentiment analysis is considered an emerging topic recently. Decision-makers, companies, and service providers as well-considered sentiment analysis as a valuable tool for improvement. This research paper aims to obtain a dataset of tweets and apply different machine learning algorithms to analyze and classify texts. This research paper explored text classification accuracy while using different classifiers for classifying balanced and unbalanced datasets. It was found that the performance of different classifiers varied depending on the size of the dataset. The results also revealed that the Naive Byes and ID3 gave a better accuracy level than other classifiers, and the performance was better with the balanced datasets. The different classifiers (K-NN, Decision Tree, Random Forest, and Random Tree) gave a better performance with the unbalanced datasets.

1. Introduction

The recent widening expansion of social media has changed communication, sharing, and obtaining information [1–4]. In addition to this, many companies use social media to evaluate their business performance by analysing the conversations' contents [5]. This includes collecting customers' opinions about services, facilities, and products. Exploring this data plays a vital role in consumer retention by improving the quality of services [6,7]. Social media sites such as Instagram, Facebook, and Twitter offer valuable data that can be used by business owners not only to track and analyse customers' opinions about their businesses but also that of their competitors [8–11]. Moreover, these valuable data attracted decision-makers who seek to improve the services provided [8,9,12].

In this research paper, several research papers that studied Twitter's data classification and analysis for different purposes were surveyed to investigate the methodologies and approaches utilized for text classification. The authors of this research paper aim to obtain open-source datasets then conduct text classification experiments using machine learning approaches by applying different classification algorithms, i.e., classifiers. The authors utilized several classifiers to classify texts of two versions of datasets. The first version is unbalanced datasets, and the second is balanced datasets. The authors then compared the classification accuracy for each used classifier on classifying texts of both datasets.

2. Literature Review

As social media websites have attracted millions of users, these websites store a massive number of texts generated by users of these websites [13–20]. Researchers were interested in investigating these metadata for search purposes [16,17,21–24]. In this section, a number of research papers that explored the analysis and classification of Twitter metadata were surveyed to investigate different text classification approaches [25] and the text classification results.

Researchers of [26] investigated the user's gender of Twitter. Authors noticed that many Twitter users use the URL section of the profile to point to their blogs, and the blogs provided valuable demographic information about the users. Using this method, the authors created a corpus of about 184000 Twitter users labeled with their gender. Then authors arranged the dataset for

*Corresponding Author: Said Salloum, University of Sharjah, UAE. Tel: +971507679647 Email: ssalloum@sharjah.ac.ae

experiments as following: for each user; they specify four fields; the first field contains the text of the tweets and the remaining three fields from the user's profile on Twitter, i.e., full name, screen name, and description. After that, the authors conducted the experiments and found that using all of the dataset fields while classifying Twitter user's gender provides the best accuracy of 92%. Using tweets text only for classifying Twitter user's gender provides an accuracy of 76%. In [27], the authors used Machine Learning approaches for Sentiment Analysis. Authors constructed a dataset consisting of more than 151000 Arabic tweets labeled as "75,774 positive tweets and 75,774 negative tweets". Several Machine Learning Algorithms were applied, such as Naive Bayes (NB), AdaBoost, Support vector machine (SVM), ME, and Round Robin (RR). The authors found that RR provided the most accurate results on classifying texts, while AdaBoost classifier results were the least accurate results. A study by [28] interested as well in Sentiment Analysis of Arabic texts. The authors constructed the Arabic Sentiment Tweets Dataset ASTD, which consists of 84,000 Arabic tweets. The number of tweets remaining after annotation was around 10,000 tweets. The authors applied machine learning approaches using classifiers on the collected dataset. They reported the following: (1) The best classifier applied on the dataset is SVM, (2) Classifying a balanced set is challenging compared to the unbalanced set. The balanced set has fewer tweets than the unbalanced set, which may negatively affect the classification's reliability. In [29], the author investigated the effects of applying preprocessing methods before the sentiment classification of the text. The authors used classifiers and five datasets to evaluate the preprocessing method's effects on the classification. Experiments were conducted, and researchers reported the following findings: Removing URL has no much effect, Removing stop words have a slight effect, Removing Numbers have no effect, Expanding Acronym improved the classification performance, and the same preprocessing methods have the same effects on the classifier's performance, NB and RF classifiers showed more sensitivity than LR and SVM classifiers. In conclusion, the classifier's performance for sentiment analysis was improved after applying preprocessing methods. A study by [30] investigated Twitter geotagged data to construct a national database of people's health behavior. The authors compared indicators generated by machine learning algorithms to indicators generated by a human. The authors collected around 80 million geotagged tweets. Then Spatial Join procedures were applied, and 99.8% of tweets were successfully linked. Then tweets were processed. After that, machine learning approaches were used and successfully applied in classifying tweets into happy and not happy with high accuracy. In [31] explored classifying sentiments in movie reviews. The authors constructed a dataset of 21,000 tweets of movie reviews. Dataset split into train set and test set. Preprocessing methods applied, then two classifiers, i.e., NB and SVM, were used to classify tweets text into positive or negative sentiment. The authors found that better accuracy achieved using SVM of 75% while NB has 65% accuracy. Researchers of [32] used Machine Learning methods and Semantic Analysis for analyzing tweet's sentiments. Authors labeled tweets in a dataset that consists of 19340 sentences into positive or negative. They applied preprocessing methods after that features were extracted; authors applied Machine Learning approaches, i.e., Naïve Bayes, Maximum Entropy, and Support Vector Machine (SVM) classifiers after that Semantic Analysis were applied. The authors found that Naïve Bayes provided the best accuracy of 88.2, the next SVM of 85.5, and the last is Maximum entropy of 83.8. The authors reported as well that after applying Semantic Analysis, the accuracy increased to reach 89.9. In [33], the authors analyzed sentiments by utilizing games. Authors introduced TSentiment, which is a web-based game. TSentiment used for emotion identification in Italian tweets. TSentiment is an online game in which the users compete to classify tweets in the dataset consists of 59,446 tweets. Users first must evaluate the tweet's polarity, i.e., positive, negative, and neutral, then users have to select the tweet's sentiment from a pre-defined list of 9 sentiments in which 3 sentiments identified for the positive polarity, 3 sentiments identified for negative polarity. Neutral polarity is used for tweets that have no sentiment expressions. This approach for classifying tweets was effective.

A study by [34] examined the possibility of enhancing the accuracy of predictions of stock market indicators using Twitter data sentiment analysis. The authors used a lexicon-based approach to determine eight specific emotions in over 755 million tweets. The authors applied algorithms to predict DJIA and S&P500 indicators using Support Vectors Machine (SVM) and Neural Networks (NN). Using the SVM algorithm in DJIA indication, the best average precision rate of 64.10 percent was achieved. The authors indicated that the accuracy could be increased by increasing the straining period and by improving the algorithms for sentiment analysis. authors conclude that adding Twitter details does not improve accuracy significantly. In [35], the authors applied sentiment analysis on around 4,432 tweets to collect opinions on Oman tourism, they build a domain-specific ontology for Oman tourism using Concept Net. Researchers constructed a sentiment lexicon based on three existing lexicons, SentiStrength, SentWordNet, and Opinion lexicon. The authors randomly divide 80% of the data for the training set and 20% for testing. The researcher used two types of semantic sentiment, Contextual Semantic Sentiment Analysis, and Conceptual Semantic Sentiment Analysis. Authors applied Naïve Base supervised machine learning classifier and found that using conceptual semantic sentiment analysis expressively improves the sentiment analysis's performance. A study by [36] used sentiment analysis and subjectivity analysis methods to analyze French tweets and predict the French CAC40 stock market. The author used a French dataset that consists of 1000 positive and negative book reviews. The author trained the neural network by using three input features on 3/4 of the data, and he tested on the remaining quarter. The achieved accuracy 80% and a mean absolute percentage error (MAPE) of 2.97%, which is less than the work reported by Johan Bollen. The author suggested adding more features as input to improve the performance. In [37], the authors examined the relationship between Twitter's social emotion and the stock market. Researchers collected millions of tweets by Twitter API. Researchers retrieved the NASDAQ market closing price in the same period. The authors applied the correlation coefficient. Authors conclude that emotion-related terms have some degree of influence on the stock market overall trend, but it did not meet standards that can be used as a guide to stock-market prediction. While at the same time, there was a fairly close association between positive, negative, and angry mood-words. Particularly sad language tends to have a far greater influence on the stock market than other groups. In [38], the authors investigated telecommunications companies' conversation on

social media Twitter ('indihome,' in Indonesia). The authors collected 10,839 raw data for segmentation. The authors collected data: over 5 periods of time in the same year. Authors found that most of the tweets (7,253) do not contain customers' perception toward Indihome. Only 3,586 tweets are containing the perception of customers toward Indihome. Most of the data contained perception reveal that the customers have the negative perception (3,119) on Indihome and only 467 tweets contain positive perceptions; the biggest number of negative perceptions relate to the first product, the second relates to a process, third relate to people, and fourth relate to pricing. Researchers of [39] examined prevalence and geographic variations for opinion polarities about e-cigarettes on Twitter. Researchers collected data from Twitter by pre-defined seven keywords. They classified the tweets into four categories: Irrelevant to e-cigarettes, Commercial tweets, organic tweets with attitudes (supporting or against or neutral) the use of e-cigarettes, and the geographic locations information city and state. Researchers selected six socio-economic variables from Census data 2014 that are associated with smoking and health disparities. Researchers classified the tweets based on a combination of human judgment and machine-learning algorithms, and two coders classified a random sample of 2000 tweets into five categories. The researcher applied a multilabel Naïve Bayes classifier algorithm; the model achieved an accuracy of 93.6% on the training data. Then the researcher applied the machine learning algorithm to a full set of collected tweets and found the accuracy of the validation data was 83.4%. To evaluate the socio-economic impact related to public perception regarding e-cigarette use in the USA, researchers calculated the Pearson correlation between prevalence and percentage of opinion polarities and selected ACS variants for 50 states and the District of Columbia. In [40], the authors Investigated the link between any updates on certain brands and their reaction. Researchers gathered geographic locations based on the data to see consumer distribution. Researchers collected Twitter data by using the REST API. In total, 3,200, from ten different profiles, then used sentiment analysis to differentiate between clustered data expressed positively or negatively then resampled the result in an object model and cluster. For every answer, the researcher has been evaluated for the textual sentiment analysis from the object model. Researchers used AFINN based word list and Sentiments of Emojis to run comprehensive sentiment analysis; for the data that not existed in the word list, researcher added a separated layer to an analysis by using emoji analysis on top of sentiment analysis, and authors did not see any difference in the level of accuracy when applying this extra layer. The researcher found some Sentiment Analysis weaknesses related to the misuse of emoji, the use of abbreviated words or terms of slang, and the use of sarcasm. In [41], the authors proposed an application that can classify a Twitter content into spam or legitimate. Auhtors used an integrated approach, from URL analysis, Natural Language Processing, and Machine Learning techniques. Auhtors analyzed the URL that derived from the tweets, then convert URLs to their long-form, then compare URLs with Blacklisted URLs, then compare them with a set pre-defined expressions list as spam; the presence of any of these expressions can conclude that the URL is spam. After cleaning data, the stemmed keywords are compared with the per set of identified spam words and, if a pre-defined expressions list are found in the tweet, then the user is classified as spam. Six features were used for classification. The training set has 100 instances with six features and a label. The author used Naïve-Bayes algorithm. Authors manually examined 100 users and found (60 were legitimate and 40 were spam) then the sampled checked by the application and the result presented that 98 were classified correctly.

3. Proposed Approach

In this work, the authors implemented and evaluated different classifiers in classifying the sentiment of the tweets. It's by utilizing RapidMiner software. Classifiers were applied on both balanced and unbalanced datasets. Classifiers used are Decision Tree, Naïve Bayes, Random Forest, K-NN, ID3, and Random Tree.

4. Experiment Setup

In this section, the dataset is described as well as the settings and evaluation techniques are used in the experiments have been discussed. The prediction for the tweet category is tested twice—the first time on an unbalanced data set and the second time on a balanced dataset as below.

- **Experiments on the unbalanced dataset:** Decision Tree, Naïve Bayes, Random Forest, K-NN, ID3, and Random Tree classifiers were applied on six unbalanced datasets.
- **Experiments on the balanced dataset:** In this experiment, the challenges related to unbalanced datasets were tackled by manual procedures to avoid biased predictions and misleading accuracy. The majority class in each dataset almost equalized with the minority classes, i.e., many positive, negative, and neutral, practically the same in the balanced dataset as represented in Table 3.

4.1. Dataset Description

We obtained a dataset from Kaggle, one of the largest online data science communities in this work. It consists of more than 14000 tweets, labeled either (positive, negative, or neutral). The dataset was also split into six datasets; each dataset includes tweets about one of six American airline companies (United, Delta, Southwest, Virgin America, US Airways, and American). Firstly, we summarized the details about the obtained datasets, as illustrated in Table 1 below.

Table 1: Summary of obtained Dataset

	American Airline Companies					
	Virgin America	United	Delta	Southwest	US Airways	American
Number of Tweets	504	3822	2222	2420	2913	2759
Positive Tweets	152	492	544	570	269	336
Negative Tweets	181	2633	955	1186	2263	1960
Neutral Tweets	171	697	723	664	381	463

4.2. Dataset Cleansing

In this section, the authors described the followed procedure in the dataset preparation. The authors utilized RapidMinor software for tweet classification. Authors followed the methods described below:

1) Splitting the dataset into a training set and test set.

2) Loading the dataset, i.e., excel file into RapidMinor software using Read Excel operator.

3) Applying preprocessing by utilizing the below operators.

- Transform Cases operator to transform text to lowercase.
- Tokenize operator to split the text into a sequence of tokens.
- Filter Stop words operator to remove stop words such as: is, the, at, etc.
- Filter Tokens (by length) operator: to remove token based on the length, in this model, minimum characters are 3, and maximum characters are 20 any other tokens that don't match the rule will be removed.
- Stem operator: to convert words into base form.

4.3. Dataset Training

Each of the datasets was divided into two-part. The first part contains 66% of the total number of tweets of the data set, and it is used to train the machine to classify the data under one attribute, which is used to classify the tweets to either (positive or Negative or Neutral). The remaining 34% of tweets were used to classify tweets' attribute to (positive or Negative or Neutral), i.e., test set.

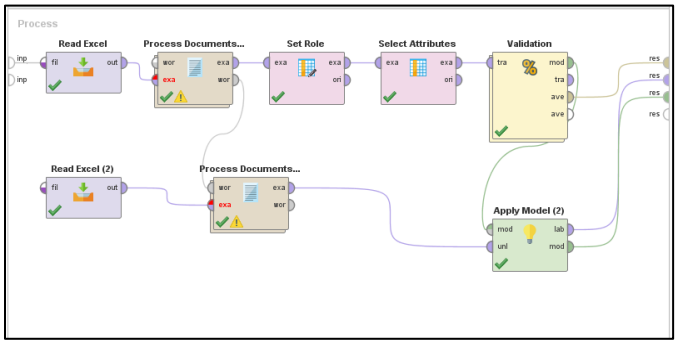


Figure 1: Summarization of the Process Model

4.4. Dataset Classifying

In this section, the authors described the steps in the tweet's classification techniques.

- Set Role operator is used to allow the system to identify sentiment as the target variable,
- Select Attributes operator is used to removing any attribute which has any missing values.
- Then in the validation operator, the dataset is divided into two parts (training and test). We used Two-thirds of the dataset to train the dataset and the last one-third to evaluate the model.
- Different machine learning algorithms are used for training the dataset (Decision Tree, Naïve Bayes, Random Forest, K-NN, ID3, and Random Tree).
- For testing the model, the Performance operator utilized to measure the performance of the model.

5. Experiment Results and Discussion

This section presented the experiment results in terms of accuracy level of prediction for each classifier on both types of datasets (balanced, unbalanced) and a comparison between the two experiments.

5.1. Experiment results for an unbalanced dataset

Figure 2 and Table 2 present the accuracy results of the utilized classifiers on the datasets.

Table 2: Accuracy results on unbalanced dataset

	Accuracy					
	Virgin America	United	Delta	Southwest	US Airways	American
Dataset	504	3822	2222	2420	2913	2759
Training set	333	2523	1467	1597	1923	1821
Test set	171	1299	755	823	990	938
Decision Tree	31.86%	72.03%	42.08%	50.46%	82.72%	68.98%
Naïve Bayes	32.74%	72.38%	42.28%	51.01%	82.72%	72.21%
Random Forest	31.86%	72.03%	42.08%	50.46%	82.72%	68.98%
K-NN	39.82%	11.66%	35.27%	50.46%	82.72%	69.43%
ID3	32.74%	72.38%	42.28%	51.01%	82.72%	72.21%
Random Tree	31.86%	72.03%	42.08%	50.46%	82.72%	68.98%

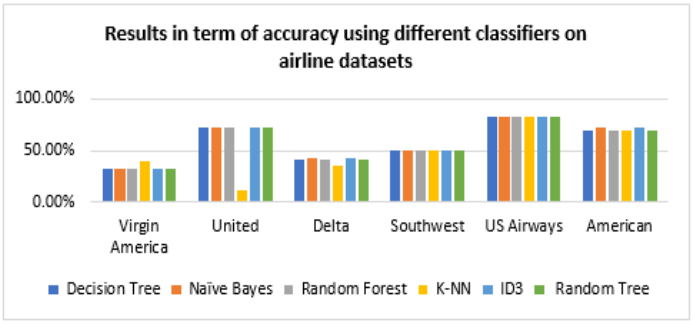


Figure 2: Accuracy results on unbalanced airline datasets using different classifiers

In some datasets, the classifier's accuracy results were very high, while it was low in others. All classifier's performance on the US airways dataset and United dataset provided the best accuracy due to the dataset's size, which was the largest. Naïve Bayes classifier, Decision Tree, and ID3 were mostly better than other classifiers and were given almost the same accuracy level. The classifiers with Virgin America dataset reported the lowest accuracy level due to the dataset's size, which is very small.

5.2. Experiment results for a balanced dataset

Decision Tree, Naïve Bayes, Random Forest, K-NN, ID3, and Random Tree classifiers were applied on the five obtained balanced datasets. (United, Delta, Southwest, and US Airways). The dataset for each was divided into two parts. The first part contains 66% of the total number of tweets of the data set, and it is used to train the machine to classify the data under one attribute, which is used to classify the tweets as either positive, Negative, or Neutral. The remaining 34% of tweets were used to classify tweets' attributes into (positive, Negative, or Neutral), i.e., test set.

Table 3: Number of tweets before and after balancing.

	Number of instances		Percentage		
	Total tweets before balancing	Total tweets after balancing	Positive	Negative	Neutral
United	3822	8276	33%	33%	34%
Delta	2222	2635	33%	33%	34%
Southwest	2420	5518	33%	33%	33%
US Airways	2913	6608	33%	33%	33%
American	2759	5924	34%	34%	33%

After applying different algorithms on the five balanced datasets, the performance, i.e., accuracy results, were reported in Table 4 and Figure 3 below:

Table 4: Accuracy results on the balanced dataset

	Accuracy					
	Virgin America	United	Delta	Southwest	US Airways	American
Dataset	8276	2635	5518	6608	5924	8276
Training set	5464	1743	3642	4363	3911	5464
Test set	2812	892	1876	2245	2013	2812
Decision Tree	35.06%	34.63%	34.35%	35.06%	33.98%	35.06%
Naïve Bayes	97.65%	36.99%	65.48%	97.65%	61.20%	97.65%
Random Forest	35.06%	34.63%	34.35%	35.06%	33.98%	35.06%
K-NN	38.79%	32.77%	35.32%	38.79%	39.47%	38.79%
ID3	97.65%	36.99%	65.48%	97.65%	61.20%	97.65%
Random Tree	35.06%	34.63%	34.35%	35.06%	33.98%	35.06%

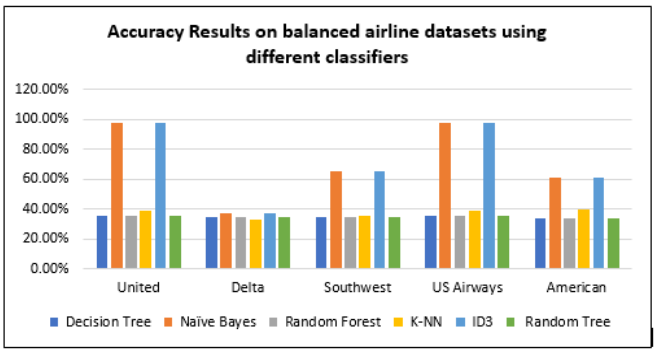


Figure 3: Accuracy results on balanced airline datasets using different classifiers

5.3. Comparison between two experiments results for each classifier

While comparing results between the performance of the classifiers on balanced and unbalanced datasets, it was found the following as seen in Figure 4 below:

5.3.1. Naive Byes and ID3

Gave the best accuracy than other classifiers in the two experiments. The accuracy level with the balanced datasets higher than unbalanced ones. In the unbalanced datasets, the maximum accuracy for both classifiers was 82.7%. In the balanced dataset, the accuracy reached 97.6%; these results confirm that these two classifiers are the best compared to the other selected classifiers in the two experiments:

5.3.2. K-NN and Decision Tree

Show better performance with the unbalanced datasets, and the difference is so apparent. The maximum accuracy with the balanced datasets is 39.4%, while it reached 82.7 % with the unbalanced datasets.

5.3.3. Random forest and Random Tree

It shows better performance with the unbalanced datasets, and the difference is so apparent. The maximum accuracy with the balanced datasets around 35%, while it reached 82.7% with the unbalanced datasets.

In conclusion, Naive Bayes and ID3 gave a better accuracy level than other classifiers, and the performance was better with the balanced datasets. The different classifiers (K-NN, Decision Tree, Random Forest, and Random Tree) gave a better understanding of the unbalanced datasets.

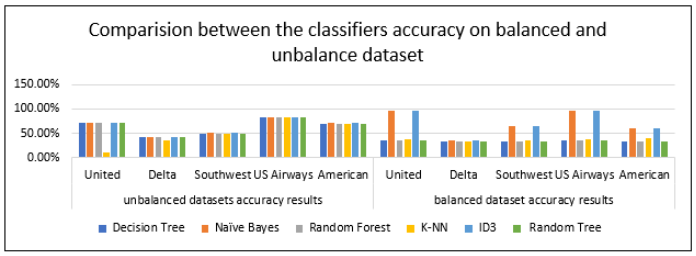


Figure 4: Accuracy results of classifiers on balanced and unbalanced datasets

6. Conclusions

Social media websites are gaining very big popularity among people of different ages. Platforms such as Twitter, Facebook, Instagram, and Snapchat allowed people to express their ideas, opinions, comments, and thoughts. Therefore, a huge amount of data is generated daily, and the written text is one of the most common forms of the generated data. Business owners, decision-makers, and researchers are increasingly attracted by the valuable and massive amounts of data generated and stored on social media websites. Sentiment Analysis is a Natural Language Processing field that increasingly attracted researchers, government authorities, business owners, services providers, and companies to improve products, services, and research. In this research paper, the authors aimed to survey sentiment analysis approaches. Therefore, 16 research papers that studied Twitter's text

classification and analysis were surveyed. The authors also aimed to evaluate different machine learning algorithms used to classify sentiment to either positive or negative, or neutral. This experiment aims to compare the efficiency and performance of different classifiers that have been used in the sixteen papers that are surveyed. These classifiers are (Decision Tree, Naïve Bayes, Random Forest, K-NN, ID3, and Random Tree). Besides, the authors investigated the balanced dataset factor by applying the same classifiers twice on the dataset, one on the unbalanced and the other, after balancing the dataset. The targeted dataset included six datasets about six American airline companies (United, Delta, Southwest, Virgin America, US Airways, and American); it consists of about 14000 tweets. The authors reported that the classifier's accuracy results were very high in some datasets while low in others. The authors indicated that the dataset size was the reason for that. On the balanced dataset, the Naïve Bayes classifier, Decision Tree, and ID3 were mostly better than other classifiers and have given the almost same level of accuracy. The classifiers with Virgin America dataset reported the lowest level of accuracy due to its small size. On the unbalanced dataset, results show that the Naive Byes and ID3 gave a better level of accuracy than other classifiers when it's applied on the balanced datasets. While (K-NN, Decision Tree, Random Forest, and Random Tree) gave a better understanding of the unbalanced datasets.

Conflict of Interest

The authors declare no conflict of interest

Acknowledgment

This is a part of project done in British University in Dubai.

References

[1] S.A. Salloum, C. Mhamdi, B. Al Kurdi, K. Shaalan, "Factors affecting the Adoption and Meaningful Use of Social Media: A Structural Equation Modeling Approach," International Journal of Information Technology and Language Studies, **2**(3), 2018.

[2] M. Alghizzawi, S.A. Salloum, M. Habes, "The role of social media in tourism marketing in Jordan," International Journal of Information Technology and Language Studies, **2**(3), 2018.

[3] S.A. Salloum, W. Maqableh, C. Mhamdi, B. Al Kurdi, K. Shaalan, "Studying the Social Media Adoption by university students in the United Arab Emirates," International Journal of Information Technology and Language Studies, **2**(3), 2018.

[4] S.A. Salloum, M. Al-Emran, S. Abdallah, K. Shaalan, Analyzing the arab gulf newspapers using text mining techniques, 2018, doi:10.1007/978-3-319-64861-3_37.

[5] F.A. Almazrouei, M. Alshurideh, B. Al Kurdi, S.A. Salloum, Social Media Impact on Business: A Systematic Review, 2021, doi:10.1007/978-3-030-58669-0_62.

[6] Alshurideh et al., "Understanding the Quality Determinants that Influence the Intention to Use the Mobile Learning Platforms: A Practical Study," International Journal of Interactive Mobile Technologies (IJIM), **13**(11), 157–183, 2019.

[7] S.A. Salloum, K. Shaalan, Adoption of E-Book for University Students, 2019, doi:10.1007/978-3-319-99010-1_44.

[8] S.A. Salloum, M. Alshurideh, A. Elnagar, K. Shaalan, "Mining in Educational Data: Review and Future Directions," in Joint European-US Workshop on Applications of Invariance in Computer Vision, Springer: 92–102, 2020.

[9] S.A. Salloum, M. Alshurideh, A. Elnagar, K. Shaalan, "Machine Learning and Deep Learning Techniques for Cybersecurity: A Review," in Joint European-US Workshop on Applications of Invariance in Computer Vision, Springer: 50–57, 2020.

[10] S.A. Salloum, R. Khan, K. Shaalan, "A Survey of Semantic Analysis Approaches," in Joint European-US Workshop on Applications of Invariance in Computer Vision, Springer: 61–70, 2020.

[11] K.M. Alomari, A.Q. AlHamad, S. Salloum, "Prediction of the Digital Game Rating Systems based on the ESRB."

[12] S.A. Salloum, A.Q. AlHamad, M. Al-Emran, K. Shaalan, A survey of Arabic text mining, Springer, Cham: 417–431, 2018, doi:10.1007/978-3-319-67056-0_20.

[13] C. Mhamdi, M. Al-Emran, S.A. Salloum, Text mining and analytics: A case study from news channels posts on Facebook, 2018, doi:10.1007/978-3-319-67056-0_19.

[14] A.S. Alnaser, M. Habes, M. Alghizzawi, S. Ali, "The Relation among Marketing ads, via Digital Media and mitigate (COVID-19) pandemic in Jordan The Relationship between Social Media and Academic Performance: Facebook Perspective View project Healthcare challenges during COVID-19 pandemic View project," Dspace.Urbe.University, (July), 2020.

[15] M. Alshurideh, B. Al Kurdi, S. Salloum, "Examining the Main Mobile Learning System Drivers' Effects: A Mix Empirical Examination of Both the Expectation-Confirmation Model (ECM) and the Technology Acceptance Model (TAM)," in International Conference on Advanced Intelligent Systems and Informatics, Springer: 406–417, 2019.

[16] M. Alghizzawi, M. Habes, S.A. Salloum, M.A. Ghani, C. Mhamdi, K. Shaalan, "The effect of social media usage on students'e-learning acceptance in higher education: A case study from the United Arab Emirates," International Journal of Information Technology and Language Studies, **3**(3), 2019.

[17] M. Habes, S.A. Salloum, M. Alghizzawi, C. Mhamdi, "The Relation Between Social Media and Students' Academic Performance in Jordan: YouTube Perspective," in International Conference on Advanced Intelligent Systems and Informatics, Springer: 382–392, 2019.

[18] M. Habes, S.A. Salloum, M. Alghizzawi, M.S. Alshibly, "The role of modern media technology in improving collaborative learning of students in Jordanian universities," International Journal of Information Technology and Language Studies, **2**(3), 2018.

[19] B.A. Kurdi, M. Alshurideh, S.A. Salloum, Z.M. Obeidat, R.M. Al-dweeri, "An empirical investigation into examination of factors influencing university students' behavior towards elearning acceptance using SEM approach," International Journal of Interactive Mobile Technologies, **14**(2), 2020, doi:10.3991/ijim.v14i02.11115.

[20] S.A. Salloum, M. Al-Emran, M. Habes, M. Alghizzawi, M.A. Ghani, K. Shaalan, "Understanding the Impact of Social Media Practices on E-Learning Systems Acceptance," in International Conference on Advanced Intelligent Systems and Informatics, Springer: 360–369, 2019.

[21] M. Alghizzawi, M.A. Ghani, A.P.M. Som, M.F. Ahmad, A. Amin, N.A. Bakar, S.A. Salloum, M. Habes, "The Impact of Smartphone Adoption on Marketing Therapeutic Tourist Sites in Jordan," International Journal of Engineering & Technology, **7**(4.34), 91–96, 2018.

[22] S.F.S. Alhashmi, S.A. Salloum, S. Abdallah, "Critical Success Factors for Implementing Artificial Intelligence (AI) Projects in Dubai Government United Arab Emirates (UAE) Health Sector: Applying the Extended Technology Acceptance Model (TAM)," in International Conference on Advanced Intelligent Systems and Informatics, Springer: 393–405, 2019.

[23] M. Alghizzawi, M. Habes, S.A. Salloum, The Relationship Between Digital Media and Marketing Medical Tourism Destinations in Jordan: Facebook Perspective, 2020, doi:10.1007/978-3-030-31129-2_40.

[24] R.S. Al-Maroof, S.A. Salloum, A.Q.M. AlHamadand, K. Shaalan, A Unified Model for the Use and Acceptance of Stickers in Social Media Messaging, 2020, doi:10.1007/978-3-030-31129-2_34.

[25] K.S.A. Wahdan, S. Hantoobi, S.A. Salloum, K. Shaalan, "A systematic review of text classification research based ondeep learning models in Arabic language," Int. J. Electr. Comput. Eng, **10**(6), 6629–6643, 2020.

[26] J.D. Burger, J. Henderson, G. Kim, G. Zarrella, "Discriminating gender on Twitter," in Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 1301–1309, 2011.

[27] D. Gamal, M. Alfonse, E.-S.M. El-Horbaty, A.-B.M. Salem, "Twitter Benchmark Dataset for Arabic Sentiment Analysis," International Journal of Modern Education and Computer Science, **11**(1), 33, 2019.

[28] M. Nabil, M. Aly, A. Atiya, "Astd: Arabic sentiment tweets dataset," in Proceedings of the 2015 conference on empirical methods in natural language processing, 2515–2519, 2015.

[29] Z.J. and G. Xiaolini, "Comparison Research on Text Pre-processing Methods on Twitter Sentiment Analysis," Digital Object Identifier, 2017, doi:10.1109/ACCESS. 2017. 2672677.

[30] Q.C. Nguyen, D. Li, H.-W. Meng, S. Kath, E. Nsoesie, F. Li, M. Wen, "Building a national neighborhood dataset from geotagged Twitter data for indicators of happiness, diet, and physical activity," JMIR Public Health and Surveillance, **2**(2), e158, 2016.

[31] A. Amolik, N. Jivane, M. Bhandari, M. Venkatesan, "Twitter sentiment analysis of movie reviews using machine learning techniques," International Journal of Engineering and Technology, **7**(6), 1–7, 2016.

[32] G. Gautam, D. Yadav, "Sentiment analysis of twitter data using machine learning approaches and semantic analysis," in 2014 Seventh International Conference on Contemporary Computing (IC3), IEEE: 437–442, 2014.

[33] M. Furini, M. Montangero, "TSentiment: On gamifying Twitter sentiment analysis," in 2016 IEEE Symposium on Computers and Communication (ISCC), IEEE: 91–96, 2016.

[34] A. Porshnev, I. Redkin, A. Shevchenko, "Machine learning in prediction of stock market indicators based on historical data and data from Twitter sentiment analysis .," 2013 IEEE 13th International Conference on Data Mining Workshops, 440–444, 2013, doi:10.1109/ICDMW.2013.111.

[35] V. Ramanathan, "Twitter Text Mining for Sentiment Analysis on People ' s Feedback about Oman Tourism," 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC), 1–5, 2019.

[36] V. Martin, "Predicting the french stock market using social media analysis," Proceedings - 8th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2013, 3–7, 2013, doi:10.1109/SMAP.2013.22.

[37] Q. Li, B. Zhou, Q. Liu, "Can twitter posts predict stock behavior?: A study of stock market with twitter social emotion," Proceedings of 2016 IEEE International Conference on Cloud Computing and Big Data Analysis, ICCCBDA 2016, 359–364, 2016, doi:10.1109/ICCCBDA.2016.7529584.

[38] Indrawati, A. Alamsyah, "Social network data analytics for market segmentation in Indonesian telecommunications industry," 2017 5th International Conference on Information and Communication Technology, ICoIC7 2017, **0**(c), 2017, doi:10.1109/ICoICT.2017.8074677.

[39] H. Dai, J. Hao, "Mining social media data for opinion polarities about electronic cigarettes," Tobacco Control, **26**(2), 175–180, 2017, doi:10.1136/tobaccocontrol-2015-052818.

[40] A. Husnain, S.M.U. Din, G. Hussain, Y. Ghayor, "Estimating market trends by clustering social media reviews," Proceedings - 2017 13th International Conference on Emerging Technologies, ICET2017, **2018**–**Janua**, 1–6, 2018, doi:10.1109/ICET.2017.8281716.

[41] K. Kandasamy, P. Koroth, "An integrated approach to spam classification on Twitter using URL analysis, natural language processing and machine learning techniques," 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, SCEECS 2014, 1–5, 2014, doi:10.1109/SCEECS.2014.6804508.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1734-1740 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Machine Vision Approach for Underwater Remote Operated Vehicle to Detect Drowning Humans

Yaswanthkumar S K*, Keerthana M, Vishnu Prasath M S

Department of Electrical & Electronics Engineering, Kumaraguru College of Technology, Coimbatore, 641049, India

ARTICLE INFO

Article history:
Received: 15 September, 2020
Accepted: 29 December, 2020
Online: 30 December, 2020

Keywords:
Remote Operated Vehicle
Image Processing
Drowning People
Semi-Autonomous

ABSTRACT

Today, Drowning is the 3rd major cause for unintentional injury death accounting for 7% of deaths of all injury deaths. Drowning is a state of suffocation when water or other fluids accumulate the lungs, resulting in respiratory impairment, ultimately leading to death. The predominant problem during rescue operations of such accidental drowning is to locate or track the person underwater, facilitated with the help of our bare eye sight, which makes the process too cumbersome. Also, in case of moving water bodies such as rivers and sea, the process of tracking the person becomes too difficult. Thus, there is pressing need for development of an engineered solution to solve this problem. This research involves development of such a robotic technology which will facilitate the work of locating the position of the drowning person underwater. This method comprises of Image processing along with GPS tagging embedded on a Remote Operated Vehicle maneuvering underwater at the site of accident to detect the exact location of the drowning person along with the GPS coordinates in order to ease the process of rescue. This robotic system was tested on a real time environment in order to demonstrate the efficacy of the system under practical circumstances. Thus, this system can be used at places where emergency rescue operations are needed.

1. Introduction

Drowning accounts for 7% of all deaths that occur due to injury. It is also the 3rd major cause for unintentional injury death. As drowning is an unexpected event, there are neither any precautionary steps to prevent it, nor any existing technology to initiate rescue operations rapidly. According to stats from 2016 nearly 350000 people are reported dead due to drowning. In India, Drowning is the second largest cause of death for children below age of 15. The situation is even worse at Russia. The drowning mortality rate is of more than 3.9 per 10000 individuals.

Nowadays, few assistive technologies are available in the market to provide safety for humans in or under water. But these solutions are based on scenarios where people must use this technology before nearing any water bodies. But since Accidental Drowning are unpredictable and the underwater environment varies from place to place a solution adaptable to all water bodies is the main focus of this work. Development of this technology can be used not only on rescue operations of accidental drowning, but also to mitigate the risk of loosing track of people during natural calamities like heavy floods or Tsunamis. Thus, there is a urgent need to focus on development of a solution for the above mentioned scenario.

Our research work is to embedded Image Processing methodologies to detect drowning people underwater and to use GPS coordinates of their exact location with the use of a Robotic Remote Operated Vehicle which can by controlled autonomously or semi autonomously to navigate through water bodies during the rescue process. Also, this system uses a on board camera, live camera feed can also be surfaced over radio communications to the rescue team, which would prove to be great advantage to the people on the field for the rescue operations. All the existing methods are either not cost-effective or have installation issues at underwater environments.

2. Survey on Previous Researches

This paper is an extension of work [1] originally presented in 2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), where a different method for the detection of humans drowning underwater was illustrated.

In [2], is a research work that involves development of an automated system that detects and tracks objects that are of potential interest for human video annotators. By pre-selecting salient targets for track initiation using a selective attention algorithm, we reduce the complexity of multi-target tracking, in

*Corresponding Author: Yaswanthkumar S K, Email: yaswanthsk4@gmail.com

particular of the assignment problem. Detection of low-contrast translucent targets is difficult due to variable lighting conditions and the presence of ubiquitous noise from high-contrast organic debris ("marine snow") particles.

In [3], the paper presents an attentional selection system for processing video streams from remotely operated underwater vehicles (ROVs). The system identifies potentially interesting visual events spanning multiple frames based on low-level spatial properties of salient tokens, which are associated with those events and tracked over time. Using video cameras, it is possible to make quantitative video transects (QVTs) through the water, providing high-resolution data at the scale of the individual animals and their natural aggregation patterns.

In [4], this paper explains an agent system is strategy to enhance the underwater manipulation. If the location of the agent can be measured, the end effector is able to be place to any position. To implement this system, the method of an agent vehicle localization is proposed. The method uses the sonar images of moving agent obtained by forward-looking sonar. To detect the location of the agent in the sonar images, the convolutional neural network is applied. Through field experiment, we confirm the proposed method can detect and track the agent in the successive sonar images.

In [5], This paper presents methods applied for automated detection of fish based on cascade classifiers of Haar-like features created using underwater images from a remotely operated vehicle under ocean survey conditions. The images are unconstrained, and the imaging environment is highly variable due to the moving imaging platform, a complex rocky seabed background, and still and moving cryptic fish targets. Several Haar cascades are developed from the training set and applied to the validation and test video images for evaluation.

In [6], This paper presents a robotics vision-based system for an underwater pipeline and cable tracker. This system is introduced to improve the level of automation of underwater remote operated vehicles (ROVs) operations which combines computer vision with an underwater robotics system to perform target tracking and intelligent navigation. This study focuses on developing image processing algorithms and nonlinear control method for tracking the pipeline or cable. By using the adaptive sliding mode controller, the ROV can easily track the cable or pipeline.

In [7], is a proposed system of a multi-step and all-round underwater image processing specially for the underwater images taken in succession to improve the image quality, remove the dynamic interference and reconstruct the image. For images with complex texture, the inpainting result brings out much smoothness, at the cost of massive details. For images with low resolution, the transition of color near the edge of the vacancy is abrupt

In [8], The paper presents a wavelet-based fusion method to enhance the hazy underwater images by addressing the low contrast and color alteration issues. The publicly available hazy underwater images are enhanced and analyzed qualitatively with some state-of-the-art methods. Green color dominates the Image in the evaluated set of images. In the results, it can be observed that the natural color profiles of Image are preserved by correcting the dominated color and it was hard before to distinguish the color of the water and objects

In [9], it shows that man-made objects in the image can be distinguished from natural objects. This paper proposes that a control strategy to adjust image processing parameters, and sequence multiple applications of tools, is especially important for digital processing of underwater images. This paper highlights the key difficulties in processing underwater images such as obscuration of objects of interest by fish / sand / coral / algae in the water and change of illumination, color balance and scale of fish images due to refraction, absorption in water and large range of distances of observation and different sizes of fish

3. Schematic & Working of ROV

As shown in Figure 1. The devised system consists of a on board camera embedded to the ROV to communicate the live video feed to the operator and also for the video to be processed by the on-board Processor to detect for humans underwater. The main system consists of sub units such as:

- Transmission & Receiving Unit
- Processor Unit
- Sensing Unit
- Maneuvering Unit
- Powering Unit

Each of which will be discussed on the following sections.

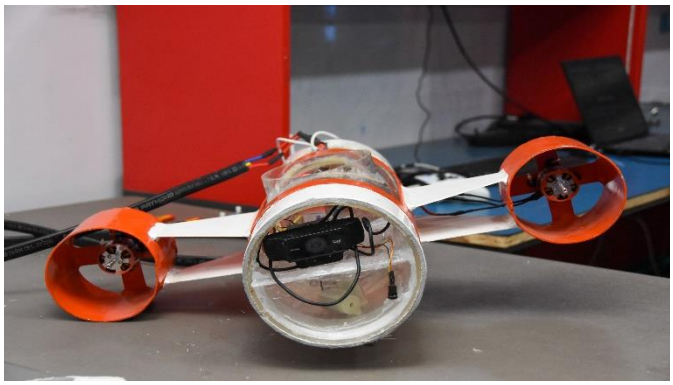


Figure 1: Hardware assembly of ROV

3.1. Transmission & Receiving Unit

A). Introduction

The transmission Unit consists of a Two axis controller to control the movement of the ROV along with a Screen or a monitor to view the live feed from the camera on-board which is interfaced to a Raspberry Pi through USB communication. A PC Joystick controller with TWO axis was used to provide the control signals as a input ls to the motor to control the movement of the ROV. Common PC monitor was used along with the raspberry PI to view the live feed from the ROV's on-board camera. The Receiver Unit Consists of a Raspberry PI which is on-board in the ROV. The Video data from the camera is given as an input to this raspberry pi from which is then transferred to the other raspberry pi through wireless communication.

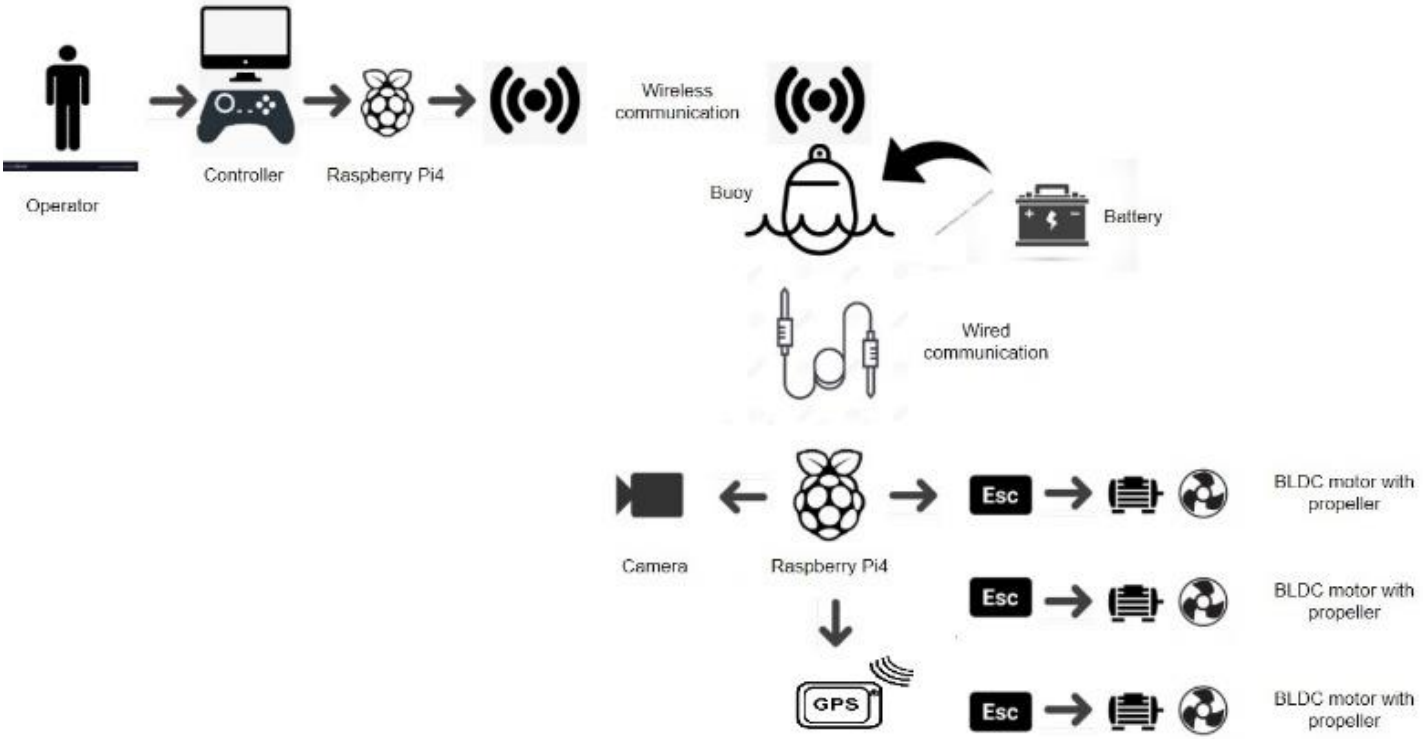


Figure 2: Schematic Diagram

B). Working

The control signals from the operator are provided as joystick input through USB communication to Raspberry PI on the Transmitter side. These signals are then transmitted to the other Raspberry PI on the Receiver side through WIFI Communication. This transfer of data can also be performed using RF transmission or MQTT via internet which will provide further distance range between the operator and the ROV. The operator can view the live feed in the monitor, to control the ROV effectively. The video feed from the USB Camera is also transmitted wirelessly through WIFI Communication from the Receiver Raspberry Pi to Transmitter Raspberry Pi.

3.2. Processor Unit

A). Introduction

The Processor unit consists of a Raspberry Pi (receiving unit) on-board in the ROV. This Receives the control signals from the Operator which is then used to control the movement of the ROV. This also receives data from the on-board camera and on-board GPS for further transmission.

B). Working

As the PWM from the Transmitter unit are received by the Processor Unit (receiving unit), these signals are the compared with the on-board accelerometer to measure the difference in the values. These values are again corresponded to respective PWM signals and then is given as an input to the respective Electronic Speed Controllers (ESC's) which control the spin of the motors. The Processor also performs the function of receiving the input data from camera and GPS on the ROV and transfer them to the operator.

3.3. Sensing Unit

A). Introduction:

The sensing unit consists of a IP68 waterproof high megapixel camera to be interfaced with the Processor. This will provide the video data for Image Processing and as well as send the live data from underwater to the operator. It also consists of a waterproof GPS chip to be enclosed in an air tight chamber of ROV. This provides the GPS coordinates of the ROV underwater.

B). Working

The camera transmits the data to the processor through USB communication, which is then used as the input for Image Processing algorithm and is also transmitted to the operator. The GPS module sends the signal to raspberry pi through SPI protocol which is also sent to the operator through wireless Communication.

3.4. Maneuvering Unit

A). Introduction

Maneuvering unit consists of 60 Amps ESC's, each individual for respective 500 kV BLDC motors. As the PWM control signals are given as an input into ESC, the speed of the motor varies accordingly the position and moment of the ROV also varies.

B). Working

As the operator provides the input signals through the JOYSTICK controller, this input data from the Transmitter Raspberry pi, reaches the Processor Raspberry pi (receiver unit). Those input signals are then provided as an input to the respective ESC after being compared with the current accelerometer value &

then the required amount of power will be delivered to the respective motor which is now being controlled by ESC.

3.5. Powering Unit

ROV is completely powered through wires as it is being tested for its buoyancy when battery pack is being placed inside the air tight enclosure of the ROV.

4. Imaging Methodology to Detect Drowning Humans in a - underwater environment

Image Processing has overcome the limits of its own discipline. Today, it is possible to process a live video feed right after few trained models. Due to this rapid increase in its uses, this technology is pressured to produce strong methodologies which would offer a wide variety of solutions for real time problems. Our methodology involves usage of Faster Region Convoluted Neural Network (FRCNN) algorithm.

4.1. Hardware Specification

The hardware specifications are illustrated in table 1.

Table 1: Hardware Specification

System	Dell G3 3579
GPU	Nvidia GeForce GTX 1050Ti
CPU	Intel Core i7-87507H
Memory (RAM)	8192 MB
Display Memory (VRAM)	4018 MB
Operating System	Windows 10 64-bit

4.2. Dataset

As only humans are to be detected underwater a dataset consisting of 800 images is used as dataset. These 800 images comprise a variety of underwater images with varying light intensity, water color, and humans in it. This dataset also contains images of halfway human pictures along with overlapping images too.

A). Data Preprocessing:

Reading

Using the matplotlib library on a python script is used to reading all the images from the dataset. This helps us to know whether all the all the images from the dataset are loadable and could be read by the algorithm.

Re-sizing

As larger images would take longer time to process, all the images collected are res-sized to be less than 200KB in size.

B). Data Labelling

The datasets must be labelled, since labelled images must be provided as input to train the training algorithm. In order to label the datasets, the images are first loaded onto the LabelImg tool. Now, humans in each image are to be bounded by drawing bounding boxes around them as shown in Figure 3. Then, the labelled data of all these images are to be stored in an XML format. Each image file would create individual XML files containing the File name, path, the label and coordinates for the bounding boxes. These are also called annotation files which will be used to train the algorithm.

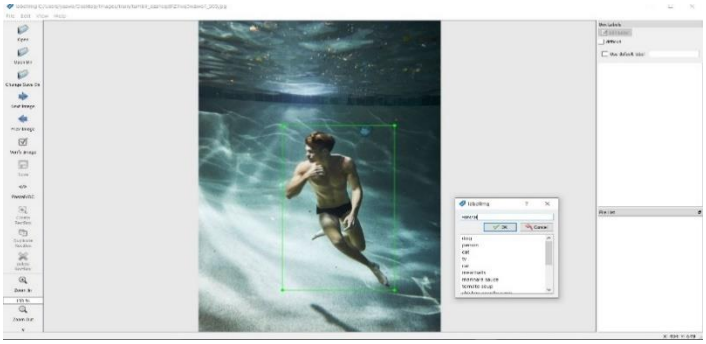


Figure 3: Labelling the datasets

4.3. Implementation

In order to bring out the best outcome, two algorithms were used 1). Faster RCNN and Yolo V3. Both these algorithms were implemented and trained using the TensorFlow Object Detection API. This makes the training process faster as TensorFlow object detection API comes with the model called ZOO which is a collection of models and algorithms.

A). Configuring the training algorithm

In order to detect the best algorithm which can be used, both the Faster RCNN and YOLO algorithms were configured to the same values.

Data Slicing

Both the algorithms were trained and tested using the same data sets. So, for both the algorithms 75% of the images and 25% of the images were used as training datasets and testing datasets respectively.

Classes

Since, humans were the only object to be detected by the algorithms, the class for these algorithms was set to 1.

Learning Rate

Learning rate is a parameter which allows us to control the model's response to estimated error each time the model weights are updated. The learning rate was set to a default value of 0.0002 while training both the algorithms.

Batch Size

The batch size corresponds to the consumption of VRAM, after looking into Table 1. The value set of batch size of the algorithms is set as 4.

4.4. Training Procedure

Figure 4 shows the training process of Faster RCNN algorithm along with the classification loss and the time taken after each iteration.

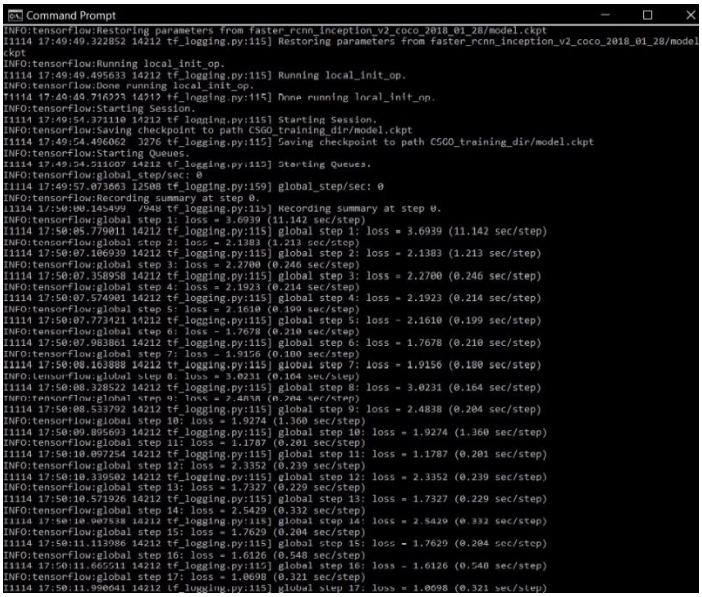


Figure 4: Training Process of Faster RCNN

As you can see, Classification loss is high when the training processes started. But after each step, there is a gradual decrease in the classification loss as the algorithm kept learning in each step. Figure 4. shows the exact decrease in classification loss during training process compared along with the testing process.

4.5. Performance Metrics

In order to compare the results of these two models, accuracy of these two models are to be contrasted. To calculate the accuracy few parameters are to be considered as metrics:

1. True Positive TP: A human is present in the Image and, the algorithm detects the human.
2. False Positive FP: No human is present in the Image, but the algorithm detects human.
3. False Negative FN: A human is present in the Image, but the algorithm does not detect the human.
4. True Negative TN: No human is present in the Image, and nothing is being detected.

A). Accuracy:

It is a measure to check whether the model or the algorithm is trained correctly to detect humans in an underwater environment or not. The formula to calculate the accuracy as follows:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

B). Precision

It is a ratio of positively predicted images that contain humans in it to that of which actually contains humans in it. The formulae to calculate precision as follows:

$$\text{Precision} = \frac{TP}{(TP+FP)}$$

C). Recall

It is the ratio of images that actually contain human in it to that of which were predicted positively. The formula to calculate recall as follows:

$$\text{Recall} = \frac{TP}{(TP+FN)}$$

D). F_1 Score

It is the measure of accuracy of the model combining both precision and recall. A model or algorithm is considered to be perfect if it's F_1 Score is 1. The formula to calculate F1 Score as follows:

$$F_1 \text{ Score} = 2 * \left(\frac{\text{Percision} * \text{Recall}}{\text{Persion} + \text{Recall}}\right)$$

5. Results

The results for both the algorithms are summarized in this section. The analysis of results as follows:

5.1. Faster RCNN Results

Figure 5, Figure 6 & Figure 7 are a collection of test data from 3 different samples under varying light conditions and other environmental factors for faster RCNN algorithm.



Figure 5: RCNN Output 1

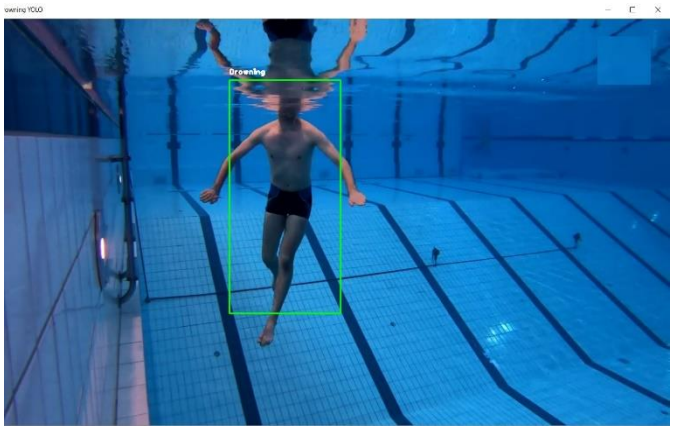


Figure 6: RCNN Output 2

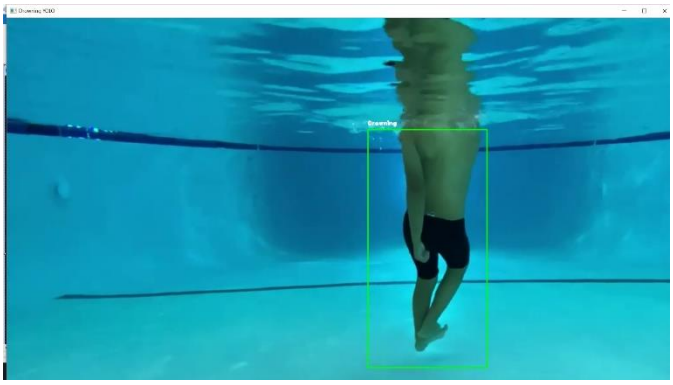


Figure 7: RCNN Output 3

The number of TP, TN, FP, FN from RCNN is tabulated in the table 2. These are further used to analysis the algorithms.

Table 2: RCNN Output Tabulation

#	True Positives	True Negative	False Positive	False Negative
Total	148	14	9	29

5.2. Yolo V3 Results

Figure 8, Figure 9 & Figure 10 are a collection of test data from 3 different samples under varying light conditions and other environmental factors for faster RCNN algorithm.

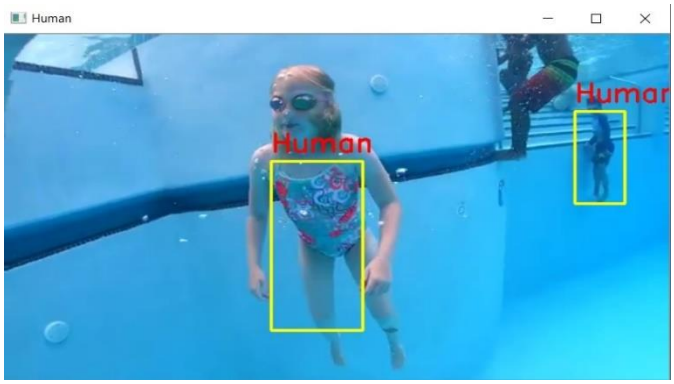


Figure 8: Yolo V3 Output 1

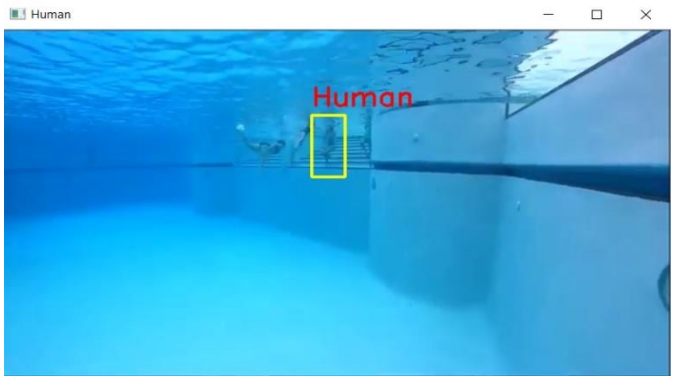


Figure 9: Yolo V3 Output 2

The number of TP, TN, FP, FN from YoloV3 is tabulated in the table 3. These are further used to analysis the algorithms.

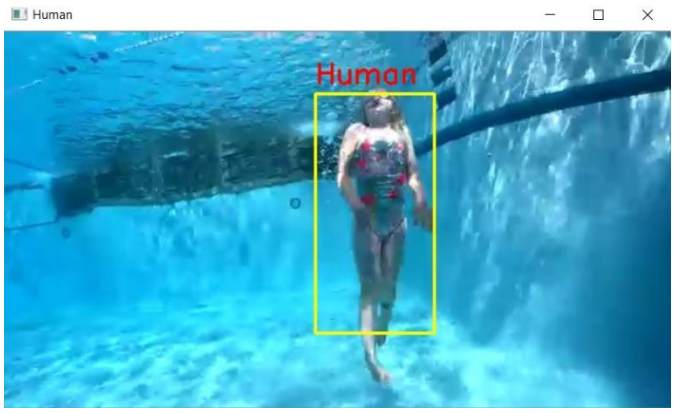


Figure 10: Yolo V3 Output 3

Table 3: YoloV3 Output Tabulation

#	True Positives	True Negative	False Positive	False Negative
Total	125	12	3	60

5.3. Analysis of Results

Table 4, provides us the collection of results containing the number of TP, TN, FP, FN parameters obtained by Faster RCNN and Yolo V3 algorithms.

Table 4

#	True Positives	True Negative	False Positive	False Negative
F - RCNN	148	14	9	29
YOLO V3	125	12	3	60

From this table, the accuracy of Faster RCNN and YOLO V3 algorithms are 81% and 68.5% respectively. Clearly, Faster RCNN is more Accurate than YOLO V3 in detecting humans in a underwater environment. It is also to note that Faster RCNN was able to consistently detect Humans if overlapping occurs. While, YOLO V3 failed to detect humans during overlapping frames.

Table 5

Algorithms	Precision	Recall	F_1 Score
Faster RCNN	0.94	0.83	0.881
YOLO V3	0.97	0.67	0.792

Table 5, Shows us that YOLO V3 is more precise compared to that of Faster RCNN, but then Faster RCNN has a high recall value than that of YOLO V3. To compare both the models, F_1 Score was calculated which also points out that Faster RCNN algorithm is the Best fitted algorithm to detect humans underwater. The results contribute to some preliminary findings to the growing field of underwater robotics in real time environments.

6. Future Work

The experimentation has been tested only on clear water bodies. The implementation of this ROV under real time circumstances such as lakes and river where the opacity will be translucent or opaque, it is better to use the thermal imaging as the

video input to this image processing models. As the body temperature of the aquatic species in that water body will be the similar, it will be easy to differentiate human body using thermal sensors and video feed of thermal imaging which will be performed in the near future.

7. Conclusion

This paper illustrates, an idea to develop a robotic system which will detect drowning people underwater at the times of an emergency. It also concludes that Faster RCNN is the best fit algorithm which can be used for this practical application. In our further research we plan to our work from a single ROV into a swarm of ROV's collectively working at a site of emergency to detect every drowning human underwater. We are also planning to enable full autonomous nature of the ROV in order to minimize the reaction time and communication time between the ROV and the Operator.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like to thank Forge Accelerator for providing the necessary support to complete this research work. We also thank Mr. Ravi Shankar, Mr. Koushal & Miss. Poovika for contributing to this research work by building the ROV from Scratch.

References

[1] S.K. Yaswanthkumar, O.K. Praveen, R. V. Rohit, "Autonomous Utility Vehicle (Auvs) Based Emergency Human Drowning Detection System Using Sonar and Thermal Detection Methods," in 2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics, DISCOVER 2019 - Proceedings, 2019, doi:10.1109/DISCOVER47552.2019.9007992.

[2] D. Walther, D.R. Edgington, C. Koch, "Detection and tracking of objects in underwater video," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, doi:10.1109/cvpr.2004.1315079.

[3] D.R. Edgington, K.A. Salamy, M. Risi, R.E. Sherlock, D. Walther, C. Koch, "Automated event detection in underwater video," in Oceans Conference Record (IEEE), 2003, doi:10.1109/oceans.2003.178344.

[4] J. Kim, S.C. Yu, "Convolutional neural network-based real- Time ROV detection using forward-looking sonar image," in Autonomous Underwater Vehicles 2016, AUV 2016, 2016, doi:10.1109/AUV.2016.7778702.

[5] G. Cutter, K. Stierhoff, J. Zeng, "Automated detection of rockfish in unconstrained underwater videos using haar cascades and a new image dataset: Labeled fishes in the wild," in Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2015, 2015, doi:10.1109/WACVW.2015.11.

[6] R. Trahan, R. Kiang, "An analysis of the die testing process using taguchi techniques and circuit diagnostics," in Proceedings - International Test Conference, 1992, doi:10.1109/TEST.1992.527832.

[7] C. Cai, Y. Zhang, T. Liu, "Underwater image processing system for image enhancement and restoration," in 2019 IEEE 11th International Conference on Communication Software and Networks, ICCSN 2019, 2019, doi:10.1109/ICCSN.2019.8905310.

[8] A. Khan, S.S.A. Ali, A.S. Malik, A. Anwer, F. Meriaudeau, "Underwater image enhancement by wavelet based fusion," in USYS 2016 - 2016 IEEE 6th International Conference on Underwater System Technology: Theory and Applications, 2017, doi:10.1109/USYS.2016.7893927.

[9] R.E. Blake, "Digital image processing of underwater images," in 2006 IEEE US/EU Baltic International Symposium, BALTIC 2006, 2006, doi:10.1109/BALTIC.2006.7266195.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1741-1749 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Evaluation of some Machine Learning Algorithms for the detection of Android Applications Malware

Olorunshola Oluwaseyi Ezekiel[*,1], Oluyomi Ayanfeoluwa Oluwasola[2], Irhebhude Martins[3]

[1]*Cyber Security Department, Faculty of Computing, Air Force Institute of Technology, Kaduna, 800283, Nigeria*

[2]*Electrical Electronics Engineering Department, Air Force Institute of Technology, Kaduna, 800283, Nigeria*

[3]*Computer Science Department, Nigerian Defence Academy, Kaduna, 800283, Nigeria*

ARTICLE INFO

Article history:
Received: 08 October, 2020
Accepted: 25 December, 2020
Online: 30 December, 2020

Keywords:
Mobile phone
Malware
Algorithms
Operating System
Benign

ABSTRACT

Android Operating system (OS) has been used much more than all other mobile phone's OS turning android OS to a major point of attack. Android Application installation serves as a major avenue through which attacks can be perpetrated. Permissions must be first granted by the users seeking to install these third-party applications. Some permissions can be subtle escaping the attentions of the users. Some of these permissions can have adverse effects like spying on the users, unauthorized retrieval and transference of the data and so on. This calls for the need of a heuristic method for the identification and detection of malware. In this discourse, testing of classification algorithms including Random forest, Naïve Bayes, Random Tree, BayesNet, Decision Table, Multi-layer perceptron (MLP), Bagging, Sequential Minimal Optimization (SMO)/Support-Vector Machine (SVM), KStar and IBK (also known as K Nearest Neighbours classifier (KNN)) was carried out to decide which algorithm performs best in android malware detection. Two dataset was used in this study and were gotten from figshare. They were trained and tested in the Waikato Environment for Knowledge Analysis (WEKA). The performance metrics used are Root Mean Square Error (RMSE), Accuracy, Receiver Operating Curve (ROC), False positive rate, F-measure, Precision and recall. It was discovered that the best performance with an accuracy of 99.4% was the multi-layer perceptron on the first dataset. Random Forest has the best performance with accuracy, 98.9% on the second dataset. The implication of this is that MLP or random forest can be used to detect android application malwares.

1. Introduction

This work is an extension of [1] in which this work considers another dataset in carrying out the research. This research answers the questions from the comments received about the biasness of the algorithms used.

There has been a considerable increase in the usage of mobile phones year on year. Mobile phones have served as a medium of communication through calls and short mailing services (SMS) [2]. Technological advancement has increased the usage of mobile phones for various purposes. The operating system (OS) of mobile phones have evolved from java to Symbian with the current ones being android OS, windows OS and Apple OS – iOS [3]. The usage of mobile phones spans from individuals to organizations as well. This study focuses on Android OS. Mobile phones with Android OS comprises of some components which are; the

Android OS, the middleware and key applications [4]. The android OS has the biggest and fastest growing smartphone OS market, hence it has become a practical target for cyber criminals [5]. Large support to third party applications is also gotten from android. These third party applications can easily be downloaded on the popular Google Play store [6]. The developers of the Android OS do not develop third-party applications; it is done by software developers. Hence the name of the third party applications [6]. In [7], there are over 2.7 billion mobile users globally, and in 2022, 258.2 billion apps are projected to be downloaded. Android phones are useful for crucial functions ranging from financial transfers to e-commerce transactions

[*]Corresponding Author: Oluwaseyi Ezekiel Olorunshola, Electrical Electronics Engineering Department, Air Force Institute of Technology, Kaduna, 800283, Nigeria, +2348068341127 & seyisola25@yahoo.com

(goods payments) and so on [8]. It also allows other features, such as photo editing, game play, social media and messaging, internet searching, typing, listening, reading, and so on. Any of all these procedures are feasible due to the implementation of third-party apps.

The growth of android application malware has been made possible by the free application market for the Android OS [9]. The most severe and important threat to the Android OS arises as a result of malicious applications especially from unauthorized applications market that can easily be installed on any phone [9]. Android applications depend on users to grant permission to work optimally. Such permissions can forewarn users to the applications' secrecy and/or security [10]. Android malware can impersonate a legitimate program, but with intrusive permissions for which an agreement is required from the users before installation can be done [10]. The malware executes activity(s) that are malicious through these permissions. SMS spy pro is an example, masking itself as a tip calculator but forwarding all sent SMS and received SMS by the user to a third party [11]. Android malware can access and steal user data such as browsing history, SMS, location of GPS, phone records, e-mails, other application information (Like Facebook log-in details) and more [12]. Tasks such as snapping and sending photographs, sending unauthorized e-mails, conducting bank transfers without user awareness can also be executed [2]. The degree to which both malicious and innocuous applications can execute their functions is primarily dependent on the form of permissions that users grant them.

Considering the sensitivity of the data that can be colleted from users via the malwares combined with the kind of damage that can be done by these malwares (for example, spying, identity theft, extortion, logic bomb) [13], these malwares need to be identified effectively. Due to zero-day attacks, the conventional malware detection techniques that uses the malware signature to store them in a repository has proved to be ineffective [9]. Zero-day attacks are based on new flaws found by cyber attackers that the creator is unaware of and thus has not issued a patch [14]. Therefore, until the developers are aware, the attackers make the most of these vulnerabilities. Android malware with its corresponding permission requests must be found [9]. This study is focused on Olorunshola and Oluyomi's recommendations in [1]. An evaluation of the classification algorithms with regard to their performance in the identification of malicious software based on their manifested permission, command signature, API call signature and intents, is the aim of this study. This study aims to verify the conclusions drawn by Olorunshola and Oluyomi by training and testing the algorithms with additional datasets in this research. Two android datasets based on manifest permission, command signature, API call signature and intent are obtained from figshare created by [15] to achieve this aim. The algorithms are trained and tested on the datasets. The remaining portion of this work is structured as follows: review of literature, methodology, result and discussion, conclusion and future work.

2. Literature Review

2.1. History of Android

Android was created using the name of Android Inc company by Rich Miner, Nick Sears, Chris White, and Andy Rubin in 2003 [16]. Android began as a camera OS intended to enhance connectivity between personal computers (PCs) and cameras, allowing communication wireless between cameras and PCs. After a couple of months of the company's formation, the OS was adapted towards making mobile cell phones smarter [17]. Android was purchased by Google in August 2005. Android OS was created using the Linux kernel to make it free for mobile phone manufacturers [16,18]. The first version of Android OS (version 1.0) was released in November 2007, and the novel Android handset, T-Mobile, was deployed in September 2008. While it has a lot of feedback and appraisal, it ran with Android 1.0 OS, and it has other Google apps such as play store, charts, YouTube, etc. built on it. [16,18]. As the year passes by, android OS has evolved into improved models. Android 10 is the latest edition of Android, launched in September 2019 [19, 20].

The stack layers of Android apps contain applications, application framework, libraries and the Linux kernel, with each layer closely incorporated to offer stronger smartphone applications. Interoperability has also been provided for to make it easier for design developers to work and develop their applications at a cost-friendly level [4]. It made it possible for anyone to obtain due to the open source nature of the Android framework thus growing its users over the years. Android sales stood at 115 million units in 2011, and Android OS accounted for around 52.5 percent of the overall number of mobile phones [21]. The Apple App Store reportedly has 2.2 million applications [7]. Recently, about 68.71 percent of people use android OS while 29.60 percent use iOS, 0.30 percent use Windows devices, while 0.06 percent use Symbian [22]. The heavy use of android OS is attributed to several factors that include the reduced price relative to iOS, open source enabling third-party software installation unlike iOS. The sum total of applications available in the Android store (play store) as of 2017 was around 2,800,000 [22]. In March 2018, the applications in the play store crossed 3.600,000 [23].

2.2. Security in Android

IOS as a closed framework that offers more data protection and is more pristine than Android which is an open source. In the play store, developers can host the applications they have built giving rise to Android vulnerabilities [22]. The security enhancements in Android were clarified in [24]. They are public vs. private components, components open implicitly, permission for broadcast intent, permission for content provision, secure APIs, service hooks, pending intents, permission protection level, URIs permission. Security refinements such as these, aid the protection of the data of users from manipulation and abuse.

2.3. Android Malware

The skill to debone the code used in the development of Android applications and the open source nature of the applications also enabled the easy encoding of malicious codes and increased number of attacks [25]. In [26], Networks recorded that mobile malware rose to 155 percent in 2011, with a rise of 614 percent from March 2012 to March 2013, of which 92 percent of the malware is Android. When android devices are upgraded to the latest OS, about 77 percent of threats to android can be eliminated. Just one threat infected about 5,000 devices in the third quarter of 2018. This accelerated through phishing using download and

installation of a voice message application that is fake thereby sends all the user's data to a server or using text messages [27]. Report has it that 2,47 million of the latest mobile malware were discovered in 2013, indicating a 197 percent increase over 2012 [28].

From a simplistic SMS sending Trojan, Android malware has advanced to more futuristic codes that can corrupt other applications, encrypt user data, gain root privileges, download and install more applications that are malicious bypassing the knowledge of the users, and load a payload from a remote server to cellular phones [9,29]. In [29], the authors gave a complete malware for android analysis looking at the past, observing the present and predicting the potential android malware cases into the future. In [30], the author gave a thorough analysis of the models of mobile malware dissemination and explored potential future developments as well.

2.4. The dangers of android malware

Users are exposed to multiple attacks and may suffer from them if an android OS is compromised with malware. According to [25]. Many of the threats they face include but are not restricted to:

- Personal information theft which can lead to theft of identity
- Loss of privacy
- Monitoring of the users
- Financial loss through ransomware
- Remote operation of the phones

2.5. Android malware detection

Detection of Android malware can broadly be divided into;

- Signature-based: Detection using signature-based detection uses the malware's identity to determine the genuineness of the malware. The downside of this method is some form of attacks may be eluded. E.g., byte code level transformation [28].
- Machine-learning focused: The detection based on machine-learning uses an analytical method that includes extracting features from the application's behaviour [28]. These behaviours (permission request, API calls) are then developed into a dataset and then a machine learning algorithm is used to test and train. The measurement metrics used in the evaluation include: precision, accuracy, false positive, to mention a few.

Several approaches to android malware detection have been put forward; the use of applications' requested permissions, specific application programming interface for detection, the use of the applications' fundamental code (since most codes for malware are modifications of existent codes), sandboxing, components encapsulation, optional access control or signing of application [31,32].

This study is aimed at the usage of manifest permission, API call signature, intents and command signature for the android malware detection. The rights defined by the developer in the application to allow system interrelationship: to access system components (such as camera and GPS) or modules of other applications is permission [33]. There are four types of android permission with two classifications, according to [2,34]. These are:

- Signature permissions: These applications are authorized only if it is certified by the creator of the stipulated consents. It is used for component access restriction to a limited collection of application controlled and trusted only by the creator.
- System permissions: Once the application satisfies the signature criteria, these applications are authorized. Applications requesting these permissions are pre-installed by an advanced customer or system manufacturer.

The two permissions (Signature and System) are booked for applications that have signatures with keys accessible only to the creator of the firmware. This is not available for implementations from third party application.

- Normal permissions: These are software that do not need the involvement of the customer. They are immediately accepted.
- Dangerous permissions: These are software that prompts the user what system facility will be used before it can perform its function. Prior to installation, they must be issued by the consumers. [2, 34].

In [34] and [35], the studies titled "PScout: Analyzing the Android Permission Specification" and "Android Permissions Demystified" evaluated the different android OS permission specification. In [2], the author also examined certain motives for some evasive consents. Some are broadcast theft, malicious service launch, hijacking of activity, malicious activity launch, hijacking of service.

2.6. Review of Past Works

Some research has been conducted in the detection of malware, part of which includes:

A Naïve Bayes algorithm model for android malware detection system was proposed by Shang, Li, Deng, & He in [36]. By using the new permissions for malware and training permissions impact as the weight, the detection accuracy was enhanced. In order to strengthen the model, a detection model of information theory and permissions based on the enhanced Naive Bayes algorithm was used also. As a rate of detection for non-malicious applications, the proposed model earned 97.59 percent.

In [37], the authors proposed a hybrid intelligent model which evolves and uses support vector support (SVM). Genetic algorithm (GA) and particle swarm optimization (PSO) were used to resolve the optimization challenge in the SVM. This helped boost the classifier's accuracy. With the use of GA, 95.60 percent accuracy was obtained as the highest.

In [38], the authors performed a fixed evaluation of applications for android, tested the existence and occurrence of words that ae key in the manifest file of the applications, and established static feature sets from a dataset of 400 applications to yield superior results for malware detection. KNN and SVM classification algorithms were applied getting an average accuracy of 79.08 percent and 80.50 percent respectively.

In order to detect malware, Shohel Rana, Gudla, & Sung in [39] improved and analyzed certain algorithms through the implementation of a static analysis that is classifier based. When random forest was used on the dataset, 94.33 percent was obtained as the accuracy.

In [31], the authors proposed a technique of malware detection by first extracting the function call graphs from android applications and then embedding the function call graphs with a direct map feature. This was inspired by a linear-time graph kernel. They used the structural features of 12,158 samples of malware. SVM classifier was trained to differentiate between malicious and benign applications. 89% of the malware was detectable using this technique with minimal false alarms pin-pointing the code structures which are malicious within the android applications.

In [28], the authors extracted a weighted contextual API dependency graphs as program semantics which was used to create the feature sets. Authors also introduced graph similarity metrics to discover identical application behaviours for the purpose of discovering zero-day malware. In their work, a system called DroidSIFT was implemented to evaluate 2200 malwares and 13500 benign samples. It was able to correctly classify 93% malware instances while the detection of zero-day malware had a low false negative rate of 2% and false positive rate of 5.15%.

In [25], the authors presented a method of detecting malware in android by examining the manifest files of the android application since there are differences that are significant in the manifest files of benign and legitimate applications. Manifest files carry important information of the application like the name of the package, the version number of the application, the intent filter (action, category and priority), the API level, required permission, this approach was discovered to be effective by the use of real samples of android malwares. A total of 365 samples were used and the total correct detection was 90.0%.

In [5], the authors presented an automated detection system for the android platform (AMDA) where the behaviour analysis of the application was utilized. The features of both malicious and benign applications were first extracted to form the dataset used. Various algorithm in WEKA which include, Navies Bayes, decision tress and so on were used for the analysis.

In [40], the authors proposed a technique of applying the Bayesian classifier in a novel way. 1,000 samples each for both benign and malicious application was gotten. The accuracy obtained was 0.921 and true positive rate of 0.906.

In [41], the authors proposed a machine learning approach to the detection of malware in android using 200 features. These features were taken out of both the static and dynamic analysis of android applications. The deep learning technique used had an accuracy of 96.5%.

In [42], the authors used 17 supervised learning techniques to analyse adware in android devices and compared the result. It was discovered that Random Forest was the best classifier with an accuracy of 0.9838 and a false positive rate of 0.017.

This research was inspired by the recommendations in [1]. Their research study comprises of commonly and hardly used machine learning algorithms and it was found that multilayer perceptron (MLP) outperforms the other algorithm having an accuracy of 0.994. It was therefore recommended that other dataset should be put into consideration to confirm if MLP is dataset-specific.

3. Methodology

3.1. Environment Description

This research was conducted using the Waikato Environment for Information Analysis (WEKA) version 3.8.3 which was developed in [43]. Ten algorithms out of the available algorithms in WEKA were used in this research.

3.2. Dataset Description

From publicly available datasets, two datasets were used in the study. The first was gotten from figshare which was developed in [15]. The dataset contains a total of 215 characteristics which were extracted from 3,799 applications; 1,260, malicious and 2,539, benign. The second was also developed by Yerima in [44]. The dataset contains 215 attributes with 15,036 instances of which 5,560 was malware and 9,476 was benign.

Table 1: Description of the Android Malware Datasets

Description	Dataset	
	First	Second
Instances	3,799	15,036
Malware	1,260	5,560
Benign	2,539	9,476
Attributes	215	215

Attributes/characteristics are the applications' features that were extracted. The extracted features were gotten from the manifest permission, the API call signature, the commands signature and the Intents. Some features extracted under the manifest permission are SEND_SMS, RECORD_AUDIO, READ_PHONE STATE, WRITE_SMS, USE_CREDENTIALS, GET_ACCOUNTS, MANAGE_ACCOUNTS, RECEIVE_SMS. Some features under the API call signature are IBinder, Binder, createSubprocess android.os.IBinder, URLClassLoader. Some features under intents are android.intent.action.ACTION_POWER_DISCONNECTED, android.intent.action.NEW_OUTGOING_CALL, android.intent.action.CALL_BUTTON. The command signature has only the following features /system/app, mount, /system/bin, chmod, chown and remount.

3.3. Classification Algorithms application

The training and testing was carried out with the 10 algorithms chosen in the WEKA environment. The algorithms are Random forest, Naïve Bayes, Random Tree, BayesNet, Multi-layer perceptron (MLP), Decision Table, Sequential Minimal Optimization (SMO), Bagging, KStar and IBK (also known as K Nearest Neighbours classifier (KNN)). The 10-folds cross validation and 66% split were used to train and test in the condition. In the 10-folds cross-validation, the dataset is divided

into 10 folds or parts wherein each fold is divulged in almost equal sizes as the full dataset while the 9 remaining is used for training. This procedure is repeated for each fold, making it ten times. The mean for all the 10 folds is then calculated for each performance metric. The 66% split divides the dataset into 66% and 34%. It uses the 66% part of the data for training while using the 34% for testing. The 66% split was used to validate each algorithm.

3.4. Graphical Representation of the Methodology

Figure 1 shows the graphical representation of the how the research was carried out.

3.5. Performance Evaluation

Below are the performance metrics used in evaluating the algorithms.

- Accuracy: This calculates the rate of applications correctly classified, taking into account true positive, false negative, true negative and false positive [45].

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$

where TP = True positive
FN = False Negative
FP = False Positive
TN = True Negative

- False Positive Rate (FPR). This calculates the rate incorrectly classified instances as benign. A low FP-rate implies that the classifier is good. [46]

$$\text{FPR} = \frac{FP}{FP + TN}$$

- Precision, Recall and F-measure. Precision is the ratio of instances that are positively predicted among the retrieved instances. Recall is the ratio of instances which were positively predicted among all the instances and F-measure is the harmonic mean of recall and precision.

A high F-measure is essential since both precision and recall are desired to be at high levels [47,48]. Therefore

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FP}$$

$$\text{F-Measure} = 2\,x\,\frac{Precision}{Precision + Recall}$$

- Receiver Operating Characteristic (ROC) curve. A high ROC value means that the algorithms are good. The true positive rate is plotted against the false positive rate [46].

ROC Curve is Plot of FPR(x) vs TPR

where TPR is True Positive Rate

- Root mean square error (RMSE). This is the predicted error's standard deviation. The error between the training and the testing dataset is the predicted error. A low RMSE means that the classifier is good [49,50].

$$\text{RMSE} = \sqrt{1 - r^2}\; x\; SD$$

where SD = Standard Deviation

r = Predicted error

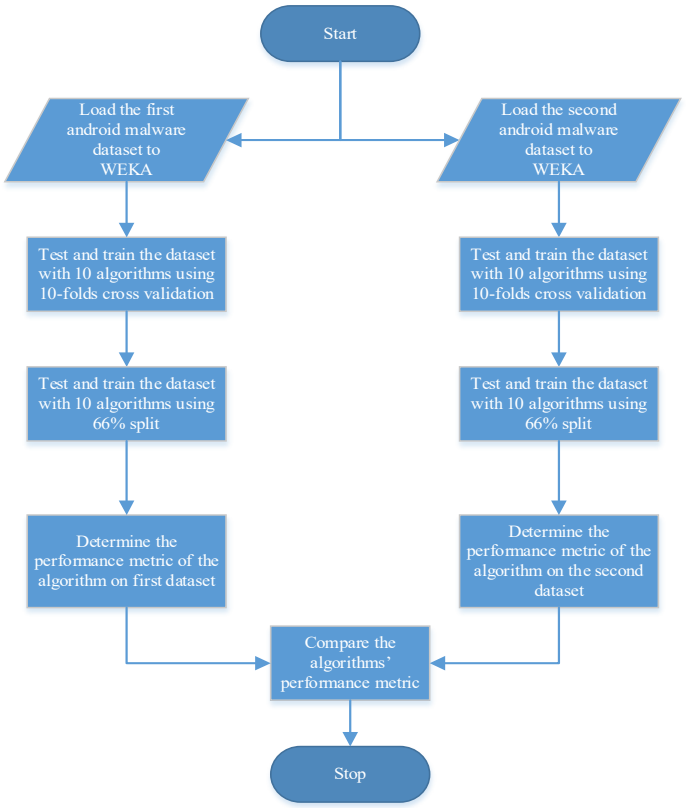


Figure 1: A graphical representation of the methodology

3.6. Algorithm Evaluated

Below are the description of two of the algorithms that stood during this research work

- Multilayer perceptron (MLP) is a perceptron that join up with extra perceptron, packed in many layers to figure out compound problems. Every perceptron in the input layer that is first layer to the left sends outputs to all the unseen layers that is the second layer perceptrons and all the second layer perceptrons send outputs to the output layer to the right that is the final layer [51].
- Random Forest: An additional dimension of randomness was introduced to bagging to create random forest in 2001. The algorithm creates tree predictors, and for all the trees in the forest, each tree depends on the random vector differently and autonomously sampled with the same dispersal. The best of these randomly chosen predictors is used in [52, 53].

4. Result and Discussion

This sections discusses the result gotten from the analysis done.

4.1. Analysis of the result

The dataset is made up of both malicious and innocuous applications. After the 10 algorithms have been trained and tested

with the 10 folds cross validation and 66 percent split in the WEKA environment, the obtained results is discussed in this section. Table 2 and table 3 show the performance metrics of dataset 1 and dataset 2 respectively. The performance metrics used are; accuracy, RMSE, false positive, ROC, precision, recall and f-measure. Figure 1 shows a display of the best performing algorithm in WEKA using the second dataset.

4.1.1. Discussion on the first dataset

While performance was optimal for all algorithms when trained and tested with the first dataset as seen in table 2, judging by accuracy, the top performing algorithm is the multilayer perceptron (MLP) having an accuracy of 99.4% when trained and tested under 66% split as previously stated in [1]. Under the 66% split, the MLP has the lowest false positive rate of 0.006 measured and also the highest recall (0.994), f-measure (0.994) and RMSE (00764).

Random Forest performs best measuring with respect to precision (0.993), accuracy (0.993), recall (0.993), f-measure (0.993) and ROC when trained and tested under the 10 folds cross validation [1]. MLP evaluated under the 66% split is, however, the best performing algorithm overall.

4.1.2. Discussion on the second dataset

In the second dataset, as seen in Table 3, while all algorithms perform well when evaluated with the dataset, the random forest had an accuracy of 98.9% when trained and tested under 10 folds cross validation making it the algorithm that performed best in terms of accuracy. The IBK has 0.013 as the lowest false positive rate. In comparison, the random forest has a reduced FPR of 0.016 and has the highest precision (0.989), recall (0.989), f-measure (0.989) and ROC (0.998). Kstar had the lowest RMSE of 0.0984.

Table 2: Showing the result of the performance metrics of the algorithms on the first datase

	Accuracy		**False Positive**		**Precision**		**Recall**		**F-Measure**		**ROC**		**RMSE**	
	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split
MLP	0.989	**0.994**	0.012	**0.006**	0.989	**0.994**	0.989	**0.994**	0.989	**0.994**	0.999	0.999	0.0938	**0.0764**
SMO	0.991	0.991	0.011	0.009	0.991	0.991	0.991	0.991	0.991	0.991	0.99	0.991	0.096	0.0923
IBK	0.992	0.991	**0.008**	0.01	0.992	0.991	0.992	0.991	0.992	0.991	0.995	0.996	0.0853	0.0851
KStar	0.992	0.991	0.01	0.01	0.992	0.991	0.992	0.991	0.992	0.991	0.999	**1**	**0.0811**	0.0823
Random forest	**0.993**	0.99	0.013	0.018	**0.993**	0.99	**0.993**	0.99	**0.993**	0.99	**1**	**1**	0.0913	0.0985
Bagging	0.986	0.983	0.02	0.023	0.986	0.983	0.986	0.983	0.986	0.983	0.999	0.998	0.1115	0.1186
Random Tree	0.972	0.972	0.033	0.029	0.972	0.972	0.972	0.972	0.972	0.972	0.97	0.972	0.1662	0.1669
NaiveBayes	0.958	0.96	0.045	0.039	0.959	0.961	0.958	0.96	0.959	0.96	0.994	0.995	0.1852	0.1821
Decision Table	0.938	0.947	0.066	0.059	0.939	0.947	0.938	0.947	0.938	0.947	0.977	0.979	0.2074	0.2032
BayesNet	0.927	0.93	0.046	0.046	0.937	0.938	0.927	0.93	0.929	0.931	0.992	0.993	0.2448	0.2482

Table 3: Showing the result of the performance metrics of the algorithms on the second dataset

	Accuracy		**False Positive**		**Precision**		**Recall**		**F-Measure**		**ROC**		**RMSE**	
	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split	10 Folds	66% Split
MLP	0.972	0.97	0.037	0.034	0.972	0.97	0.972	0.97	0.972	0.97	0.992	0.992	0.1621	0.1676
IBK	0.988	**0.985**	**0.013**	**0.018**	0.988	**0.985**	0.988	**0.985**	0.988	**0.985**	0.994	0.99	0.1032	0.1167
KStar	0.988	**0.985**	0.014	**0.018**	0.988	**0.985**	0.988	**0.985**	0.988	**0.985**	**0.998**	**0.998**	**0.0984**	**0.1113**
SMO	0.978	0.972	0.027	0.035	0.978	0.972	0.978	0.972	0.978	0.972	0.976	0.969	0.147	0.1673

Random forest	**0.989**	0.984	0.016	0.024	**0.989**	0.984	**0.989**	0.984	**0.989**	0.984	**0.998**	0.997	0.1058	0.121
Bagging	0.978	0.973	0.029	0.034	0.978	0.973	0.978	0.973	0.978	0.973	0.996	0.995	0.1333	0.1465
Random Tree	0.973	0.968	0.03	0.037	0.973	0.968	0.973	0.968	0.973	0.968	0.972	0.965	0.164	0.1791
NaiveBayes	0.834	0.836	0.122	0.126	0.865	0.863	0.834	0.836	0.834	0.839	0.949	0.839	0.3972	0.3957
Decision Table	0.922	0.907	0.077	0.093	0.924	0.909	0.922	0.907	0.922	0.908	0.976	0.962	0.2322	0.2524
BayesNet	0.828	0.829	0.124	0.128	0.862	0.861	0.828	0.829	0.831	0.831	0.928	0.925	0.4082	0.4081

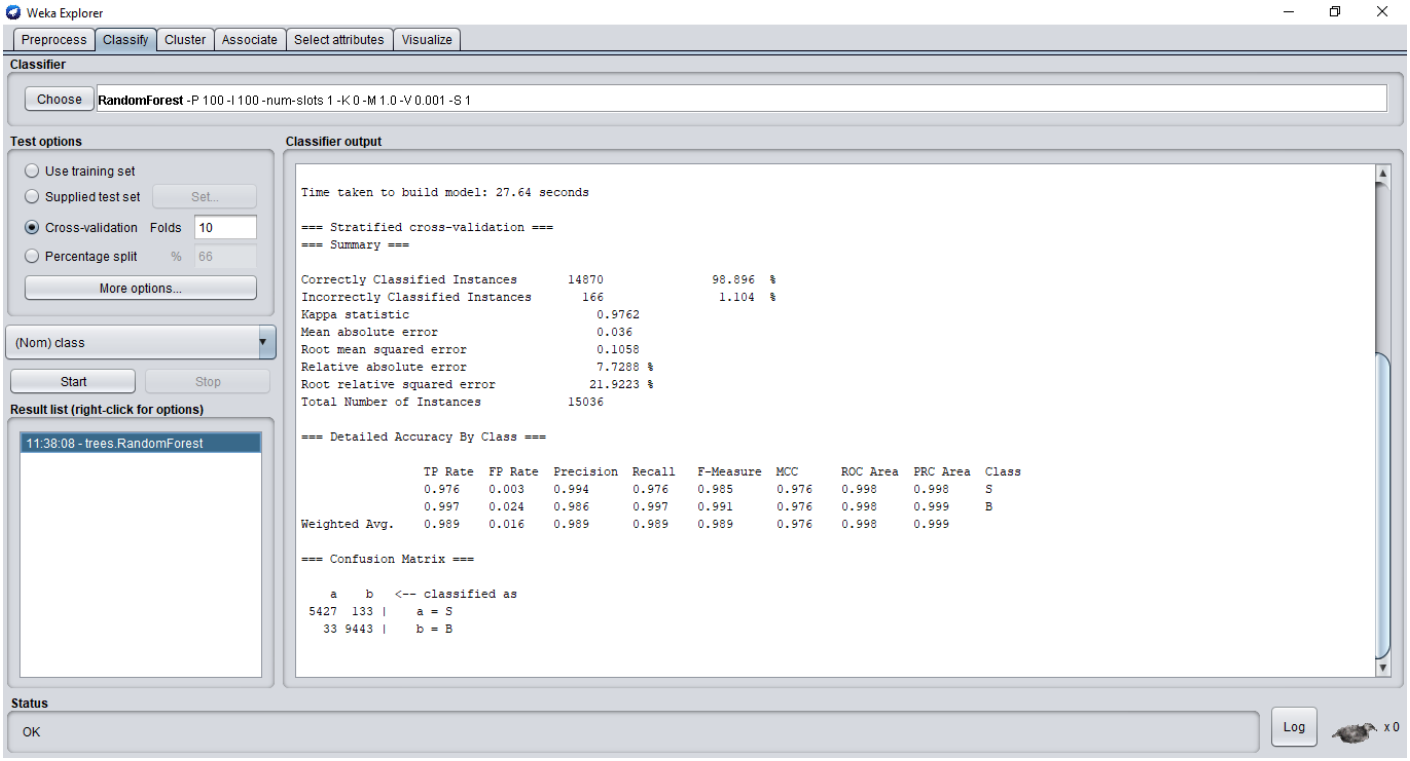


Figure 2: A display of Random forest in the environment using the second dataset.

Under the 66% split, KStar and IBK performs best with the same accuracy (98.5%), false negative (0.018), precision (0.985), recall (0.985) and f-measure (0.985). Kstar performs best in terms of ROC and RMSE with values of 0.998 and 0.1113 respectively. However, the overall best performing algorithm is Random forest tested under the 10 folds cross validation.

Although MLP was found to have the best performance under the 66% split when tested with the first dataset, it is seen that it was not able to maintain this performance when tested on the second dataset, this shows that MLP is data-set specific. This is as a result of the increased instances in the second dataset. The performance metrics of MLP was not maintained as a result of the increase in the instances. It is noteworthy in the classification of the first dataset that the analysis done under the 10 folds cross validation shows that random forest performs best with an accuracy of 99.3% having a difference of 0.001 when compared to MLP. The performance metric of random forest was further confirmed when analysis was done on the second dataset. It was observed that random forest performs best when trained and tested under the 10 folds cross validation. This therefore backs up the result of the first dataset analysis proving that random forest is not data specific.

5. Conclusion and Recommendations

Malicious applications has increasingly become sophisticated and complex with advancement in technological innovation. Models and methods for improved detection of such malware have been actively developed by various researchers in the field. The goal of this research study was to evaluate/appraise classification algorithms for malware in android detection. After the analysis in the WEKA domain, it was found that the performance metric of MLP was diminished with increased instances, making it dataset-specific. Random forest was found to perform better than the other algorithms with the first and second dataset when trained and tested under the 10 folds cross validation, having an accuracy of 99.3% and 98.9% respectively. Therefore, the contribution of this research is to determine that random forest

is the best performing algorithm in the detection of Android malware. Random forest is being proposed for use as the classification algorithm when building an android application for the purpose of android malware detection.

Further studies can consider selecting the best features of the dataset through feature selection or optimization for better performance. Other classification techniques under the reinforcement and unsupervised learning can be trained and tested. The studies should consider using Random forest as the classification and detection algorithm when developing anti-malwares, so as to provide a safe space for android users.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] O.E. Olorunshola, A.O. Oluyomi, "ANDROID APPLICATIONS MALWARE DETECTION: A Comparative Analysis of some Classification Algorithms," in 2019 15th International Conference on Electronics, Computer and Computation, ICECCO 2019, Institute of Electrical and Electronics Engineers Inc., 2019, doi:10.1109/ICECCO48375.2019.9043284.

[2] E. Chin, A.P. Felt, K. Greenwood, D. Wagner, "Analyzing Inter-Application Communication in Android," in Proceedings of the 9th international conference on Mobile systems, applications, and services, ACM: 239–252, 2011, https://doi.org/10.1145/1999995.2000018

[3] OS Timeline, Mobile operating system - Wikipedia, 2020.

[4] A. Developers, "What is android," 2011,

[5] K.J. Abela, D.K. Angeles, J.R.D. Alas, R.J. Tolentino, M.A. Gomez, "An Automated Malware Detection System for Android using Behavior-based Analysis AMDA," in International Journal of Cyber-Security and Digital Forensics (IJCSDF), Citeseer: 1–11, 2013.

[6] K. Nadia, The risks of third-party app stores, 2020.

[7] Ian Blair, Mobile App Download and Usage Statistics (2020) - BuildFire, 2020.

[8] P. Tracy, Smartphone Banking Definition & Example | InvestingAnswers, 2020.

[9] S. Verma, S.K. Muttoo, "An Android Malware Detection Framework-based on Permissions and Intents," Defence Science Journal, **66**(6), 618–623, 2016, DOI : 10.14429/dsj.66.10803.

[10] A.P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, D. Wagner, "Android Permissions : User Attention , Comprehension , and Behavior," in Proceedings of the eighth symposium on usable privacy and security, ACM: 3, 2012, https://doi.org/10.1145/2335356.2335360.

[11] M. Vennon, "Android malware: Spyware in the Android Market," Technical Report, SMobile Systems, 2010.

[12] G. Thompson, Android Malware Steals Data From Apps, Records Audio, Captures Screenshots - Binary Defense, 2020.

[13] Webroot, The Dangers of Hacking and What a Hacker Can Do to Your Computer | Webroot, 2020.

[14] Norton, Zero-day vulnerability: What it is, and how it works, 2020.

[15] S.Y. Yerima, Android malware dataset for machine learning 1 (Version 1), 2018, doi:https://doi.org/10.6084/m9.figshare.5854590.v1.

[16] J. Callaham, The history of Android OS: its name, origin and more, 2019.

[17] J. Alabaster, Android founder: We aimed to make a camera OS | PCWorld, IDG News Service, 2013.

[18] W. Bellevue, T-Mobile Unveils the T-Mobile G1 — the First Phone Powered by Android | T-Mobile Newsroom, 2008.

[19] Android version history, Android version history, Wikipedia, 2019.

[20] Android 10 | Android, 2019.

[21] M. Grace, Y. Zhou, Q. Zhang, S. Zou, X. Jiang, "RiskRanker : Scalable and Accurate Zero-day Android Malware Detection Categories and Subject Descriptors," in Proceedings of the 10th international conference on Mobile systems, applications, and services, ACM: 281–294, 2012, https://doi.org/10.1145/2307636.2307663

[22] D. Kavinya, R.S. Priya, R. Deebika, "Vulnerabilities attacks on mobile operating systems (Android versus iOS): A review," International Journal of Advance Research, Ideas and Innovations in Technology, **4**(5), 214–216, 2018.

[23] Statista, Number of available applications in the Google Play Store from December 2009 to June 2019, Google Play Store: Number of Available Apps 2009-2019, 2019.

[24] W. Enck, M. Ongtang, P. McDaniel, "Understanding Android Security," IEEE Security & Privacy, **7**(1), 50–57, 2009, https://doi.org/10.1109/MSP.2009.26

[25] R. Sato, D. Chiba, S. Goto, "Detecting Android Malware by Analyzing Manifest Files," Proceedings of the Asia-Pacific Advanced Network, **36**(17), 23–31, 2013, http://dx.doi.org/10.7125/APAN.36.4.

[26] J. Networks, Juniper networks third annual Mobile threats report, 2013.

[27] Alexandre Mundo Alguacil, C. Beek, C. Castillo, T. Dunton, J. Fokker, S. Grobman, T. Hux, N. Minihane, L. Munson, E. Peterson, M. Rivero, T. Roccia, C. Schmugar, R. Samani, R. Sims, D. Sommer, B. Sun, McAfee Labs Threats Report: December 2018, 2018, doi:10.1016/s1361-3723(19)30004-1.

[28] M. Zhang, Y. Duan, H. Yin, Z. Zhao, "Semantics-Aware Android Malware Classification Using Weighted Contextual API Dependency Graphs," in Proceedings of the 2014 ACM SIGSAC conference on computer and communications security, ACM: 1105–1116, 2014, https://doi.org/10.1145/2660267.2660359.

[29] C. Castillo, Android Malware Past , Present , and Future, 2011.

[30] S. Peng, S. Yu, A. Yang, "Smartphone malware and its propagation modeling: A survey," IEEE Communications Surveys and Tutorials, **16**(2), 925–941, 2014, doi:10.1109/SURV.2013.070813.00214.

[31] H. Gascon, F. Yamaguchi, D. Arp, K. Rieck, "Structural Detection of Android Malware using Embedded Call Graphs Categories and Subject Descriptors," in Proceedings of the 2013 ACM workshop on Artificial intelligence and security, 45–54, 2013, https://doi.org/10.1145/2517312.2517315.

[32] L. Davi, A. Dmitrienko, A. Sadeghi, M. Winandy, "Privilege Escalation Attacks on Android," in International conference on Information security, Springer, Berlin, Heidelberg: 346–360, 2010, https://doi.org/10.1007/978-3-642-18178-8_30

[33] F. Di Cerbo, A. Girardello, F. Michahelles, S. Voronkova, "Detection of malicious applications on android OS," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **6540 LNCS**(January 2010), 138–149, 2011, doi:10.1007/978-3-642-19376-7_12.

[34] K.W.Y. Au, Y.F. Zhou, Z. Huang, D. Lie, "PScout: Analyzing the Android Permission Specification," in Proceedings of the 2012 ACM conference on Computer and communications security, ACM: 217–228, 2012, https://doi.org/10.1145/2382196.2382222.

[35] A.P. Felt, E. Chin, S. Hanna, D. Song, D. Wagner, "Android Permissions Demystified," in Proceedings of the 18th ACM conference on Computer and communications security, ACM: 627–637, 2011, https://doi.org/10.1145/2046707.2046779.

[36] F. Shang, Y. Li, X. Deng, D. He, "Android malware detection method based on naive bayes and permission correlation algorithm," Cluster Computing, **21**(1), 955–966, 2017, doi:10.1007/s10586-017-0981-6.

[37] W. Ali, "Hybrid Intelligent Android Malware Detection Using Evolving Support Vector Machine Based on Genetic Algorithm and Particle Swarm Optimization," IJCSNS International Journal of Computer Science and Network Security, 19(9), September 2019.

[38] M. Kakavand, M. Dabbagh, A. Dehghantanha, "Application of machine learning algorithms for android malware detection," ACM International Conference Proceeding Series, (February 2019), 2018, doi:10.1145/3293475.3293489.

[39] M. Shohel Rana, C. Gudla, A.H. Sung, "Evaluating machine learning models for android malware detection - A comparison study," ACM International Conference Proceeding Series, (March 2019), 17–21, 2018, doi:10.1145/3301326.3301390.

[40] S.Y. Yerima, S. Sezer, G. Mcwilliams, I. Muttik, "A New Android Malware Detection Approach Using Bayesian Classification," in 2013 IEEE 27th international conference on advanced information networking and applications (AINA), IEEE: 121–128, 2013.

[41] Z. Yuan, Y. Lu, Z. Wang, Y. Xue, "Droid-Sec : Deep Learning in Android Malware Detection," in ACM SIGCOMM Computer Communication Review, ACM: 371–372, 2014, https://doi.org/10.1145/2619239.2631434.

[42] J.Y. Ndagi, J.K. Alhassan, "Machine learning classification algorithms for adware in android devices: A comparative evaluation and analysis," in 2019 15th International Conference on Electronics, Computer and Computation, ICECCO 2019, Institute of Electrical and Electronics Engineers Inc., 2019,

doi:10.1109/ICECCO48375.2019.9043288.
[43] G. Holmes, A. Donkin, I.H. Witten, "Weka: A machine learning workbench," 357–361, 1994.
[44] S. Yerima, Android malware dataset for machine learning 2, 2018.
[45] J. Brownlee, Failure of Classification Accuracy for Imbalanced Class Distributions, 2020.
[46] J. Brownlee, ROC Curves and Precision-Recall Curves for Imbalanced Classification, 2020.
[47] Metacademy, F measure, 2020, Jun. 2020.
[48] Metacademy, Precision and recall, 2020.
[49] S. Glen, RMSE: Root Mean Square Error, StatisticsHowTo.com, 2020.
[50] S. Holmes, RMS Error, 2000.
[51] Y. Taguri, N. Gilboa, S. Gigi, I. Kober, A. Farhi, S. Awrabi, J. Freeman, Perceptrons & Multi-Layer Perceptrons: the Artificial Neuron - MissingLink, Dec. 2020.
[52] L. Breiman, "Random Forests," Machine Learning, 45(1), 5–32, 2001.
[53] A. Liaw, M. Wiener, "Classification and Regression by randomForest," R News, **2**(3), 18–22, 2002.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1750-1758 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Touristic's Destination Brand Image: Proposition of a Measurement Scale for Rabat City (Morocco)

Abdellatif Elouali[*,1], Smail Hafidi Alaoui[1], Noura Ettahir[2], Abderrazzak Khohmimidi[1], Nadia Motii[1], Keltoum Rahali[3], Mustapha Kouzer[3]

[1]*Management Sciences Research Team, Faculty of Law, Economics and Social Sciences (FSJES) Agdal, Mohammed V University in Rabat, 10090, Morocco*

[2]*Management Sciences Research Team, High School of Technology Salé, Mohammed V University in Rabat, 10090, Morocco*

[3]*Laboratory of Agrophysiologie, Biotechnology, Environment and Quality, Faculty of Science, Ibn Tofail University, Kénitra, 14000, Morocco*

ARTICLE INFO

Article history:
Received: 14 October, 2020
Accepted: 21 December, 2020
Online: 30 December, 2020

Keywords:
Brand Image
Touristic destination
Measurement scale
Morocco

ABSTRACT

The objective of this article is to highlight the existing relationship between the tourist's destination brand image of the city of Rabat (Morocco) and its choice by foreign tourists to be visited once again. We suggest a reliable measurement scale able to measure the dimensions and associations of the brand image most memorized by tourists, so as to further attract the researchers and stakeholders curiosity in the tourism sector which face, every day, new challenges of attractiveness and sustainability of tourist destinations in the new Moroccan regionalization context. The survey is carried out on a sample of 454 foreign tourists in the city of Rabat because of the importance that this city requires in terms of overnight stays, that is to say nearly one million in 2018 according to the statistics of the Moroccan Tourism Observatory. The results obtained from the survey show that the functional and abstract associations jointly constitute the brand image of the city of Rabat in the foreign tourist's memory and that the abstract associations are more significant in the destination's choice.

1. Introduction

Today, tourism in all its forms is marked by strong competition between tourist destinations at national and international levels. The emergence of a large number of tourist destinations has compelled tourism managers to adopt new managerial approaches [1]. Faced with this situation, it is no longer easy to attract so many tourists because these people are more demanding and seek quality or even total satisfaction regarding a precise destination. Henceforth, in order to market their tourist sites, tourism professionals will have an interest in mastering territorial marketing in order to position themselves clearly on the most buoyant markets by adopting strategies with a political brand trend [2].

Indeed, likewise several destinations, Morocco is positioned as a country with a purely tourist vocation with 13 million non-resident tourists in 2019, according to the Moroccan Tourism Observatory, this is because of the country's geographical position, its natural and human resources, and its civilizational and historical heritage. Nevertheless, Morocco has been facing a fierce competition over the last decade with the emergence of countries such as Turkey with 39 million foreign tourists in 2018, Egypt with more than 8 million foreign tourists in 2017 and Tunisia with more than 8 million foreign tourists in 2018 (the Moroccan Tourism Observatory).

This is why our study focused on the study of the brand image of the tourist destination with the aim to analyze the measurement scale of the variable of the brand image of Rabat city tourist destination in order to define the different dimensions linked to it regarding the functional and abstract brand image. Consequently, we will present the concept of the Brand Image of the Tourist Destination and the multiple measurement instruments related to this concept in order to answer the following central questions: By which forms of perceived images can the attractiveness of Rabat city tourist destination be improved? And what are the associations with the brand image that have a greater impact on the choice of destination by tourists?

2. Territorial marketing: theoretical foundations

Territorial marketing research has experienced unprecedented growth over the last ten years. What all this research has in common is the application of marketing

[*]Corresponding Author: Abdellatif ELOUALI, Avenue le Prince Héritier BP. 227-Salé, Maroc+212663173006 – abdellatif.elouali@gmail.com

principles to the sphere, the territory, the city, and the destination [2]. In fact, several designations exist for territorial marketing: city marketing or public service marketing, urban marketing.

Territorial marketing does not only include communication strategies to promote the territory, but all the strategies that allow the attraction of tourists in relation to other territories. Territorial marketing is therefore based on effective, coherent and well-thought-out communication. The territorial brand plays a crucial role in this communication approach. From this point of view, the territorial brand is a means to attract not only tourists, but also investors, businesses and residents, who are also involved in the process of tourism development of a territory [3].

Although it is different from the enterprise, the territory will gain in adopting the strategies of the enterprise, especially the marketing strategy to be able to attract tourists in a context of strong competition [4].

3. The brand image: a multidimensional concept

The brand represents an intangible asset in order to attract new customers and subsequently building customer loyalty [1]. The brand fills numerous functions designed to support consumption. In this way, the brand carries a value and brings certain notoriety [5]. The brand is a product communication tool. For the consumer, the brand confers a reference point with which he can identify [6]. It makes it possible to identify and make the city desirable in the eyes of visitors [7]. At the level of a tourist destination, the brand takes the form of a name and a logo. For the destination, the brand carries its identity and protects its reputation [1]. The territorial brand is used as a lever of attractiveness of a tourist territory, within the framework of territorial marketing [8].

In [9], the author explains that the mark is both a signifier and a signified. It is a signifier because it enables the goods to be distinguished. Distinction does not only derive from sight, but also from other senses: touch, hearing, taste, smell. But as a signifier, the trade mark evokes a sense in the consumer. In [10], the author reports the characteristics of a brand at the territorial level, arguing that the territorial brand refers to the perception of the reputation of the territory by internal and external actors. The territorial brand therefore aims to convey the meaning of the city object. Furthermore, the city itself has always been considered a mark by nature, inasmuch as its name enables it to distinguish itself from others. Moreover, the name of a city already evokes its reputation and image in the mind. The territorial brand therefore encompasses all perceptions relating to a place and its inhabitants [11].

3.1. The notion of brand versus the notion of brand image

A differentiation is to be made between the brand and the brand image. The latter corresponds to the mental image that emerges in the individual when he hears, sees or evokes the brand [12]. In [13], the author explains that brand image corresponds to "all the associations and impressions that a consumer has in his memory about a brand. In other words, the brand image designates everything that comes to mind when talking about a brand. It can be tangible elements such as the goods or services bearing the brand, but it is also possible to associate the brand image with abstract elements from the producer's communication or from other competing brands [14]. As a result, the brand image could be assimilated to consumer representations or perceptions of a brand. It has emotional, rational and memory dimensions. However, for the consumer to be able to remember an attribute of the trade mark, he must have already experienced it or at least have been exposed to it on many occasions. The brand image is therefore the result of experiences resulting from the use of the product or exposure to it [15].

Within an enterprise context, the brand image is the perception of the brand by all stakeholders and the public [16]. From the marketing point of view, the touristic territory is assimilated to a product for sale and is associated with the brand. The territory becomes a brand that has its own image to promote through branding, but the brand image could also be considered as the characteristics that allow the individual to evaluate the brand in relation to others. The characteristics taken into account are in most cases concrete attributes [17].

The brand image of a tourist destination corresponds to its mental representation by the consumer. From this point of view, symbolic meanings are associated with the characteristics of the destination in question. This representation varies according to the experience or communications made by marketers and/or acquaintances who have previously experienced the tourist destination [18].

3.2. Brand image psychological approach

The branding process formation is split into four phases [19]: The first phase acts to express the process of the image perception by the individual's senses. It is primarily the physical characteristics of the product and the consumers' expectations and interests that affect or not the consumers' attention. If the consumer doesn't grasp the stimulus of the product by the senses, then the perception is not possible and ultimately the purchase will not occur.

During the second phase the consumer decodes the stimulus and interprets it according to his learning. This phase is unique to each individual.

The third phase deals with the perception's mental representation. The perceived image embodies a meaning and is somehow translated.

It is at the fourth phase that the individual can, through his evocation capacity, make a judgment, express an opinion, feel about the perception or simply take to purchase decision. The image itself can be divided into a desired image, transmitted image and perceived image [20].

3.3. Related associations to the brand image

In order for a brand to trigger a consumer's positive attitude, the consumer must develop strong and positive associations with the brand. It is therefore essential to create a strong image [14]. The structure of an image depends on the key attributes of the tourism product as well as the associations with the brand. However, the attributes and associations to the brand are derived from the different information retained in the consumer's mind after a consumption experience. Interactions at tourist destinations structure a complex image of the destination. The associations derive from this fact, from the confrontation between the image of the destination as perceived by the visitor before the experience, and the one developed after the tourism experience. Moreover, this complex image corresponds to the brand image perceived by visitors [21].

Branding consists of associations linking the individual to a brand. These associations can be functional or abstract. Functional associations group together attributes situations of use, benefits of using the product. Abstract associations, on the other hand, involve psychological phenomena such as affects, the attribution of a symbol to the trade mark or the emotions that the individual experiences in relation to the trade mark [18].

The functional or symbolic association with the image depends on the intensity and frequency of the positive experiences that the consumer has with the product [6].

4. Perception of destination's brand image as a communication strategy

The strong growth of the tourism industry has prompted many destinations to position themselves on this market [22], not only capital cities (Paris, London, Vienna, ...) but also cities that have developed a scientific (Cambridge) or cultural (New York, Aix-en-Provence, Dijon, ...) image, art cities such as Venice or emerging cities that have developed their own identity (Prague, Istanbul, Seville) [23].

Indeed, the competition between territories and even between cities has triggered the rise to an action mode which assumes the rise of reflections on attractiveness at several levels. Territories and destinations are the first willing to produce discourses to attract but also to retain a certain category of investors, tourists or residents.

In present day, the challenge for tourist destinations consists in differentiating themselves by developing a specific identity. They will carry out communication campaigns which may be reinforced by heritage images (culture, local customs, festivals, architecture, nature, etc.) intended to stage and perpetuate local traditions and culture [24].

Tourism communication therefore plays a capital role in the construction of the touristic destination image. Nevertheless, the multiplication of communication campaigns by the different territorial strata and the media coverage of the various events can blur the image of destinations.

Besides, it is important to study the effectiveness of communication between various territorial entities towards an urban destination in the light of the theory of social representations. Indeed, the development of destinations brands becomes compulsory for the cities' survival in the face of international competition. In order to differentiate themselves, they must develop their own identity brand and consequently create a good image.

Let us point out that touristic destinations are no longer perceived as mere entry points, boarding locations or transit points during a trip, but also, they are seen as eye-catching sites (natural or cultural resources) and attractions (appropriate development for tourism of these resources) [25] of fully-comprehensive destinations.

Previously, destinations promoted themselves thinking that the creation of tourist attractions would suffice to attract tourists (European Travel Commission). Nevertheless, exposed to tourism development, to the multiplication of communication campaigns at the different territorial levels (national, regional and city), as well as to the media coverage of the various events (terrorist attacks, strikes; riots in some countries; Pandemic, for example (Covid-19), the images and representations of destinations may be blurred.

In other words, the image keeps a very important place in the choice of a destination, and the different information sources play a crucial role in the tourists' minds regarding this image formation [26]-[28]. Consequently, the communication issued by destinations and touristic organizations influences the determination of the traveler's attitudes and behavior as well as in the formation of the tourist's image of a destination [29].

Typically, researches related to tourist destinations deal with the destination's image, which is an essential element in the choice of a touristic destination [27], as it provides information on the specific and unique attributes related to the destination before the potential stay [28].

As a result, the destination's image can, through its communication with tourists, generate a strong and distinctive image compared to its competitors. Moreover, many studies distinguish two types of images resulting from this formation process: the induced image resulting from the information of tourism actors (advertising and promotion of the destination and tourism organizations) and the organic image resulting from the consumers exposure to non-tourist information sources, such as the media outside the tourism sector [27], [28].

The organic image is often rooted in the tourists' minds because it is made up of shared representations which convey stereotypes [30].

5. Measurement scales of the Brand Image of the tourist destination

The measurement of brand image is based on the associations that consumers attribute to the brand. The stronger are the associations, the stronger is the brand equity. The associations should be strong and varied [31]. The measurement of destination branding is complex due to the fact that it is relative and dynamic and, therefore, varies over time. Thus, the models developed so far to measure the brand image of a tourist territory are based on tangible and intangible elements. Tangible elements include price, accommodation, different tourist services. The intangible elements refer to the ideas and sensations that consumers experience when they see or hear about the destination [32].

The measurement of the image of the destination is based on tourist activities and their needs, including transport, various amenities, food, services and travel prices. The image can also be measured through the characteristics of the destination and its inhabitants, such as their hospitality, the political stability of the country, economic development and environmental management. The image can also be measured through the attributes of the destination and its inhabitants. These different elements make it possible to gauge the image of a destination not only in the eyes of tourists, but also in the eyes of the whole world [33].

5.1. Quantitative and qualitative approaches

The quantitative and qualitative approaches make it possible to measure the brand image of a destination. Four scales have been indicated by [16] to measure brand image: the attitude scale, the ranking scale, purchasing intentions and turnover evolution. The attitude scale refers to the attractiveness of the

brand. The rating scale considers brand preference, as well as the characteristics that distinguish the destination from its competitors. Purchase intentions, on the other hand, provide information on the consumer's conviction about the brand. The evolution of turnover gives indications concerning the volume of purchases of the brand and, hence, the realization of the purchase intention [16].

5.2. Direct and indirect approaches

The measurement of brand image can be done indirectly through the study of perceptions, or directly, through the analysis of preferences and direct measurements. For the indirect measurement approach, it is possible to base the measurement on mental representation and the individual's ability to remember the brand. The consumer recognition or recalling of the brand allows the assessment of the brand's strength. The stronger is the brand equity, the stronger and more positive are the associations. Moreover, brand image is considered strong in this indirect measurement approach when consumers associate many attributes with it. In other words, the indirect approach to brand measurement refers to the measurement of brand awareness. The latter has six dimensions: brand acceptance, brand memorization, first brands that come to the consumer's mind, dominant brand (only one brand memorized), brand awareness and brand opinion [16].

6. Psychometric approach

Branding image is a multidimensional concept, particularly since it is often measured indirectly through its sources and its consequences. Unlike the various techniques already mentioned, psychometric measurements display the advantage of being more practical, more operational and above all more direct. In fact, it is a question of requesting consumers' opinions, through a pre-established questionnaire. In other words, satisfaction's surveys should be carried out among tourist consumers. The main measurement scale developed in this respect is the initiative of [34]. Nevertheless, the representations' mental extraction from the consumer's mind requires qualitative methods that are more rigorous than metric scales.

7. Brand's image other measures

7.1. Explicit measures

When the consumer proceeds to a purchase decision, he tries to remember the advertising message and what the brand represents. It is above all the level of learning of the message that varies the result of this consumer research in classic or so-called Territorial Marketing. The effectiveness of an advertisement and the product or territory brand will be evaluated on the basis of the words memorized. This is an explicit measure, as it focuses on a single person and on what that person has retained from an explicit message.

Sometimes there may be a distortion between what the consumer has memorized, and the information given. In this case, the person does not correctly restitute what he or she has "learned" from the advertisement but adds other incorrect associations or fractions.

7.2. Implicit measures

Implicit measurement is useful to advertisers to determine what consumers know about their brand, regardless of the information source. Different types of measurement can be distinguished [18]:

- Memorization: measuring the effect of advertising exposure on brand memorization;
- Spontaneous notoriety: consists of measuring of advertising effectiveness by asking consumers to name brands for a product category.

8. Attitude study

Consists of determining the perception of the brand and not what consumers know about it. In this method, they make a judgment on the basis of the various criteria of the brand.

The aim of these pre-tests is to determine what makes consumers change their brand preferences. We strive to single out the elements that build consumer's loyalty. Furthermore, classical theories indicate that there are strong and positive associations in the consumer's mind that influence the general positive attitude towards the brand [35], [36]. The literature then suggests that the strength of the image, i.e. its ability to translate itself into utility and loyalty behavior, depends mainly on four attributes according to [37]:

- The strength of associations;
- The dominance of associations;
- The valence of associations;
- The cohesion of the associations.

Therefore, in [37] the author described that only associations considered by the consumer as positive, strong, unique, and coherent, build the strength of the image. For instance, they affect the overall preference for the brand or a re-purchase or recommendation behavior. By analogy, and mentioning in [38], we suggest that attitudinal fidelity to a destination depends on the strength and valence of its image. Ultimately, it is expected that a tourist with strong and positive associations with the destination brand would be more inclined to revisit it and recommend it to others.

After our review of the literature, it appears that there is no agreement on the dimensions that form the brand image of the tourist destination. In addition, many of the brand features of the touristic destination have remained vague and lacking of clarity.

This generalism causes a certain ambiguity between certain concepts close but different from the brand image of the tourist destination, which causes a problem in terms of its determinants and its scales of measurement. Therefore, the design of a tool (scale) to measure the image of the tourist destination presents a real challenge in order to guide future research and support the professionals in order to better manage their management system of the brand image of the touristic destination.

9. Research methodology

In order to develop a standardized measurement tool, we will adopt the advocated approach known as the Churchill's paradigm [39]. The objective of this methodological approach is incorporating the knowledge of measurement theory and the appropriate techniques for improving it into an automatic procedure.

Similarly, Churchill's paradigm offers the possibility of rigorously designing measurement instruments such as multi-scale questionnaires. To this end, we deemed it appropriate to use, first of all, a qualitative approach to explore the characteristics of the brand image of the touristic destination.

In our field investigation, we used the semi-directive interview for data collection. This method is favored given its great flexibility and the wealth of information it can produce [40].

According to the literature, we have elaborated an interview guide that includes about the respondent's identification, age, gender and the determinants of the brand image of the touristic destination; the reputation of the touristic destination and ultimately the expectations of tourists towards (Rabat city).

The purpose of all these questions is to understand how the touristic destination is to understand how the brand image of the touristic destination is perceived by foreign tourists, and consequently to identify the different characteristics of the brand image of the touristic destination. Ten interviews were conducted with random tourists in Rabat city with average time duration of twenty (20) minutes. The population sample was characterized by female predominance (06 women and 04 men). The average age span is 28 years.

Afterwards, we have used a quantitative method. Indeed, in order to proceed in a methodological logic, we have formulated items based on all the items used in the different studies examined in the literature and the attributes collected after processing the information from the interviews (the results of the qualitative analysis will be dealt with in another research paper). Hence, a first version of the questionnaire was drawn up, then it was examined by two teacher-researchers in management sciences, and finally it was tested with 16 tourists from the city of Rabat.

The comments collected made it possible, on one hand, to improve the questionnaire, and on the other hand, to underline the latest version of the questionnaire which consists of 26 items on a five-point Likert's scale. Also, we also used exclusively affirmative sentences (positive statements) following Devellis' recommendations. The data were collected in paper format. 454 usable questionnaires were collected. The representative population of the control sample is composed of 64% men and 36% women. The majority of the tourists are situated in the age group of 25 to 60 years old.

The following figures present the various information about our sample population (gender, age, country of origin):

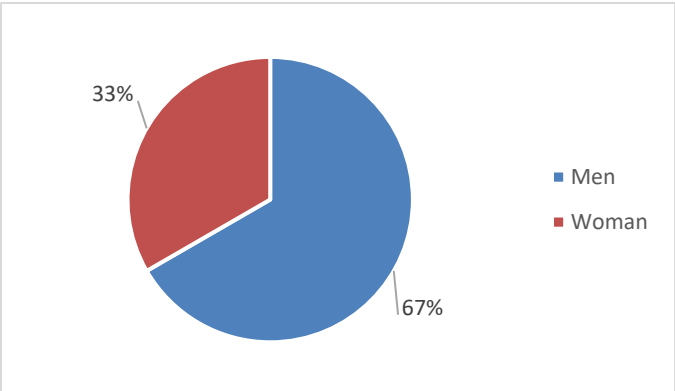


Figure 1: Tourists distribution according to the gender

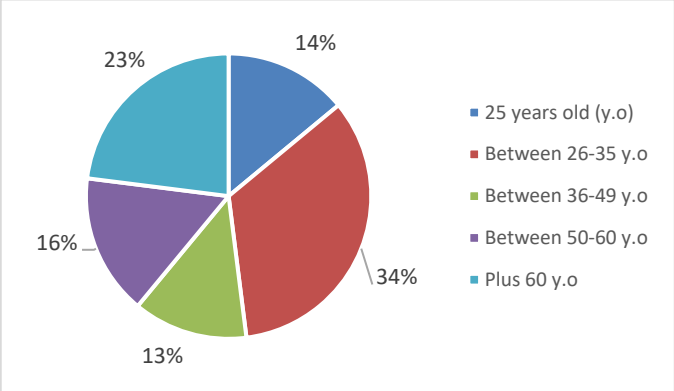


Figure 2: Tourists dispatching according to age

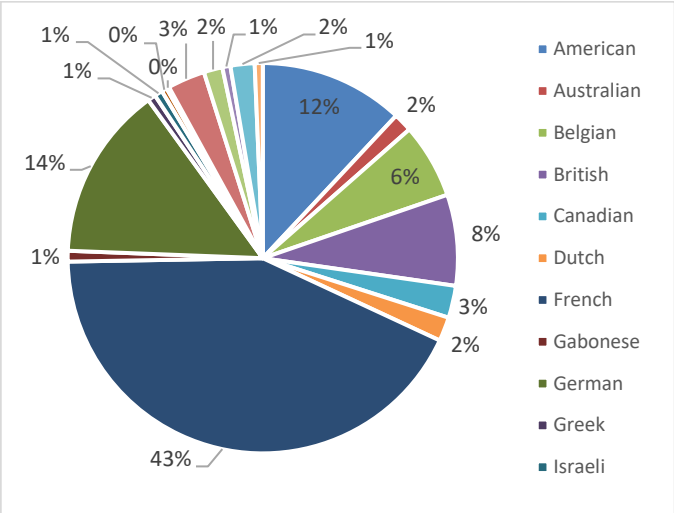


Figure 3: Tourists' dispatching according to their nationality

10. Exploratory Factor Analysis

To elaborate our measurement scale, we have gone through several steps. First, we have verified the content validity of the scale for measuring the brand image of the tourist destination. Content validity tests aims to check whether the questionnaire's different items are a good representation of the concept under consideration [41]. For this purpose, as explained before, we reviewed the academic literature to accompany and guide us in the questionnaire's elaboration and to reach a measuring scale of the brand image of the touristic destination beforehand. Indeed, the large number of items constituting the measurement scale (16 items), their heterogeneity and the way in which they can be taken out of the dimensions, offer the researcher a choice between three types of analysis' approaches:

- Independently, to examine each item by selecting only one item per theme;
- Aggregate all the items into a single scale;
- Undertake a factor analysis to identify the different underlying dimensions.

We have decided to introduce all items in a single scale without distinguishing between dimensions at the questionnaire level, in order to test the stability of the factor structure of the measurement scale we will use. The factor structure and the psychometric qualities of the scale of the tourist destination brand image were analyzed using SPSS (Statistical Social Sciences Package) 23 software, the data were subjected to

principal component factor analysis with Varimax rotation in order to test the dimensionality of the construct.

The Principal Component Analysis (PCA) is considered desirable in the development phase of a questionnaire in order to decrease the number of items and retain only keep those that allow the phenomenon to be characterized, particularly in order to identify the main factors [42].

The items purification deemed unsatisfactory was performed using the following elimination parameters: the rejection of items with a factor score of less than 0.3 and those with a high factor score on several factors [43]. The internal consistency reliability of the scale and its different dimensions was assessed by the Cronbach's alpha analysis as well as the correlation level for dimensions with at least two items.

11. Results and discussions

We undertook an exploratory approach to determine the dimensions of the tourism destination's brand. Firstly, we have prepared the number of variables, types of variables and finally the sample size.

Secondly, we looked at inter-item correlations. Then, we have measured the adequacy of sampling (KMO) and Bartlett's sphericity Test. The results showed that the KMO index is 0,758. It is qualified as good. This index shows that correlations between items are of good quality. After this, the result of the Bartlett's sphericity test is significant since $p < 0, 0005$ (Table 1).

Table 1: KMO Index and Bartlett's Test

Sampling precision measurement Kaiser-Meyer-Olkin	0,758
Approximate Khi-two	3664,384
Bartlett's sphericity test ddl	120
Bartlett's significance	0,000

In order to extract the appropriate of factors for our scale, we first have analyzed the table of total variance explained, and we have noticed that 4 components have an eigenvalue higher than 1. The 4 factors explain 66.38% of the variance (Table 2). These factors are considered significant.

Table 2: Initial eigenvalues

Component	Initial eigenvalues		
	Total	% of the variance	% cumulated
1	5,277	32,983	32,983
2	2,595	16,219	49,202
3	1,496	9,350	58,552
4	1,253	7,831	66,383
5	0,994	6,211	72,594
6	0,723	4,518	77,112
7	0,670	4,188	81,299
8	0,540	3,378	84,677
9	0,453	2,834	87,511
10	0,442	2,766	90,276
11	0,401	2,503	92,780
12	0,359	2,247	95,026
13	0,274	1,713	96,739
14	0,192	1,200	97,940
15	0,176	1,103	99,043
16	0,153	0,957	100,000

Secondly, and for more certainty, we have created a graphic from the eigenvalue using SPSS 23 (Figure 4) through the collapse layout and after examination of the Cattell's Elbow rupture. We have noticed that there is a change after the third factor. For this purpose, we retain three (03) factors that lie before the abrupt change in the slope of the Cattell's Elbow break for the analysis, since this criterion is precise than the eigenvalue criterion.

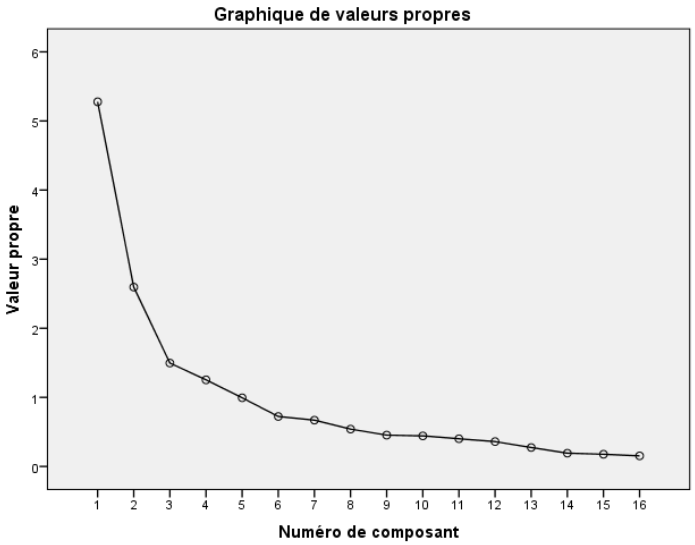


Figure 4: Initial eigenvalues

Table 3: Total variance explained through three dimensions (Extraction method: Principal Component Analysis)

Component	Initial eigenvalues		
	Total	% of the variance	% cumulated
1	3,537	32,153	32,153
2	2,125	19,320	51,473
3	1,327	12,064	63,537
4	0,901	8,193	71,731
5	0,808	7,345	79,076
6	0,630	5,731	84,807
7	0,463	4,210	89,017
8	0,382	3,470	92,488
9	0,340	3,087	95,575
10	0,303	2,754	98,328
11	0,184	1,672	100,000

The exploratory analysis in Table3, shows three (03) factors according to [44], which leads to the selection of the number of factors with a value of 1or greater. A clear factorial solution appears, without overlap. Factors are easily interpretable. In addition, the selected items appear to represent the brand image of the touristic destination in terms of content.

According to the component matrix (Table 4), the first factor consists of 5 items that represent the abstract brand image [45]. The latter characterizes the destination which is Rabat city as an adventures' destination judged by the foreign tourists interviewed, it is unique by its diverse touristic sites, A Sports destination because of the number of sporting events organized, namely the Rabat International Athletics Meeting, the Rabat International Marathon, the Hassan II Golf Trophy ,a city ranked second in the Maghreb region in terms of life quality after Tunis (capital of Tunisia), with a green belt of 1,063 hectares, Rabat

city managed to double the global area of green spaces for each individual given the standard of the World Health Organization set at 10 square meters per person (Tests garden), Nazhat Ibn Sina, Hilton forest, Maamora,... and ultimately, an events' destination (sporting, artistic and cultural).

In this regard, the author [46] argues that it is likely not all associations to the trademark have the same impact on behavior: associations involving symbolic benefits that are also abstract, would have a stronger influence than associations involving functional benefits.

The second factor represents the functional brand image, translated by the aspect of belonging to the territory and the local products, is composed of three items. These are related to the Peaks and Scents, Region City of the Rabat-Salé-Kénitra and the cultural activities namely Mawazine and the Rabat Jazz Festival for instance.

The third "Human and Intangible" factor is composed of three items. Rabat is a friendly city, a welcoming city, a wealthy city in terms of Historical Monuments (Mohammed-V Mausoleum, Oudayas Kasbah, Chellah, Rabat Saint-Pierre Cathedral, Hassan Tower, Mohammed VI Museum …)

Table 4: Components Matrix after rotation

	Component		
	1	2	3
IMA6	0,840		
IMA5	0,837		
IMF10	0,777		
IMA3	0,658		
IMF8	0,564		
IMF4		0,810	
IMA1		0,793	
IMF5		0,673	
IMA2			0,796
IMF9			0,769
IMF7			0,610

Table 5: Main Components' Analysis (Varimax Rotation) of the Rabat Destination Branding Scale (n=454)

Items	Components			Factors
	1	2	3	
IMA6. Adventures' destination	0,840			Brand image Marque Symbolic, Abstract
IMA5. Unique destination	0,837			
IMF10. Sports destination	0,777			
IMA3. Green city	0,658			
IMF8. Events (athletic, artistic, cultural).	0,564			
IMF4. Spices and scents		0,810		Belonging to the Territory and Local Products
IMA1. Region's city of Rabat Salé Kénitra		0,793		
IMF5. Cultural activities		0,673		
IMA2. Friendly city			0,796	Human capital
IMF9. Welcoming city			0,769	
IMF7. Historical monuments			0,610	Human, Capital Intangible

The three factors explain 63,537% (Table 3) of the total variance explained for a KMO of, 758 (Table 1). The internal consistency reliability for each component is mentioned in Table 5. After identifying the items that constitute our measurement scale (Table 5). We presently want to verify whether this scale is stable over time and whether it makes it possible to properly measure the brand image of the destination we have identified. So, we are going to perform an internal consistency analysis.

Table 6 displays the Cronbach's Alpha Index values for each component of the destination brand image. We observe that the value of the 1st component «Symbolic Brand Image» is 0,817, which is excellent, since it exceeds the minimum required threshold of 0,70 [47].

This beacon is arbitrary, but widely accepted by the scientific community. Consequently, we can claim that we obtain, for this scale composed of five elements, a satisfactory internal coherence. Similarly, the value of the 2nd component «Belonging to the Territory and Terroir Product» is 0,764, which is excellent. While the 3rd component's value «Human and intangible capital» is 0,673, which is acceptable.

Table 6: Cronbach's alpha coefficients

Components	Cronbach's alpha values
Symbolic brand image	0,817
Belonging to the territory and local product	0,764
Human and intangible capital	0,673

After examining the tables of total statistics of the elements corresponding to each dimension of the measurement scale of the Brand Image of Rabat Destination (Appendix 1), we have noticed that the Cronbach's alpha values of the first dimension improves (0,819 instead of 0,817) if we suppress item IMF8. Also, the Cronbach's alpha value of the third dimension improves (0,705 instead of 0,673) if we discard item IMF7. So our measurement scale of the Brand Image of the Final Destination of Rabat is presented as follows:

Table 7: City of Rabat Brand Image measurement scale

Factors	Items	Cronbach's Alpha
Symbolic brand image (abstract or functional)	IMA6. An adventures' destination	0,819
	IMA5.A unique destination	
	IMF10.Sports destination	
	IMA3. Green city	
Belonging to the territory and local products.	IMF4.Spices and scents	0,764
	IMA1.Region city of Rabat Salé Kénitra	
	IMF5. Cultural activities	
Human and intangible capital.	IMA2.Friendly city	0,705
	IMF9.Welcoming city	

11.1. Study results and implications

The results obtained show that functional and abstract associations jointly form the brand image of Rabat city in the foreign tourists' memory and that abstract associations are more significant in the choice of the destination. These results confirm that the brand image is constituted of several attributes linked to the associations, whether functional or abstract (the belonging to the territory and local product 0.764). The respondents consider that they are part of this destination and that local products remain a strong argument in the choice of destination. These data make it possible to mobilize a different approach in the management of the Rabat destination which prioritizes a diversified offer for foreign tourists capable of enriching this dimension of belonging and preference for local products and improving their quality.

The resources of the Rabat city destination mentioned in this study represent important distinctive and specific qualities in the overall satisfaction for the visit to a touristic location[48]. Moreover, the attributes linked to associations are a real means in order to develop and strengthen the brand image and, consequently, improving the attractiveness of the destination[49]. The results of this study represent a real reference for the managers of the Rabat destination and the Moroccan tourist sector for new prospects for the segmentation of the touristic offer and the opening up of new markets.

Conclusion

The aim of this research is suggesting a measurement scale of the touristic destination image in a Moroccan setting (Rabat city). To achieve this goal, we have conducted two empirical investigations. First, we have conducted a qualitative exploratory study in the form of ten semi-directive interviews, during July 2019, with a multi-nationality tourists visiting Rabat. Each interview lasted about twenty minutes. Then we have performed a quantitative study among 454 foreign tourists in Rabat during the time period from October to December 2019, and our results show:

- The existence of two different tourist's profiles based on the brand image they develop in their memory, a functional or abstract image and that the abstract brand image (Rabat unique city, city of adventure,0,819 of Cronbach's alpha) is more significant than the functional image (sports city...).These results are more compliant with those of [47] who suggested that not all associations with the brand have the same impact on tourist's behavior, with so-called symbolic or abstract associations having a stronger influence than those relating to functional benefits;
- The brand image, whether functional or abstract, constitutes an enormous potential for the Rabat destination (factor of belonging to the territory and terroir's product) of our model, 0,764);
- Our measurement scale suggests three (03) complementary factors, 9 items to measure the brand image of Rabat city. The three factors explain 63,537% (Table 3) of the total variance explained for a KMO of 0,758 (Table 1);
- At the managerial level, this study will enable the stakeholders of the national tourism sector to draw the necessary conclusions to segment the tourist's market at the level of Rabat city and at the national level according to the identified behaviors and preferences.

Likewise, all scientific research, our study displays some limitations which constitute as many threads of investigation. In fact, the scale has only been tested on one category of foreign tourists - those of Rabat city - and it is therefore appropriate to generalize the measure to several cities. Such an approach can improve the predictive validity of our scale at the national level.

Certain other research perspectives related to the results can be suggested. It would be interesting to study the structural relationships between the 3 factors of our model and to test their impact on other concepts such as attachment to the destination.

References

[1] M. Abakouy, M. Khatib, "Positionnement marketing et développement de la ville de Casablanca en tant que destination touristique," 13, 2016.

[2] B. Meyronin, J. Gayet, G. Collomb, Marketing territorial: enjeux et pratiques, Vuibert, Paris, 2015.

[3] A.E. Khazzar, H. Echattabi, "Les pratiques du marketing territorial dans le contexte marocain : Eléments de réflexion," **16**(1), 14, 2016. https://doi.org/10.48396/IMIST.PRSM/mjeim-v2i1.8989

[4] P. Marcotte, L. Bourdeau, E. Leroux, "Branding et labels en tourisme : réticences et défis," Management Avenir, **47**(7), 205–222, 2011.

[5] M.-U. Proulx, D. Tremblay, "Marketing territorial et positionnement mondial," Geographie, economie, societe, **8**(2), 239–256, 2006. https://www.cairn.info/journal-geographie-economie-societe-2006-2-page-239.htm

[6] C. Camelis, "L'influence de l'expérience sur l'image de la marque de service," Vie sciences de l'entreprise, **182**(2), 57–74, 2009. https://doi.org/10.48396/v2i1.8989

[7] F. Smaoui, "Les déterminants de l'attachement émotionnel à la marque : Effet des variables relationnelles et des variables relatives au produit," 27, 2008.

[8] F. Cusin, J. Damon, "Les villes face aux défis de l'attractivité. Classements, enjeux et stratégies urbaines," Futuribles, **367**, 25–46, 2010, doi:10.1051/futur/36725.

[9] T. Albertini, D. Bereni, G. Luisi, "Une approche comparative des pratiques managériales de la Marque Territoriale Régionale," Gestion et management public, **5**(2), 41–60, 2017. https://www.cairn.info/revue-gestion-et-management-public-2017-2-page-41.htm

[10] C. Lai, I. Aimé, La marque, 2016.

[11] H. Azouaoui, A. Ismaili, "L'identité de la marque de territoire et décision de localisation des entreprises : Approche par la littérature," 21, 2015.

[12] E. Belkadi, "Marketing Territorial de Casablanca : Etude de l'Image de Marque [Place Marketing: The Brand image of Casablanca]," **13**(3), 11, 2015.

[13] É. Tortochot, G. Alt, Design(s): de la conception à la diffusion, Bréal, Rosny-sous-Bois, 2004.

[14] A.-C. Marchat, C. Camelis, "L'image de marque de la destination et son impact sur les comportements post-visite des touristes," Gestion et management public, **5/3**(1), 43, 2017, doi:10.3917/gmp.053.0043.

[15] Malaval Philippe, Pentacom: communication corporate, interne, financière, marketing b-to-c et b-to-b / Philippe Malaval,... Jean-Marc Décaudin,... ; avec la collaboration de Christophe Bénaroya,... Jacques Digout,..., 3e édition, Pearson, Paris, 2012.

[16] A. ABYRE, Y. ALLAOUI, "Savoir mesurer l'image de marque pour pouvoir l'améliorer," Management Research, 7.

[17] J.-L. Leignel, E. Ménager, S. Tablonsky, Performance durable de l'entreprise. Piliers, évaluation, leviers - Jean-Louis Leignel,Emmanuel Ménager,Serge Tablonsky, Dec. 2020.

[18] H. Berrehail, E.H. Boukalkoul, "Concept de l'image de marque," . . May, **5**, 19, 2016.

[19] P.S. Manhas, L.A. Manrai, A.K. Manrai, "Role of tourist destination development in building its brand image: A conceptual model," Journal of Economics, Finance and Administrative Science, **21**(40), 25–29, 2016, doi:10.1016/j.jefas.2016.01.001.

[20] J. Lacœuilhe, "L'attachement à la marque : proposition d'une échelle de mesure," Recherche et Applications En Marketing, **15**(4), 61–77, 2000.

[21] G. Prayag, S. Hosany, K. Odeh, "The role of tourists' emotional experiences and satisfaction in understanding behavioral intentions," Journal of Destination Marketing & Management, **2**, 118–127, 2013, doi:10.1016/j.jdmm.2013.05.001.

[22] G. Dunne, S. Flanagan, J. Buckley, "Towards an Understanding of

International City Break Travel," International Journal of Tourism Research, **12**, 409–417, 2009, doi:10.1002/jtr.760.
[23] V. Minghetti, F. Montaguti, Cities to Play: Outlining Competitive Profiles for European Cities, Springer, Vienna: 171–190, 2010, doi:10.1007/978-3-211-09416-7_10.
[24] M. Gigot, "La patrimonialisation de l'urbain," 9, 2012.
[25] F.D. Grandpré, "Attraits, attractions et produits touristiques : trois concepts distincts dans le contexte d'un développement touristique régional," Téoros. Revue de recherche en tourisme, **26**(26–2), 12–18, 2007.
[26] T. Vo Thanh, V. Kirova, R. Daréous, "L'organisation d'un méga-événement sportif et l'image touristique de la ville hôte : une perspective par le concept de transfert d'image," Téoros : revue de recherche en tourisme, **33**(1), 87–98, 2014, doi:https://doi.org/10.7202/1036722ar.
[27] H. Kim, J.S. Chen, "Destination image formation process: A holistic model," Journal of Vacation Marketing, **22**(2), 154–166, 2016, doi:10.1177/1356766715591870.
[28] C. Dubois, M. Cawley, S. Schmitz, "The tourist on the farm: A 'muddled' image," Tourism Management, **59**(C), 298–311, 2017.
[29] P. Cottet, M.-C. Lichtlé, V. Plichon, J.-M. Ferrandi, "Image d'hospitalité des villes touristiques : le rôle de la communication," Recherches en Sciences de Gestion, **108**(3), 47, 2015, doi:10.3917/resg.108.0047.
[30] S. Zenker, E. Braun, S. Petersen, "Branding the destination versus the place: The effects of brand complexity and identification for residents and visitors," Tourism Management, **58**, 15–27, 2017, doi:10.1016/j.tourman.2016.10.008.
[31] H. Guizani, P. Valette-Florence, "Proposition d'une mesure psychométrique du capital client de la marque," Marche et organisations, **N° 12**(2), 11–41, 2010.
[32] P. Scorrano, M. Fait, L. Iaia, P. Rosato, "The image attributes of a destination: an analysis of the wine tourists' perception," EuroMed Journal of Business, **13**(3), 335–350, 2018, doi:10.1108/EMJB-11-2017-0045.
[33] H. Zhang, Y. Wu, D. Buhalis, "A model of perceived image, memorable tourism experiences and revisit intention," Journal of Destination Marketing & Management, **8**, 326–336, 2018, doi:10.1016/j.jdmm.2017.06.004.
[34] B. Yoo, N. Donthu, Developing a Scale to Measure the Perceived Quality of an Internet Shopping Site (PQISS), Springer International Publishing, Cham: 471–471, 2015, doi:10.1007/978-3-319-11885-7_129.
[35] C. Viot, "David Aaker: Efficacité publicitaire, capital marque, comportement du consommateur et lien marketing-finance," 21, 1994.
[36] K.L. Keller, "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," Journal of Marketing, **57**(1), 1–22, 1993, doi:10.1177/002224299305700101.
[37] S. Changeur, Le territoire de marque : proposition et test d'un modèle basé sur la mesure des associations des marques, These de doctorat, Aix-Marseille 3, 1999.
[38] K. ZENAIDI, J.-L. Chandon, "Une nouvelle mesure du capital marque par une approche formative," Revue Tunisienne Du Marketing, **1**(1), 8–21, 2009.
[39] G.A. Churchill, "A Paradigm for Developing Better Measures of Marketing Constructs," Journal of Marketing Research, **16**(1), 64–73, 1979, doi:10.2307/3150876.
[40] M.B. Miles, A.M. Huberman, Analyse des données qualitatives, De Boeck Supérieur, 2003.
[41] L. Dialogues, Market - 3ème édition - Études et recherches en... - Yves Evrard, Bernard Pras, Elyette Roux - Dunod, Dec. 2020.
[42] A.B. Costello, J. Osborne, "Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis," 7, **10**, 2005, doi:10.7275/JYJ1-4868.
[43] P. Roussel, F. Wacheux, Management des ressources humaines : méthodes de recherche en sciences humaines et sociales / sous la direction de Patrice Roussel et Frédéric Wacheux, de Boeck. Bruxelles, 2005.
[44] H.F. Kaiser, "The varimax criterion for analytic rotation in factor analysis," Psychometrika, **23**(3), 187–200, 1958, doi:10.1007/BF02289233.
[45] J.L. Crompton, "Motivations for pleasure vacation," Annals of Tourism Research, **6**(4), 408–424, 1979, doi:10.1016/0160-7383(79)90004-5.
[46] M. Korchia, "Une nouvelle typologie de l'image de marque," 16ème congrès international de l'Association Française du Marketing, 18, 2001.
[47] J.C. Nunnally, Psychometric theory, 2nd Edition, McGraw-Hill, New York, 1978.
[48] A.F. Lecompte, M. Trelohan, M. Gentric, M. Aquilina, "Putting sense of place at the centre of place brand development," Journal of Marketing Management, 2017.
[49] R.H. Taplin, "The value of self-stated attribute importance to overall satisfaction," Tourism Management, **33**(2), 295–304, 2012, doi:10.1016/j.tourman.2011.03.008.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1759-1777 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Designing of Institute's Educational Mascots for Brand Identity

Nop Kongdee*, Suparada Prapawong, Manissaward Jintapitak

Department of Animation and Visual Effects, College of Arts, Media and technology, Chiang Mai University, Chiang Mai, 50200, Thailand

ARTICLE INFO

Article history:
Received: 05 November, 2020
Accepted: 21 December, 2020
Online: 30 December, 2020

Keywords:
Character Design
Mascot Design
Toy and Souvenir
Representative
Graphic Identity
Advertising
3D Printing
Institute Marketing
Thai Cultural
Thai Character Design
University Advertising
Graphic Design
Symbol analysis
Concept Arts
Arts and Design

ABSTRACT

The purpose of this study is to show how to design the character used as the symbol or Mascot of the educational institution and how to apply the design results into advertising and Brand identity design such as printing art, poster, digital media, toy, and souvenir. In this study, the ideas of mascot design have been leaded to present a contemporary design under the conceptual framework from local culture, organizational culture, theories of society, arts and design through the analysis and synthesis of the organization, which is CAMT from Chiang Mai University, as the model of this study. CAMT has 7 bachelor's degree majors, namely Animation and Visual effects (ANI), Software Engineering (SE), Digital Film (DF). Modern Management and Information Technology (MMIT), Digital Game (DG), Digital Industry Integration (DII) and Knowledge Innovation Management (KIM). These Mascot design group were applied into Art product such as poster, Line sticker and end up with Toy product as a final stage, it was using digital 3D printing for action figure product in prototype level. This study including Toy creating concept designing and how the outcome product affected to Mascot identity, institution brand and the target groups of study which are widely from generation X to Z, mostly age range 18-23 years old. As a result, it founded that this Mascot package advertising becomes the distinctive character creating the new image of CAMT and the enhanced of Toy design product that raise CAMT remembered as a contemporary and digital organization in Chiang Mai University. Accordingly, it produced the impact results to receivers in many dimensions involving trend, creative behaviors such as Fan Art design, Doujinshi, or Caricature, Cosplay, Arts, and craft also Subculture from the public user.

1. Introduction

This paper is an extension of work originally presented in 2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering [1]. In the lights of arts and designing, a mascot[1] is an effective public relation, promotion and marketing means of building brand images. A mascot is an extension of a logo in order to create the deeper dimension of image by building graphic identities or brand identities, it bears the responsibility of lifting the spirits of the group as well [2] which is highly popular nowadays. The building of a brand can rely on different techniques. However, most of the time, it involves the design of products such as shirts, hats or even collectibles that are memorable and tangible.

As for this project, it concerns the creation of new mascots based on the rebranding of College of Arts, Media, and Technology (CAMT[2]) Organization, Chiang Mai University,

Thailand. This college offers courses that are relating to Digital Contents in fields of Sciences and Arts focusing on innovation and digital technology. This project involves the designing of the new mascots for CAMT that correspond to the vision of the organization that emphasizes on digital innovation and advancement. Thus, the direction for the designing of the mascots/characters focuses on modernity and concurrence with the vision of CAMT and images of Chiang Mai University while

[1] Mascot is any person, object, animal used to represent a group of common identity mostly use for professional sport team also involve with advertising and marketing of organization or team.

*Corresponding Author: Nop Kongdee, Email : nopmerzy@gmail.com

building new images that focus on the dimension if arts and designing of Chiang Mai province simultaneously. Thus, it is apparent that the scope for the design corresponds to culture, images and perspectives that are overlapping on one another in various dimensions. Nevertheless, a new and meaningful image, which is the core outcome of this project, depends on the designing method, analysis and definition.

Marketing and promotional extensions an important element that leads to the memorization of brand identities. To achieve this requires images to consist of multi-dimensional perspectives that do not go beyond the conceptual framework, such as the design of envelope, letterhead, graphics on shirts and even graphic posters in the forms of digital media and digital image communication (stickers of Line application), as well as the design of toys and souvenirs, which are included as parts of the creation in this project. It is expected that the designed mascots will lead to the added value to the brand and the wide publication of promotional media and images in unique manners.

2. Literature Review of Mascot and Brand

In [3], it is stated that brand identity is how a company is being identified. The consistency of this brand identity is formed by its features like culture, vision, personality, positioning, presentations, relationships and other meaning beliefs followed by the entity. In term of commerce, brand identity is required to create costumers' awareness and remembrance in order to get access and chosen from consumers. However, there are many methods of brand identity expression which mascot is one of the potential paths.

It is quoted that the origin of the word "Mascot" is derived from "Mascara" or mask in the etymological dictionary [4], Corominas (1980, 1985) collected the first identifiable written use of the word "Mascota" or mascot in the year 1233. The word "Mascot" comes from the provincial language "Mascoto" which means magical things or amulets. Afterwards, it has been abbreviated to be "Masco" which roughly means which that relates to supernatural [5]. This is the linguistic and meaning background of the word "Mascot" which leads to the perception that "Mascot" is the representative of mask and magic.

"Mascot" would be a real person or animal or a character (or a group) used continuously and in different poses and attitudes to represent, advertise or promote a product, a brand, a company or institution, an event or a specific cause [6]. Therefore, mascot could be made of and could be representation of things for gaining awareness and recall and also could make the product alive and cause positive attitude for consumers.

"Mascot is being heralded as the magic bullet in the contemporary marcomms armoury [7]. Visual representation of a brand by using brand mascots is a technique which has been applied by organizations for many years. As the visual 'ambassador' of a brand, the mascot's goal is to strengthen the brand identity. The visual properties which are translated by these mascots can cause the distinctiveness of a brand [8]. It could be concluded that mascot is one of the most powerful tools for communication of brand's identity as we could see many famous brands with their identities' representation by mascots for instance: Michelin Man by Michelin, Mr. Peanut by Planters, The Laughing Cow by The Laughing Cow, Mickey Mouse by The Walt Disney Company, Jolly Green Giant by B&G Foods, etc. Therefore, it has assured that mascot is appropriate with intimate image with consumers setting organizations [9]. Another reaffirmation is that Mascot usage in Thailand tends to be usage for public relations mainly in events [10]. It has been mentioned that 60% of tourists in Shiga Province, Japan has been increased by the Hikonyan mascot. All this information insist that mascot is a very important instrument for communication, especially for brand identity representation.

3. Analysis of the Symbol

"Meaning" is defined that while everything has its meaning, meanwhile, everything does not have a chance to form its meaning by itself. Human is the entire meanings former which is not under human nature in an individual context to create meaning for things. It takes at least two or more humans to perform to make it meaningful where perception of meaning of things is based on experiences from the five senses. Since the past, interpretation and meaning transmission have been hidden in many visual languages [11]. Symbol is one of the hidden visual languages which is significantly influence interpretation and perception of human for a long time.

In general, mascot is a mean to create brand identity, and has the objective to enable people to remember an organization through a unique character in addition to logo design because a mascot is created within the similar framework of art concepts, such as the use of colors, shapes and forms. A mascot is one of the channels representing semiotics which denote sign and symbol that has been produced and sent or transmitted to indicate how the meaning of representation was formed, including its understanding of a thing or a process of giving such meaning. Semiology was a tool that could make individuals understand the formation of meaning and its application [12]. Semiotic theory is an "Interpretive" theory that can be applied to most aspects of everyday life although most people would not realise it.

However, a character has the complicated story, as well as emotions, characteristics, identities, style and overall image through the characteristics of the character. Meanwhile, logo design has the limitation; but to create a mascot requires creator to consider the campaigns that the mascot is used for and how to make people remember the mascot. In addition, mascot power can build harmony and unity in an organization such as the use of a mascot of sport teams to promote the organization and tourism of the local area. For instance, the use of the raccoon mascot of American football team of Tennessee, USA. Raccoon is a local animal of Tennessee State; thus, it is used as the mascot to express

[2] CAMT or College of Arts, Media and Technology is a Digital School located at Chiang Mai University, Thailand.

cultural identity. Therefore, mascot creation is a mean to present organizational culture and images based on the concept of artistic marketing or 'Graphic Identity, which is a branch of marketing strategy. In the education and academic society, it is known that in Thailand, animals are used for representing, which is relating to the contemporary society, and has endless development in accordance with the ever-changing trend in the global society. As for Chiang Mai University, the representing symbol is an elephant. However, the University has no exact or clear mascot. However, based on the definition of the symbol, it can be said that the keyword that is significant for the creation of core identity of Chiang Mai University is a violet elephant, which has been used for a long period of time although there has never been any exact or official mascot. Therefore, it can be said that a symbol needs people to drive it and time for cultural absorption until it is memorized and defined eventually.

In this case study, College of Arts, Media and Technology (CAMT) is an organization of the university. Thus, the culture of the university dominates the identity of this organization. Consequently, this organization has 2 logos of 2 different types. The first logo is the core logo of the university that is jointly used by all faculties, with the name changed in accordance with each faculty, and the defined logo that is especially designed for the organization (as shown in Figure.1) in order to express the definition of the organization that is contemporary and specific through the use of art compositions and designing.



Figure 1: Logos of College of Arts, Media and Technology (CAMT): Core Logo and Defined Logo

The aforementioned logos reflect different definitions. The core logo of the university is composed of the symbolic animal and violet color as the focal points of the local, which create identities of the university. Meanwhile, the shapes and forms of components in the logo clearly reflect the Thai style. Meanwhile, the defined logo of CAMT is completely different from the former because it reflects the modernity of the organization. The definition of CAMT is the emphasis on inventions through digital contents. This definition concurs with the courses that this organization opens; hence, the logo is designed to be in a modern form that is minimized, without any implication or symbol of Thai culture used at all. However, the status as being Thai has been defined in the name that is another defined image. In other words, the status as being Thai or Lanna Thai has already been inserted in the contexts of the organization already. However, images or representing symbol that are shown have the aim to reflect the status as a contemporary organization. In addition, the use of color of the organization (faculty) reflects the status as a luxurious and international organization. This logo has been developed from the previous logo that was defined 10 years ago (as shown in Figure 2).



Figure 2: Previous logo of College of Arts, Media and Technology (CAMT)

It is apparent that the contexts of CAMT have changed and the logo has been minimized with the different aim and in a different manner. Originally, the concept of the logo was to make people memorize the alphabets, but the current logo emphasizes on making people memorize an image with contexts. The letter 'Can has the form like a beehive or geometric form, which concurs with the previous definition. Nevertheless, the contexts of CAMT are flowing and changing along with the changes in different eras and the expansion of scientific and art concepts that still create the complexity of the organization.

Since symbols are changed in accordance with societal ideas and trends that serve as driving forces, mascot designing has been changed accordingly. Originally, there had been 3 mascots to represent the Department of Animation and Games, the Department of Software Engineering and the Department of Modern Management (MMIT) because there were 3 departments back then, which were divided into 6 branches in the bachelor level, as mentioned before. Thus, an interesting question is whether the contemporary design of mascot to reinforce the images of an organization should be for the timeline of 10 years in advance, together with the analysis on the nature and ideas of university students nowadays and in the future (Generations Y-Z) in the scope of organizations of university students or not. However, the implication of CAMT from previous mascot still has core definitions that include freshness, curiosity, contemporaneity and ethicality of students still remain in the symbol. Therefore, it is necessary to use the previous mascot as the basis for the analysis and development for designing a new mascot.

3.1. Previous Mascot

Originally, the previous mascots of CAMT had been designed with the focus on the education and the teaching and learning that was clearly separated by departments. Thus, the designed characters were students in the uniform that precisely reflect the characteristics of the 3 departments, namely, Animation and Game, Modern Management and Software Engineering in the bachelor level (as shown in Figure 3).

Analyzed in the aspect of arts, the 3 previous characters were designed to have color theme which was like an old logo that reflects moods through facial expressions of a board character[3], the characteristics of which are reflected through facial expression and shape and form of the eyes. As for the proportion, each of the characters have the proportion of a kid, with a big head, big hands

and a small body. This style of design is called Super Deformation (SD)[4], which is the deformation of a human body into a downsized body like a deformed body of a small kid. The advantage of this style is that the designed characters are trendy in accordance with the drawing style of Japanese graphic novel that is widely popular. Consequently, the characters are eye-catching and easy to remember.

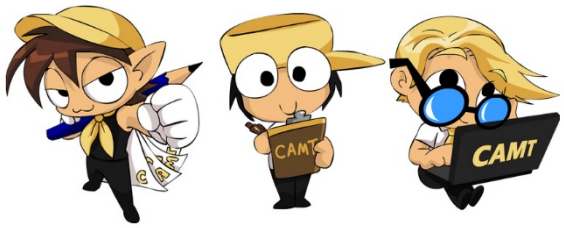


Figure 3: The previous Mascots of College of Arts, Media and Technology (CAMT)

Nevertheless, the mascots with the old definitions and policies have been used for more than 10 years since the establishment of the organization. With new policies and directions, there has been the need for extension and development of these characters so that they are contemporary while reflecting the identities of CAMT and CMU simultaneously.

3.2. Data Identification

College of Arts, Media and Technology has the educational approach that focuses on complexity which includes both sciences and arts, taught in a faculty that has learning method that is changed in accordance with eras, with the emphasis on innovation and digital technology as key principles. Therefore, the analysis and definition of the identities of the university are not fixed because the concepts and directions change annually in accordance with eras and trends. However, some keywords of the organization still remain. For instance, an organization that emphasizes on agriculture and technology may use lines and colors in organic art form in order to enable people to understand the definitions immediately. Alternatively, the Faculty of Political Science and Public Administration uses a scale and a lion as symbols that are related to Thai beliefs and symbolic definition which are combined together to form identities of the organization. As for CAMT, the symbol is designed to reflect the identities CAMT. However, such identities are under the domination of the identities of Chiang Mai University. Thus, there is the combination between identities of CAMT and identities of CMU, only for the identities from the social definition and graphic identities.

Chiang Mai University is an old complex university that provides courses of several subjects including medicine, sciences and arts in several branches, and extends and develops new curricula that fit modern era. From the view of the society, Chiang Mai University is a key University in the northern region of Thailand. At present, the university has the policy that focuses on digital technology and graduates of new generation with the competencies for diverse jobs, ethicality and promptness to the international level. However, such driving also leads to the change in the view of identities. In other words, this leads to the status as a contemporary university that accepts more foreign students, with the emphasis on linguistic aspect and tour system for foreigners in the university, or even the expansion of territory and branches in the northern region of Thailand, and the use of digital systems in the university, which turn the university Smart University, such as vehicle tracking system, game application for introducing places, class attendance (check-in) system with QR code and online method. Therefore, it is apparent that the current definition of the university is concurrent with the policy of CAMT because they are overlapping with each other and go in the same direction.

In addition, the definitions in the two aforementioned parts are also dominated by Lanna Thai status. It can be said that the Thai status and identities in terms of culture, thinking and traditions have control over the overall identity that cannot go out of this frame of social concept or definition. Consequently, the unique identity is created from the combination of several aspects such as media, people, policies, trends and politics, all of which are combined together. Therefore, the identity from the combination of the aforementioned aspects is a basis and component for designing a mascot with a clear scope that is different from a character that is a mascot of a private organization that has a broader scope. Nevertheless, this is no disadvantage. Rather, it enables the creation and designing of a character become clearer. However, there are several factors that are relevant, especially the concepts of the artist and the society which are key factors that lead to the building of the Character Style that are influenced by histories and the society as well.

3.3. Cultural Appropriation and Domination

Cultural Appropriation[5] is an approach concerning arts and society and is a key factor for the creation of art works. Therefore, it has both direct and indirect influences on the designing of mascots and characters based on experiences and societal perceptions of the artists or designers, as well as social trends, approaches referred to in a particular era and the implantation of ideas. For example, in this case, it is the creation of characters in the style of manga, which is a result from the great influences that Japanese manga works explored in Thailand to current. As for the designing, the thinking system of the artists focuses on applied art or extended arts, with the adherence to the origin. Therefore, the new identity reflects the combination among manga, original Thai arts and each artist's ideas. Not only does Thai status in arts and designing refer to lines and drawing but it also means identities, actions and stories that are added to a particular character. Thus, a character can be deemed to have identities that are based on Thai culture although it has culturally mutated or developed, which is called the 'mixed blood' or 'mongrel'. Nevertheless, based on the

[3] Board Character, according to the principle of character designing, is a character the characteristics and moods of which are mainly expressed with its look such as Scooby-Doo, 1969, Ruby J.

[4] Super Deformation Character (SD) is the style where by a character is deformed to look like a child, which originates from Japanese 'mangas'; and is sometimes called "Chibi" meaning a small thing, a petite person or a small kid.

approach of Culture Appropriation, the aforementioned phenomenon may be deemed a wrongdoing like plagiarism or copying of styles from any other culture in order to build and claim as our culture. Thus, the definitions of culture in various approaches are unclear. On the contrary, identity that is especially from the development of the aforementioned basic approach is accepted by people in the society and eventually becomes a symbol. Thus, the borrowing of culture that is caused from the domination over thinking occurs all the time, with the society acting like the judge and artists acting as presenters [13]. However, goals, objectives and approaches that newly occur have more significant implication. Hence, as for mascot designing at present, artists design their works to reflect the influences from mangas because of the aforementioned factors. It is not surprising that a newly designed character will look like a manga character. In addition, the key overlapping definitions, combined with styles, can create specific identity that serves as representation[6] of that society or organization as well. Therefore, cultural burrowing can occur all the time as a normal event in the society and in the field of designing.

3.4. Character Population and Demand

First of all, the target audiences for the mascot designing in this project are students in the age range of 15 – 25 years (Generation x) at present. From the survey, it is discovered that the characters that are popular are in the style of Manga[7] and have features of a kid (Manga Cute Characters). The advantage of this approach is the popularity of the style among the target audiences and audiences in other age ranges because cute characters do not mean manga style only; rather, it also means western characters such as Bug Bunny (W. Brothers, 1940). This character is used in western cartoon. It should be noted that this character is not originally created to serve as a symbol or logo but for producing animations for children audience. However, since the popularity of this character rises by dint of the expansion of media, Bug Bunny has become a symbol of western cartoon. Furthermore, Mickey Mouse (W. Disney, 1928) and Elza (C. Buck, 2013) have made Disney's cartoon characters memorable and great. In addition, may other characters from movies have also been popular and a societal factor. Examples of such characters are super hero group from Marvel Comics (1939), which have created trends that affect behaviors and thinking. Apart from memorable identities and looks, other attributes of these characters such as characteristics and thinking of the characters can generate new trends and thinking of people in the society. They can also trigger the building of social groups in accordance with personal preferences, which is concurrent with the concept of freedom of thinking nowadays. Looking back to Asian society that has manga characters as the foundations for thinking, one can say that manga characters were originally created to disseminate Japanese culture and thinking. However, with the popularity, manga characters have expanded to Thailand and become imprinted into many Thai artists during their childhood. Consequently, such artists design their works in the directions that they have been imprinted and in accordance with the popularity that evolves in each era.

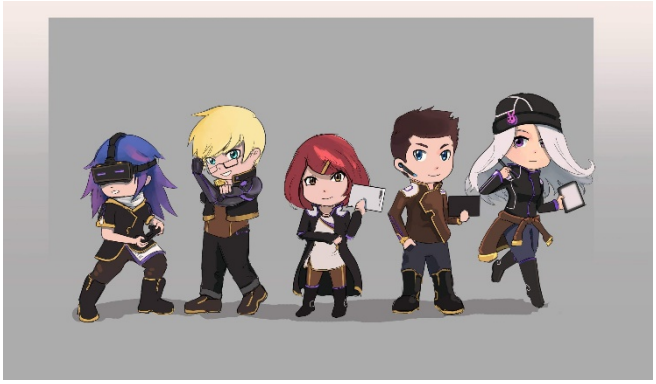


Figure 4: The Experimental Designs of Mascots:
DG – SE – MMIT - Dii – ANI draft 1

The mascots designed in this project belong to an educational organization. Thus, the outfits should be a uniform that reflects Thai culture in the improvised style and politeness in the society, mixed with the identities that the university wants to present, namely, digital technology, modernity and future world, with the basics of the identities from the identities of the university. In this case, violet, which is the color of the university, is used as a component of the design. Thus, in this designing experiment, culture domination and character population play prominent roles. As a result, there is a clear direction for the possible design.

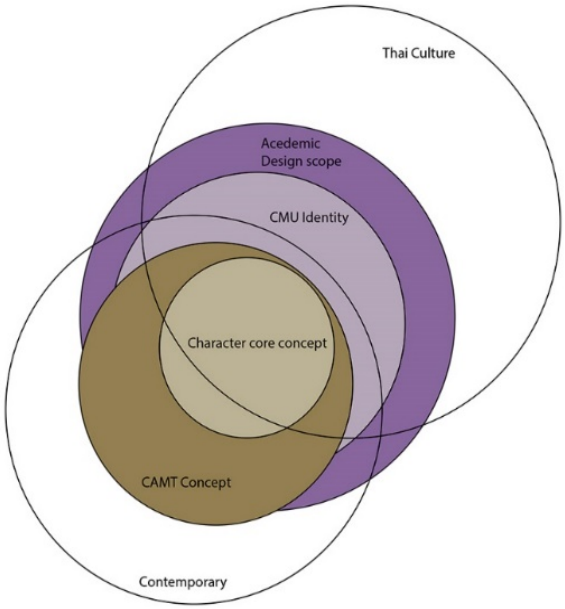


Figure 5: CAMT Mascot Design Scope, the overlapping area of contents analysis to finding of core concept design.

[5] Cultural Appropriation is the influenced of the artistìs background. In this case, it refers to Manga arts which applied to character design harmonious with Thai culture, CMU and CAMT.

[6] Representation in this case it intended merging of semiology and concept idea to presented something in specific purposed which approached to the meaning of Art work as a core concept.

[7] Manga are the Japanese comic or graphic novels style and become popular cultural in worldwide level since late 19th century.

Concerning the concept for designing the characters to represent the educational institute in this case, or CMU (as shown in Figure5), it is apparent that the scopes of the concepts are overlapping. Thai culture has partly taken part in the determination of the concepts of the characters because CAMT also has concepts in the international level. It cannot adhere to only Thai culture. Rather, foreign cultures have also partly played roles. At the same time, certain concepts that are the scopes of an educational organization, such as uniform, knowledge pursuit (curiosity), intelligence and friendliness. Such concepts belong to CMU and CAMT itself. The scopes of CMU have covered all the characters. At the same time, concepts of CAMT are slightly different from CMU because its identities are different from those of other faculties and are harmonious. In addition, such identities include contemporaneity that regards the directions of creation with all overlapping conceptual frameworks and scopes that cover the concepts of characters. Thus, it is apparent that the contemporary status and the Thai status are mixed together in the even ratio, with the concepts of characters as the cores for the designing and creation. In addition, feedback from audiences and users of the mascots is also taken into account. The most important aspect is the outcomes from the designing, which are regarded as the priority purpose.

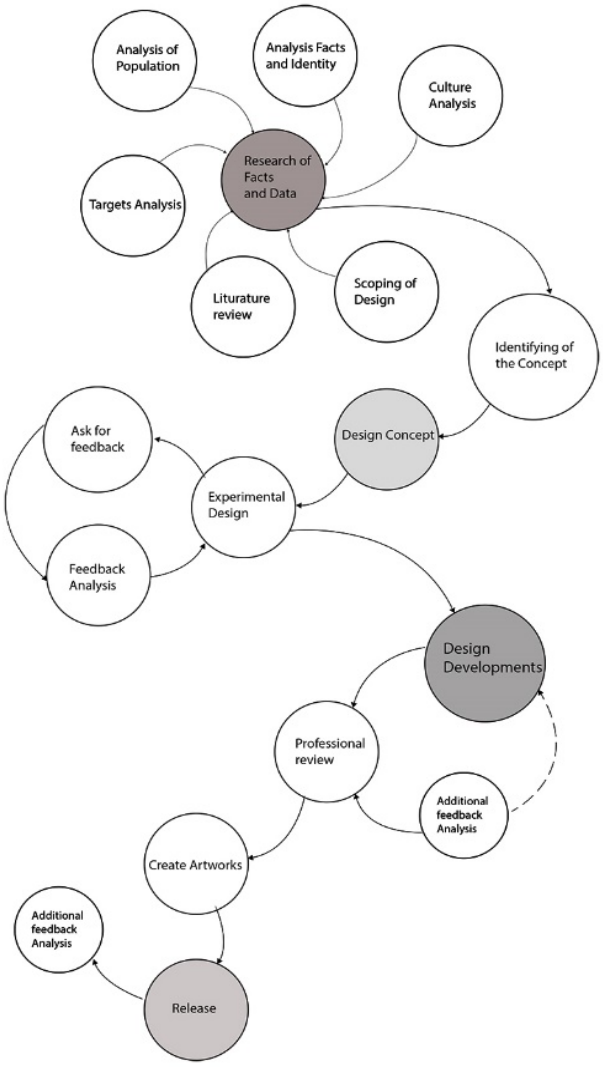


Figure 6: Design frameworks that specific for this project that require the audiences create an idea together with the artists.

4. Methodology

As for this study, the framework of the designing process consists of 4 phases that are 4 key steps, as follows. 1) Pre-production is the step that includes data research and analysis and determination for definitions and scope of the outputs and directions for designing, and study on the experiments by other artists. 2) Production is the step of the development of the designs that build images and patterns of the characters in ways that concur with principles of arts and designing and the concepts of contemporary design. Thus, this step incorporates experiments and creation of images in different dimensions. 3) Post-production is the step that includes the creation of media to present the outcomes from mascot designing, printed media to display in the building of CAMT, and promotional media including Line sticker and action figures. 4) Development and Conclusion is the step in which data are stored and products that have been released are refined in special occasions in order to add values to the brand in the consistent manner. In addition, new products are launched and new experiments in more diverse styles are conducted in accordance with social behaviors in each particular era.

In addition to product development, each of the aforementioned steps also includes the experiments that are based on the feedback from experts and some audience. The target audiences are divided into different age ranges, from people of generation Y or high school and university students to working adults. The designs of the mascots in this project must be neutral as possible whilst satisfaction and recognition should be supported by and concurrent with opinions and critics in social media as well.

5. Character Concept Identifying

The design definitions of each characters, Animation and Visual effects (ANI), Software Engineering (SE), Digital Film (DF). Modern Management and Information Technology (MMIT), Digital Game (DG), Industry Integration (DII) and Knowledge Innovation Management (KIM). The main concepts are contemporary and trendy including clothes and attitudes. The characters have their own distinctive characteristics, but work as a team. The same as CAMT concept that each major are related each other. The visions are combining which the words Honest, Discipline and Sharing (HOD)[8] which comes from the core values of CAMT (CMU CEO)[9] defining the current identity of CAMT. These definition and keywords are applied into Mascot mind and characteristic as a concept identity.

The definition of the Mascot of Animation and Visual Effect (ANI) is that this character has a high level of artistic features with individual thinking in digital technology-related arts. Therefore, the key instruments for creating the works in this project is the tablet pen painting plate, with the emphasis on creativity that allows the character to have high flexibility in terms of its appearances, which mean the character will not strictly adhere to any principle or rule.

As for the Mascot of Software Engineering (SE), this character is defined to be a scientific person who emphasizes on complex and systematic thinking with ability to create new software and

has literacy in special languages or digital ones (programming). Hence, this character relies mainly on left side brain (IQ) and truth.

As for the Mascot of Modern Management and Information Technology (MMIT), this character is defined as a person with leadership, innovativeness, capability of managing team members and planning, and ability to think and analyze, and intuition for forecasting things in the society through digital tools. This character is thus classified to have administrative skills and instant confidence for sound judgment.

In addition, the Mascot of Digital Game (DG) is defined separately from Software Engineering, but the emphasis is put on the creation of digital games. Thus, this character is both artistic and scientific and have more commitment than others. At present, there is a competition called E-sport. Hence, this character incorporates competitiveness, innovativeness and creativity.

Meanwhile, the Mascot of Digital Industry Integration (DII) is defined as a character who is based on facts and truth in the society, with the characteristics that are changing in accordance with the social trend. This character can learn from the issues in the digital industry; and thus, has diverse characteristic and ideas that are consistently changed. This character has diverse knowledge and ability to gather all knowledge in order to create new knowledge that is concise and easy to use.

The Mascot of Digital Film (DF) is defined as a character with creativity like the ANI Mascot, but this character focuses on the creation of cinematographic media, with combination between reality and imagination, and confidence. In addition, this character has capabilities and uniqueness in the field of entertainment arts because it has most prominent look of all members in the team. This character can be a director and an actor; thus, this character has as high confidence as MMIT Mascot, but in terms of entertainment only.

The definition of the Mascot Knowledge Innovation Management (KIM) is that it is the character that is regarded as the source of knowledge for the team because this subject is in Master Degree and Doctoral Degree levels only. The character is the eldest and most experienced of the team, with the characteristics to provide suggestions and advice to others in order to create new things from available information. Therefore, this character is prudent and does not show the emotion clearly. The special capability of this character is to analyze and file a great amount of information.

From the definitions of the characters, each character can be clearly and individually identified with values and policies of CAMT. Consequently, each character has specific definitions that are concurrent with the image and look of the character and different from another character depending on the subject. This is the concept for the designs of all the characters that has been combined with contemporaneity.

5.1. First Look of Design

After the data management is complete, in the Production Step, there will be the experiments of the art works by creating the mascots starting from forms, along with color designing, all of which are based on the definitions of the characters. In the image, it is apparent that gadgets that are related to each subject are added. For example, the ANI will hold a tablet and is floating all the time in order to show the implication of imagination and freedom in terms of arts. The experiments are still based in the feedback from the experts, which will be followed by the refinement of the designs because the designed characters are still not artfully perfect and are too complicated for the memorability.



Figure 7: CAMT early mascot part 2 design from artist A with digital gadget design e.g. tablet, gaming keyboard or smart phone.

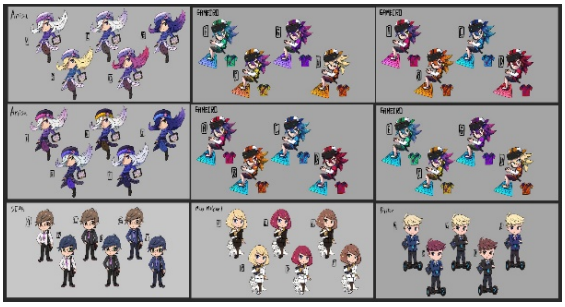


Figure 8: CAMT early Mascot design color alternative experiment

The using of color and shape, the trendy costume is also the key component by using the concept of Science fiction (Sci-Fi). The concept is that the costume design should indicate the futurism and the digital technology infused with CMU attributes and Lanna culture for representing the overall CAMT identity. In this process, there are experiments of color composition in various patterns (as shown in Figure7) as the attempts to clarify the identities of the mascot e.g. using red to portray danger and using blue to portray serenity.

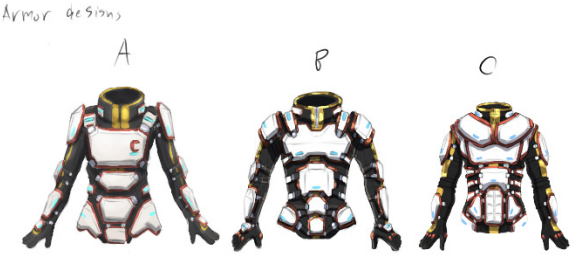


Figure 9: Early Digital CAMT Costume Design, Armor design

Concerning the designing of costumes of characters (as shown in Figure 9), initially, it is apparent that the look of the costume is too hard, which reduces the gentleness of the characters. In addition, the stories of characters are not agreeable. Thus, it is necessary that fabric materials are used as much as possible in order to create the costume that can be used in real life.

[8] HOD the definition of person who works at CAMT and part of CMU CEO representative words with honest, Discipline and Sharing used for personal mind at CAMT should have.

[9] CMU CEO the CAMT core values comprised with Community, Morality, Unity, Creativity, Excellence and Open for Challenge.

Furthermore, the proportion of the characters is changed from SD (Superhero Scale) in order to present charisma, confidence and implication of leadership, as well as the results from the analysis of the popular characters. As for this case, the characters from animation and game media are used. It is also apparent that characters from Ragnarok Online [10] Game have been highly popular for the past 20 years.



Figure 10: The example of ANI character transformation, *SD* re-scale design process.

Apart from rescaling, it is also vital to focus on the gestures of characters so that the identities and habits of each character will be presented more clearly. To do this, it is necessary to design the sketches in order to promote the mascots simultaneously by analyzing the characteristics of each character. For example, the Mascot of Digital Game (DG) (as shown in Figure 11) is reserved, loves playing games, and is not talkative. However, this character is keen on computer use and programming. This character is obsessed with some specific matters only. Such characteristics and habits are from the analysis of people of generation X who have complexity in thinking, individualism and introversion in high levels. However, with the light of the identities of CAMT people, this character also has optimism, which adds cheerfulness to this character until new identities are built. As a result, this character has gestures that show consistent movement and high responsiveness because most gamers have responsiveness rate that is faster than that of non-gamers. This is because of the main activities of the character.



Figure 11: The alternative of Character posing Design of CAMT DG Mascot

Unlike DG Mascot, Digital Industry Integration (DII) Mascot (as shown in Figure 12) is confident and capable of getting along with others easily. As a new character because the branch has been established this year, the character is designed to be still and prudent, with systematic thinking for management and confidence, but not in a great level as the confidence of MMIT and DF characters. In other words, this character can talk to, meet with and learn from other people all the time as reflected through the habit of the character whereby the character wears Bluetooth all the time with the promptness to discuss matters with people. At the same time, the character has still and stable gestures, with less movements than that of DG character.

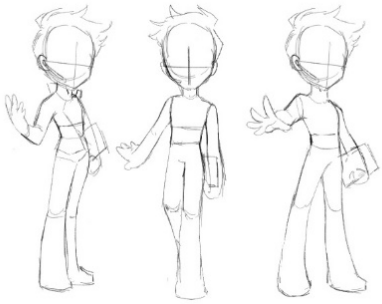


Figure 12: The alternative of Character posing design of CAMT DII Mascot: Confident Gestures.

Therefore, the creation of new postures that is carried out simultaneously with the rescaling is a significant step for adding personalities to each character. For instance, SE (Software Engineering) character (as shown in Figure 13) the uniqueness of which is the Smart Digital Arm. This character is the smartest one of the team, with expertise in language literacy and knowledge of programming. Hence, the gestures of this character represent confidence and uniqueness of the character. Initially, the postures emphasize on the Smart Digital Arm that is the unique feature of the character. Furthermore, the postures make the character look like a superhero but the child-like SD proportion make the character not much different from others in the team.



Figure 13: The alternative of Character posing design of CAMT SE Mascot

5.2. Detailing and Identifying

After designing the posing of characters, the next step is detailing and finishing the costumes by adding lines and details from the postures of the mascots. In this step, it is possible to create the graphics that can be actually used. In other words, there is the production of concept art which will be promoted in the step of publicity, which will shorten the period of time for media production. Nevertheless, there is a limitation that comes from a combination of organization synthesis. Besides, another limitation is that the costumes including hairs and faces have to be different and be infused with the gimmick[11] and the symbol. These are used to enhance the understanding of the receivers, e.g. Using of Digital suit look mixed Traditional Thai culture on fabric

[10] Ragnarok online is MMORPG game developed by Gravity interactive Company in Korea. The contents are based on manhwa Ragnarok by Lee Myung-jin. The game is the one of the most popular in the world and character design population.

form and texture with which can be seen in the details of the character, MMIT (as shown in Figure 14).



Figure 14: Detailing and Re-posing of design art concept, MMIT and ANI Characters

After the drawing process, more details are added to the characters. Each character is presented on different sides (in Character Sheet), together with expressions (in Expression Sheet) with details of gadget of each character in order to extend the story and details of the character in the future. In this step, the key action is to determine colors used for creating graphic identity without any deviation. The color system used is RGB that is concurrent with the color of the organization.

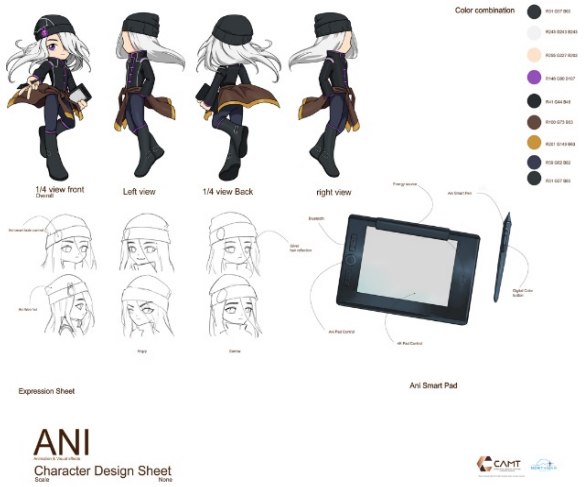


Figure 15: Character design Sheet of CAMT ANI (Animation and Visual Effects) Mascot

15, the ANI (Animation and visual effects) character is based on the personalities of the students of this branch and the body of knowledge from the analysis, the results from which reveal that the character is highly individual, especially in terms of art, or computer art, reserved, and outgoing, not for speaking but for presentation of arts and opinions. Therefore, the character is designed to have long silver hair, which reflects the obsession of independence. Thus, the character can fly, which is like the characteristics that is flexible, highly independent and sometimes introversive. The character carries a digital tablet as the important gadget that is used for drawing all the time. The eyes of the character are in violet color, which is a gimmick because violet is the color of CMU, which is inserted in the concept of the character. Meanwhile, the costume of the character is in the same direction as CAMT suit, with color tone limited to brown color and the character seems to be floating all the time.

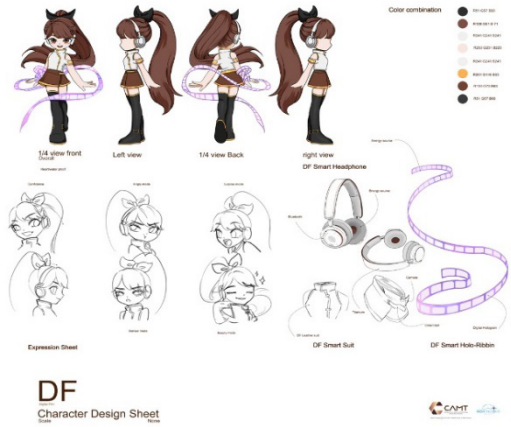


Figure 16: Character Sheet of CAMT DF (Digital Film) Mascot

Digital film) character is the combination of digital art and communication arts relating to entertainment art, especially cinematography. Therefore, the character has individualism as the ANI character and confidence to present self-identities because of capabilities and knowledge of acting and producing for films. Hence, this character is girly but confident, good looking and taking care of herself all the time, in terms of personalities and beauty. This is the only character of the team that is related to front stage. The accessory that represents this character is the Hologram Film Ribbon that is a traditional film roll. It symbolizes that although the character is in digital era, she still has knowledge of cinematography history. In addition, this symbol is easy to understand for propel in general, based on basic perception. In other words, when people see the symbol, they can immediately define the character. This character wears a miniskirt, but the miniskirt is also in the form of the contemporary uniform of the university, with the senses of fashion and sex appeal added. At the same time, the colors used are still related to the color of the organization, or brown in different shades.



Figure 17: Character Sheet of CAMT DG (Digital GAME) Mascot

[11] Gimmick, in this case is the way to input symbol identity into the design, including color, facial or accessories

DG (Digital Game) character is not much different from ANI one in terms of art statement with high levels of confidence and introversion. The differences are the love to competition and the capability of programming, which make this character inferior to ANI one in terms of arts. However, this character has complex thinking relating knowledge of digital interactive. This character is also inspired by e-sports that turn some game players to game producers. Hence, the character is highly obsessed and self-oriented. The key tool of this character is a traditional joystick that is linked to Oculus [2]. This tool is the key gimmick that presents a symbol that can hasten people's understanding on the concepts of the character. Meanwhile, the costume of the character has colors that are concurrent with the color of the organization. The costume is also designed to be agreeable to e-sport game team suit.



Figure 18: Character Sheet of CAMT DII (Digital Industry Integration) Mascot



Figure 19: Character Sheet of CAMT KIM (Knowledge Innovation Management) Mascot

DII (Digital Industry Integration) character is a male character and the only character of the team is has high confidence and skills of speaking and presentation, as well as concern with personality, which is reflected through the face and hairstyle that look polite and slick. The body of knowledge in this branch is related to the building of bodies of knowledge that can be implemented to contemporary industries in order to always catch up with current trends. Therefore, the character is confident and highly capable of communicating with others. The personal gadgets of this character are a smart Bluetooth and a notepad that the character carries all the time. This can manifest the implication of speaking and analytical skills, high are the key skills in this character. The costume of the character is adapted from suit jacket that makes the character looks more agile and active than other characters in the team.

Knowledge Innovation Management) character is the eldest character of the team. This character works at the back of the stager as the consultant of the team. This character collect and file all bodies of knowledge in every subject in the world. Thus, he is designed to have some gray hair as highlight, which is a tangible symbol to represent an aged person. Another unique point of this character is that he sits all the time, with the characteristics of a leadership who does not disclose himself or is the person secretly working at the back to ensure the success of the team. The significant skills of this character include the dissemination of knowledge, and the provision of advice in the fields of academy. The key gadget is the smart throne that is a tool to show and to store data like a hard disk. In addition, this gadget allows the character to communicate with others all the time. It is deemed as one of the most versatile digital gadgets because it can be used for creating, storing, and communicating. This throne is designed to have unique symbolic uniqueness of this character because it has violet color that is the thematic color of Chiang Mai University.

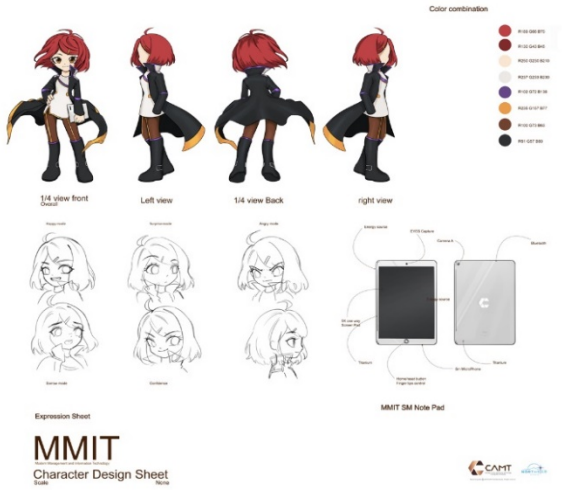


Figure 20: Character Sheet of CAMT MMIT (Modern Management and Information Technology) Mascot

As for MMIT (Modern Management and Information Technology) Mascot, this character is the leader of the team. She is capable of managing things quickly and efficiently. This character is also the key planner of the group and is also as capable of communication as DII character but has a greater level of leadership. Therefore, the character is designed to have the facial expression that reflects her confidence, with gestures that show high agility and delicacy, which are presented through the costume that consists of the cloak that waving all the time. The character has red hair which reflects boldness and confidence at the same

[12] Oculus is a VR game gadget from American technology company that used for visual reality in game specific support hardware and software in digital game.

time. Also, MMIT character has the key gadget of this character is the Smart Light Notebook that can store a lot of data but less than that of KIM character. However, the gadget of this character allows the character to have agility like a note pad does. Nevertheless, the concept of this character is inspired by the overall personality of the graduates of this branch, and the body of knowledge that has recently been developed.

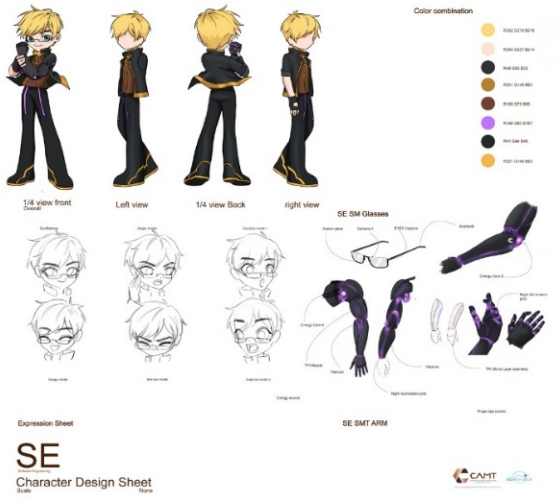


Figure 21: Character sheet of CAMT SE (Software Engineering) Mascot

Software Engineering) Mascot is deemed as the smartest character of the team, and has the knowledge of programming, production and language. This is the only character of the team that has mixed blood (a Thai – American person). The weakness of this character is the skills of arts. Even though being the smartest one, this character has to rely on ideas of ANI character as well. Therefore, SE character and ANI character are the closest friends of each other. The key tool of this is the Smart Digital Arms with AI at the finger tips, which allows the quick creation of programming works. This Smart Digital Arm was built by SE character himself, with color of CMU reflected through the violet light emitted from the tool. This gimmick has implication that can be widely understood with general perception in vision of color and arts.



Figure 22: Character Size Comparison when the character assemble as a team group.

6. Graphic for Advertising Product

The use of mascots for promoting an educational organization has more limitations than the use for a private organization in several aspects, including ethical images, definitions of education and organizational culture, all of which have to be taken into accounts and lead to the consideration of the types of media used. All media types, both physical and digital ones, are related to target groups and images of the organization. For instance, it may be deemed that the medium that is the symbol for memorization is a keychain souvenir because a general function of a keychain is to be a memorable symbol of an event or a place. However, at present, some functions may be added, such as to add a USB port flash drive to the keychain for multifunctional purpose. Likewise, designing the graphics that have mascots as the core element, it is necessary to present the images and tones that make the organization look like a lively and attractive place for every student like the students' second home that has modernity, and contemporaneity. In addition, the definitions of the term 'welcome' are reflected through the facial expressions of the characters. All the aforementioned concepts generate keywords that will be applied to all products under the principle of graphic identity in combination with the principle of graphic design in order to blend perception into arts perfectly without the need to sacrifice the identities of the organization. The building of such images needs the release of images in phases with the goals to allow audiences to send feedback and to enable all audience to perceive a little bit of new images of the organization at a time, starting from the display of the characters in the forms of posters in social media in special occasions such as New Year's Day, or in an event that has become a hot topic in the current society, such as the COVID-19 circumstance in 2020.

6.1. Poster Design

. rather, it includes the change of the head of the Facebook page of the organization, public announcements and question what the designed characters are representing. The chosen release day is the day for the celebration of 11th Anniversary of the University. This may have impacts on people in and outside of the University. In addition, this scheme al so makes students feel attached to the subjects they are majoring.



Figure 23: The Example of Character Poster Design

In 2020, there had been the outbreak of COVID-19 disease that caused Chiang Mai to be locked down for a certain period of time. At that moment, the professions that were worrisome were doctors and nurses who played key roles in taking care of patients, with the close encounter to the incurable disease. Therefore, the team designed a poster which was used in an online campaign to show the characters in the outfits of doctors and nurses in order to show the feelings of care and appreciation to doctors and nurses who worked hard. This design expressed the message that enabled people seeing the poster aware of the concern by the University with the situation, through the characters in the well-blended manner whilst expressing the perspectives that the organization had towards the country.



Figure 24: A Digital Poster using of Mascot on COVID-19 period in Thailand that also promoting a brand identity and policy at the same time.

Concerning the use of the poster design to deliver a message, at present the technique that has been found most effective currently is the creation of graphics that can be shown online. However, it is merely a one-way communication technique that can create awareness. Meanwhile, the use of mascots of the University to deliver messages to audience is a good way that reduces tension in certain situations along with presenting the images of the organization in new dimensions. Especially, in case of an educational organization, the delivered messages will be more reliable and accepted. All the aforementioned principles are related to social structure.

6.2. Line Sticker Design

To design characters to be used in social media communication is included in the marketing plan for building a brand in the modern era. This technique is to use characters that express emotions in the conversation to promote the key issue. In this case, the chosen media is Line[13] chat with is a popular mean of communication of people nowadays and is widely used in Thailand. Therefore, the team has created the mascots of CAMT to show emotions, appearances and symbols of human beings. This enables the users to choose the suitable characters that they like to complement their communications. Thus, designs of the gestures of the characters are related to currently popular messages such as 'congratulations', 'work hard' or 'thank you', all of which are still within the frameworks for designing the mascot of the organization without communicating messages that are too in-depth, allowing people of various age ranges, from high school students to middle-aged people, to use the designed characters and signs. In addition, the designed appearances and the used colors make users feel positive because the characters do not look too serious, which makes them fit the current trend.

The use of these character in digital communication (Line application in this case) has boosted the awareness of identity in deep and cognition because of each character has a differences personality for example, when they use MMIT emotion it can feel the leadership personality somehow from the face and body language, and that can show the identity of organization while using these symbols. The gender of the characters is also affected the user via using image or sticker chat, it makes the user more attractive on chat, same as using new patterns of language in

. Each version is designed to have a concept that is different from each other. The first version is for expressing common emotions such as 'thank you' and 'good' in 'good day'. At the same time, stickers in the other version are designed to be more specific with the works of the organization or professions, with new concepts and modern words, including slangs added, in order to make the designed characters become contemporary in the aspects of appearances, images and notions. In addition, facial expressions of the designed characters can look hilarious as well.



Figure 25: The example of Line Sticker CAMT mascot version 1-2 that already in the market.

-19 disease; hence, the research team has designed Line stickers in Thai language in order to use the designed mascots to promote and reinforce certain messages to the society. The mascots in this special version wear the outfits of doctors and nurses to show care and concerns with people in the society, which will spread positive vibe. Therefore, it is apparent that that the designed characters can be used for conveying diverse meanings and can be transformed into different styles whilst carrying the identities of the organization and their own identities at the same time, e.g., the KIM character, who is the source of knowledge or the team leader, wears a doctor uniform and has gray hair whilst other characters wears uniforms of PAs and nurses. At the same time, each character still shows his/her personal characteristics. For instance, the DF character, even though wearing a nurse uniform, shows cute and lively personalities as well. And that it makes people smile even in bad situation, this is a power of Mascot.



Figure 26: The examples of Line Sticker CAMT mascot COVID-19 version

[13] Line is a new communication application managed by a Japan company called Line Corporation. This application allows people to make free calls and free messaging. Invented by Naver Corporation Korea, it is the most popular

6.3. Architectural Decoration

The plan for the presentation of identities of the organisation is to increase the number of the woks consistently. One of the significant symbols is the L-shape building with glass wall at the main entrance. One thing that is indispensable is the decoration of the building, which will create the atmosphere and emphasize on the identities of the organization. The team designs big posters showing the designed characters to be displayed on the glass wall of the entrance in order to present new images of the organization and to renovate the building because the building is heated by the sunlight from the northern side. Pasted on the glass wall, the stickers will serve as blinds instead of curtains. The effects that happen are that the appearances of the building are completely changed because the colors and looks of the designed characters make the building of the organization look more exciting, active and delightful. The outcomes from this work lead to the discovery that to use bright and lively looks of the characters makes people visiting the building feel lively. These effects are like the ambience of Akihabara area in Japan, where some buildings are decorated with manga characters. Even though such manga characters are for advertising purposes, they make people in the space feel lively. Therefore, the decoration of the building, together with the design of landscape to have the building encompassed with a green garden, will also make students feel fresh and lively. At the same time, the decoration of the building builds unity in the organization. Also, the building of CAMT is the only building in Chiang Mai University, or even in Chiang Mai Province, that is decorated with images of contemporary characters in order to create the positive perception perfectly.



Figure 27 The use of Mascot for decoration, College of Arts Media and Technology, CMU.

7. Toy and Souvenir Product

As for the designing of characters, one of the indispensable elements is the creation of 3D models in order that the images and appearances of the designed characters will be more concrete. This concept of the creation of 3D models has led to the idea to create souvenirs to reinforce the identities of the organization through digital symbols and digital way of life. At present, many people, with artists and designers included, prefer to collect toys in the types of action figure, statue, Figma[4]. Therefore, the team has an idea to create souvenirs that are 3D models of the designed mascot characters with the framework to present the images of the organization along with the creation of aestheticism from small toys. The creation process starts with the creation of the characters in a computer before using a resin 3D printer which has the greater resolution than aplastic (PLA) 3D printer. The concept of this work is also based on some marketing principles concerning the presentation of images and symbols of the organization simultaneously.

7.1. Concept Design

This project is the creation of the set statue contained with 7 characters of CAMT in the sizes and scale were in accordance with the standard of Nendoroid[5], with the heights ranking between 5-15 cm, in the forms of SD characters [15]. Nendoroid is defined as a popular brand of small plastic figure replicas of manga and anime characters [16]. This is not too big statue but not too small based on population toys in the market. Mostly, for toys collector in worldwide so it become a standard size, scale and proportions. As a comparison of Nendoroid in the market place, there are the same scale and can replace the face and emotion but it lacks of group assembler product and light up system. This is an idea to create a difference type of Nendoroid figure. In this case, the works are created as statue figures, in order to set the gestures and looks to be presented that go in the same direction. The scopes of the design are as follows.

. Each collector must collect all the 7 models in order to get a complete logo.

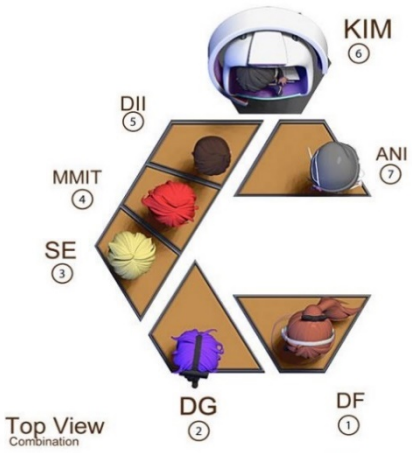


Figure 27: A Top view of 7 pieces of mascot combined into a CAMT logo

- The base is supported by a plate which is included in the model of KIM character, which is the team leader.
- The base of each model consists of the lighting system that related to the main plate, called LED light up base, which makes the display look more dimensional.
- The models are displayed in the stepped manner, with DF and DG characters sitting on the lowest steps whilst other characters are standing on the upper more steps. Thus, all characters will be seen from all angles without blocking one another.

Some characters, such as ANI, are floating all the time. Thus, the model of such a character consists of a metal pole that makes

[14] Figma Toy is a series of Japan toys invented by Max factory and distributed by Good Smile Company. The product focuses on highly articulated hand-sized figure of popular anime figure.

[15] Nendoroid is a series of plastic figures created by Japanese Good smile company in 2006. It becomes popular due to cute appearance. The figure has a function of articulation in appropriate scale for collectors.

the model look floating. In the future, the pole many be created with transparent material such as clear acrylic so that it will look more realistic. All the aforementioned ideas are complemented with the marketing design in order to make collectors desire to collect all the models in the set-in order to have a complete logo of the organization. In addition, the set of the models can serve as a lamp. As for toy designing, the most important elements are sizes and perspectives of the work in 3D view. Therefore, sizes are first set in the program of 3D image creation. The framework that is especially dedicated to this work is designed.



28: A Side view show the levels of the model steps

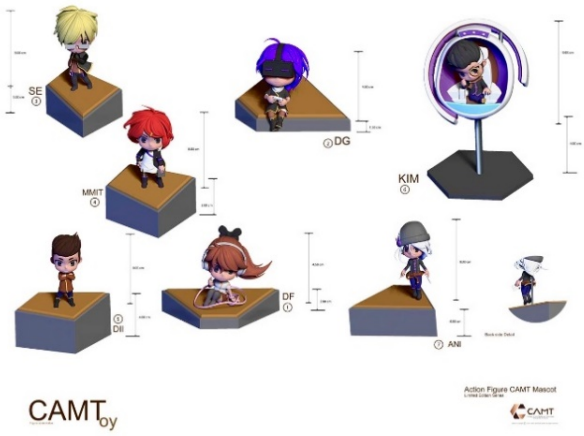


Figure 29: The separated details of each character

7.2. Methodology

The production of the toys can be divided into 3 main steps, namely, Pre-production, Production and Post-production, with the details of activities as follows.

- Pre-production is the step in which the initial design of each character and the concept for the assembly of the toys (Toy Mechanic Combination) are drafted. Also, details concerning used materials, postures, communication of the meanings and perspective are tested in the 3D space. In addition, the proportion and function for lighting systems can be adjusted in the step of the sketching in the 3D space. Furthermore, there is also the research for information of toys of the types of Statue and Nendoroid. If the size is bigger that 15 cm, the model set will need a great space, but a purpose of this model set is to be placed in any place such as at the space over the console of a car or a computer desk, which has no great space. Therefore, each model is designed to be 5-10 cm in accordance with the scale of each character.
- Production is the step that directly involves with the creation of the model set. This step starts with the creation of 3D model prototypes and the number of polygon is not taken into account. Thus, the design can have high quality subdivision, which will not affect the software for the printing. In this step, the focus is on the parts with joints that will be seriously taken into account because of the significant effects on the printing out. The consideration is also put on the designing of the colors for each character. Finally, the lighting system is added to the printed out model.
- Post-production is the step that still incorporates adjustment but the adjustment is more like the finishing of the details of each work in order to create the mold that will be used for mass production. This step also includes the taking of photographs to present the works and to get feedback which is used as statistical data of marketing demand. This step also includes activities in the production step if there is any suggestion from audience or expert. Most of the attained suggestions concern the use of mixed materials and the colors of the product.

From all the 3 steps as mentioned before, it is apparent that there are 2 keys activities in the production, which are the creation of products in 3D space and the printing out which require many experiments and adjustments.

This activity of the creation of the model is the experiment to determine the perspective and composition whereby all the models can be seen. From Figure 30, it is apparent that the SE character model is swapped with the DG character model in the final product. This is because SE model is in the standing posture whilst DG model is in the sitting posture. Hence, DG model is placed in the central area instead of SE model.

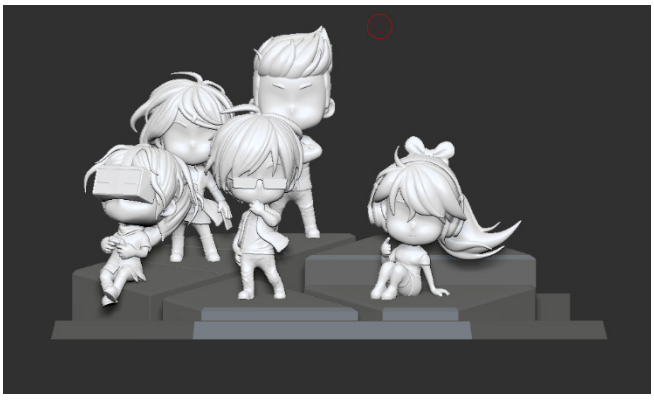


Figure 30: The First Draft 3D model in 3D Software.

This step also includes an important activity, which is the designing of separable parts such as faces, hair styles and bodies in order to let each model look seamless as possible. For instance, an SE model consists of 3 separable parts, namely, hair, head and body, which are easy for the assembly and the addition of details.



Figure 31: The test 3D model printing using resin separated 3 pieces.

. From the experiment, it is discovered that resin is very fragile especially in the parts of hair and small pieces in the model. Therefore, in this case, glasses are made of transparent plastic instead of resin.

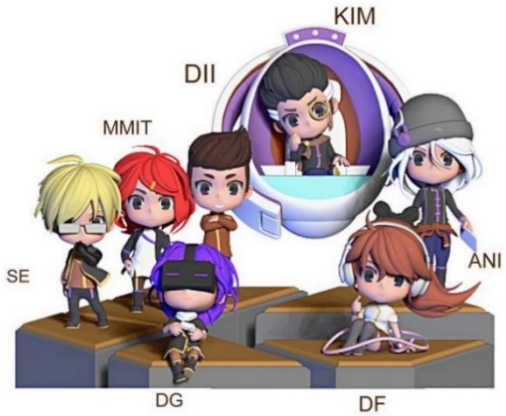


Figure 32: The Combination of the Mascot set in 3D vision.

After the test 3D printing of SE models, it has to make sure that when the character combine together, it in the right position and proportion. We use symmetrical in a feeling of mass and space, as you can see that KIM is blended to the right but on the left, there are 3 character standing on the step to make the composition look actually balance. The based system also designed in this step, in order to move on to printing a prototype step.

7.4. Printing a Product

This step involves the production of the models that have passed the test printing in 3D space already. All the pieces are put together, and the bases to support the works are also designed. The designing of the base to support each model is an important action in peeling the mold off the work. After the mold is peeled off, the work has to be scrubbed in order that the surface is smooth. In addition, a UV LED lamp is used for making the resin material hard and have smooth surface as expected. Instead of the lamp, sunlight can be used as well. Resin printing is different from PLA because it can show more details. In this case, a UV LED lamp is used for creating the prototypes, by using UV light to cause curing reaction on resin which makes resin hard.

. As shown in Figure 33, put together on the plate, the bases of the models do not fit well with one another. Therefore, the pieces of works have to be scrubbed with sand paper before the coloring process. Another problem is that the metal pole for holding the model is too high, disabling the model to stand by itself. In this step, there will be slight adjustment to the base. In the following step, the installation of lighting system will be focused on.



32: The Printed Resin Prototypes of Some Characters



Figure 33: The testing of LED light up system

As for the installation of lighting system, the switch is placed at the back of the plate to allow the connection with the adaptor for the light-up display. All the works in all the aforementioned steps have been described in posts on social media in order to attract feedback from message receivers. The attained feedback is satisfying to a certain degree but the light is too sharp and too bright. Therefore, the material of the base that cover the light bulb is changed to translucent one in white color in order to accumulate the light that hit the object. The opacity level of the material is 60 percent to deficit the light.

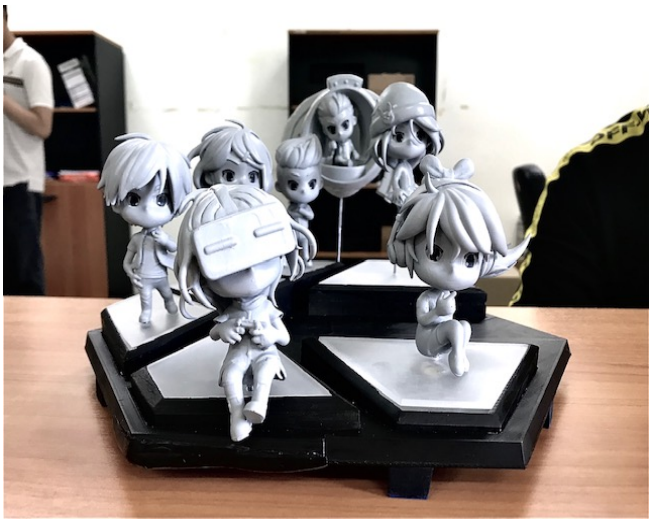


Figure 34: The Printed Prototypes Put Together on Plate in the Shape of the Logo of the Organization

7.5. Product Development

Based on suggestions and feedback from the experts and the audiences, the bases are modified to be in light gray color and the positions are changed, with KIM model pushed back of the plate. In case where a collector does not want to display all the models together, each model can be displayed separately. In addition, the KIM model is modified to show less seams because this model is the biggest and heaviest of all.



Figure 34: A prototype of the assembled product: picture for commercial released in October, 2020, on social media

After the product development leads to the satisfying outcomes, photographs of products are taken and shown on social media in order to get feedback in the Post-Production step. The results show that the satisfaction level is greater than 80 percent, but there are still requests for colored products, which are still in the course of experiment on different materials and colors.

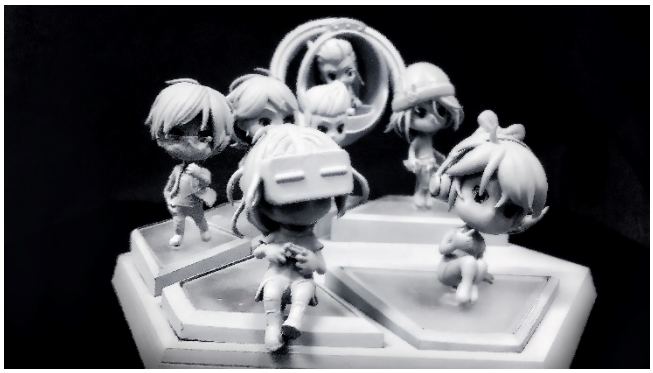


Figure 35: A prototype of the combination product and this picture is for commercial released in October, 2020 on social media.

7.6. Prototype

This final step involves the presentation of prototype products on social media. The presented prototype products are monotonously in the grayish white color. The products are presented in black and white color tone so that audiences can see the forms of all models, the lighting system of the base of each model and effects on the products. This style of presentation is popular among toy producers to present their works in a way to enable audiences to use additional imagination from experiences in primary perception and concept prior to the colored version.



Figure 36: A prototype of the combination product with LED light up test, released in October 2020 on social media.

In certain cases, consumers or audiences may want to see the prototypes in gray color, which can be optional. However, some audiences may p want to see just the colored version in order to consider the features of the products. Nevertheless, this style of prototype presentation will make audience consistently follow-up the updates of the products and give feedback. Afterwards, there is the primary virtual coloring so that audiences can imagine on the final products.

7.7. Product Publicity for Feedback

After the prototypes of the works are created, the next step is to create digital visualized images of the colored works to be presented to the audiences for their feedback. The colored works are presented with the concept of the overall product set in order to show clear identities of the designed mascots, which is the extension of the mascot designing.



Figure 37: A prototype in color version from digital visualization published for feedback

The issue relating to the works in this set is that they are just prototypes. This means the works can be modified in many other steps. There have been critics on material that the used material should be the mixed one that is composed of silicone because nowadays, silicone has been added as a component of materials to be used for creating works in the type of toy or statue in order to create the moods and dimensions of the surface that look more realistic. The extension of the product aims at the same goal, which is to extend the brand of the educational organization that is a choice for the promotion to make the brand sustainable.



Figure 38: A close-up shot of products of colored version

8. Results

As for the creation of elements of brand identity of the educational organization with various digital design-oriented media, namely, digital image, digital communication icon (sticker massage), poster, building decoration and toy (collectible) products, different feedback and opinions concerning different issues have been given by audiences and the experts.

In this project, many new characters (mascots) have been designed to have features different from original ones designed by artists/students. This means the characters in this project are re-versioned to have contemporaneity as other artists/students deem preferable. Theoretically, this style of drawing is called 'Dojinshi' [17] in Japanese language, which means the artists who are fond of certain characters to certain degrees re-create such characters in their own styles. Also, this style is called 'fan-art'.

The aforementioned fan-art phenomenon can be seen as the expression of feedback in the concrete and tangible manner.



Figure 39: Transformation of style of a character (ANI) into teen look can attract more favorable feedback.

Nevertheless, the designing of mascots for educational organizations still faces with certain limitations to designing, on the bases of target groups, organizational culture, spatial culture, and ethicality, which, all together, form the main scopes of the designing. Consequently, the designing cannot go beyond such scopes. This incident reflects the key difference between free comics and education comics because the designing has still had to embrace all the concept keywords. As a result, positive feedback and satisfaction are attained because more than 60% of the audiences (students and staff) deem that the designed mascots (characters) are related to themselves. Even though the uniform is designed to be more contemporary, the used symbolic colors of the organization, i.e., violet and brown, are implication that are hidden in the compositions of the designing and are comprehensible among target audiences.

The unique brand at CMU is a marketing concept for the creation in this project. The creation of the mascots that look contemporary and the extensive designing of other products in order to add values to the brand have led to the uniqueness of the University which results from this creation because the University is the only educational organization that applies platform and media usage in industrial level through the digital processing method in order to create versatile and tangible works. The ultimate result is qualitative and positive feedback that reflects a high level of satisfaction.

To create toy product is a key idea that enhances the strength of the brand of the organization because it reflects the versatile knowledge and skills acquire from learning. The concept of the University is to be a governmental university that is under regulations and conceptual frameworks but has been developed in the digital new normal era that follows the post-modern era. The designed products are tangible and create aestheticism while letting audiences absorb brand identities. This idea leads to a new direction of artwork creation that can be extended to industrial production. Furthermore, this experiment has received positive feedback from audiences in social media, which can be implied that it has market demand in a good level.

In techniques of creating these toys set, the using of digital methodology in ordinary which are follow step 1) 3D creating 2) Printing 3) Mass product creating is the same, but the idea of adding more option into the product such as light up figure or combination of set is the core of concept campaign if comparing with others Nendoroid in the marketplace and it keep develop and added more function into the product. However, the use of Toy

product in brand identity is a good idea of memorizing by touching, playing, and watching especially in institute's educational organization.

9. Conclusions and Discussion

From feedbacks from 3 groups of audiences, namely, CMU students, officer, staff and people in general, it is apparent that there are certain ideas that can be extensively applied to Institute's education brand designing and identity creation, with the main goal being to create tangible uniqueness for the organization through digital platform and digital media, in the form of products that people in the digital contemporary era can comprehend well an touched. Most of organization need a symbol present their identity. Logo is a good example, but it is lack of life and emotional in deep then that the idea of creating the Mascot was born. In this study, the scope idea does not include Digital Animation Film or game or any advance media, but it is the next step in advance once the brand is growing up and popular at certain level in social. Therefore, this study emphasize on to a new way of using Mascot design to enchanced and added value of the University brand in complicated cultural. It need many research of people mind, comments, feedback, population statices and demands, as well as analysis on behaviors of visitors who visited to the University and current students, which contributed to the creation of the new brand identity platform and the improvement of art identity in a future mindset. Key findings from the results analysis and presentation of the artwork product testing are as follows.

First, the designing of mascots for an education institute requires the connection among different dimensions of organizational culture, from the personal level that relies on the in-depth analysis on attitudes and specific personalities, the level of organizational image of the University, the level of provincial culture and the level of the country, under the conceptual framework that has limitations of dressing, ethicality, and purposes and policies of the organization. This is a contemporary study and creation project that creates new identities of the organization, which are reflected through characteristics of mascots (characters) that can be extensively designed as other products, based on the approaches of marketing strategies and advertising, all of which are systematically and sophisticatedly inter-related to one another. In addition, the trend of design is a factor for the creation. The results from the analysis point out that all the symbols designed in this project will be changed again in next ten years in order to sustain the values of the brand.

Secondly, the key factor for character designing is the target groups that are potential customers in term of marketing. Thus, it is significant to focus on students in order to study on trends and art population. From the study, it is discovered that 80 percent of the samples are interested in the styles of comic characters in animations and games rather than general photographs whilst 20 percent of the samples are interested in fashion in portrait style. It is also discovered that the cosplay and Korean pop (K-POP) fashion styles are most influential. Hence, the characters are designed in accordance with the aforementioned demands and trends, in combination with the spirits of digital era in order that they will have the ultimate looks that fit the trends while reflecting the organizational culture simultaneously.

Thirdly, the process of the creation of characters for an educational organization has conceptual framework that has more limitations than organizations of other types, including ethics, morals, and local culture. As for this case, it is Chiang Mai University which is a symbol of the city. Hence, it is necessary to create specific concept from analysis and synthesis. The aspect that must not be overlooked is the combination between organizational culture and local culture that is composed of ideas of people from local communities. Therefore, the outcomes are based on contemporary ideas that are more powerful than conservative perspective alone.

The fourth point of the findings is that toy product is a significant factor that affects imaginations for the audiences, people and provide beyond personal experience as we know that toys can make you imagine and more [18]. With this power, it is a powerful instrument for running promotional and marketing activities to promote brand identities perfectly. In general, people in this modern era, especially teenagers of generation Y people and younger people of following generations prefer to collect such products or any other products that features comic characters. From the analysis on the data of the market of Japan which known as the comic city as a culture, it is discovered that comic characters are used for presenting the tourism, and there have been new characters created until this style of art has become an element of culture since the post-modern era until the Bro-postmodern era when this art style is combined with the elements of the digital era. All the aforementioned cultural features have been absorbed by Thai youths. Consequently, modern artists can create manga-like art works, which can be regularly seen at present. From the analysis, it can be seen that this art style has also incorporated the features of traditional Chinese arts, especially the Chinese brush line drawing that has been developed to be digital line drawing that enables the creation of art works with new styles and contents. Eventually, the art works have been used as the bases for producing tangible mass products that, in turn, increase the popularity of the characters on which they are based. As for this project, the aforementioned product is the designed toy that also serves as a tool for measuring marketing outcomes, and effects in terms of art and culture simultaneously.

The final point from this study is that a perfect brand identity package should consist of tangible and intangible media that are contemporary, which will increase the values of brand and enable the brand to have versatility in terms of social statuses and positions so that the brand can effectively communicate with people in the current era. The findings from the analyses on consumer behaviors, products, trends, and positive ideas that have led to aesthetic perception which involve the awareness of artful beauty from sight and other senses. Hence, the design package for this study has incorporated digital art, and physical art and craft, which can reinforce the images of the organization in recognition and memories of audience in the easily comprehensible and quick manner, in addition to the use of ordinary graphic image or the logo of the organization.

These mascot design set and graphic brand identity project has been used in officially in Chiang Mai university, Thailand and its keep inspired to the people who visited CAMT as a school of new arts, media and technology at all the time.

References

[1] N. Kongdee, S. Prapawong and J. Manissaward, "The Designing of Institute's Educational Mascots," in 2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), 287-291, 2020, doi: 10.1109/ECTIDAMTNCON48261.2020.9090707.

[2] L.L. Adams , "The Mascot Researcher: Identity, Power, and Knowledge in Fieldwork," Journal of Contemporary Ethnography, **28**(4), 331-363, 1999. https://doi.org/10.1080/17430437.990688

[3] S. Mindray, A. Manolica, T. Roman, "Building Brands Identity," Procedia Economics and Finance, **20**, 393-403, 2015, doi: 10.1016/s2212-5671(15)00088-X.

[4] P. Medrano-Bigas, The Forgotten Years of Bibendum. Michelin's American Period in Milltown: Design, Illustration and Advertising by Pioneer Tire Companies (1900-1930), Ph.D Thesis, University of Barcelona, 2015: English translation, 2018.

[5] W. Nunthaporn, Graphic Design for Campaign of No Sweet for Thai Youths. M.A. Thesis, Chulalongkorn University, 2010.

[6] P. Medrano-Bigas, The Forgotten Years of Bibendum. Michelin's American Period in Milltown: Design, Illustration and Advertising by Pioneer Tire Companies (1900-1930), Ph.D Thesis, University of Barcelona, 2015: English translation, 2018.

[7] S. Brown, S.P. McCabe, Brand Mascots: And Other Marketing Animals, Routledge, 2014.

[8] V. Daniel, Does your mascot match your brand's personality?, M.S. Thesis, University of Twente, 2014.

[9] W. Nunthaporn, Graphic Design for Campaign of No Sweet for Thai Youths. M.A. Thesis, Chulalongkorn University, 2010.

[10] P. Prueangphong, Mascot Design for Cultural Events of Chiang Mai, M.F.A. Thesis, Chulalongkorn University, 2014

[11] N. Tongprasert, Meaning Design, Urgent Tag Co.,Ltd., 2016.

[12] W. Chiemvisudhi, "The sign and idenity signification through the officil Olympic game emblem," Veridian E-Journal, Silpakorn University, **10**(1), 2011-2092, 2017. https://doi.org/10.5749/12.2.0124

[13] C. Bacchilega, Postmodern Fairy Tales: Gender and Narrative Strategies, Philadelphia: University of Pennsylvania Press, 1997.

[14] S. Kapidzic, S.C. Herring, "Gender, Communication, and Self-Presentation in Teen Chatrooms Revisited: Have Patterns Changed?," Journal of Computer-Mediated Communication, **17**(1), 53-54, 2011. https://doi.org/10.1111/j.1083-6101.2011.01561.x

[15] S. Seibundo, My First doll: Clothing Patterns Creating in Nendoroid Doll Size, Seibundo Shinkosha, 2019.

[16] Y. Sone, Japanese Robot Culture: Performance, Imagination, and Modernity. Palgrave Macmillan, 2017.

[17] S. Seibundo, My First doll: Clothing Patterns Creating in Nendoroid Doll Size, Seibundo Shinkosha, 2019.

[18] S. Akiko, "Emerging 2.5-dimensional Culture: Character-oriented Cultural Practices and Community of Preferences as a New Fandom in Japan and Beyond," Journal Article Mechademia: Second Arc, **12**(2), 124-125, 2020. https://doi.org/10.5749/mech.12.2.0124